Tail and quantile estimation for real-valued β-mixing spatial data

Tchamiè Tchazino, Sophie Dabo-Niang, Aliou Diop

Tail and quantile estimation for real-valued β-mixing spatial data

Tchamiè Tchazino[1], Sophie Dabo-Niang[2] and Aliou Diop[3]

1 Institut de Mathématiques et de Sciences Physiques IMSP-UAC, Porto-Novo, Benin.

2 Université de Lille, CNRS, UMR 8524-Laboratoire Paul Painlevé, INRIA-MODAL, F-59000 Lille, France.

3 Université Gaston Berger, Laboratoire LERSTAD, Saint-Louis, Senegal.

Abstract

This paper deals with extreme-value index estimation of a heavy-tailed distribution of a spatial dependent process. We are particularly interested in spatial rare events of a $\beta-$mixing process. Given a stationary real-valued multidimensional spatial process $\left\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right\}$, we investigate its heavy-tail index estimation. Asymptotic properties of the corresponding estimator are established under mild *mixing* conditions. The particularity of the tail proposed estimator is based on the spatial nature of the sample and its unbiased and reduced variance properties compared to well known tail index estimators. Extreme quantile estimation is also deduced. A numerical study on synthetic and real datasets is conducted to assess the finite-sample behaviour of the proposed estimators.

keywords: Asymptotic normality; β-mixing; extreme value index; bias correction; spatial dependence; functional estimation.

1 INTRODUCTION

Extreme value theory is around a strong dynamic growth in recent years motivated by the large number of applications in various and varied fields. The literature on statistical inference in extreme value theory develops sophisticated statistical tools for extreme events modeling towards several directions.
Estimation of the extreme-value index and extreme quantiles has been investigated rather extensively in the recent literature. We refer particularly to Bobbia et al. [2019, 2021], Daouia et al. [2020], Bassene [2016], Opitz [2015], Ndao et al. [2014], Goegebeur and Guillou [2013], Resnick et al. [1998], Hsing [1991], Hill [1975] to name a few. Since the seminal tail index estimator of Hill [1975] under the independent hypothesis, few works were dedicated to dependent data. In one hand, Resnick et al. [1998] discussed the consistency of Hill's estimator in the case of some classes of heavy-tailed stationary and dependent processes. On the other hand, almost all the extensions of Hill estimator are developed for time-dependent data (see e.g. Chavez-Demoulin and Guillou [2018] for more details) despite the numerous situations where data are of spatial dependency nature. In fact, in many fields, data are nowadays collected with geographical positions such as oceanography, epidemiology, forestry survey, economy and many others. The study of these kinds of data or any characteristic of such data cannot be done without spatial statistical analyzes by taking into account the geographical positions and possibly spatial dependency. Spatial analysis is a general term to describe a technique that uses the spatial information in order to better handle the dependency of observed geo-localized data in an inference.
For modelling extreme spatial processes, the reader may refer to Bopp et al. [2021], Sharkey and Winter [2019], Opitz [2016], Bassene [2016], Basrak and Tafro [2014], Thibaud et al. [2013], Davison et al. [2012], Blanchet and Davison [2011], Turkman et al. [2010] among others. In particular, for tail index estimation, Basrak and Tafro [2014] considered the extremal behaviour of spatial moving averages and moving maxima on a regular spatial discrete grid while Bassene [2016] extended the previous works to a more general context under strongly conditions. Tail index estimation is important in many extreme value theory problems in particular when estimating extreme quantiles (see Bolancé and Guillen [2021], Velthoen et al. [2021], Chavez-Demoulin and Guillou [2018], Bassene [2016], Goegebeur et al. [2014] among others). Goegebeur et al. [2014] and Bassene [2016] proposed Weissman extreme quantile estimators for $\beta-$mixing non-spatial process (resp. $\alpha-$mixing spatial process) from tail estimation. Velthoen et al. [2021] proposed recently a gradient boosting procedure to estimate a conditional generalized Pareto quantiles while Bolancé and Guillen [2021] introduced a new method to estimate longevity risk based on the kernel estimation of extreme quantiles.
Let $\{X_i, i \in \mathbb{Z}\}$ be a real and measurable process with X_i having the same distribution as a random variable X defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We provide $\mathbb{R}$ with the metric $d(\cdot, \cdot)$. We assume that the condition of the regular variation of the X's tail probability is given by:

$$\forall x > 0 \qquad \mathbb{P}(X > x) = x^{-\frac{1}{\gamma}} L(x), \tag{1}$$

where $\gamma > 0$ is an unknown parameter and $L(\cdot)$ is a slowly varying function.
The unknown γ is the parameter designating the tail index of extreme distribution.

In a series of observations, since we are interested in extreme or unusual values, it is essential to find a method for identifying and collecting such values. One of the most widely used methods for collecting extreme values in extreme statistics can be summarized as follows.

- Let observations $X_1, X_2, \ldots, X_n$; ordered as $X_{(1)} \geq X_{(2)} \geq, \ldots, \geq X_{(n)}$,
- Choose an intermediate sequence k_n such that $1 \leq k_n < n$; $k_n = o(n)$ as $n \to \infty$,
- the well known Hill [1975] estimator of the tail index γ is:

$$\gamma_n = \frac{1}{k_n} \sum_{i=1}^{k_n} \log \left(\frac{X_{(i)}}{X_{(k_n+1)}} \right). \tag{2}$$

This estimator has generated a lot of interest among many researchers in particular when the random variables of interest are independent and identically distributed. In this contribution, we aim to estimate the extreme index in the case of spatial dependent distributed random variables.
In the spatial context, the usual methods of counting the first kth extremes observations will not be applicable. Nevertheless, there are a collection of methods adapted to the spatial framework, the one used in this document transforms the spatial data into a triangular table (see Robinson [2011]) and then applies the previous method to these triangular data.
This spatial ordering has been first used in the extreme value index estimation by Bassene [2016] who expanded the work of Basrak and Tafro [2014] that was limited to bivariate spatial domain.
Indeed Bassene [2016] considers $\left\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right\}$ a real and measurable strongly-mixing spatial process where $X_{\mathbf{i}}$ has the same distribution as X defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Under certain conditions Bassene [2016] built the estimator of the tail index γ:

$$\gamma_{\mathbf{n}} = \frac{1}{\hat{k}_{\mathbf{n}}} \sum_{i=1}^{\hat{k}_{\mathbf{n}}} \log \left(\frac{X_{(i)}}{X_{(\hat{k}_{\mathbf{n}}+1)}} \right) \tag{3}$$

where $X_{(1)} \geq X_{(2)} \geq \cdots \geq X_{(\hat{\mathbf{n}})}$ is the order statistic associated with the observations $\left\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right\}$; and $(\hat{k}_{\mathbf{n}})$ is an intermediate sequence such that $1 \leq \hat{k}_{\mathbf{n}} \leq \hat{\mathbf{n}}$ (see the definition of $\hat{\mathbf{n}}$ below).
In this work, we propose a functional estimation of the extreme tail index for $\beta-$mixing spatial process. Indeed, we wish to extend the estimator proposed by Chavez-Demoulin and Guillou [2018] for $\beta-$mixing time series process to the spatial framework thus generalizing that of Bassene [2016] to a larger class of process and unbiased estimators.
Chavez-Demoulin and Guillou [2018]) considered a β-mixing time series $\{X_i, i \in \mathbb{N}\}$ and built an estimator of the tail index γ:

$$\hat{\gamma}_k(K) = T_K(Q_n) = \int_0^1 \log \left(\frac{Q_n(t)}{Q_n(1)} \right) d(tK(t)), \tag{4}$$

where $Q_n(t) = X_{n-\lfloor kt \rfloor, n}$ $0 < t < n/k$, is the quantile function and K is a function with support in $(0, 1)$.

Under the β-mixing conditions on the time process, the regularity of the function K and other conditions [Chavez-Demoulin and Guillou, 2018], asymptotic properties of this estimator have been proved by the authors. The main objectives of this work is to extend the previous estimator in a context of spatial context. The proposed estimator is not a simple extension, it also gives some improvements of Chavez-Demoulin and Guillou [2018] in terms of reduction of estimation mean squared error.
This paper is organized as follows. Section 2 presents the estimator of the tail index and a bias correction method while Section 3 deals with the extreme quantile estimator. In order to study the finite sample performance of our estimators, we proposed some examples of spatial processes in Section 4. We also propose finite sample properties of the estimates with simulated study and a real data application in Section 5. The proofs of the main results are presented in Section 6.

2 Functional estimation of the extreme value index

2.1 Methodology

We want to build a spatial version of the γ estimator defined in (4).
It is not easy to transcribe asymptotic properties in this spatial framework due to the fact that there is no natural order in $\mathbb{Z}^N$, $N \geq 2$.
Let $\mathbb{Z}^N$; $N \geq 1$ be a Euclidean N-dimensional space of the point indices and $\left\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right\}$ a real and measurable spatial process where $X_{\mathbf{i}}$ has the same distribution as X defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We assume that the condition of the regular variation of the X's tail probability stated in equation (1) is satisfied; i. e. X belongs to the Fréchet attraction domain of parameter $\alpha = 1/\gamma$; where $\gamma > 0$ is an unknown parameter and $L(\cdot)$ is a slowly varying function i.e. verifying:

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1, \ \forall \ t > 0. \tag{5}$$

The unknown γ parameter is the parameter designating the tail index. In the following we are interested in the non-parametric estimation of this parameter $\gamma > 0$.
Let $\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{Z}^N$ be a site and consider the notation of the rectangular domain (see Bassene [2016]):

$$\mathtt{I}_{\mathbf{n}} = \{\mathbf{i} = (i_1, \ldots, i_N); 1 \leq i_k \leq n_k; k = 1, \ldots, N\}; \tag{6}$$

with lexicographical order; in the sense that $\mathbf{i} = (i_1, \ldots, i_N) \leq \mathbf{j} = (j_1, \ldots, j_N) \Leftrightarrow i_1 \leq j_1$ or $i_1 = j_1$ and $i_2 \leq j_2$ or, $\ldots, i_k = j_k$ and $i_N \leq j_N$, $k = 1, \ldots, N-1$. Consider a sample $(X_{\mathbf{i}})_{\mathbf{i}\in \mathtt{I}_{\mathbf{n}}}$ of dependent variables verifying the relationship (1). In this triangular ordering, the observations become $(X_i)_{1\leq i\leq \hat{\mathbf{n}}}$ where each index $i = 1, \ldots, \hat{\mathbf{n}} = n_1 \times n_2 \times \ldots \times n_N$, is identified by a site $\mathbf{i}$ in the region $\mathtt{I}_{\mathbf{n}}$ (see Robinson [2011]).
Let $\mathtt{I}_n = \{i; 1 \leq i \leq \hat{\mathbf{n}}\}$ and for simplicity set $\hat{\mathbf{n}} = n$. In the following all limits are considered for $n \to \infty$.

Let $(k_n)_n$ be a sequence of elements such that $1 \leq k_n \leq n$. As in the case of non-spatial data, we will assume that $(k_n)_n$ is an intermediate sequence of integers such that:

$$k_n \to \infty \quad \text{and} \quad k_n = o(n), n \to \infty. \tag{7}$$

Let $Q_n(t) = X_{n-\lfloor k_n t \rfloor,n}$ $0 < t < n/k_n$, be the quantile function. It is clear that $\forall\ 0 < t \leq 1$, $Q_n(t) \geq Q_n(1)$.
Let $z : [0,1] \longrightarrow \mathbb{R}$ be a measurable function; consider the functional:

$$T_K(z) = \begin{cases} \int_0^1 \log\left(\frac{z(t)}{z(1)}\right) d(tK(t)) \text{ if this integral is defined and finite,} \\ \\ 0 \quad \text{otherwise.} \end{cases}$$

Then, we extend Chavez-Demoulin and Guillou [2018] estimator defined in (4) to the spatial context. We base the inference on the sample fraction k_n. The non-parametric functional estimator of γ for spatial data is as follows:

$$\hat{\gamma}_{k_n}(K) = T_K(Q_n) = \int_0^1 \log\left(\frac{Q_n(t)}{Q_n(1)}\right) d(tK(t)), \tag{8}$$

where K is any function with support in $(0,1)$.

Remark 2.1 *Under the differentiability conditions on the function K, the estimator in Bassene [2016] (spatial version of Hill [1975]) is a special case of this functional estimator for $K \equiv 1$).*

2.2 Asymptotic properties of the tail index estimator

This section gives asymptotic properties of the tail index estimator. We establish asymptotic normality under the following assumptions.
Condition C_K: Let K be a function such as $0 < \int_0^1 K(t)dt < \infty$. Suppose that K is continuously differentiable on $(0,1)$ and that there is $M > 0$ and $\tau \in [0, 1/2)$ such that $|K(t)| \leq Mt^{-\tau}$.
Condition C_M (mixing condition): Let $\sigma_X(T) = \sigma(\{X_{\mathbf{i}}, \mathbf{i} \in T\})$ denote the σ-field generated by $\{X_{\mathbf{i}}, \mathbf{i} \in T\}$ for $T \subset \mathbb{Z}^N$. For any subsets T_1 and T_2 of $\mathbb{Z}^N$, the β-mixing coefficient between $\sigma_X(T_1)$ and $\sigma_X(T_2)$ is defined by

$$\tilde{\beta}(T_1, T_2) = \sup \frac{1}{2} \sum_{j=1}^{J} \sum_{s=1}^{S} |P(A_j \cap B_s) - P(A_j)P(B_s)|, \tag{9}$$

where the supremum is taken over all partitions $\{A_j\}_{j=1}^J \subset \sigma_X(T_1)$ and $\{B_s\}_{s=1}^S \subset \sigma_X(T_2)$ of $\mathbb{Z}^N$. Let $\mathcal{R}(b)$ denote the collection of all finite disjoint unions of cubes in $\mathbb{Z}^N$ with total volume not exceeding b. Then, let

$$\beta(a,b) = \sup\left\{\tilde{\beta}(T_1, T_2); d(T_1, T_2) \geq a; T_1, T_2 \in \mathcal{R}(b)\right\}, \quad a, b > 0, \tag{10}$$

where $d(T_1, T_2) = \inf\{\|x - y\|; x \in T_1, y \in T_2\}$. We assume that there exist a nonincreasing function β_1 with $\lim_{a\to\infty} \beta_1(a) = 0$ and a nondecreasing function g (that may be unbounded) such that

$$\beta(a,b) \leq \beta_1(a)g(b); \quad a, b > 0. \tag{11}$$

Condition C_A: (Second order condition).
There exist a constant ρ <0 and a rate function $\mathcal{A}(\cdot)$ verifying $\mathcal{A}(x) \to 0$ when $x \to \infty$ such that for $t > 1$,

$$\lim_{x\to\infty} \frac{\log(\mathrm{U}(tx)/\mathrm{U}(x)) - \gamma \log t}{\mathcal{A}(x)} = \frac{t^\rho - 1}{\rho}, \tag{12}$$

where U is the quantile function with regular variation defined by $U = (1/(1-F))^{\leftarrow}$ ($\leftarrow$ refers to the continuous generalized inverse on the left).
Condition C_R: (regularity)
There is $\epsilon > 0$, a function $r : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$.
Set $\mathcal{I}(\mathbf{j}) = \{\mathbf{i}; j_k(p+q)+1 \le i_k \le j_k(p+q)+p; k = 1, \ldots, N\}$ a collections of disjoints sites of size p^N ($p^N = o(n)$, $p = p_{\mathbf{n}} \to \infty$, $q = q_{\mathbf{n}} \to \infty$, $q/p \to 0$) and separate at list by q.

(a) $\frac{\beta(q,p^N)}{p^N} n + p^N \frac{\log^2 \hat{k}_{\mathbf{n}}}{\sqrt{\hat{k}_{\mathbf{n}}}} \to 0$;

(b) $\frac{n}{p^N \hat{k}_{\mathbf{n}}} \mathrm{Cov}\left(\sum_{\substack{\mathbf{i}\in\mathcal{I}(\mathbf{j}) \\ \mathbf{j}\in \mathtt{I}_{\mathbf{n}}}} 1_{\{X_{\mathbf{i}} > F^{\leftarrow}(1-\hat{k}_{\mathbf{n}} x/n)\}}, \sum_{\substack{\mathbf{i}\in\mathcal{I}(\mathbf{j}) \\ \mathbf{j}\in \mathtt{I}_{\mathbf{n}}}} 1_{\{X_{\mathbf{i}} > F^{\leftarrow}(1-\hat{k}_{\mathbf{n}} y/n)\}} \right) \to r(x,y),$

$\forall\, 0 \le x, y \le 1+\epsilon$;

(c) there exists a constant C such that $\forall\ 0 \le x < y \le 1+\epsilon$:

$$\frac{n}{p^N \hat{k}_{\mathbf{n}}} E\left[\left(\sum_{\substack{\mathbf{i}\in\mathcal{I}(\mathbf{j}) \\ \mathbf{j}\in \mathtt{I}_{\mathbf{n}}}} 1_{\{F^{\leftarrow}(1-\hat{k}_{\mathbf{n}} y/n) < X_{\mathbf{i}} \le F^{\leftarrow}(1-\hat{k}_{\mathbf{n}} x/n)\}} \right)^4 \right] \le C(y-x).$$

Remark 2.2 *The hypotheses C_A is classical in extreme value theory (see Drees [2000], Drees et al. [2003], De Haan and Ferreira [2016]). The $\beta-$mixing random fields condition C_M is given in Kurisu et al. [2021], Bradley [1993], Dedecker et al. [2007]) while condition C_R is an extension to spatial context of the one-dimension regular condition given in Drees [2000]). Condition C_K is technical to achieve our goals. We need the additional assumptions (like C_2 and C_3 in Drees et al. [2003]) if one wants to obtain the regularity condition C_R. Particularly the condition $C_R-(a)$ is verified if (11) hold.*

Theorem 2.1 *Let $\left\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right\}$ be a stationary spatial process with continuous distribution function and verifying conditions C_M, C_A and C_R. Set (k_n) an intermediate sequence such that $\sqrt{k_n}\mathcal{A}(b(n/k_n)) \longrightarrow \lambda < \infty, \quad n \to \infty$. We have:*

$$\sqrt{k_n}\left[\hat{\gamma}_{k_n}(K) - \gamma - \mathcal{A}(b(n/k_n)) \int_0^1 t^{-\rho} K(t) dt\right] \xrightarrow{d} \gamma \int_0^1 \left[t^{-1} W(t) - W(1)\right] d(tK(t)), \tag{13}$$

where $(W(t))_{t\in[0,1]}$ is a centered Brownian motion, r is the covariance defined in C_R and

$$b(t) = F^{\leftarrow}(1-t^{-1}) \qquad t > 1, \tag{14}$$

$$F^{\leftarrow}(y) = \inf\{x, F(x) \geq y\} \qquad 0 \leq y \leq 1,$$

F being a distribution function and therefore subject to regular variation :

$$\bar{F}(b(t)) \sim t^{-1}, \qquad \bar{F} := 1 - F. \tag{15}$$

Remark 2.3 *The result of this theorem is an extension of Theorem 1 of Chavez-Demoulin and Guillou [2018]. The particularity lies in the multivariate indexation ($\mathbf{i} \in \mathbb{Z}^N$, $N \geq 1$) of the considered process (spatial in our context).*
If we assume that the observations $(X_{\mathbf{i}}, \mathbf{i} \in \mathcal{I}_{\mathbf{n}})$ are i.i.d. and $K(t) = 1$ for all $t \in (0,1)$, we find the asymptotic normality of the Hill [1975] estimator (see De Haan and Resnick [1998], Resnick and Stărică [1995], Mason [1982]...).

The following corollary specifies the asymptotic bias and variance.

Corollary 2.1 *Under the condition of Theorem 2.1, we have:*

$$\sqrt{k_n}\left(\hat{\gamma}_{k_n}(K) - \gamma\right) \xrightarrow{d} \mathcal{N}\left(\lambda \mathcal{AB}(K),\ \mathcal{AV}(K)\right). \tag{16}$$

where $\mathcal{AB}(K) = \displaystyle\int_0^1 t^{-\rho} K(t) dt$

and $\mathcal{AV}(K) = \gamma^2 \displaystyle\int_0^1 \int_0^1 \left[\frac{r(t,s)}{ts} - \frac{r(t,1)}{t} - \frac{r(1,s)}{s} + r(1,1)\right] d(tK(t)) d(sK(s))$.

This tail estimator being asymptotically biased with bias $\left(\lambda \displaystyle\int_0^1 t^{-\rho} K(t) dt\right)$. The aim of the next section if to reduce it.

2.3 Bias correction

2.3.1 Existing works and some remarks

Before stating the bias reduction contribution, let us recall the existing works on this direction.

- Firstly, let us analyse the estimator of extreme value index proposed by Chavez-Demoulin and Guillou [2018] (in the case where $N = 1$) according to ρ denoted by $\hat{\gamma}_{k_n}(K_{\Delta^*})$ with the function $K_{\Delta^*}(t) = \left(\frac{1-\rho}{\rho}\right)^2 - \frac{(1-\rho)(1-2\rho)}{\rho^2} t^{-\rho}$. This gives an unbiased estimator and a variance $\mathcal{AV}(\hat{\gamma}_{k_n}(K_{\Delta^*})) = \gamma^2 \left(\frac{1-\rho}{\rho}\right)^2$. One can remark that this variance increases and is closer to γ^2 for large values of $|\rho|$ and very high for low values of $|\rho|$. According to these remarks we conclude that this estimator is useful for large value of $|\rho|$.

- Secondly, how does the associated quantile estimator behave for large values of $|\rho|$? The quantile estimator proposed by Chavez-Demoulin and Guillou [2018] is given by;

$$\begin{aligned} \hat{x}_{p,\xi} &= X_{n-k,n} \left(\frac{k}{pn}\right)^{\hat{\gamma}_{k_n}(K_{\delta^*})} \exp\left\{ -\frac{(1-\rho)(1-2\rho)}{\rho^2} \right. \\ &\quad \left. \times \left[\hat{\gamma}_{k_n}(K_1) - \hat{\gamma}_{k_n}(K_{2,\rho})\right] \frac{\left(\frac{k}{pn}\right)^{\rho} - 1}{\rho} \right\}, \end{aligned}$$

which gives for large values of $|\rho|$

$$\hat{x}_{p,\xi} = X_{n-k,n}\left(\frac{k}{pn}\right)^{\hat{\gamma}_{k_n}(K_{\hat{\delta}*})},$$

since $\frac{\left(\frac{k}{pn}\right)^{\rho}-1}{\rho} \to 0$ when $\rho \to -\infty$. Then one get directly a Weismann-type estimator without the correction term.

2.3.2 Our proposition

In this section, we propose a bias reduction method, useful for low values of $|\rho|$ and for the corresponding quantile estimator. Knowing that the bias of the considered class of estimators depends on the function K, we will choose an optimal one, i.e. making both the bias almost zero and minimizing the variance. The particularity of our approach compared to those of Chavez-Demoulin and Guillou [2018] and that of Goegebeur and Guillou [2013] lies in the consideration of a more wide class of functions K.
For that purpose, consider two functions K_1 and K_2 verifying C_K and set:

$$K_{(\alpha,\beta)}(t) = \alpha K_1(t) + \beta K_2(t);\ \ \alpha,\ \ \beta\ \ \in\ \ \mathbb{R}^*, \tag{17}$$

such that $\frac{1}{\alpha} + \frac{1}{\beta} = -1$ (this condition imposed on α and β is for technical reason) and $\int_0^1 K_{(\alpha,\beta)}(t)dt > 0$. Let us now evaluate the bias of the estimator $\hat{\gamma}_{k_n}(K_{(\alpha,\beta)})$. We have

$$\begin{aligned}\frac{\lambda}{\sqrt{k_n}}\mathcal{AB}\left(K_{(\alpha,\beta)}\right) &= \frac{\lambda}{\sqrt{k_n}}\int_0^1 t^{-\rho}K_{(\alpha,\beta)}(t)dt \\ &= \frac{\lambda}{\sqrt{k_n}}\left(\alpha\mathcal{AB}(K_1) + \beta\mathcal{AB}(K_2)\right).\end{aligned}$$

Since it is question of bias reduction, let us find the values of α and β for which the bias is close to 0. We then obtain the system of equations:

$$\begin{cases} \alpha\mathcal{AB}(K_1) + \beta\mathcal{AB}(K_2) = 0, \\ \\ \qquad\qquad \frac{1}{\alpha} + \frac{1}{\beta} = -1. \end{cases}$$

The solution of this system is given by $S = (\alpha^*, \beta^*) \in \mathbb{R}^{*2}$ where $\alpha^* = \frac{\mathcal{AB}(K_2)-\mathcal{AB}(K_1)}{\mathcal{AB}(K_1)}$ and $\beta^* = \frac{\mathcal{AB}(K_1)-\mathcal{AB}(K_2)}{\mathcal{AB}(K_2)}$. The resulting function K_S leads to an estimator of γ of asymptotically zero bias as stated by the following result.

Corollary 2.2 *Under the hypotheses of Theorem* 2.1*, and assuming that K_1 and K_2 satisfy the C_K condition, we have:*

$$\sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_S) - \gamma\right) \xrightarrow{d} \mathcal{N}\left(0,\ \mathcal{AV}(K_S)\right). \tag{18}$$

Let $\mathcal{C}_\gamma = \{\hat{\gamma}_{k_n}(K_S), K_1\ \ K_2$ verifying $C_K\}$ be the class of asymptotically unbiased estimators of γ. It remains to construct an estimator with minimum asymptotic variance.

In the case of i.i.d. random variables, Goegebeur and Guillou [2013] (Theorem 2 and Corollary 4) established that the couple of functions $(K_{1,\rho}, K_{2,\rho})$ verifying the condition C_K for which the variance is minimal is given by: $(1,\ (1-\rho)t^{-\rho})$. Although we have not yet established the optimal variance in our context, let us take the same functions $(1,\ (1-\rho)t^{-\rho})$ as in the i.i.d. case. Thus, the function K_{S^*} of "optimal" variance from the $(1,\ (1-\rho)t^{-\rho})$ of the form (17) is given by

$$K_{S^*}(t) = \frac{\rho^2}{1-2\rho} - \frac{\rho^2}{1-\rho}t^{-\rho}. \tag{19}$$

It is easy to see that $\hat{\gamma}_{k_n}(K_{S^*}) \ \in \ \mathcal{C}_\gamma$ is of minimal variance.

Remark 2.4 *Under the assumption of i.i.d. observations, the tail estimator with the optimal choice $\hat{\gamma}_{k_n}(K_{S^*})$ is better than that of Goegebeur and Guillou [2013] (Corollary 4) for low values of $|\rho|$ (see Figure 6.1).*

Indeed, the obtained variance is $\mathcal{AV}(K_{S^*}) = \gamma^2 \frac{\rho^6}{(1-\rho)^2(1-2\rho)^2} \underset{\rho\to 0}{\to} 0$ compare to $\mathcal{AV}(K_{\Delta^*}) = \gamma^2 \left(\frac{1-\rho}{\rho}\right)^2 \underset{\rho\to 0}{\to} \infty$ of Goegebeur and Guillou [2013].

By closely analyzing Figure 6.1, we can exhibit a better estimator $\hat{\gamma}_{k_n}(K_{S\Delta})$ which is a compromise between that of Goegebeur and Guillou [2013] and the one proposed above. Let $K_{S\Delta}$ be defined as

$$\begin{cases} K_{S\Delta} = K_{\Delta^*} \ \text{ if } \ \rho \le \bar{\rho}, \\ \\ K_{S\Delta} = K_{S^*} \ \text{ if } \ \rho \ge \bar{\rho}, \end{cases} \tag{20}$$

where $\bar{\rho}$ is the intersection point of the two solutions variance of the equation:

$$\mathcal{AV}(K_{S^*}) = \mathcal{AV}(K_{\Delta^*}).$$

The dichotomy resolution allows us to find an approximate value of $\bar{\rho} = -3.67$.
In the following we focus on the study of $\hat{\gamma}_{k_n}(K_{S^*})$ instead of $\hat{\gamma}_{k_n}(K_{S\Delta})$ since $\hat{\gamma}_{k_n}(K_{\Delta^*})$ is widely studied by our predecessors (e.g Goegebeur and Guillou [2013]).
Note that the function $K_{S^*}(t)$ is function of the unknown parameter ρ. So, to overcome this problem let $\tilde{\rho}$ be a value from ρ in (19) or a point estimator (or canonical) of ρ. We have

$$K_{\tilde{S}^*}(t) = \frac{\tilde{\rho}^2}{1-2\tilde{\rho}} - \frac{\tilde{\rho}^2}{1-\tilde{\rho}}t^{-\tilde{\rho}}, \tag{21}$$

The following corollary is then stated

Corollary 2.3 *Under the assumptions of the Theorem* 2.1*, and assuming that K_1 and K_2 satisfy the condition C_K, we have:*

$$\sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_{\tilde{S}^*}) - \gamma\right) \xrightarrow{d} \mathcal{N}\left(\lambda \frac{\tilde{\rho}^3(\tilde{\rho}-\rho)}{(1-\tilde{\rho})(1-\rho)(1-2\tilde{\rho})(1-\rho-\tilde{\rho})},\ \mathcal{AV}(K_{\tilde{S}^*})\right). \tag{22}$$

Remark 2.5 *The two corollaries of this section coincide when $\tilde{\rho} = \rho$.*
This Corollary 2.3 is similar to Corollary 3 *in Chavez-Demoulin and Guillou [2018], with a particular difference in the behaviour of the asymptotic bias. Indeed, the bias in our corollary decreases to* 0 *while the one in Corollary* 3 *of Chavez-Demoulin and Guillou [2018] explodes for $|\tilde{\rho}|$ rather small. Thus we propose to choose the previously constructed $K_{S\Delta}$ which makes a compromise.*

Certainly, this estimator is biased but it provides a particular interest for the control (reduction) of the bias because a judicious choice of $\tilde{\rho}$ (since ρ is a parameter that controls the convergence speed) would allow us to considerably reduce the bias unlike Corollary 2.1. On the other hand, since it is possible to cancel this bias (for $\tilde{\rho} = \rho$) if we replace ρ by one estimator $\hat{\rho}_{k_n,\rho}$ which is a function of an intermediate sequence $(k_{n\rho})_{n\in\mathbb{Z}^N}$, we obtain the following theorem:

Theorem 2.2 *Let $\left\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right\}$ be a stationary spatial process with continuous distribution function and verifying conditions C_M, C_R and C_A. Let $\hat{\rho}_{k_n,\rho}$ be an estimator of ρ consisting of probability, depending on an intermediate sequence $(k_{n,\rho})_{n\in\mathbb{N}}$. Let (k_n) be an intermediate sequence such that $\sqrt{k_n}\mathcal{A}(b(n/k_n)) \longrightarrow \lambda < \infty$, $n \to \infty$. We have:*

$$\sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_{\hat{S}^*}) - \gamma\right) \xrightarrow{d} \mathcal{N}\left(0,\ \mathcal{AV}(K_{S^*})\right), \tag{23}$$

*where $K_{\hat{S}^*}$ is in the form (*19*) by replacing ρ by $\hat{\rho}_{k_n,\rho}$.*

Note that $K_{\hat{S}^*}$ is a function of $\hat{\rho}_{k_n,\rho}$ and n, then we need an additional condition on the term $\sqrt{k_{n,\rho}}\mathcal{A}(b(n/k_{n,\rho}))$. Although the function $K_{\hat{S}^*}$ is of the form (19), it cannot be written in the form (17) (where K_1 and K_2 are functions of $\hat{\rho}_{k_n,\rho}$, α and β are functions of ρ functions). So it is necessary to estimate ρ.
Gomes et al. [2002] have proposed a possible set of values for $\hat{\rho}_{k_n,\rho}$ also used in De Haan and Ferreira [2016] and in Chavez-Demoulin and Guillou [2018]:

$$\hat{\rho}_{k_n,\rho} := \frac{-4 + 6S^{(2)}_{k_n} + \sqrt{3S^{(2)}_{k_n} - 2}}{4S^{(2)}_{k_n} - 3} \ \text{ with } \ S^{(2)}_{k_n} \in \left]\frac{2}{3}, \frac{3}{4}\right[, \tag{24}$$

where

$$S^{(2)}_{k_n} := \frac{3}{4}\frac{\left[M^{(4)}_{k_n} - 24\left(M^{(1)}_{k_n}\right)^4\right]\left[M^{(2)}_{k_n} - 2\left(M^{(1)}_{k_n}\right)^2\right]}{\left[M^{(3)}_{k_n} - 6\left(M^{(1)}_{k_n}\right)^3\right]^2},$$

with

$$M^{(\alpha)}_{k_n} := \frac{1}{k_n}\sum_{i=1}^{k_n}\left(\log X_{n-i+1,n} - \log X_{n-k_n,n}\right)^\alpha, \quad \alpha \in \mathbb{N}. \tag{25}$$

In this ρ domain we have the following corollary.

Corollary 2.4 *Let $\left\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right\}$ be a stationary spatial process with continuous distribution function and verifying the conditions C_M, C_R and C_A. Let us also assume*

that the condition C_K is satisfied. Let $\hat{\rho}_{k_{n,\rho}}$ be an estimator of ρ where the intermediate sequence $(k_{n,\rho})_{n\in\mathbb{Z}^N}$ is such that $\sqrt{k_{n,\rho}}\mathcal{A}(b(n/k_{n,\rho})) \longrightarrow \infty$. Let (k_n) be another intermediate sequence such that $\sqrt{k_n}\mathcal{A}(b(n/k_n)) \longrightarrow \lambda < \infty, \quad n \to \infty$. We have:

$$\sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_{\hat{S}^*}) - \gamma\right) \xrightarrow{d} \mathcal{N}\left(0,\ \mathcal{AV}(K_{S^*})\right), \tag{26}$$

where $K_{\hat{S}^*}$ is in the form (19), replacing ρ by $\hat{\rho}_{k_{n,\rho}}$.

This corollary is similar to Corollary 4 in Chavez-Demoulin and Guillou [2018]; the difference lies in the spatial nature of the process studied here. Here the condition on $(k_{n,\rho})_{n\in\mathbb{N}}$ is such that $\sqrt{k_{n,\rho}}\mathcal{A}(b(n/k_{n,\rho})) \longrightarrow \infty$ is necessary to ensure the probability consistency of the estimator $\hat{\rho}_{k_{n,\rho}}$ of ρ.
In practice, it is usual to consider K as a kernel but this does not prevent from using the proposed function $K_{S\Delta}$. Note that the originality of the proposed approach, does not only lies in the spatial nature of the studied dependent process, but also on a wide choice of functions K and assumption C_K, reducing the asymptotic bias and variance, different from those of our predecessors Chavez-Demoulin and Guillou [2018], De Haan and Ferreira [2016], Goegebeur and Guillou [2013] where the function K is a kernel.
Estimating γ is a (necessary) step in estimating extreme quantile, so in the next section we apply the tail index estimate to quantile inference.

3 Estimation of extreme distribution quantile

Since one of the main purpose of statisticians is to predict future events, the tail index estimation of the previous section is useful to reach this aim. Thus in this section, we are interested in estimating the extreme quantile of order p:

$$x_p = U(1/p), \quad p \to 0. \tag{27}$$

From condition C_A, one obtain

$$\frac{U(ux)}{U(u)} = x^{\gamma}\exp\left\{\mathcal{A}(u)\frac{x^{\rho}-1}{\rho}\right\} + \mathrm{o}(1).$$

By setting $ux = 1/p$ and $u = Y_{n-\lfloor tk_n\rfloor,n}$ where $Y_i \sim$ Pareto (standard Pareto distribution) and as $X_{n-\lfloor tk_n\rfloor,n} = U(Y_{n-\lfloor tk_n\rfloor,n})$, an approximation of x_p is

$$\begin{aligned} x_p &\simeq X_{n-\lfloor tk_n\rfloor,n}\left(\frac{1}{pY_{n-\lfloor tk_n\rfloor,n}}\right)^{\gamma}\exp\left\{\mathcal{A}(Y_{n-\lfloor tk_n\rfloor,n})\frac{\left(\frac{1}{pY_{n-\lfloor tk_n\rfloor,n}}\right)^{\rho}-1}{\rho}\right\} \\ &\simeq X_{n-\lfloor tk_n\rfloor,n}\left(\frac{1}{pb(n/k_n)}\right)^{\gamma}\exp\left\{\mathcal{A}\left(b(n/k_n)\right)\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\rho}-1}{\rho}\right\}, \end{aligned} \tag{28}$$

where the last step follows from replacing $Y_{n-\lfloor tk_n\rfloor,n}$ by its expected value $b(n/k_n)$. This approximation is only accessible if one replace γ and $\mathcal{A}\left(b(n/k_n)\right)$ by their estimators. It should be noted that the term $\mathcal{A}\left(b(n/k_n)\right)$ is seen as the moderator or corrector of

the quantile estimator since we find the Weissman-type estimator if n is large enough that is, if $\mathcal{A}\left(b(n/k_n)\right) \to 0$ (see Weissman [1978]).
Chavez-Demoulin and Guillou [2018] proposed an estimator for $\mathcal{A}\left(b(n/k_n)\right)$ that we adapt to our context. These authors estimated $\mathcal{A}\left(b(n/k_n)\right)$ by :

$$-\frac{(1-\xi)(1-2\xi)}{\xi^2}[\hat{\gamma}_{k_n}(K_1) - \hat{\gamma}_{k_n}(K_{2,\xi})],$$

where ξ is a negative real number or an estimator of ρ. So by replacing γ and $\mathcal{A}\left(b(n/k_n)\right)$ with their estimators in the relationship (28), we get the estimator:

$$\begin{aligned}\hat{x}_{p,\xi} &= X_{n-\lfloor tk_n \rfloor,n}\left(\frac{1}{pb(n/k_n)}\right)^{\hat{\gamma}_{k_n}(K_{\hat{S}^*})} \exp\left\{-\frac{(1-\xi)(1-2\xi)}{\xi^2}\right. \\ &\left.\times[\hat{\gamma}_{k_n}(K_1) - \hat{\gamma}_{k_n}(K_{2,\xi})]\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\xi}-1}{\xi}\right\}. \end{aligned} \tag{29}$$

The convergence in distribution of this quantile estimator have been established in the literature under certain appropriate assumptions. We extend this asymptotic result to the spatial context.

Theorem 3.1 *Let $\left\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right\}$ be a stationary spatial process with continuous distribution function and verifying the conditions C_M, C_R and C_A. Let $\hat{\rho}_{k_n,\rho}$ be a consistent estimator of ρ, depending on an intermediate sequence $(k_n)_{n\in\mathbb{N}}$. Let (k_n) be an intermediate sequence such that $\sqrt{k_n}\mathcal{A}(b(n/k_n)) \longrightarrow \lambda < \infty, \quad n \to \infty$ and suppose that $p = p_n$ such that $\frac{1}{pb(n/k_n)} \to \infty$, $\frac{log\left(\frac{1}{pb(n/k_n)}\right)}{\sqrt{k_n}} \to 0$ and $n^{-a}logp \to 0$ for all $a > 0$, then we have:*

$$\frac{\sqrt{k_n}}{log\frac{1}{pb(n/k_n)}}\left(\frac{\hat{x}_{p,\xi}}{x_p} - 1\right) \xrightarrow{d} \mathcal{N}\left(0,\ \mathcal{AV}(K_{S^*})\right), \tag{30}$$

where ξ is a negative parameter $\tilde{\rho}$ or an consistent estimator in probability $\hat{\rho}$, such that $|\hat{\rho} - \rho| = O_{\mathbb{P}}(n^{-\epsilon})$ for $\epsilon > 0$

4 Some examples of heavy tailed spatial processes

In this section, we give some examples of spatial processes satisfying heavy tailed assumptions.

4.1 Brown-Resnick process

Let $\left\{Z_s; s \in \mathbb{R}^N\right\}$ be a stationary Spatial Gaussian process with mean 0 and unit variance. The following hypothesis is assumed on φ (see Davis et al. [2013]).

Hypothesis H_φ: there exist a sequence $(S_n)_{n\in\mathbb{N}}$ of constants; $S_n \to 0$ as $n \to \infty$ and a nonnegative function δ such that

$$(1 - \varphi(S_n h)) \log n \to \delta(h), \quad \text{as } n \to \infty.$$

The following proposition gives the construction of the Brown-Resnick process. It is an entension to spatial case of Proposition 2.2 in Davis et al. [2013].

Proposition 4.1 *(Davis et al. [2013])*
Let $\left\{Z_s^j; s \in \mathbb{R}^N\right\}$, $j = 1, \ldots, n$, *be* n *independent copies of* Z_s. *Let* $\{\xi_j, j \in \mathbb{N}\}$ *denote points of a random Poisson measure on* $(0, \infty]$ *with with intensity* $\xi^{-2}d\xi$. *Suppose* H_φ *hold, then the random field* $\left\{\eta_s^n; s \in \mathbb{R}^N\right\}$ *defined for any* $n \in \mathbb{N}^*$ *by (where* Φ *denotes the standard normal distribution function)*

$$\eta_s^n = \bigvee_{j=1}^{n} - \frac{1}{log\left(\Phi(Z^j_{S_n s})\right)}, \quad s \in \mathbb{R}^N, \tag{31}$$

converges weakly in the space of continuous functions in $\mathbb{R}^N$ *towards a stationary Brown-Resnick process*

$$\eta_s = \bigvee_{j=1}^{\infty} \xi_j \exp\left(W_s^j - \delta(s)\right), \quad s \in \mathbb{R}^N, \tag{32}$$

where δ *is defined in* H_φ *and* $(W_s^j)_{s\in\mathbb{R}^d}$, $j \in \mathbb{N}$, *are independent copies of a stationary Brownian process* $(W_s)_{s\in\mathbb{R}^d}$ *such as* $W_0 = 0$, $\mathbb{E}(W_s) = 0$, $cov(W_{s_1}, W_{s_2}) = \delta(s_1) + \delta(s_2) - \delta(s_1 - s_2)$, *($s_1 - s_2$ is defined component wise).*

The bivariate distribution function of $(\eta_s)_{s\in\mathbb{R}^d}$ can be expressed in closed form (based on Davis et al. [2013]) for $x_1, x_2 > 0$ as

$$F(x_1, x_2) = \exp\left\{-\frac{1}{x_1}\Phi\left(\frac{\log\frac{x_2}{x_1}}{2\sqrt{\delta(h)}} + \sqrt{\delta(h)}\right) - \frac{1}{x_2}\Phi\left(\frac{\log\frac{x_1}{x_2}}{2\sqrt{\delta(h)}} + \sqrt{\delta(h)}\right)\right\},$$

and the tail dependence coefficient is given by:

$$\chi(h) = \lim_{u\uparrow 1} \mathbb{P}\left(\eta_{s_1} \geq F^{\leftarrow}_{\eta_{s_1}}(u) \mid \eta_{s_2} \geq F^{\leftarrow}_{\eta_{s_2}}(u)\right), \quad h = s_1 - s_2.$$

4.2 Auto-regressive (AR) spatial process

Let $(\xi_i)_{i\in\mathbb{Z}}$ be a family of an i.i.d. variables with Lebesgue density function f_ϵ (resp. distribution function F_ϵ) continuous and verifying L_1-Lipschitz condition

$$\int |f_\epsilon(\epsilon + u) - f_\epsilon(\epsilon)| d\epsilon = \mathrm{O}(u), \;\; u \searrow 0.$$

Suppose that

$$1 - F_\epsilon(\epsilon) \sim q\epsilon^{-1/\gamma} l(\epsilon) \text{ and } F_\epsilon(-\epsilon) \sim (1-q)\epsilon^{-1/\gamma} l(\epsilon), \;\; \epsilon \to \infty,$$

where $l(\cdot)$ is a slowly varying function and $q \in]0, 1[$, $\gamma > 0$. Let us consider the stationary solution of the equation of an AR(1) time-process $\{X_i, i \in \mathbb{R}\}$

$$X_i = \theta X_{i-1} + \xi_i, \;\; \theta \in]0, 1[.$$

The regularity condition is checked with

$$r(x, y) = \min(x, y) + \sum_{m=1}^{\infty} \left(r_m(x, y) + r_m(y, x)\right),$$

where $r_m(x,y) = \min(x, y\theta^{\frac{m}{\gamma}})$ with γ the tail index (see Chavez-Demoulin and Guillou [2018]).
Starting from this process, we build the spatial process $\left\{Y_{(i,j)}; i,j \in \mathbb{Z}\right\}$ defined as follows

$$\forall\ i,j \in \mathbb{Z}, \quad Y_{(i,j)} = X_{\min(i,j)}. \tag{33}$$

This process is stationary, with tail index γ.

4.3 Log-Laplace spatial process

Let $(Z = \{Z_s, s \in S \subset \mathbb{R}^2\})$ be a centered weakly Gaussian spatial process with covariance function $(\Sigma(h) = \operatorname{cov}(Z_{s+h}, Z_s), (s+h) \in S)$. Let Y be a non negative random variable such that Y^2 has an exponential distribution with scale $\lambda = 2$ and independent of Z. In the literature, the distribution of the random variable Y is called Rayleigh's law of parameter 1. Its probability distribution function is given by $f_Y(y) = e^{-0.5y^2}$. The random field

$$X = \{X_s, s \in S\} = \{YZ_s, s \in S\},$$

is called Laplace spatial process (see for example Opitz [2016, 2015]). If Z is a standard Gaussian process, then X is said to be Laplace Standard process. Its covariance function is equal to $2\Sigma(h)$ and that of the conditional distribution of X given $Y = y$ is Gaussian with covariance $y^2\Sigma(h)$.

If X is a standard Laplace process then $U = \{U_s = e^{X_s},\ s \in S \subset \mathbb{R}^2\}$, is a process whose marginals are log Laplace of tail index $\gamma = 1$ (see Opitz [2015]) and density

$$f_U(t) = 1/(2t^2),\ \text{ if } t \geq 1;\ f_U(t) = 1/2;\ \text{ if } 0 \leq t < 1;\ f_U(t) = 0;\ \text{ otherwise.}$$

5 Finite sample properties

In this section, we illustrate the finite-sample performance of the proposed estimator using simulated and real datasets.

5.1 Simulation study

Throughout this section, for illustration purpose, we simulate a Log-Laplace process as above such that the tail index is equal to $\gamma = 1$ (see Figure 6.2).

We aim to estimate the index $\gamma = 1$ of the simulated log-Laplace process. For this purpose, we generate $N = 100$ samples of log-Laplace process and we consider a range of the highest values k $(k = 1, \cdots, 1000)$ to be considered in order to display the best estimator, i.e. the one of minimum error. We have computed three families of estimators; that of Hill given by (2), Chavez-Demoulin and Guillou [2018] given by (4) ignoring the spatial nature of the data and the one proposed in this article given by (8). This will allows to appreciate each estimators but also validate our contribution (8).
For this simulation study, we have used the empirical form of the estimator (20) that is:

$$\hat{\gamma}_{k_n}(K) = \frac{1}{k_n}\sum_{i=1}^{k_n} \log\left(\frac{X_{(i)}}{X_{(k_n+1)}}\right) \times \left(K\left(\frac{i}{k_n}\right) + \frac{i}{k_n} \times K'\left(\frac{i}{k_n}\right)\right), \tag{34}$$

where K' is the first derivative of the function K.
We essentially verified the performance of our estimator (8) and that of Chavez-Demoulin and Guillou [2018] given by (4) according to some appropriate choice of the values of ρ (see Section 2.3). The results (see Table 6.1) of the simulations allows us to conclude that the proposed estimator (8) is more useful than Chavez-Demoulin and Guillou [2018] (4) for low values of ρ. Thus, based on these results we recommend the use of the estimator (8) in which the function $K(\cdot)$ is given by (20) realizes the compromise between the two estimators (8) and (4).

Note that each estimate in Table 6.1 is the average of the best (in terms of MSE) estimates over the 100 samples. The choice of the optimal sample fraction K_{op} is done using a data-driven method that is the average of the values k resulting from the best estimators (in terms of MSE) over the 100 samples.
We also investigate the performance of the associate extreme quantile estimate q_n such that $\mathbb{P}(X_i \geq q_n) = p$, when p is very small. In our simulation, we present the predicted quantile for $p = 10^{-3}$, where the theoretical quantile $q_n = 80.995$ (true value given by the function *qllaplace* under the package "*LaplacesDemon*"). The results are shown in Table 6.2.
One can notice from these results that sometimes the predicted quantiles beyond the true value, which means that the risk of occurrence of other extremes values is not taken into account. In addition, regarding our estimator (last column), it overperforms compared to Hill [1975]'s estimator that is independent of ρ, then a judicious choice of ρ allows to take into account the occurrences of risky values so that to get better predictions.
As mentioned in the previous sections, our estimator realizes a trade-off between $\gamma(K_\Delta)$ ($\rho \leq -3.67$) and $\gamma(K_S)$ ($\rho \geq -3.67$).

5.2 Real data application

In this section, we propose to illustrate the behaviour of the proposed methodology on rainfall data from 1559 stations in the West Africa Region on first September 2019 (available on NASA website). The mesh was made over the West Africa Region by varying the longitude in the interval $[-17, 16]$ and the latitude in $[4.5, 20]$.
The spatial aspect is materialized by longitude and latitude. The observed process has a minimum value of 0 and a maximum value of 65.94 with a mean of 6.94 and a median of 4.94. It can be seen that the mean is very close to the minimum observation and very distant from the maximum observation. The median shows the extremal aspect of the observations quite large. Indeed 50% of the observations are concentrated in $[0, 4.94]$ and the rest scattered in $[4.94, 65.94]$; this aspect is visible on Figure 6.4. We can thus think of the existence of a heavy tail on the right, hence be interested of an extreme data study. Figure 6.4 shows a grouping by similarity (color gradient) of the data on the geographical level. This makes us think of spatialized data or spatial dependence of data. Figure 6.4 gives the spatial representation of the data. This figure illustrates two main aspects of the data: the spatial one materialized by the longitude and latitude and the extreme behavior visualized by the color gradient which shows very few large observations (light blue, yellow and red). It also illustrates the spatial

dependence marked by the grouping of data by similarity (size of observations) and according to geographical positions; we can notice the cluster of colors (materializing the value of the observations): dark blue (the most frequent observations), the highest observations in red and dark red (rare) surrounded by the observations more or less high (see Figure 6.5).

We carry out, in Figure 6.6, visual checks of whether the heavy-tailed assumption makes sense for this sample of data (composed of observations denoted Y_i). The boxplot and histogram of the Y_i both give descriptive evidence of heavy right tail. To further confirm that the heavy-tailed framework is appropriate, we drew a quantile-quantile plot of the weighted log-spacings within the top of the data against the quantiles of the unit exponential distribution. Formally, let $Y_{1,n} \leq Y_{2,n} \leq \cdots \leq Y_{n,n}$ denote the order statistics of the sample $(Y_1, \cdots, Y_n)$. Let $Z_{i,n} = i\log(Y_{n-i+1,n}/Y_{n-i,n})$, $1 \leq i \leq n-1$, denote the weighted log-spacings computed from the consecutive top order statistics. It is known that, if a distribution is heavy-tailed with tail index γ then, for low i, the $Z_{i,n}$ are approximately independent copies of an exponential random variable with mean γ (see e.g. Beirlant et al. [2006]). Figure 6.6, bottom row, right panel gives a quantile-quantile plot of the $Z_{i,n}$ for $1 \leq i \leq \lfloor n/5 \rfloor$ versus the exponential distribution. The relationship in this quantile-quantile plot is approximately linear, which constitutes further evidence that the heavy tail assumption makes sense. We can well notice that on average 76.97% of the observations are closed to the mean while very few observations (less than 6.41%) are very far from the majority.
The results are summarized in Table 6.3.

Figure 6.7 and Figure 6.8 show the boxplots of the $N = 1000$ realizations of Hill's estimator and $\hat{\gamma}_{k_n}(K_{\hat{S}^*})$ for $\rho = -3.67$ while Figure 6.9 and Figure 6.10 show the corresponding boxplots for $\rho = -5$. In view of these results we can clearly see that the proposed estimator is more efficient. Indeed the boxplots show that there is more concentration around the median for our estimator than for Hill's estimator.
To better evaluate the performance of our estimators, we used a bootstrap method. Indeed we resampled 100 random samples (with replacement) from the initial sample (with $n = 1559$). On each sample, we run the previous algorithm. We then obtain a vector of 100 estimators and we return the mean. The results are presented in Table 6.4 and Figure 6.11. This figure shows the boxplots of the $N = 100$ bootstrap samples of Hill's estimator and $\hat{\gamma}_{k_n}(K_{\hat{S}^*})$ for $\rho = -3.67$ while Figure 6.12 shows the boxplots for the associated quantile estimators. The corresponding boxplots for $\rho = -5$ are given in Figure 6.13 and Figure 6.14. These results confirm the performance of the proposed methodology.

Conclusion

The estimation of the extreme distribution tail index discussed in this article generalizes that of Bassene [2016] for more efficiency and that of Chavez-Demoulin and Guillou [2018] to the spatial case. The asymptotic properties of the proposed tail index estimators (biased and unbiased) have been established under certain conditions, in particular

the β-mixing condition compare to the α-mixing in Bassene [2016]. The originality of the considered framework lies in the spatial nature of the dependent process studied but also on a wide choice of functions K and assumption C_K, reducing the asymptotic bias and variance of the estimators, compare to our predecessors [Chavez-Demoulin and Guillou, 2018, De Haan and Ferreira, 2016, Goegebeur and Guillou, 2013]. We also proposed an asymptotically normal quantile estimator similar to the one proposed by Chavez-Demoulin and Guillou [2018]. Future directions may include considering models with exogenous variables or space-time processes, with a number of potential applications.

6 Appendix

Proofs of the main results

To establish the proofs of the main results, we adopt Robinson [2011]'s notation of the spatial locations (for seek of simplicity). That is the process $\left\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right\}$ is written as $\{X_i, 1 \leq i \leq n = n_1 \times n_2 \times \cdots \times n_N\}$ using for instance a triangular array notation and a lexicographic ordering. For this notation the mixing conditions C_M and C_R (regularity) are written as:

Condition C'_M (mixing condition): Let's $(l_n)_{n\in\mathbb{N}^*}$ be a sequence of integers such that $1 \leq l_n \leq n$; set $\mathcal{B}_m^j = \sigma(X_i, m \leq i \leq j)$ be σ-fields generated by the random variables $(X_i)_i$ with $m \leq i \leq j$. the β-mixing condition is given by:

$$\beta(l_n) := \underset{m\in\mathbb{N}^*}{sup}\,\mathbb{E}\left[\underset{A\in\mathcal{B}_{l_n+m+1}^{+\infty}}{sup}\left|\mathbb{P}(A|\mathcal{B}_1^m) - \mathbb{P}(A)\right|\right] \underset{l_n\to\infty}{\longrightarrow} 0 \tag{35}$$

See Drees [2000] for a discussion on the β-mixing and examples.

Condition C'_R: (regularity)
There is $\epsilon > 0$, a function $r : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, and (l_n) defined above is such that $l_n = \mathrm{o}(n/k_n)$; and when $n \to \infty$

(a') $\frac{\beta(l_n)}{l_n} n + l_n \frac{\log^2 k_n}{\sqrt{k_n}} \to 0$;

(b') $\frac{n}{l_n k_n}\mathrm{Cov}\left(\sum_{i=1}^{l_n} 1_{\{X_i > F^{\leftarrow}(1-k_n x/n)\}}, \sum_{i=1}^{l_n} 1_{\{X_i > F^{\leftarrow}(1-k_n y/n)\}}\right) \to r(x,y)$,

$\forall\, 0 \leq x, y \leq 1+\epsilon$;

(c') there exists a constant C such that :

$$\frac{n}{l_n k_n} E\left[\left(\sum_{i=1}^{l_n} 1_{\{F^{\leftarrow}(1-k_n y/n) < X_i \leq F^{\leftarrow}(1-k_n x/n)\}}\right)^4\right] \leq C(y-x) \;\forall\; 0 \leq x < y \leq 1+\epsilon.$$

6.1 Proof of Theorem 2.1

To establish the proof of the theorem, we need the following proposition.

Proposition 6.1 *Let* $\left\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{Z}^N\right\}$ *be a* β*-mixing stationary spatial process with a distribution function* F*; verifying* C_A *and* C_R *and* K *a function verifying* C_K*. Let* (k_n) *be an intermediate sequence such that* $\sqrt{k_n}\mathcal{A}(b(n/k_n)) \to \lambda$*, as* $n \to \infty$*. For all* $\epsilon > 0$*, by Skorohod construction, there exist a function* $\tilde{\mathcal{A}} \sim \mathcal{A}$ *and a Gaussian centred process*

$(W(t))_{t\in[0,1]}$ *with covariance function* r *such that, as* $n\to\infty$,

$$\sup_{t\in(0,1]} t^{1/2+\epsilon}\left|\sqrt{k_n}\left(\log\left(\frac{Q_n(t)}{U(b(n/k_n))}\right)+\frac{\gamma}{\int_0^1 K(s)ds}\log t\right)-\gamma t^{-1}W(t)\right.$$
$$\left.-\sqrt{k_n}\tilde{\mathcal{A}}(b(n/k_n))\frac{t^{-\rho}-1}{\rho}\right|\xrightarrow{a.s}0. \tag{36}$$

Proof of Proposition 6.1
Suppose the relationship (12) (from the condition C_A) hold. By applying Theorem B.2.18 in De Haan and Ferreira [2006], we get:
$\forall\,\epsilon,\ \delta>0\ \exists\ u_0=u_0(\epsilon,\delta)$ such that $\forall\ ux\geq u_0$;

$$\left|\frac{\log(\mathrm{U}(ux)/\mathrm{U}(u))-\gamma\log(x)}{\tilde{\mathcal{A}}(u)}-\frac{x^\rho-1}{\rho}\right|\leq\epsilon x^\rho\max(x^\delta,x^{-\delta}). \tag{37}$$

Set $X_i=U(Y_i)$ where Y_i follows a standard Pareto distribution. Then $(Y_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^N}$ is stationary and $\beta-$ mixing satisfying the regular variation (C_R). Then, since $Q_n(t)=U(Y_{n-\lfloor k_nt\rfloor,n})$ and according to Theorem 2.1 in Drees (2003) and under Skorohod construction, there exists a centred Gaussian process $(W(t))_{t\in[0,1]}$ with covariance function r such that for all $\epsilon>0$

$$\sup_{t\in(0,1]} t^{1/2+\epsilon}\left|\sqrt{k_n}\left(t\frac{Y_{n-\lfloor k_nt\rfloor,n}}{b(n/k_n)}-1\right)-t^{-1}W(t)\right|\to 0,\ \text{ a.s}, \tag{38}$$

as $n\to\infty$. The inequality (37) gives, for all $n>n_0(\epsilon,\delta)$:

$$\left|\log(Q_n(t))-\log\left(\mathrm{U}(b(n/k_n))\right)-\gamma\log\left(\tfrac{1}{b(n/k_n)}Y_{n-\lfloor k_nt\rfloor,n}\right)-\tilde{\mathcal{A}}(b(n/k_n))\frac{\left(\frac{1}{b(n/k_n)}Y_{n-\lfloor k_nt\rfloor,n}\right)^\rho-1}{\rho}\right|$$
$$\leq\epsilon\left|\tilde{\mathcal{A}}(b(n/k_n))\right|\left(\tfrac{1}{b(n/k_n)}Y_{n-\lfloor k_nt\rfloor,n}\right)^{\rho+\delta}.$$

So,

$$t^{1/2+\epsilon}\left|\sqrt{k_n}\left(\log\left(\frac{Q_n(t)}{\mathrm{U}(b(n/k_n))}\right)+\frac{\gamma}{\int_0^1 K(s)ds}\log(t)\right)-\gamma t^{-1}W(t)\right.$$
$$-\sqrt{k_n}\tilde{\mathcal{A}}(b(n/k_n))\frac{t^{-\rho}-1}{\rho}+\sqrt{k_n}\tilde{\mathcal{A}}(b(n/k_n))\frac{1}{\rho}\left(t^{-\rho}-\left(\frac{1}{b(n/k_n)}Y_{n-\lfloor k_nt\rfloor,n}\right)^\rho\right)$$
$$\left.-\gamma\left\{\sqrt{k_n}\left(\log\left(\frac{1}{b(n/k_n)}Y_{n-\lfloor k_nt\rfloor,n}\right)+\frac{1}{\int_0^1 K(s)ds}\log(t)\right)-t^{-1}W(t)\right\}\right|$$
$$\leq\epsilon\sqrt{k_n}\left|\tilde{\mathcal{A}}(b(n/k_n))\right|t^{1/2+\epsilon}\left(\frac{1}{b(n/k_n)}\right).$$

Since $\frac{1}{b(n/k_n)}Y_{n-\lfloor k_nt\rfloor,n}\geq 1$, by choosing $\delta\in(0,-\rho)$ the right term tends to 0 when $\epsilon\to 0$. Thus under the convergence (38) the proof of the Proposition 6.1 is obtained as in De Haan et al. [2016].

Proof of Theorem 2.1: From Proposition 6.1, we deduce that:

$$\begin{aligned}\sqrt{k_n}\left\{\hat{\gamma}_{k_n}(K)+\frac{\gamma}{\int_0^1 K(s)ds}\int_0^1 \log(t)d(tK(t))\right\} &= \gamma\int_0^1 (t^{-1}W(t)-W(1))d(tK(t))\\ &\quad+\sqrt{k_n}\tilde{\mathcal{A}}(b(n/k_n))\int_0^1 \frac{t^{-\rho}-1}{\rho}d(tK(t))\\ &\quad+\mathrm{o}(1)\int_0^1 t^{-1/2-\epsilon}d(tK(t)).\end{aligned}$$

Using an integration by part, we can write

$$\int_0^1 \log(t)d(tK(t)) = -\int_0^1 K(s)ds \quad \text{and} \quad \int_0^1 \frac{t^{-\rho}-1}{\rho}d(tK(t)) = \int_0^1 t^{-\rho}K(t)dt.$$

Hence

$$\begin{aligned}\sqrt{k_n}\left\{\hat{\gamma}_{k_n}(K)-\gamma-\tilde{\mathcal{A}}(b(n/k_n))\int_0^1 t^{-\rho}K(t)dt\right\} &= \gamma\int_0^1 (t^{-1}W(t)-W(1))d(tK(t))\\ &\quad+\mathrm{o}(1)\int_0^1 t^{-1/2-\epsilon}d(tK(t)).\end{aligned}$$

By taking $0<\epsilon<1/2-\tau$, we get the convergence of $\displaystyle\int_0^1 t^{-1/2-\epsilon}d(tK(t))$. Thus, this ends the proof.

6.2 Proof of Corollary 2.1

Indeed, the term $\mathcal{A}(b(n/k_n))\displaystyle\int_0^1 t^{-\rho}K(t)dt$ outcome from (13) is the bias of the estimator; and as $\sqrt{k_n}\mathcal{A}(b(n/k_n)) \longrightarrow \lambda$ then we get the asymptotic bias $\lambda\mathcal{AB}(K) = \lambda\displaystyle\int_0^1 t^{-\rho}K(t)dt$.
The variance $\mathcal{AV}(K)$ is obtained from the Gaussian centered process $(W(t))_{t\in[0.1]}$ covariance function r; that is

$$\mathcal{AV}(K) = \gamma^2\mathrm{E}\left[\left(\int_0^1 \left[t^{-1}W(t)-W(1)\right]d(tK(t))\right)^2\right].$$

The proofs of Corollary 2.2 and Corollary 2.3 are straightforward and follow the same lines as the Corollary 2.1.

6.3 Proof of Theorem 2.2

Let

$$\begin{aligned}K_{S^*}(t) &= \frac{\rho^2}{1-2\rho}-\frac{\rho^2}{1-\rho}t^{-\rho} := \alpha K_{1,\rho}(t)+\beta K_{2,\rho}(t) \quad \text{and}\\ K_{\hat{S}^*}(t) &= \frac{\hat{\rho}^2_{k_n,\rho}}{1-2\hat{\rho}_{k_n,\rho}}-\frac{\hat{\rho}^2_{k_n,\rho}}{1-\hat{\rho}_{k_n,\rho}}t^{-\hat{\rho}_{k_n,\rho}} := \hat{\alpha}K_{1,\hat{\rho}_{k_n,\rho}}(t)+\hat{\beta}K_{2,\hat{\rho}_{k_n,\rho}}(t),\end{aligned}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are consistent estimators of α and β respectively.
Let us first, give the following decomposition

$$\sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_{\hat{S}^*})-\gamma\right)=\sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_{S^*})-\gamma\right)+\sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_{\hat{S}^*})-\hat{\gamma}_{k_n}(K_{S^*})\right). \tag{39}$$

According to Corollary 2.2, the first term on the right converges to Gaussian distribution. So it remains to prove that the second term tends to 0 in probability. The proof follows the same lines as those of Theorem 2 in Chavez-Demoulin and Guillou [2018]. The difference in our approach is in managing the assumptions of the C_R and C_K conditions since we assumed that K is not necessarily a kernel function. We have:

$$\begin{aligned}
\sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_{\hat{S}^*})-\hat{\gamma}_{k_n}(K_{S^*})\right) &= \sqrt{k_n}\left\{\int_0^1\log\frac{Q_n(t)}{Q_n(1)}d(tK_{\hat{S}^*}(t))-\int_0^1\log\frac{Q_n(t)}{Q_n(1)}d(tK_{S^*}(t))\right\}\\
&= \sqrt{k_n}\left\{\hat{\alpha}\int_0^1\log\frac{Q_n(t)}{Q_n(1)}dt-\alpha\int_0^1\log\frac{Q_n(t)}{Q_n(1)}dt\right.\\
&\quad\left.+\hat{\beta}\int_0^1\log\frac{Q_n(t)}{Q_n(1)}d(tK_{2,\hat{\rho}_{k_n,\rho}}(t))-\beta\int_0^1\log\frac{Q_n(t)}{Q_n(1)}d(tK_{2,\rho}(t))\right\}\\
&= (\hat{\alpha}-\alpha)\sqrt{k_n}\left\{\int_0^1\log\frac{Q_n(t)}{Q_n(1)}dt-\int_0^1\log\frac{Q_n(t)}{Q_n(1)}d(tK_{2,\rho}(t))\right\}\\
&\quad+(\hat{\alpha}-\alpha+\hat{\beta}-\beta)\sqrt{k_n}\int_0^1\log\frac{Q_n(t)}{Q_n(1)}d(tK_{2,\rho}(t))\\
&\quad+\hat{\beta}\sqrt{k_n}\left\{\int_0^1\log\frac{Q_n(t)}{Q_n(1)}d(tK_{2,\hat{\rho}_{k_n,\rho}}(t))-\int_0^1\log\frac{Q_n(t)}{Q_n(1)}d(tK_{2,\rho}(t))\right\}\\
&=: T_1+T_2+T_3.
\end{aligned} \tag{40}$$

Let us evaluate the three terms.
Using Corollary 2.1 and the fact that $\hat{\alpha}$ and $\hat{\beta}$ are consistent estimators of α and β respectively, we have $T_1=\mathrm{o}_{\mathbb{P}}(1)$ and $T_2=\mathrm{o}_{\mathbb{P}}(1)$.
It remains to deal with **Term** T_3.
Noting that $\log\frac{Q_n(t)}{Q_n(1)}=\log\frac{Q_n(t)}{\mathrm{U}(b(n/k_n))}-\log\frac{Q_n(1)}{\mathrm{U}(b(n/k_n))}$, the Proposition 6.1 gives for all $\epsilon\in(0,1/2)$

$$\begin{aligned}
\log\frac{Q_n(t)}{Q_n(1)} &= \frac{\gamma}{\int_0^1K(s)ds}(-\log(t))+\frac{\gamma}{\sqrt{k_n}}[t^{-1}W(t)-W(1)]\\
&\quad+\tilde{\mathcal{A}}(b(n/k_n))\frac{t^{-\rho}-1}{\rho}+\frac{\mathrm{o}(1)}{\sqrt{k_n}}t^{-\epsilon-1/2}.
\end{aligned}$$

Thus we get:

$$
\begin{aligned}
& \sqrt{k_n}\left\{\int_0^1 \log\frac{Q_n(t)}{Q_n(1)} d(tK_{2,\hat{\rho}_{k_n,\rho}}(t)) - \int_0^1 \log\frac{Q_n(t)}{Q_n(1)} d(tK_{2,\rho}(t))\right\} \\
= \;& \gamma\sqrt{k_n}\left\{\frac{1}{\int_0^1 K_{2,\hat{\rho}_{k_n,\rho}}(t)dt}\int_0^1(-\log(t))d(tK_{2,\hat{\rho}_{k_n,\rho}}(t))\right. \\
& \left. -\frac{1}{\int_0^1 K_{2,\rho}(t)dt}\int_0^1(-\log(t))d(tK_{2,\rho}(t))\right\} \\
& +\gamma\left\{\int_0^1[t^{-1}W(t)-W(1)]d(tK_{2,\hat{\rho}_{k_n,\rho}}(t)) - \int_0^1[t^{-1}W(t)-W(1)]d(tK_{2,\rho}(t))\right\} \\
& +\sqrt{k_n}\tilde{\mathcal{A}}(b(n/k_n))\left\{\int_0^1\frac{t^{-\rho}-1}{\rho}d(tK_{2,\hat{\rho}_{k_n,\rho}}(t)) - \int_0^1\frac{t^{-\rho}-1}{\rho}d(tK_{2,\rho}(t))\right\} \\
& +\mathrm{o}(1)\left\{\int_0^1 t^{-\frac{1}{2}-\epsilon}d(tK_{2,\hat{\rho}_{k_n,\rho}}(t)) - \int_0^1 t^{-\frac{1}{2}-\epsilon}d(tK_{2,\rho}(t))\right\} \\
=: \;& A+B+C+D.
\end{aligned}
$$

The term A converges by using an integration by part.
The term B is

$$
\begin{aligned}
B = \gamma\left\{\right. & \textstyle\int_0^1[t^{-1}W(t)-W(1)]\left(K_{2,\hat{\rho}_{k_n,\rho}}(t)-K_{2,\rho}(t)\right)dt \\
& \left. + \textstyle\int_0^1[t^{-1}W(t)-W(1)]t\left(K'_{2,\hat{\rho}_{k_n,\rho}}(t)-K'_{2,\rho}(t)\right)dt\right\}.
\end{aligned}
$$

Let us consider $\epsilon\in(0,1)$ and $\tilde{\rho}$ a random value between ρ and $\hat{\rho}_{k_n\rho}$. We have

$$
\begin{aligned}
& \left|\int_0^1[t^{-1}W(t)-W(1)]\left(K_{2,\hat{\rho}_{k_n,\rho}}(t)-K_{2,\rho}(t)\right)dt\right| \\
\leq \;& \int_0^1\left|t^{-1}W(t)-W(1)\right|\left|K_{2,\hat{\rho}_{k_n,\rho}}(t)-K_{2,\rho}(t)\right|dt \\
\leq \;& (1-\hat{\rho}_{k_n,\rho})\int_0^1\left|t^{-1}W(t)-W(1)\right|\left|t^{-\hat{\rho}_{k_n,\rho}}-t^{-\rho}\right|dt \\
& +\left|\hat{\rho}_{k_n,\rho}-\rho\right|\int_0^1\left|t^{-1}W(t)-W(1)\right|t^{-\rho}dt \\
\leq \;& (1-\hat{\rho}_{k_n,\rho})\sup_{t\in(0,1]}t^{\frac{1}{2}+\epsilon}\left|t^{-1}W(t)-W(1)\right|\sup_{t\in(0,1]}t^{\frac{1}{4}}\left|t^{-\hat{\rho}_{k_n,\rho}}-t^{-\rho}\right|\int_0^1 t^{-\frac{3}{4}-\epsilon}dt \\
& +\left|\hat{\rho}_{k_n,\rho}-\rho\right|\sup_{t\in(0,1]}t^{\frac{1}{2}+\epsilon}\left|t^{-1}W(t)-W(1)\right|\sup_{t\in(0,1]}t^{\frac{1}{4}-\rho}\int_0^1 t^{-\frac{3}{4}-\epsilon}dt \\
\overset{*}{\leq} \;& \frac{4}{1-4\epsilon}\left|\hat{\rho}_{k_n,\rho}-\rho\right|(1-\hat{\rho}_{k_n,\rho})\sup_{t\in(0,1]}t^{\frac{1}{2}+\epsilon}\left|t^{-1}W(t)-W(1)\right|\sup_{t\in(0,1]}(-\log t)\,t^{\frac{3}{4}-\tilde{\rho}} \\
& +\frac{4}{1-4\epsilon}\left|\hat{\rho}_{k_n,\rho}-\rho\right|\sup_{t\in(0,1]}t^{\frac{1}{2}+\epsilon}\left|t^{-1}W(t)-W(1)\right| \\
= \;& \mathrm{o}_{\mathbb{P}}(1).
\end{aligned}
$$

The inequality $\overset{*}{\leq}$ is justified by:
set $h(t)=t^{\frac{1}{4}-\hat{\rho}_{k_n,\rho}}-t^{\frac{1}{4}-\rho}=t^{\frac{1}{4}-\tilde{\rho}}\left(t^{\tilde{\rho}-\hat{\rho}_{k_n,\rho}}-t^{\tilde{\rho}-\rho}\right)$. A Taylor expansion of the term

$\left(t^{\tilde{\rho}-\hat{\rho}_{k_n,\rho}} - t^{\tilde{\rho}-\rho}\right)$ gives

$$h(t) \simeq t^{\frac{1}{4}-\tilde{\rho}}(\rho - \hat{\rho}_{k_n,\rho})\log(t),$$

and we get

$$\sup_{t\in(0,1]} |h(t)| \leq \left|\hat{\rho}_{k_n,\rho} - \rho\right| \sup_{t\in(0,1]} t^{\frac{1}{4}-\tilde{\rho}} \left(-\log(t)\right).$$

In the same way we have:

$$\begin{aligned}
& \left|\int_0^1 [t^{-1}W(t) - W(1)]t\left(K'_{2,\hat{\rho}_{k_n,\rho}}(t) - K'_{2,\rho}(t)\right)dt\right| \\
\leq\ & \int_0^1 \left|t^{-1}W(t) - W(1)\right| t \left|K'_{2,\hat{\rho}_{k_n,\rho}}(t) - K'_{2,\rho}(t)\right| dt \\
\leq\ & -\hat{\rho}_{k_n,\rho}(1-\hat{\rho}_{k_n,\rho}) \int_0^1 \left|t^{-1}W(t) - W(1)\right| \left|t^{-\hat{\rho}_{k_n,\rho}} - t^{-\rho}\right| dt \\
& - \left(\hat{\rho}_{k_n,\rho} + \rho\right)\left|\hat{\rho}_{k_n,\rho} - \rho\right|^2 \int_0^1 \left|t^{-1}W(t) - W(1)\right| t^{-\rho} dt \\
\leq\ & -\hat{\rho}_{k_n,\rho}(1-\hat{\rho}_{k_n,\rho}) \sup_{t\in(0,1]} t^{\frac{1}{2}+\epsilon}\left|t^{-1}W(t) - W(1)\right| \sup_{t\in(0,1]} t^{\frac{1}{4}} \left|t^{-\hat{\rho}_{k_n,\rho}} - t^{-\rho}\right| \int_0^1 t^{-\frac{3}{4}-\epsilon} dt \\
& - \left(\hat{\rho}_{k_n,\rho} + \rho\right)\left|\hat{\rho}_{k_n,\rho} - \rho\right|^2 \sup_{t\in(0,1]} t^{\frac{1}{2}+\epsilon}\left|t^{-1}W(t) - W(1)\right| \sup_{t\in(0,1]} t^{\frac{1}{4}-\rho} \int_0^1 t^{-\frac{3}{4}-\epsilon} dt \\
\overset{*}{\leq}\ & -\hat{\rho}_{k_n,\rho}\frac{4}{1-4\epsilon}\left|\hat{\rho}_{k_n,\rho} - \rho\right|(1-\hat{\rho}_{k_n,\rho}) \sup_{t\in(0,1]} t^{\frac{1}{2}+\epsilon}\left|t^{-1}W(t) - W(1)\right| \sup_{t\in(0,1]} (-\log t)\, t^{\frac{3}{4}-\tilde{\rho}} \\
& - \frac{4}{1-4\epsilon}\left(\hat{\rho}_{k_n,\rho} + \rho\right)\left|\hat{\rho}_{k_n,\rho} - \rho\right|^2 \sup_{t\in(0,1]} t^{\frac{1}{2}+\epsilon}\left|t^{-1}W(t) - W(1)\right| \\
=\ & \mathrm{o}_{\mathbb{P}}(1).
\end{aligned}$$

Then we get $B = \mathrm{o}_{\mathbb{P}}(1)$.

By making an integration by part C is

$$\begin{aligned}
C &= \sqrt{k_n}\tilde{\mathcal{A}}(b(n/k_n)) \left\{ \left[t\frac{t^{-\rho}-1}{\rho}\left(K_{2,\hat{\rho}_{k_n,\rho}}(t) - K_{2,\rho}(t)\right)\right]_0^1 \right. \\
&\quad \left. + \int_0^1 t^{-\rho}\left(K_{2,\hat{\rho}_{k_n,\rho}}(t) - K_{2,\rho}(t)\right) dt \right\},
\end{aligned}$$

and as $\sqrt{k_n}\tilde{\mathcal{A}}(b(n/k_n)) \to \lambda$ and $\int_0^1 t^{-\rho}\left(K_{2,\hat{\rho}_{k_n,\rho}}(t) - K_{2,\rho}(t)\right) dt$ converges so we have $C = \mathrm{o}_{\mathbb{P}}(1)$.
Similarly, an integration by part allows us to conclude that $D = \mathrm{o}_{\mathbb{P}}(1)$.

In short, we have $T_3 = \mathrm{o}_{\mathbb{P}}(\hat{\beta})$ and since $\hat{\beta} < 1$ we have

$$\sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_{\hat{S}^*}) - \hat{\gamma}_{k_n}(K_{S^*})\right) = \mathrm{o}_{\mathbb{P}}(1).$$

This ends the proof of Theorem 2.2.

6.4 Proof of Theorem 3.1

To do this, we only need to show the asymptotic normality of $\frac{\sqrt{k_n}}{\log\frac{1}{pb(n/k_n)}}\log\frac{\hat{x}_{p,\xi}}{x_p}$.

We have the decomposition below:

$$
\begin{aligned}
\frac{\sqrt{k_n}}{\log\frac{1}{pb(n/k_n)}}\log\frac{\hat{x}_{p,\xi}}{x_p} &= \frac{\sqrt{k_n}}{\log\frac{1}{pb(n/k_n)}}\left\{\log X_{n-\lfloor tk_n\rfloor,n}+\hat{\gamma}_{k_n}(K_{\hat{S}^*})\log\frac{1}{pb(n/k_n)}-\log x_p\right.\\
&\quad\left.-\frac{(1-\xi)(1-2\xi)}{\xi^2}[\hat{\gamma}_{k_n}(K_1)-\hat{\gamma}_{k_n}(K_{2,\xi})]\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\xi}-1}{\xi}\right\}\\
&= \sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_{\hat{S}^*})-\gamma\right)+\frac{\sqrt{k_n}}{\log\frac{1}{pb(n/k_n)}}\log\frac{Q_n(t)}{\mathrm{U}(b(n/k_n))}\\
&\quad-\frac{\sqrt{k_n}}{\log\frac{1}{pb(n/k_n)}}\left\{\log\frac{\mathrm{U}(\frac{1}{p})}{\mathrm{U}(b(n/k_n))}-\gamma\log\frac{1}{pb(n/k_n)}\right\}\\
&\quad-\frac{(1-\xi)(1-2\xi)}{\xi^2}\frac{\sqrt{k_n}[\hat{\gamma}_{k_n}(K_1)-\hat{\gamma}_{k_n}(K_{2,\xi})]}{\log\frac{1}{pb(n/k_n)}}\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\xi}-1}{\xi}\\
&= \sqrt{k_n}\left(\hat{\gamma}_{k_n}(K_{\hat{S}^*})-\gamma\right)+\frac{\sqrt{k_n}}{\log\frac{1}{pb(n/k_n)}}\log\frac{Q_n(t)}{\mathrm{U}(b(n/k_n))}\\
&\quad-\frac{\sqrt{k_n}}{\log\frac{1}{pb(n/k_n)}}\tilde{\mathcal{A}}(b(n/k_n))\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\rho}-1}{\rho}-\frac{\sqrt{k_n}}{\log\frac{1}{pb(n/k_n)}}\tilde{\mathcal{A}}(b(n/k_n))\\
&\quad\times\left\{\frac{\log\,\mathrm{U}(\frac{1}{p})-\log\mathrm{U}(b(n/k_n))-\gamma\log\frac{1}{pb(n/k_n)}}{\tilde{\mathcal{A}}(b(n/k_n))}-\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\rho}-1}{\rho}\right\}\\
&\quad-\frac{(1-\xi)(1-2\xi)}{\xi^2}\frac{\sqrt{k_n}[\hat{\gamma}_{k_n}(K_1)-\hat{\gamma}_{k_n}(K_{2,\xi})]}{\log\frac{1}{pb(n/k_n)}}\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\xi}-1}{\xi}\\
&=: T_4+T_5-T_6-T_7-T_8.
\end{aligned}
$$

Let us now look at the 5 terms.
Theorem 2.2 ensures the asymptotic normality of the term T_4

$$T_4\xrightarrow{d}\mathcal{N}\left(0,\ \mathcal{AV}(K_{S^*})\right).$$

Using Proposition 6.1 (for $t=1$ and the fact that $\log(x)\sim x-1$ when $x\to 1$), we can show

$$T_5\xrightarrow{\mathbb{P}}0.$$

Indeed,

$$\begin{aligned}
& \sup_{t\in(0,1]} t^{1/2+\epsilon}\left|\sqrt{k_n}\log\left(\frac{Q_n(t)}{\mathrm{U}(b(n/k_n))}\right)-\gamma W(1)\right| \\
\leq\ & \sup_{t\in(0,1]} t^{\frac{1}{2}+\epsilon}\left|\sqrt{k_n}\left(\log\left(\frac{Q_n(t)}{\mathrm{U}\left(b\left(\frac{n}{k_n}\right)\right)}\right)+\frac{\gamma\log(t)}{\int_0^1 K(s)ds}\right)\right. \\
-\ & \left.\gamma t^{-1}W(t)-\sqrt{k_n}\tilde{\mathcal{A}}\left(b\left(\frac{n}{k_n}\right)\right)\frac{t^{-\rho}-1}{\rho}\right| \\
+\ & \sup_{t\in(0,1]} t^{\frac{1}{2}+\epsilon}\left|\sqrt{k_n}\frac{\gamma\log(t)}{\int_0^1 K(s)ds}-\gamma\left(t^{-1}W(t)-W(1)\right)-\sqrt{k_n}\tilde{\mathcal{A}}\left(b\left(\frac{n}{k_n}\right)\right)\frac{t^{-\rho}-1}{\rho}\right| \\
=\ & \mathrm{o}(1).
\end{aligned}$$

To prove that $T_6=\mathrm{o}(1)$, inequation (37)
leads to:

$$\begin{aligned}
|T_7| &\leq \frac{\sqrt{k_n}}{\log\frac{1}{pb(n/k_n)}}|\tilde{\mathcal{A}}(b(n/k_n))| \\
&\quad\times\left|\frac{\log\mathrm{U}(\frac{1}{p})-\log\mathrm{U}(b(n/k_n))-\gamma\log\frac{1}{pb(n/k_n)}}{\tilde{\mathcal{A}}(b(n/k_n))}-\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\rho}-1}{\rho}\right| \\
&\leq \frac{\sqrt{k_n}}{\log\frac{1}{pb(n/k_n)}}|\tilde{\mathcal{A}}(b(n/k_n))|\epsilon\left(\frac{1}{pb(n/k_n)}\right)^{\rho+\delta} \\
&= \mathrm{o}(1),
\end{aligned}$$

for all $0<\delta<-\rho$.
Note that the term T_8 is a function of ξ which can be a canonical value or an consistent estimator of ρ.

- If $\xi=\tilde{\rho}$ then, we have:
  $\sqrt{k_n}[\hat{\gamma}_{k_n}(K_1)-\hat{\gamma}_{k_n}(K_{2,\xi})]=\sqrt{k_n}[\hat{\gamma}_{k_n}(K_1)-\gamma]-\sqrt{k_n}[\hat{\gamma}_{k_n}(K_{2,\xi})-\gamma]=\mathrm{O}_{\mathbb{P}}(1)$ according to the Corollary 2.1. This leads to $T_8=\mathrm{o}_{\mathbb{P}}(1)$.

- If $\xi=\hat{\rho}$

$$\begin{aligned}
T_8 &= \frac{(1-\hat{\rho})(1-2\hat{\rho})}{\hat{\rho}^2}\frac{\sqrt{k_n}[\hat{\gamma}_{k_n}(K_1)-\hat{\gamma}_{k_n}(K_{2,\hat{\rho}})]}{\log\frac{1}{pb(n/k_n)}} \\
&\quad\times\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\rho}-1}{\rho}+\frac{(1-\hat{\rho})(1-2\hat{\rho})}{\hat{\rho}^2} \\
&\quad\times\frac{\sqrt{k_n}[\hat{\gamma}_{k_n}(K_1)-\hat{\gamma}_{k_n}(K_{2,\hat{\rho}})]}{\log\frac{1}{pb(n/k_n)}}\left(\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\hat{\rho}}-1}{\hat{\rho}}-\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\rho}-1}{\rho}\right).
\end{aligned}$$

  However, according to Corollary 2.1 and Theorem 2.2

$$\begin{aligned}\sqrt{k_n}[\hat{\gamma}_{k_n}(K_1) - \hat{\gamma}_{k_n}(K_{2,\hat{\rho}})] &= \sqrt{k_n}[\hat{\gamma}_{k_n}(K_1) - \gamma] - \sqrt{k_n}[\hat{\gamma}_{k_n}(K_{2,\rho}) - \gamma] \\ &\quad -\sqrt{k_n}[\hat{\gamma}_{k_n}(K_{2,\hat{\rho}}) - \hat{\gamma}_{k_n}(K_{2,\rho})] \\ &= \mathrm{O}_{\mathbb{P}}(1).\end{aligned}$$

The term T_8 becomes,

$$\begin{aligned}T_8 &= \mathrm{o}_{\mathbb{P}}(1) + \mathrm{o}_{\mathbb{P}}(1)\left\{\frac{\left(\frac{1}{pb(n/k_n)}\right)^{\hat{\rho}} - 1}{\hat{\rho}} - \frac{\left(\frac{1}{pb(n/k_n)}\right)^{\rho} - 1}{\rho}\right\} \\ &= \mathrm{o}_{\mathbb{P}}(1) + \mathrm{o}_{\mathbb{P}}(1)\int_0^{\frac{1}{pb(n/k_n)}} s^{\rho-1}(s^{\hat{\rho}-\rho} - 1)ds,\end{aligned}$$

Inspired by Chavez-Demoulin and Guillou [2018], we get
$\int_0^{\frac{1}{pb(n/k_n)}} s^{\rho-1}(s^{\hat{\rho}-\rho} - 1)ds = \mathrm{o}_{\mathbb{P}}(1)$,

which leads to the conclusion that $T_8 = \mathrm{o}_{\mathbb{P}}(1)$ and therefore we get the proof of Theorem 3.1.

6.5 Figures and tables

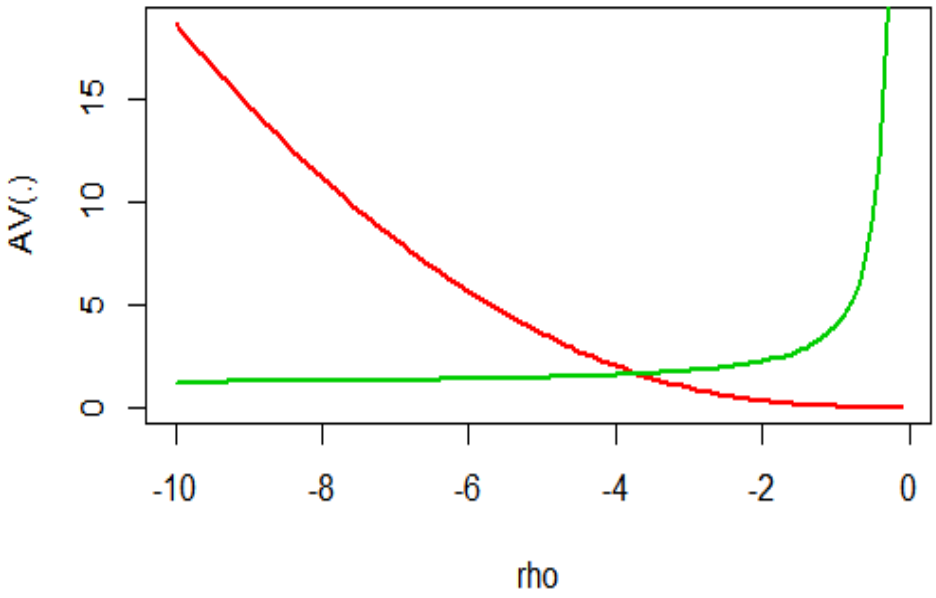

Figure 6.1 *Variances comparison for* $\gamma = 1$*;* $\mathcal{AV}(K_{\Delta^*})$ *(green) and* $\mathcal{AV}(K_{S^*})$ *(red).*

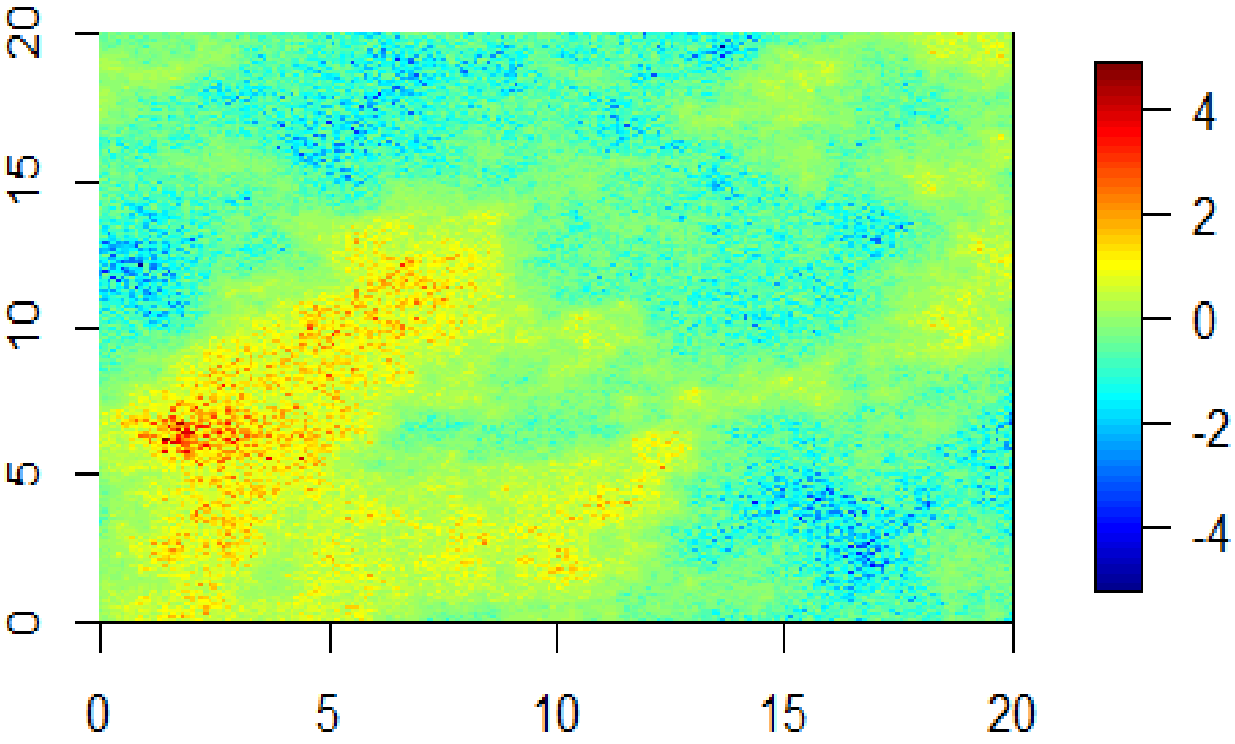

Figure 6.2 *log-Laplace field representation for* $\gamma = 1$

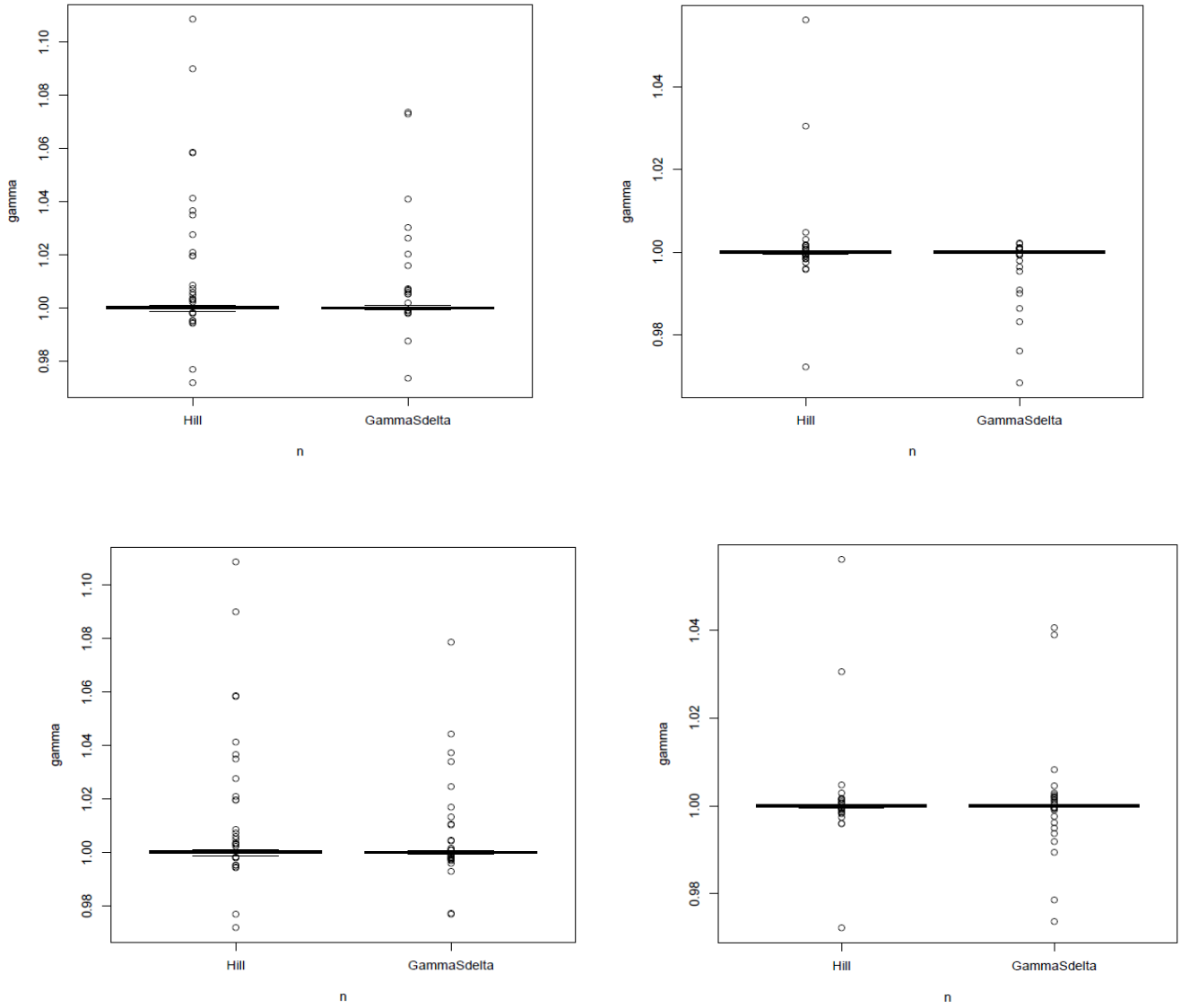

Figure 6.3 *γ Hill (left); $\gamma_{k_n}(K_{S\Delta})$ (right) in each graph for $\rho = -3.67$ on the first line and $\rho = -4$ on the second line. At left $n = 36000$ and on right $n = 40000$.*

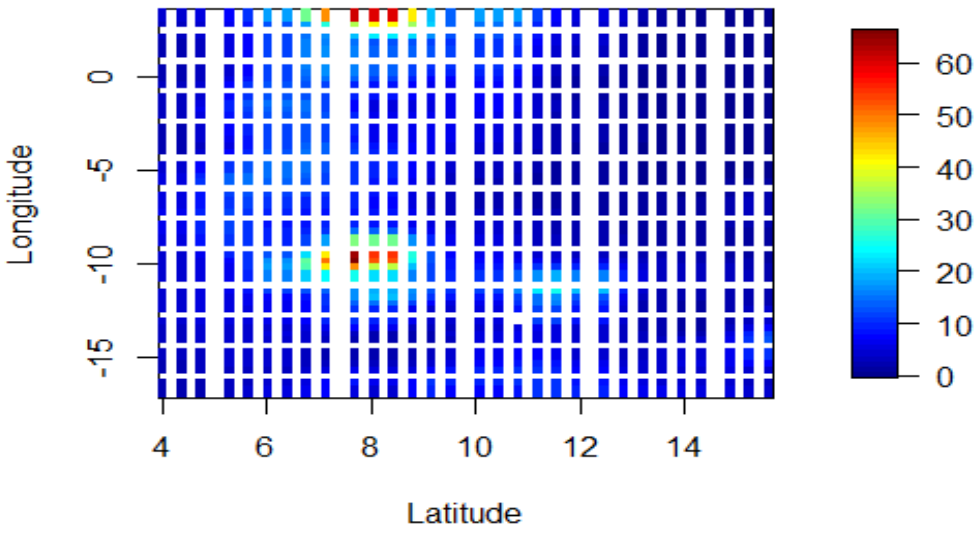

Figure 6.4 *Spatial representation of data.*

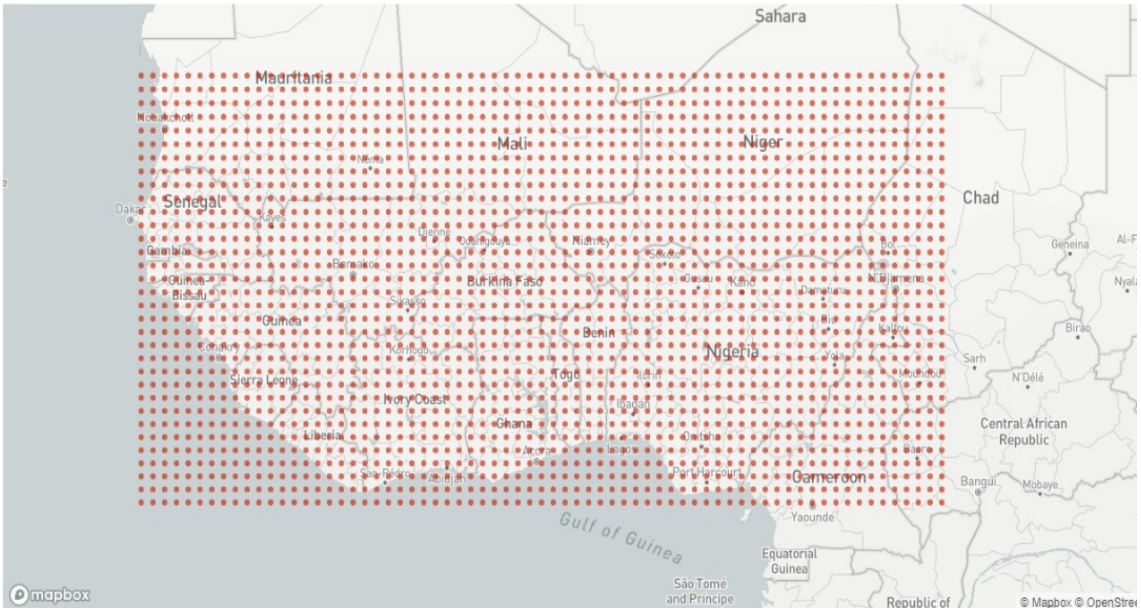

Figure 6.5 *Geographical representation of data.*

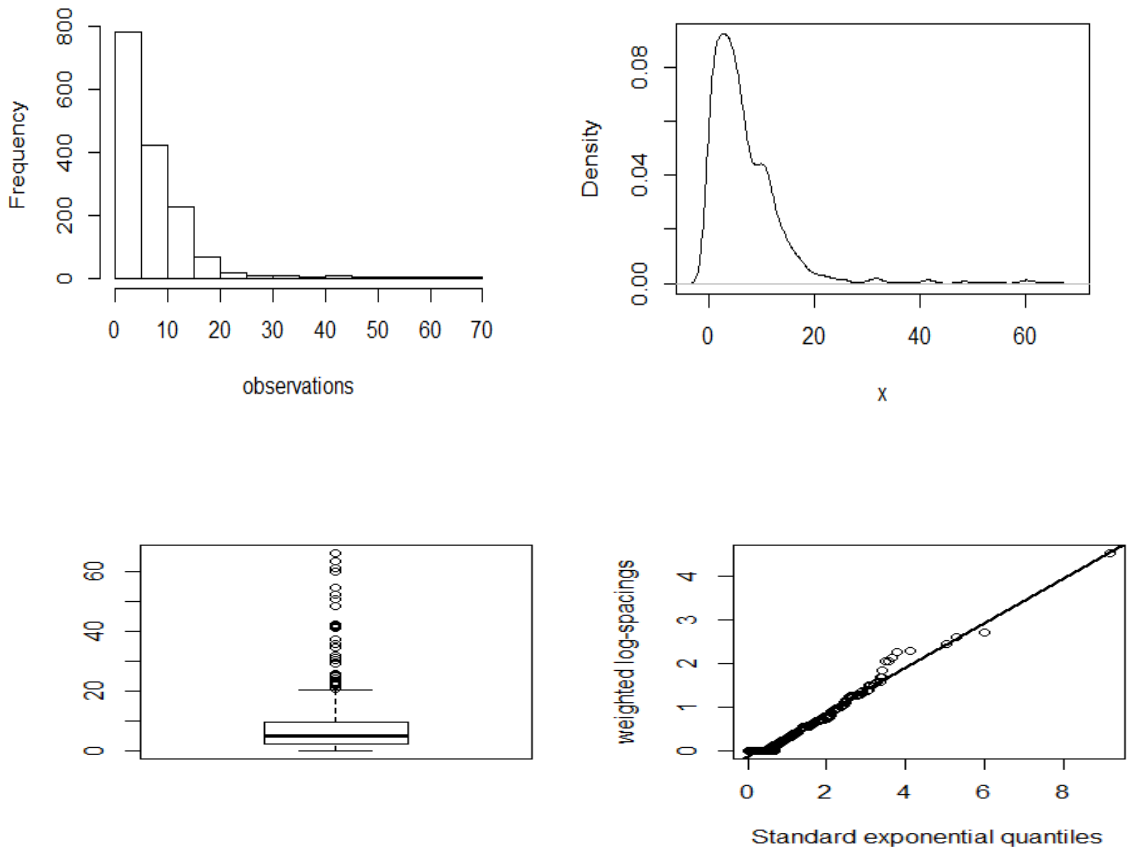

Figure 6.6 *Top row, left panel: histogram of the data. Top row, right panel: density. Bottom row, left panel: boxplot of the data. Bottom row, right panel: quantile-quantile plot of weighted log-spacings* $Z_{i,n}$ *for* $1 \leq i \leq \lfloor n/5 \rfloor$ *versus the standard exponential quantiles.*

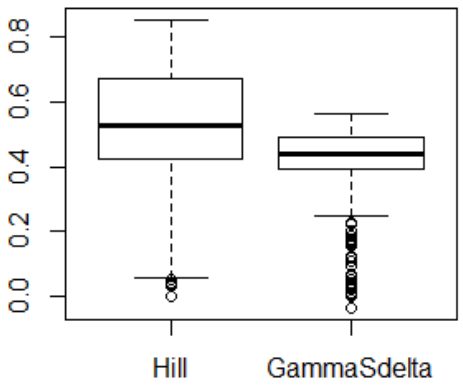

Figure 6.7 *Boxplots of the* $N = 100$ *estimates of* γ *for* $\rho = -3.67$.

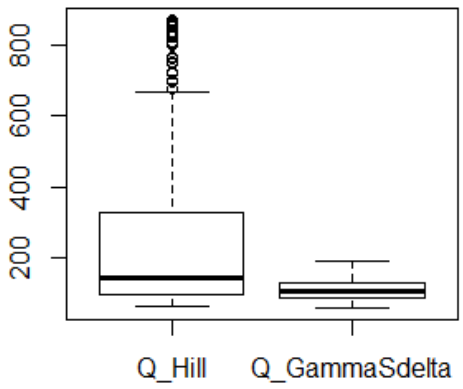

Figure 6.8 *Boxplots of the* $N = 100$ *estimates of* $q(1/1000)$ *for* $\rho = -3.67$.

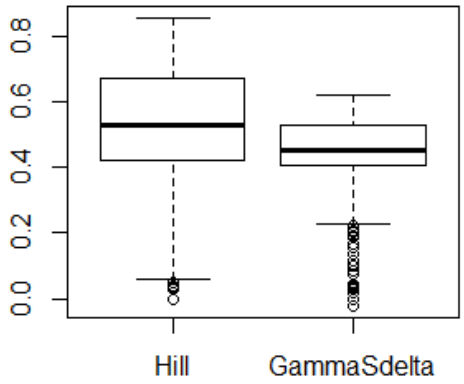

Figure 6.9 *Boxplots of the $N = 100$ estimates of γ for $\rho = -5$.*

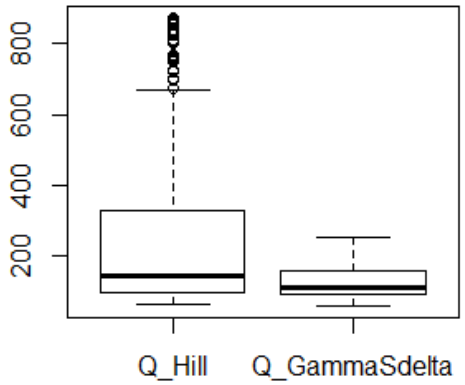

Figure 6.10 *Boxplots of the $N = 100$ estimates of $q(1/1000)$ for $\rho = -5$.*

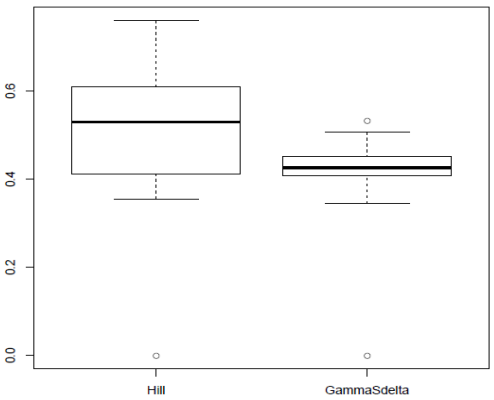

Figure 6.11 *Gammas $\rho = -3.67$.*

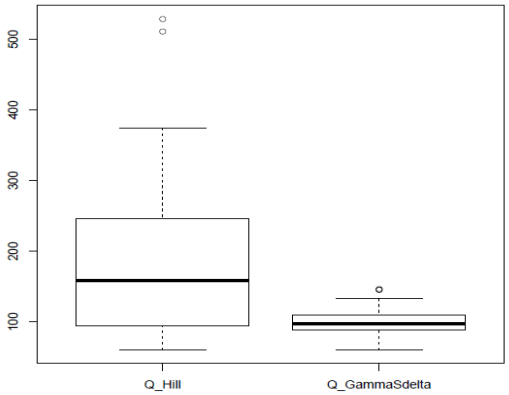

Figure 6.12 *Quantile for $\rho = -3.67$ and $p = 1/1000$.*

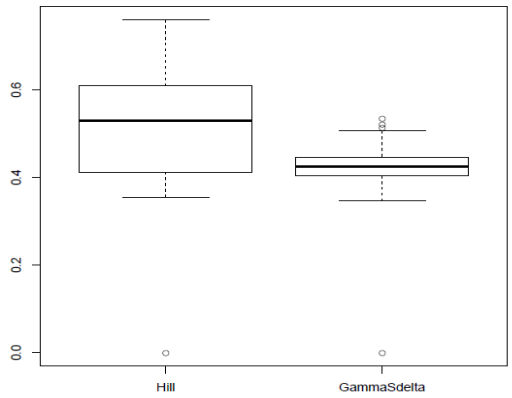

Figure 6.13 *Gamma for* $\rho = -5$.

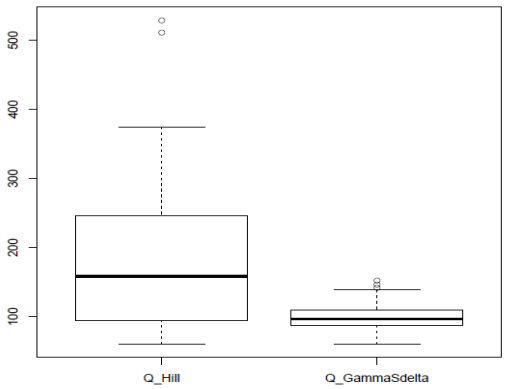

Figure 6.14 *Quantiles for* $\rho = -5$ *and* $p = 1/1000$.

	$n = 10\,000$		$n = 40\,000$	
value of ρ	Hill	$\hat{\gamma}_{k_n}(K_{S\Delta})$	Hill	$\hat{\gamma}_{k_n}(K_{S\Delta})$
$\rho = -3.5$	$[7.04.e^{-3}]$	$[5.85.e^{-3}]$	$[5.83.e^{-3}]$	$[4.15.e^{-3}]$
	$(7.06.e^{-3})$	$(5.88.e^{-3})$	$(5.86.e^{-3})$	$(4.16.e^{-3})$
	381	439	262	349
$\rho = -3.67$		$[4.79.e^{-3}]$		$[3.28.e^{-3}]$
		$(4.19.e^{-3})$		$(3.40.e^{-3})$
		440		368
$\rho = -3.8$		$[4.50.e^{-3}]$		$[3.29.e^{-3}]$
		$(4.52.e^{-3})$		$(3.29.e^{-3})$
		407		369
$\rho = -4$		$6.87.e^{-3}$		$2.53.e^{-3}$
		$(6.90.e^{-3})$		$(2.51.e^{-3})$
		425		354
$\rho = -5$		$[6.68.e^{-3}]$		$[3.54.e^{-3}]$
		$(6.71.e^{-3})$		$(3.52.e^{-3})$
		449		317
$\rho = -10$		$[7.56.e^{-3}]$		$[4.41.e^{-3}]$
		$(7.56.e^{-3})$		$(4.38.e^{-3})$
		406		293
$\rho = -20$		$[6.82.e^{-3}]$		$[7.32.e^{-3}]$
		$(6.78.e^{-3})$		$(7.27.e^{-3})$
		416		294

**[MSE]: Mean Square Error.
(sd): standard deviation
Kop: optimum k

Table 6.1 *Gamma estimators*

	$n = 10\,000$		$n = 40\,000$	
value of ρ	$\hat{q}_{\text{Hill}}$	$\hat{q}_{\gamma(K_{S\Delta})}$	$\hat{q}_{\text{Hill}}$	$\hat{q}_{\gamma(K_{S\Delta})}$
$\rho = -3.5$	80.795 [1.499] (1.493)	81.220 [0.782] (0.773)	81.011 [0.295] (0.296)	80.966 [0.128] (0.128)
$\rho = -3.67$		81.208 [0.870] (0.848)		80.998 [0.122] (0.122)
$\rho = -3.8$		81.142 [1.009] (1.003)		81.021 [0.125] (0.125)
$\rho = -4$		81.069 [0.975] (0.977)		81.018 [0.125] (0.125)
$\rho = -5$		81.064 [0.867] (0.869)		81.008 [0.124] (0.124)

** [MSE]: Mean Square Error.
(sd): standard deviation

Table 6.2 *Quantile estimators*

value of ρ	$\hat{\gamma}_{\text{Hill}}$ (k)	$\hat{\gamma}_{k_n}(K_{S\Delta})$ (k)	$q(1/1000)$		$q(1/100)$	
			Hill	$\hat{q}_{K_{S\Delta}}$	Hill	$\hat{q}_{K_{S\Delta}}$
$\rho = -3$	0.463 (465)	0.315 (303)	118.11	53.80	40.70	26.06
$\rho = -3.67$		0.428 (312)		98.93		36.91
$\rho = -4$		0.429 (324)		99.10		37.25
$\rho = -5$		0.424 (323)		99.22		36.99

Table 6.3 *Gamma estimator and quantile predictors on real data. The Hill's estimators are not function of ρ.*

value of ρ	$\hat{\gamma}_{\text{Hill}}$ (k)	$\hat{\gamma}_{k_n}(K_{S\Delta})$ (k)	$q(1/1000)$		$q(1/100)$	
			Hill	$\hat{q}_{K_{S\Delta}}$	Hill	$\hat{q}_{K_{S\Delta}}$
$\rho = -3$	0.514 (529) |0.122| $[0.27 - 0.75]$	0.309 (306) |0.040| $[0.23 - 0.39]$	179.515 |100.72| $[17.9 - 377]$	53.498 |6.487| $[40.78 - 66.21]$	49.370 |14.488| $[20.97 - 77.77]$	26.238 |3.790| $[18.80 - 33.67]$
$\rho = -3.67$		0.426 (302) |0.054| $[0.32 - 0.53]$		99.506 |16.660| $[66.85 - 132.16]$		37.033 |3.960| $[29.27 - 44.80]$
$\rho = -4$		0.425 (302) |0.054| $[0.31 - 0.53]$		99.470 |16.701| $[66.71 - 132.22]$		37.079 |3.972| $[29.29 - 44.87]$
$\rho = -5$		0.424 (304) |0.055| $[0.31 - 0.53]$		99.541 |17.405| $[65.42 - 133.65]$		37.138 |4.024| $[29.25 - 45.02]$

* |sd| = standard deviation
* [IC (5%)] = Confidence Interval

Table 6.4 *Gamma estimator and quantile predictors on real data (bootstraping). the Hill's estimetors are not function of ρ.*

References

B. Basrak and A. Tafro. Extremes of moving averages and moving maxima on a regular lattice. *Probability and Mathematical Statistics*, 34:61–67, 2014.

A. Bassene. Contribution à la modélisation spatiale des événements extrêmes. 2016.

J. Beirlant, Y. Goegebeur, J. Segers, and J. L. Teugels. *Statistics of extremes: theory and applications*. John Wiley & Sons, 2006.

J. Blanchet and A. C. Davison. Spatial modeling of extreme snow depth. *The Annals of Applied Statistics*, pages 1699–1725, 2011.

S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, and J. Dubois. Unsupervised skin tissue segmentation for remote photoplethysmography. *Pattern Recognition Letters*, 124:82–90, 2019.

S. Bobbia, R. Macwan, Y. Benezeth, K. Nakamura, R. Gomez, and J. Dubois. Iterative boundaries implicit identification for superpixels segmentation: a real-time approach. *IEEE Access*, 2021.

C. Bolancé and M. Guillen. Nonparametric estimation of extreme quantiles with an application to longevity risk. *Risks*, 9(4):77, 2021.

G. P. Bopp, B. A. Shaby, and R. Huser. A hierarchical max-infinitely divisible spatial model for extreme precipitation. *Journal of the American Statistical Association*, 116(533):93–106, 2021.

R. C. Bradley. Some examples of mixing random fields. *The Rocky Mountain Journal of Mathematics*, pages 495–519, 1993.

V. Chavez-Demoulin and A. Guillou. Extreme quantile estimation for β-mixing time series and applications. *Insurance: Mathematics and Economics*, 83:59–74, 2018.

A. Daouia, S. Girard, and G. Stupfler. Tail expectile process and risk assessment. *Bernoulli*, 26(1):531–556, 2020.

R. A. Davis, C. Klüppelberg, and C. Steinkohl. Statistical inference for max-stable processes in space and time. *Journal of the Royal Statistical Society: SERIES B: Statistical Methodology*, pages 791–819, 2013.

A. C. Davison, S. A. Padoan, and M. Ribatet. Statistical modeling of spatial extremes. *Statistical science*, 27(2):161–186, 2012.

L. De Haan and A. Ferreira. Extreme value theory: an introduction. 2006.

L. De Haan and A. Ferreira. Adapting extreme value statistics to financial time series: dealing with bias and serial dependence. 2016.

L. De Haan and S. Resnick. On asymptotic normality of the hill estimator. *Stochastic Models*, 14(4):849–866, 1998.

L. De Haan, C. Mercadier, and C. Zhou. Adapting extreme value statistics to financial time series: dealing with bias and serial dependence. *Finance and Stochastics*, 20 (2):321–354, 2016.

J. Dedecker, P. Doukhan, G. Lang, L. R. J. Rafael, S. Louhichi, and C. Prieur. Weak dependence. In *Weak dependence: With examples and applications*, pages 9–20. Springer, 2007.

H. Drees. Weighted approximations of tail processes for β-mixing random variables. *Annals of Applied Probability*, pages 1274–1301, 2000.

H. Drees et al. Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli*, 9(4):617–657, 2003.

Y. Goegebeur and A. Guillou. Asymptotically unbiased estimation of the coefficient of tail dependence. *Scandinavian Journal of Statistics*, 40(1):174–189, 2013.

Y. Goegebeur, A. Guillou, and A. Schorgen. Nonparametric regression estimation of conditional tails: the random covariate case. *Statistics*, 48(4):732–755, 2014.

M. I. Gomes, L. De Haan, and L. Peng. Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes*, 5(4):387–414, 2002.

B. M. Hill. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, pages 1163–1174, 1975.

T. Hsing. On tail index estimation using dependent data. *The Annals of Statistics*, pages 1547–1569, 1991.

D. Kurisu, K. Kato, and X. Shao. Gaussian approximation and spatially dependent wild bootstrap for high-dimensional spatial data. *arXiv preprint arXiv:2103.10720*, 2021.

D. M. Mason. Laws of large numbers for sums of extreme values. *The Annals of Probability*, pages 754–764, 1982.

P. Ndao, A. Diop, and J.-F. Dupuy. Nonparametric estimation of the conditional tail index and extreme quantiles under random censoring. *Computational Statistics & Data Analysis*, 79:63–79, 2014.

T. Opitz. Un panorama de l'analyse des valeurs extrêmes. 2015.

T. Opitz. Modeling asymptotically independent spatial extremes based on laplace random fields. *Spatial Statistics*, 16:1–18, 2016.

S. Resnick and C. Stărică. Consistency of hill's estimator for dependent data. *Journal of Applied probability*, pages 139–167, 1995.

S. Resnick, C. Stărică, et al. Tail index estimation for dependent data. *Annals of applied Probability*, 8(4):1156–1183, 1998.

P. M. Robinson. Asymptotic theory for nonparametric regression with spatial data. *Journal of Econometrics*, 165(1):5–19, 2011.

P. Sharkey and H. C. Winter. A bayesian spatial hierarchical model for extreme precipitation in great britain. *Environmetrics*, 30(1):e2529, 2019.

E. Thibaud, R. Mutzner, and A. C. Davison. Threshold modeling of extreme spatial rainfall. *Water resources research*, 49(8):4633–4644, 2013.

K. F. Turkman, M. A. Turkman, and J. Pereira. Asymptotic models and inference for extremes of spatio-temporal data. *Extremes*, 13(4):375–397, 2010.

J. Velthoen, C. Dombry, J.-J. Cai, and S. Engelke. Gradient boosting for extreme quantile regression. *arXiv preprint arXiv:2103.00808*, 2021.

I. Weissman. Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815, 1978.