Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels

Laurent Gardes, Stephane Girard

CONDITIONAL EXTREMES FROM HEAVY-TAILED DISTRIBUTIONS: AN APPLICATION TO THE ESTIMATION OF EXTREME RAINFALL RETURN LEVELS

Laurent Gardes and Stéphane Girard*

Team Mistis, INRIA Rhône-Alpes & Laboratoire Jean Kuntzmann
655, avenue de l'Europe, Montbonnot, 38334 Saint-Ismier Cedex, France.
* Corresponding author, Stephane.Girard@inrialpes.fr

Abstract – This paper is dedicated to the estimation of extreme quantiles and the tail index from heavy-tailed distributions when a covariate is recorded simultaneously with the quantity of interest. A nearest neighbor approach is used to construct our estimators. Their asymptotic normality is established under mild regularity conditions and their finite sample properties are illustrated on a simulation study. An application to the estimation of pointwise return levels of extreme rainfalls in the Cévennes-Vivarais region is provided.

Keywords – Conditional extreme quantiles, heavy-tailed distribution, nearest neighbor estimator, extreme rainfalls.

AMS Subject classifications – 62G32, 62G05, 62E20.

1 Introduction

An important literature is dedicated to the estimation of extreme quantiles, *i.e.* quantiles of order $1-\alpha$ with α tending to zero as the sample size increases. The most popular estimator was proposed in (Weissman 1978), in the context of heavy-tailed distributions, and adapted to Weibull-tail distributions in (Diebolt et al. 2008; Gardes and Girard 2005). We also refer to (Dekkers and de Haan 1989) for the general case.

When some covariate x is recorded simultaneously with the quantity of interest Y, the extreme quantile thus depends on the covariate and is referred in the sequel to as the conditional extreme quantile. In our real data study, we are interested in the estimation of return levels associated to extreme rainfalls as a function of the geographical location. In this case, x is a three-dimensional covariate involving the longitude, latitude and altitude.

Parametric models for conditional extremes are proposed in (Davison and Smith 1990; Smith 1989) whereas semi-parametric methods are considered in (Beirlant and Goegebeur 2003, Hall and Tajvidi 2000). Fully non-parametric estimators have been first introduced in (Davison and Ramesh 2000), where a local polynomial modeling of the extreme observations is used. Similarly, spline estimators are fitted in (Chavez-Demoulin and Davison 2005) through a penalized maximum likelihood method. In both cases, the authors focus on univariate covariates and on the finite sample properties of the estimators. These results are extended in (Beirlant and Goegebeur 2004) where local polynomial estimators are proposed for multidimensional covariates and where their asymptotic properties are established.

We propose here to estimate the conditional extreme quantile by a nearest neighbor approach. We refer to (Loftsgaarden and Quesenberry 1965) for the first asymptotic properties of the nearest neighbor density estimator and to (Stone 1977) for the regression case. As an illustration, in the above mentioned climatology study, the estimation of the return level at a given geographical point is based on rainfalls measured at the nearest raingauges. Once the selection of the nearest observations is achieved, extreme-value methods are used to estimate the conditional quantile. Whereas no parametric assumption is made on the covariate x, we assume that the conditional distribution of Y given x is heavy-tailed. This semi-parametric assumption amounts to supposing that the conditional survival function decreases at a polynomial rate. The conditional tail index $\gamma(x)$ drives this rate of convergence and has to be estimated before conditional extreme quantiles. In our real data study, the estimation of $\gamma(x)$ permits to assess the tail-heaviness of the rainfall distribution at each geographical point x, indicating which areas are more likely to suffer from extreme climate events.

Nearest neighbor estimators of the conditional tail-index and conditional extreme quantiles are defined in Section 2. Their asymptotic distributions are derived in Section 3 and some examples are provided in Section 4. The finite sample properties of the estimators on dependent data are illustrated in Section 5 and an application to the extreme rainfall study is presented in Section 6. Proofs are postponed to Section 7.

2 Nearest neighbor estimators

Let E be a metric space associated to a metric d. For $y > 0$ and $x \in E$, denote by $F(y, x)$ the conditional distribution function of Y given x. For instance, in the case where E is finite dimensional, each coordinate of x may represent a geographical coordinate. At the opposite, when x is a time series or a curve, E is infinite dimensional. We assume that for all $x \in E$, the conditional distribution function of Y is heavy-tailed, see also Gardes and Girard (2008). More specifically, we have for all $y > 0$,

$$1 - F(y, x) = y^{-1/\gamma(x)} L(y, x), \tag{1}$$

or equivalently, for all $\alpha \in (0, 1]$,

$$q(\alpha, x) \overset{def}{=} F^{\leftarrow}(1 - \alpha, x) = \alpha^{-\gamma(x)} \ell(\alpha^{-1}, x),$$

where $F^{\leftarrow}(1 - \alpha, x) = \sup\{y > 0,\ F(y, x) \leq 1 - \alpha\}$ denotes the generalized inverse of $F(., x)$. Here, $\gamma(.)$ is an unknown positive function of the covariate x referred to as the conditional tail index. The larger $\gamma(x)$ is, the heavier is the tail at point x. Besides, for all $x \in E$ fixed, $L(., x)$ and $\ell(., x)$ are slowly varying functions, *i.e.* for all $v > 0$,

$$\lim_{y\to\infty} \frac{L(vy,x)}{L(y,x)} = \lim_{y\to\infty} \frac{\ell(vy,x)}{\ell(y,x)} = 1. \tag{2}$$

Let $(Y_1, x_1), \ldots, (Y_n, x_n)$ be a sample of independent observations from (1). For a given $t \in E$, our aim is to build an estimator of $\gamma(t)$ and, for a sequence $(\alpha_{n,t})$ tending to 0 as n goes to infinity, an estimator of $q(\alpha_{n,t}, t)$. In the sequel, $q(\alpha_{n,t}, .)$ is referred to as a conditional extreme quantile and we focus on the case where the design points $x_1, \ldots, x_n$ are non random. Let $(m_{n,t})$ be a sequence of integers such that $1 < m_{n,t} < n$ and let $\{x_1^*, \ldots, x_{m_{n,t}}^*\}$ be the $m_{n,t}$ nearest covariates of t (with respect to the distance d). The associated observations taken from $\{Y_1, \ldots, Y_n\}$ are denoted by $\{Z_1, \ldots, Z_{m_{n,t}}\}$. The corresponding order statistics are denoted $Z_{1,m_{n,t}} \leq \ldots \leq Z_{m_{n,t},m_{n,t}}$ and the rescaled log-spacings are defined for all $i = 1, \ldots, m_{n,t} - 1$ as:

$$C_{i,n,t} \stackrel{def}{=} i(\log Z_{m_{n,t}-i+1,m_{n,t}} - \log Z_{m_{n,t}-i,m_{n,t}}).$$

Our estimators of $\gamma(t)$ are linear combinations of these rescaled log-spacings:

$$\hat{\gamma}_n(t, a, \lambda) = \sum_{i=1}^{k_{n,t}} p\left(i/k_{n,t}, a, \lambda\right) C_{i,n,t} \Bigg/ \sum_{i=1}^{k_{n,t}} p\left(i/k_{n,t}, a, \lambda\right), \tag{3}$$

where $(k_{n,t})$ is a sequence of integers such that $1 < k_{n,t} < m_{n,t}$ and the weights are defined for all $s \in (0, 1)$, $a \geq 1$, $0 < \lambda \leq 1$ by

$$p(s, a, \lambda) = \frac{\lambda^{-a}}{\Gamma(a)} s^{1/\lambda - 1}(-\log s)^{a-1}. \tag{4}$$

Note that $p(., a, \lambda)$ is the density function on $(0, 1)$ introduced as the *log-gamma distribution* by (Consul and Jain 1971). Examples of such densities are provided in Section 4 and illustrated on Figure 1. Four main behaviors can be exhibited: (i) $p(., 1, 1)$ is constant, (ii) $p(., 1, \lambda)$ is increasing for all $0 < \lambda < 1$, (iii) $p(., a, 1)$ is decreasing for all $a > 1$ and (iv) $p(s, a, \lambda)$ has an unique mode at $s = \exp\{\lambda(1-a)/(1-\lambda)\}$ for $a > 1$ and $0 < \lambda < 1$. Let us highlight that other weight functions could be considered in (3) provided that Lemma 3 in Subsection 7.1 still holds. In the same spirit as the quantile estimator proposed by (Weissman 1978), the following estimator of $q(\alpha_{n,t}, t)$ can be derived from (3):

$$\hat{q}(\alpha_{n,t}, t) = Z_{m_{n,t}-k_{n,t}+1,m_{n,t}} \left(\frac{k_{n,t}}{m_{n,t}\alpha_{n,t}}\right)^{\hat{\gamma}_n(t,a,\lambda)},$$

where $(\alpha_{n,t})$ is a sequence in $(0, 1)$. The limiting distributions of these estimators are established in the next section.

3 Asymptotic results

We first give all the conditions and notations required to obtain the asymptotic normality of our estimators. In the sequel, we fix $t \in E$ such that $\gamma(t) > 0$.

(A.1) The slowly varying function $\ell(., t)$ is normalized.

Assumption **(A.1)** is equivalent to supposing that, for $\alpha \in (0,1)$, the Karamata representation of $q(\alpha, t)$ can be simplified as:

$$q(\alpha, t) = c(t) \exp\left\{ \int_1^{1/\alpha} \frac{\gamma(t) + \Delta(v,t)}{v} dv \right\}, \tag{5}$$

with $c(t) > 0$ and where $\Delta(v,t)$ converges to 0 as v goes to infinity. Note also that condition **(A.1)** implies (1), see for instance (Bingham et al. 1987; Geluk and de Haan 1987). The next two assumptions control the rate of convergence of the function $\Delta(., t)$ to zero.

(A.2) The function $\Delta(., t)$ is regularly varying with index $\rho(t) < 0$, *i.e.* for all $v > 0$, $\Delta(vy, t)/\Delta(y,t) \to v^{\rho(t)}$ as $y \to \infty$.

Conditions **(A.1)** and **(A.2)** imply that for all $v > 0$,

$$\log\left(\frac{\ell(vy,t)}{\ell(y,t)} \right) = \Delta(y,t) \frac{1}{\rho(t)} (v^{\rho(t)} - 1)(1 + o(1)),$$

which is the so-called second-order condition classically used to establish the asymptotic normality of tail-index estimators. The second-order parameter $\rho(t)$ controls the rate of convergence of $\Delta(v,t)$ to 0 *i.e.* the rate of convergence of $\ell(vy,t)/\ell(y,t)$ to 1 in equation (2). If $\rho(t)$ is close to 0, this convergence is slow and thus the estimation of the conditional tail index and of the conditional extreme quantile are difficult.

(A.3) The function $|\Delta(., t)|$ is ultimately decreasing.

In the following, we denote by $\mathcal{V}_{n,t}$ the set $\{t, x_1^*, \ldots, x_{m_{n,t}}^*\} \subset E$. The largest oscillation of the log-quantile function with respect to its second variable is defined for all $\beta \in (0, 1/2)$ as

$$\omega_n(\beta) = \sup\left\{ |\log q(\alpha, x) - \log q(\alpha, x')| \ , \ \alpha \in (\beta, 1-\beta) \ , \ (x, x') \in \mathcal{V}_{n,t}^2 \right\}.$$

We also assume that $(k_{n,t})$ is an intermediate sequence which is a classical assumption in extreme value theory.

(B) $m_{n,t}/k_{n,t} \to \infty$ and $k_{n,t} \to \infty$ as $n \to \infty$.

We are now in position to state our asymptotic normality result for $\hat{\gamma}_n(t, a, \lambda)$.

Theorem 1. *Suppose* **(A.1)**−**(A.3)**, **(B)** *hold. If, moreover, for some* $\delta > 0$,

$$k_{n,t}^{1/2} \Delta(m_{n,t}/k_{n,t}, t) \to \xi(t) \in \mathbb{R} \text{ and } k_{n,t}^2 \omega_n(m_{n,t}^{-(1+\delta)}) \to 0 \tag{6}$$

then

$$k_{n,t}^{1/2} \left(\hat{\gamma}_n(t, a, \lambda) - \gamma(t) - \Delta\left(m_{n,t}/k_{n,t}, t\right) \mathcal{AB}(a, \lambda, \rho(t)) \right)$$

converges in distribution to a $\mathcal{N}\left(0, \gamma^2(t) \mathcal{AV}(a, \lambda)\right)$ *random variable where*

$$\mathcal{AB}(a, \lambda, \rho(t)) = (1 - \lambda\rho(t))^{-a} \text{ and } \mathcal{AV}(a, \lambda) = \frac{\Gamma(2a-1)}{\lambda \Gamma^2(a)} (2 - \lambda)^{1-2a}.$$

The first part of condition (6) is standard in the extreme-value theory. It prevents the bias of the estimate from being too large compared to the standard-deviation. The second part of the condition is due to our conditional framework. It is dedicated to the control of the variations with respect to the covariate. For instance, if the slowly varying function ℓ does not depend on the covariate, the second part of condition (6) reduces to a regularity condition on the tail-index:

$$k_{n,t}^2 \log(m_{n,t}) \sup_{(x,x')\in\mathcal{V}_{n,t}^2} |\gamma(x)-\gamma(x')| \to 0 \text{ as } n\to\infty.$$

The following result establishes that $\hat{q}(\alpha_{n,t},t)$ inherits its asymptotic distribution from $\hat{\gamma}_n(t,a,\lambda)$.

Theorem 2. *Suppose* **(A.1)**−**(A.3)**, **(B)** *and condition (6) hold. If, moreover,* $m_{n,t}\alpha_{n,t} < k_{n,t}$ *then*

$$\frac{k_{n,t}^{1/2}}{\log\left(\frac{k_{n,t}}{m_{n,t}\alpha_{n,t}}\right)}\left(\log\left(\frac{\hat{q}(\alpha_{n,t},t)}{q(\alpha_{n,t},t)}\right) - \log\left(\frac{k_{n,t}}{m_{n,t}\alpha_{n,t}}\right)\Delta\left(\frac{m_{n,t}}{k_{n,t}},t\right)\mathcal{AB}(a,\lambda,\rho(t))\right)$$

converges in distribution to a $\mathcal{N}\left(0,\gamma^2(t)\mathcal{AV}(a,\lambda)\right)$ *random variable.*

The asymptotic bias of estimators $\hat{\gamma}_n(.,a,\lambda)$ and $\hat{q}(\alpha_{n,t},.)$ are both proportional to $\mathcal{AB}(a,\lambda,\rho(t))$ while their asymptotic variances are proportional to $\mathcal{AV}(a,\lambda)$. These quantities can be controlled by an appropriate choice of a and λ, see Section 4 for a discussion on this topic. Concerning the asymptotic variance, the proportionality factor is $\gamma^2(t)$. Hence, the heavier the tail is, the larger the asymptotic variance is. Moreover, the asymptotic variance can be lower bounded since

$$\mathcal{AV}(a,\lambda)-1 = \int_0^1 (p(s,a,\lambda)-1)^2 ds \geq 0$$

which entails that $\mathcal{AV}(a,\lambda)\geq 1$ for all $a\geq 1$ and $\lambda\in(0,1]$. It is thus clear that the minimum variance estimator is obtained with the uniform distribution ($a=\lambda=1$). Let us also highlight that $\mathcal{AB}(a,\lambda,\rho(t))$ is an increasing function of $\rho(t)$. Thus, the closer $\rho(t)$ is to zero, the larger is the asymptotic bias. However, the second-order parameter $\rho(t)$ is unknown in practice making difficult the comparison of asymptotic bias associated to different log-gamma weights. We refer to (Gomes et al. 2003; Gomes et al. 2004) for estimators of the second-order parameter in the unconditional case. To overcome this problem, one can define the *mean-squared bias* as:

$$\mathcal{MSB}(a,\lambda) = \int_{-\infty}^0 \mathcal{AB}^2(a,\lambda,\rho)d\rho = \frac{1}{\lambda(2a-1)}.$$

Note that the mean-squared bias converges to 0 as a tends to infinity. It it thus not possible to define in our family a minimum mean-squared bias estimator.

4 Choice of log-gamma parameters

4.1 Nearest neighbor Hill estimator

As remarked in the previous section, the minimum variance estimator is obtained by letting $a = \lambda = 1$ in (4). This choice yields

$$\hat{\gamma}_n^{\mathrm{H}}(t) = \hat{\gamma}_n(t, 1, 1) = \frac{1}{k_{n,t}} \sum_{i=1}^{k_{n,t}} C_{i,n,t}.$$

which is an adaptation of the classical Hill estimator (Hill 1975) to our conditional framework. In the following, this estimator is referred to as the nearest neighbor Hill estimator. The asymptotic normality of $\hat{\gamma}_n^{\mathrm{H}}(t)$ is a direct consequence of Theorem 1 with $\mathcal{MSB}(1,1) = 1$ and $\mathcal{AV}(1,1) = 1$. The associated conditional quantile estimator $\hat{q}^{\mathrm{H}}(\alpha_{n,t}, t)$ admits the same limiting distribution as in Theorem 2.

4.2 Nearest neighbor Zipf estimator

The Zipf estimator, initially introduced in the unconditional case (Kratz and Resnick 1996; Schultze and Steinebach 1996), can be adapted to our framework by remarking that, for $i = 1, \dots, k_{n,t}$, the pairs $\left(\log(m_{n,t}/i), \log Z_{m_{n,t}-i+1, m_{n,t}}\right)$ are approximately distributed on a line of slope $\gamma(t)$. Then, a least-squares estimation yields the following estimator of $\gamma(t)$:

$$\sum_{i=1}^{k_{n,t}} \mu_{i,n,t} C_{i,n,t} \Bigg/ \sum_{i=1}^{k_{n,t}} \mu_{i,n,t} \ , \quad \text{where } \mu_{i,n,t} = \log(k_{n,t}!)/k_{n,t} - \log(i!)/i.$$

It can be shown that $\mu_{i,n,t}$ is asymptotically equivalent to $\log(k_{n,t}/i)$, and thus, the nearest neighbor Zipf estimator is defined as

$$\hat{\gamma}_n^{\mathrm{Z}}(t) = \hat{\gamma}_n(t, 2, 1) = \sum_{i=1}^{k_{n,t}} \log(k_{n,t}/i) C_{i,n,t} \Bigg/ \sum_{i=1}^{k_{n,t}} \log(k_{n,t}/i) \ .$$

Theorem 1 holds for this estimator with $\mathcal{MSB}(2,1) = 1/3$ and $\mathcal{AV}(2,1) = 2$. Similarly, Theorem 2 also holds for the conditional quantile estimator $\hat{q}^{\mathrm{Z}}(\alpha_{n,t}, t)$ derived from the nearest neighbor Zipf estimator.

4.3 Controlling the asymptotic mean-squared error

Following Theorem 1, the *asymptotic mean-squared error* of the estimator $\hat{\gamma}_n(t, a, \lambda)$ can be defined as

$$\mathcal{AMSE}(a, \lambda) = \Delta^2\left(\frac{m_{n,t}}{k_{n,t}}, t\right) \mathcal{MSB}(a, \lambda) + \frac{\gamma^2(t)\mathcal{AV}(a, \lambda)}{k_{n,t}},$$

One way to choose the log-gamma parameters could be to minimize the asymptotic mean-squared error. In practice, the function Δ is unknown and thus the asymptotic mean-squared error cannot be evaluated. To overcome this problem, it is possible to introduce an upper bound on $\mathcal{AMSE}(a, \lambda)$. Letting $\pi(a, \lambda) = \mathcal{MSB}(a, \lambda)\mathcal{AV}(a, \lambda)$, we obtain for all $\lambda \in (0, 1]$ and $a \in [1, a_{\max}]$,

$$\begin{aligned}\mathcal{AMSE}(a, \lambda) &= \frac{\pi(a, \lambda)}{k_{n,t}} \left\{ \frac{\xi^2(t) + o(1)}{\mathcal{AV}(a, \lambda)} + \frac{\gamma^2(t)}{\mathcal{MSB}(a, \lambda)} \right\} \\ &\leq \frac{\pi(a, \lambda)}{k_{n,t}} \left\{ \xi^2(t) + o(1) + \gamma^2(t)(2a_{\max} - 1) \right\}.\end{aligned}$$

We thus propose to consider the log-gamma parameters (a_π, λ_π) minimizing $\pi(a, \lambda)$, leading to the estimator $\hat{\gamma}_n^\pi(t, a, \lambda) = \hat{\gamma}_n(t, a_\pi, \lambda_\pi)$. Annulling the partial derivative of $\pi(a, \lambda)$ with respect to λ yields $\lambda_\pi = 4/(1 + 2a_\pi)$ whereas it is not possible to find an explicit value for a_π. A numerical optimization yields $a_\pi \approx 2.19$. Theorem 1 holds with $\mathcal{MSB}(a_\pi, \lambda_\pi) \approx 0.40$ and $\mathcal{AV}(a_\pi, \lambda_\pi) \approx 1.51$, and Theorem 2 holds for the corresponding conditional quantile estimator $\hat{q}^\pi(\alpha_{n,t}, t)$.

4.4 Discussion

The three previously introduced log-gamma densities are represented on Figure 1. The nearest neighbor Hill estimator gives the same weight to all the $k_{n,t}$ largest observations. The nearest neighbor Zipf estimator corresponds to a decreasing log-gamma density. Finally, the log-gamma density used in $\hat{\gamma}_n^\pi(.)$ has a mode in $(0, 1)$. A heavy left tail for the log-gamma distribution (4) gives large weights to large observations in (3) and yields large asymptotic variances:

$$\mathcal{AV}(2, 1) > \mathcal{AV}(a_\pi, \lambda_\pi) > \mathcal{AV}(1, 1).$$

Asymptotic bias have an opposite behavior:

$$\mathcal{MSB}(2, 1) < \mathcal{MSB}(a_\pi, \lambda_\pi) < \mathcal{MSB}(1, 1).$$

It is thus not possible to find log-gamma parameters giving rise to the best estimator both in terms of asymptotic bias and variance. However, for a given mean-squared bias, the best asymptotic variance can be computed. Letting $\mathcal{MSB}(a, \lambda) = b$, we obtain $\lambda(a, b) = 1/(b(2a - 1))$ and consequently

$$\mathcal{AV}(a, \lambda(a, b)) = b\frac{\Gamma(2a)}{\Gamma^2(a)} \left\{ 2 - \frac{1}{b(2a - 1)} \right\}^{1-2a},$$

where $a \geq \max\{1, (1 + b)/(2b)\}$ in order to ensure $0 < \lambda(a, b) \leq 1$. The *optimal asymptotic variance* for a fixed mean-squared bias b is obtained by minimizing this quantity with respect to a:

$$\mathcal{OAV}(b) = \min_{a \geq \max\{1,(1+b)/(2b)\}} \mathcal{AV}(a, \lambda(a, b)).$$

Here again, an explicit solution is not available. The graph of the function $\mathcal{OAV}$ obtained by numerical optimization is depicted on Figure 2. Some level curves of $\pi(a, \lambda)$ are also represented. It appears that $\hat{\gamma}_n^\pi$ and $\hat{\gamma}_n^{\mathrm{H}}$ can be considered as optimal estimators since, for a fixed value of the mean-squared bias, they have the optimal asymptotic variance. In contrast, the nearest neighbor Zipf estimator is not optimal. It is possible to build an estimator with the same mean-squared bias ($= 1/3$) and smaller asymptotic variance (≈ 1.85).

5 Simulation study

The asymptotic properties of our estimators are established in Section 3 under an independence assumption. This hypothesis may not be verified on rainfall data where temporal and/or spatial dependence is expected, see Section 6. The impact of temporal and spatial dependence on the bias and variance of the tail-index estimator $\hat{\gamma}_n^\pi$ is illustrated in Subsection 5.1 and 5.2 respectively.

5.1 Temporal dependence

Here, for the sake of simplicity, we do not introduce a covariate information, the conditional estimator being studied in the next paragraph. A temporal series $\{y_1, \dots, y_n\}$ with $n = 500$ is generated following the method proposed by (Fawcett and Walshaw 2007): First, a temporal series $\{f_1, \dots, f_n\}$ with standard Fréchet margins is simulated. The joint distribution of (f_i, f_{i+1}), $i = 1, \dots, n-1$ is given by a bivariate extreme-values distribution $G(u, v) = \exp\{-V(u, v)\}$ with a logistic dependence function $V(u, v) = (u^{-1/\alpha} + v^{-1/\alpha})^\alpha$, $u > 0$, $v > 0$ and $\alpha \in (0, 1]$. Note that α tunes the dependence between two consecutive observations: $\alpha = 1$ leads to the independent case while $\alpha \to 0$ corresponds to complete dependence. The following procedure is used:

1. Simulate the first observation f_1 from the standard Fréchet distribution.
2. For $i = 1, \dots, n-1$: Compute the conditional distribution of f_{i+1} given f_i and simulate f_{i+1} from this distribution.

Finally, the temporal series $\{f_1, \dots, f_n\}$ is transformed so that the margins are Burr distributed. Recall that the Burr distribution function is given for $y \geq 0$ by $1 - (1 + y^{-\rho/\gamma})^{1/\rho}$ where ρ is the second order parameter as defined in condition **(A.2)**. Here, we set $\rho = -1$ and $\gamma = 0.2$. Using this strategy, $N = 100$ temporal series with Burr margins are simulated. The estimator $\hat{\gamma}_n^\pi$ is computed on each replication, leading to N values $\{\hat{\gamma}_{n,1}^\pi, \dots, \hat{\gamma}_{n,N}^\pi\}$, with

$$\hat{\gamma}_{n,j}^\pi = \sum_{i=1}^{k} p\left(\frac{i}{k}, a_\pi, \lambda_\pi\right) i \log\left(\frac{y_{n-i+1,n}^{(j)}}{y_{n-i,n}^{(j)}}\right) \Bigg/ \sum_{i=1}^{k} p\left(\frac{i}{k}, a_\pi, \lambda_\pi\right),$$

where $y_{1,n}^{(j)} \leq \ldots \leq y_{n,n}^{(j)}$ is the $j-$th temporal series ranked in ascending order. The empirical squared bias ($\mathcal{ESB}$) and the empirical variance ($\mathcal{EV}$) defined by

$$\mathcal{ESB} = \frac{1}{N}\sum_{j=1}^{N}(\hat{\gamma}_{n,j}^\pi - \gamma)^2 \text{ and } \mathcal{EV} = \frac{1}{N}\sum_{j=1}^{N}\left(\hat{\gamma}_{n,j}^\pi - \frac{1}{N}\sum_{j=1}^{N}\hat{\gamma}_{n,j}^\pi\right)^2$$

are represented on Figures 3 and 4 as functions of the sample fraction k and for different values of the dependence coefficient $\alpha \in \{1, 0.8, 0.5, 0.2\}$. It appears on Figure 3 that, even for a strong dependence ($\alpha = 0.2$), a good choice of k leads to an estimator with small bias. Thus, focusing on the bias, the main problem in estimating the tail index with temporally dependent observations is more on the choice of the sample fraction than on the dependence degree. Turning to the variance of $\hat{\gamma}_n^\pi$, Figure 4 shows that it increases with the temporal dependence degree. Note that these remarks are consistent with the conclusions drawn in (Fawcett and Walshaw 2007).

5.2 Spatial dependence

This paragraph is dedicated to the illustration of spatial dependence consequences on the conditional tail index estimation. Using Subsection 5.1 strategy, $n_s = 10$ independent temporal series $\{s_1, \dots, s_{n_s}\}$ of size 500 with standard normal margins are simulated. The temporal dependence coefficient is fixed to $\alpha = 0.5$. Here, n_s can be interpreted as the number of gauged stations and the total number of observations is thus $n = 5000$. Spatial correlation is introduced by defining $(s'_1, \dots, s'_{n_s}) = (s_1, \dots, s_{n_s})A(\theta)$ where $A(\theta)$ is a $n_s \times n_s$ circulant matrix defined for all $\theta \in [0, 1]$ and $(i, j) \in \{1, \dots, n_s\}^2$ by:

$$A_{i,j}(\theta) = \begin{cases} 1/\delta & \text{if } \delta > (j-i) \text{ modulo } n_s, \\ 0 & \text{otherwise} \end{cases}$$

with $\delta = \lfloor n_s - (n_s - 1)\theta \rfloor$, $\lfloor . \rfloor$ denoting the integer part. Parameter θ tunes the spatial dependence between the temporal series $\{s'_1, \dots, s'_{n_s}\}$: Each series s'_j is the mean of δ temporal series taken in the set $\{s_1, \dots, s_{n_s}\}$. For instance, $\theta = 1$ leads to $\delta = 1$, $A(1)$ is the identity matrix and thus $s'_1, \dots, s'_{n_s}$ are independent. At the opposite, $\theta = 0$ yields $\delta = n_s$ and $s'_1 = \dots = s'_{n_s}$ which corresponds to the complete dependence case. An intermediate case is $\theta = 4/5$ which leads to $\delta = 2$ and $s'_1 = (s_1 + s_2)/2$, $s'_2 = (s_2 + s_3)/2$, $\dots$ when $n_s = 10$.

Finally, each series s'_j, $j = 1, \dots, n_s$ is transformed so that its margins are Burr distributed with second order parameter $\rho = -1$ and conditional tail index $\gamma_j = 0.16 + j(0.26 - 0.16)/n_s$. In order to illustrate the influence of the spatial dependence on the estimation of the conditional tail index, the empirical mean squared error (defined by $\mathcal{EMSE} = \mathcal{EV} + \mathcal{ESB}$) is computed on $N = 100$ independent replications of the n_s temporal series for $\theta \in \{1, 0.8, 0.5, 0.2\}$ and $m \in \{500, 500 \times 2, \dots, 500 \times n_s\}$. Since the effect of the sample fraction k has been investigated in Subsection 5.1, it is now fixed to the value minimizing the $\mathcal{EMSE}$. Results are represented on Figure 5. It appears that the spatial dependence coefficient θ does not influence much the estimation error. A more detailed study revealed that dependence slightly increases the bias but decreases the variance, leading to small fluctuations of the empirical mean squared error. Besides, it is also apparent that taking account of nearest neighbors permits to reduce the $\mathcal{EMSE}$.

6 Application to rainfall data

Extreme rainfall statistics are often used when a flood occurred to assess the rarity of such an event. A typical question is to estimate what is the amount of rain on one hour that is expected to be exceeded once every T years. Mathematically speaking, the problem is to estimate the T-years quantile $q(1/(365 \times 24T), .)$ of the hourly rainfall.

In (Coles and Tawn 1996), a Bayesian approach is used to model extreme precipitations at a location in south-west England. The excess distribution is represented by a Generalized Pareto Distribution (GPD) with some prior information on its parameters. (Cooley et al. 2007) take profit of the Bayesian framework to model GPD parameters with stochastic processes depending on some geographical variables. An alternative approach consists in modeling the rainfall process itself by a max-stable process. We refer to (Buishand et al. 2008;

Padoan et al. 2009) for applications to the daily rainfalls in the Netherlands and in the USA respectively. Here, we consider hourly rainfall observations at 142 stations in the Cévennes-Vivarais region (southern part of France) from 1993 to 2000. In this context, the variable of interest Y is the hourly rainfall and the covariate x is the three dimensional geographical location (x_1 is the longitude, x_2 is the latitude and x_3 is the altitude). The set of coordinates $S = \{(x_{1,j}, x_{2,j}, x_{3,j}),\ j = 1, \dots, 142\}$ of the raingauge stations is depicted on Figure 6. The total number of observations is $n = 264056$. The extreme rainfall in the Cévennes-Vivarais region has already been studied in (Bois et al. 1997). The data consisted in hourly rainfalls measured at 48 raingauge stations from 1948 to 1991. The 10-years quantile is estimated under a Gumbel assumption and using a kriging technique.

Let us first focus on the estimation of the conditional tail index $\gamma(x)$ as a function of x. To this aim, the three previously described estimators $\hat{\gamma}_n^{\mathrm{H}}$, $\hat{\gamma}_n^{\mathrm{Z}}$ and $\hat{\gamma}_n^{\pi}$ are used. All of them depend on the choice of $m_{n,t}$ and $k_{n,t}$. For the sake of simplicity, these parameters are chosen to be independent of the location t. They are selected by minimizing some dissimilarity measure between the estimators:

$$(\hat{k}, \hat{m}) = \arg\min_{k,m} \sum_{t \in S} \mathcal{D}(\hat{\gamma}_n^{\mathrm{H}}(t), \hat{\gamma}_n^{\mathrm{Z}}(t), \hat{\gamma}_n^{\pi}(t))$$

with $\mathcal{D}(u_1, u_2, u_3) = (u_1 - u_2)^2 + (u_2 - u_3)^2 + (u_3 - u_1)^2$. This heuristics is sometimes used in nonparametric estimation. It relies on the idea that, for a properly chosen pair $(\hat{k}, \hat{m})$, all three estimates should approximately give the same tail index. We refer to (Gardes et al. 2010) for an illustration of this procedure on simulated data. It appears that it behaves similarly to an Oracle method which minimizes the distance to the true function $\gamma(t)$ associated to the simulated data. Here, this procedure yields $\hat{m}/n = 25\%$ and $\hat{k}/\hat{m} = 0.1\%$. Note that the graphical representation of the estimated tail index as a function of a three dimensional covariate is not possible. The role of the altitude x_3 is illustrated on Figure 7 while the role of the planar coordinates (x_1, x_2) is represented on Figure 8. The shapes of the three curves representing the estimated tail index as a function of the altitude are qualitatively the same. The estimated tail indices are decreasing functions of the altitude till $x_3 = 400$ meters and are increasing for altitudes ranging from 400 and 1600 meters. For the visualization sake, the estimators have been computed on a regular grid over the considered geographical region. This grid is depicted on Figure 6 and involves 60×50 ungauged locations. The result are presented on Figure 8 where $\hat{\gamma}_n^{\pi}$ is represented as a function of the longitude and latitude. The large estimated values ($\hat{\gamma}_n^{\pi} \in [0.15, 0.28]$) are consistent with the credibility intervals found in (Coles and Tawn, 1996) but contradict the Gumbel assumption of (Bois et al. 1997). To compare both approaches, we focus on three stations: Deaux, Bedoin and Chateauneuf-de-Gadagne (localized from left to right by a $*$ on Figure 8). These three stations measured hourly rainfalls larger than 100 millimeters (mm). For each of the above stations, $\hat{\gamma}_n^{\pi}$ is computed without using neighbor stations, $i.e$ using a standard pointwise approach. The number of upper order statistics varies for each station such that $k_{n,t}/m_{n,t} \in \{0, \dots, 50\%\}$. Results are presented on Figure 9. It appears that, for all values of $k_{n,t}$, the estimated tail indices are larger than 0.35. The Gumbel assumption seems therefore to be unrealistic. Let us note that, when the neighborhood information is taken into account (Figure 8), the geographical smoothing leads to smaller estimated tail indices, but still larger than 0.15.

Similar results are obtained concerning the 10-years return level. It appears on Figure 10 that the considered return level is globally decreasing with the altitude. However, the observed variability indicates that altitude is not the unique factor. Indeed, one can see on Figure 10 that the Valence area of Rhône Valley does not suffer from high pointwise return levels whereas the southern part does.

The drift of the rainfall rate as a function of the altitude is in agreement with the rainfall descriptive statistics in the region (Molinié et al. 2008). Since in this region low altitude areas are flat areas and are closed to the sea, the deconvolution of physical processes involved in such an altitude-rainfall rate relationship is complex. Therefore, the enhancement of extreme rainfall rates could be either a regional specificity: the supply of warm and moist air by northward low level winds over the Mediterranean sea; or a more universal phenomena: flat areas are the most efficient in capturing the solar energy which is in turn available to involve deep convective clouds.

Unsurprisingly, the estimated return levels are higher than those found in (Bois et al. 1997), ranging from $30mm$ to $65mm$ under the Gumbel assumption. In view of our data, these return levels seem under-estimated since 7 over the 142 considered stations measured rainfalls larger than $80mm$ and 3 of them measured rainfalls larger than $100mm$.

Finally, let us emphasize that our results are obtained under both an independence and a temporal stationarity assumption. Following (Fawcett and Walshaw 2007) and our simulation study (Section 5), it seems that temporal and/or spatial dependence has little effect on the estimation bias. However, the influence of dependence on the variance prevents us from directly deriving confidence intervals from the asymptotic results established in Section 3. Concerning the temporal stationarity assumption, the short observation period (7 years) does not allow to discern any trend in the time series. However, it would be interesting to take seasonal effects into account. To this end, our further work will consist in splitting the data into homogeneous time periods. Such seasonal approaches have already been considered in Bayesian models (Coles and Pericchi 2003; Coles et al. 2003).

7 Proofs

Some preliminary results are given in Subsection 7.1. Their proofs are postponed to Subsection 7.3 while main results are proved in Subsection 7.2. For the sake of simplicity, in the sequel, we note k_t for $k_{n,t}$, Δ_t for $\Delta(m_{n,t}/k_{n,t}, t)$, α_t for $\alpha_{n,t}$ and m_t for $m_{n,t}$. Letting $J_{k_t} = \{1, \dots, k_t\}$ and $J_{m_t} = \{1, \dots, m_t\}$, we finally introduce

- $\{V_i, i \in J_{m_t}\}$ a set of independent standard uniform variables,
- $V_{1,m_t} \leq \dots \leq V_{m_t,m_t}$ the associated order statistics,
- $\{F_i, i \in J_{k_t}\}$ a set of independent standard exponential random variables.

7.1 Preliminary results

The first lemma provides a representation in distribution of the logarithm of the observations whose covariate is in the neighborhood of t.

Lemma 1. *Under* **(A.1)** *and* **(A.3)***, if $k_t/m_t \to 0$ and $k_t^2 \omega_n(m_t^{-(1+\delta)}) \to 0$ for some $\delta > 0$, then, there exists a sequence of events $(\mathcal{A}_n)$ with $\mathbb{P}(\mathcal{A}_n) \to 1$ as $n \to \infty$ such that $\{\log Z_{m_t-i+1,m_t}, i \in J_{k_t} | \, \mathcal{A}_n\}$ has the same distribution as*

$$\left\{ \log q(V_{i,m_t}, t) + O_{\mathrm{P}}(\omega_n(m_t^{-(1+\delta)})), i \in J_{k_t} \middle| \mathcal{A}_n \right\}.$$

In order to be self-contained, we quote a lemma proved in (Beirlant et al. 2002). This result provides an exponential regression model for rescaled log-spacings.

Lemma 2. *Suppose* **(A.1)***,* **(A.2)** *and* **(B)** *hold. Then, the random vector $\{i(\log q(V_{i,m_t}, t) - \log q(V_{i+1,m_t}, t)),\ i \in J_{k_t}\}$ has the same distribution as*

$$\left\{ \left(\gamma(t) + \Delta\left(\frac{i}{k_t+1}, t\right)^{-\rho(t)} \right) F_i + \beta_{i,n}(t) + o_{\mathrm{P}}\left(\Delta_t\right),\ i \in J_{k_t} \right\},$$

where $\sum_{j=1}^{k_t} \left(\frac{1}{j} \int_0^{j/k_t} u(\nu) d\nu \right) \beta_{i,n}(t) = o_{\mathrm{P}}(\Delta_t)$, for every function u defined on $(0,1)$ satisfying

$$\left| k_t \int_{(j-1)/k_t}^{j/k_t} u(\nu) d\nu \right| \leq g\left(\frac{j}{k_t+1}\right), \tag{7}$$

for some fixed positive continuous function $g(., t)$ defined on $(0,1)$ and satisfying

$$\int_0^1 \max(1, \log(1/s)) g(s) ds < \infty. \tag{8}$$

In the following lemma, an integral representation of log-gamma weights is established.

Lemma 3. *Let $a \geq 1$ and $0 < \lambda \leq 1$. There exists a function u satisfying (7) and (8) such that for all $j \in J_{k_t}$,*

$$p(j/k_t, a, \lambda) = \frac{1}{j} \int_0^{j/k_t} u(\nu) d\nu.$$

Finally, the following lemma is a simple unconditioning tool for determining the asymptotic distribution of a random variable. We refer to (Gardes et al. 2010) for a proof.

Lemma 4. *Let (X_n) and (Y_n) be two sequences of real random variables. Suppose there exists a sequence of events $(\mathcal{A}_n)$ such that $(X_n | \mathcal{A}_n) \stackrel{d}{=} (Y_n | \mathcal{A}_n)$ with $\mathbb{P}(\mathcal{A}_n) \to 1$. Then, $Y_n \xrightarrow{d} Y$ implies $X_n \xrightarrow{d} Y$.*

7.2 Proofs of main results

The following result is a consequence of Lemmas 1–3. It establishes a representation of log-spacings in terms of standard exponential random variables which is the cornerstone of the proof of Theorem 1. We refer to (Falk et al. 2004, Theorem 3.5.2) for the approximation of the nearest neighbors distribution using the Hellinger distance and to (Gangopadhyay 1995) for the study of their asymptotic distribution.

Proposition 1. *Suppose* **(A.1)**, **(A.2)**, **(A.3)** *and* **(B)** *hold. If, moreover,* $k_t^2 \omega_n(m_t^{-(1+\delta)}) \to 0$ *for some* $\delta > 0$ *then*

$$\{C_{i,n,t},\ i \in J_{k_t} | \mathcal{A}_n\} \stackrel{d}{=} \{C_{i,n,t}^{(1)} + C_{i,n,t}^{(2)},\ i \in J_{k_t} | \mathcal{A}_n\}$$

where

$$\{C_{i,n,t}^{(1)},\ i \in J_{k_t}\} \stackrel{d}{=} \left\{ \left(\gamma(t) + \Delta \left(\frac{i}{k_t+1}, t \right)^{-\rho(t)} \right) F_i + \beta_{i,n}(t) + o_{\mathrm{P}}(\Delta_t),\ i \in J_{k_t} \right\}$$

$$C_{i,n,t}^{(2)} = O_{\mathrm{P}}\left(k_t \omega_n(m_t^{-(1+\delta)}) \right) \text{ uniformly in } i \in J_{k_t}.$$

and with

$$\frac{1}{k_t} \sum_{i=1}^{k_t} p(i/k_t, a, \lambda) \beta_{i,n}(t) = o_{\mathrm{P}}(\Delta_t). \tag{9}$$

Proof of Proposition 1 – From Lemma 1,

$$\begin{aligned} \{C_{i,n,t},\ i \in J_{k_t} | \mathcal{A}_n\} &\stackrel{d}{=} \left\{ i \log \left(\frac{q(V_{i,m_t}, t)}{q(V_{i+1,m_t}, t)} \right) + i O_{\mathrm{P}}\left(\omega_n(m_t^{-(1+\delta)}) \right),\ i \in J_{k_t} | \mathcal{A}_n \right\} \\ &= \{C_{i,n,t}^{(1)} + C_{i,n,t}^{(2)},\ i \in J_{k_t} | \mathcal{A}_n\}, \end{aligned}$$

where $C_{i,n,t}^{(1)} = i(\log q(V_{i,m_t}, t) - \log q(V_{i+1,m_t}, t))$ and $C_{i,n,t}^{(2)} = i O_{\mathrm{P}}(\omega_n(m_t^{-(1+\delta)}))$. From Lemmas 2 and 3, $\{C_{i,n,t}^{(1)},\ i \in J_{k_t}\}$ has the same distribution as

$$\left\{ \left(\gamma(t) + \Delta_t \left(\frac{i}{k_t+1} \right)^{-\rho(t)} \right) F_i + \beta_{i,n}(t) + o_{\mathrm{P}}(\Delta_t),\ i \in J_{k_t} \right\},$$

with $\frac{1}{k_t} \sum_{i=1}^{k_t} p(i/k_t, a, \lambda) \beta_{i,n}(t) = o_{\mathrm{P}}(\Delta_t)$. Finally, for all $i \in J_{k_t}$, we have $C_{i,n,t}^{(2)} = O_{\mathrm{P}}\left(k_t \omega_n(m_t^{-(1+\delta)}) \right)$, and the result is proved. ■

Proof of Theorem 1 – Let us consider the random variables defined as

$$\begin{aligned} \Lambda_n^{(1)} &= k_t^{1/2} \{\hat{\gamma}_n(t, a, \lambda) - \gamma(t) - \Delta_t \mathcal{AB}(a, \lambda, \rho(t))\} \\ \Lambda_n^{(2)} &= k_t^{1/2} \sum_{i=1}^{k_t} \tilde{p}(i/k_t, a, \lambda)(C_{i,n,t}^{(1)} + C_{i,n,t}^{(2)}) - \gamma(t) - \Delta_t \mathcal{AB}(a, \lambda, \rho(t)), \end{aligned}$$

where the following normalized weights $\tilde{p}(i/k_t, a, \lambda)$, $i = 1, \dots, k_t$ have been introduced

$$\tilde{p}(i/k_t, a, \lambda) = p(i/k_t, a, \lambda) \Bigg/ \sum_{j=1}^{k_t} p(j/k_t, a, \lambda)\ .$$

Proposition 1 states that $\{\Lambda_n^{(1)} | \mathcal{A}_n\} \stackrel{d}{=} \{\Lambda_n^{(2)} | \mathcal{A}_n\}$. From Lemma 4, to prove Theorem 1, it is sufficient to show that $\Lambda_n^{(2)}$ converges in distribution to a $\mathcal{N}(0, \gamma^2(t)\mathcal{AV}(a,\lambda))$ random variable. Introducing

$$T_{1,n} = \sum_{i=1}^{k_t} p(i/k_t, a, \lambda)(F_i - 1), \quad T_{2,n} = \sum_{i=1}^{k_t} p(i/k_t, a, \lambda)\left(\frac{i}{k_t+1}\right)^{-\rho(t)}(F_i - 1),$$
$$T_{3,n} = \sum_{i=1}^{k_t} p(i/k_t, a, \lambda)\beta_{i,n}(t), \quad T_{4,n} = \Delta_t \sum_{i=1}^{k_t} p(i/k_t, a, \lambda)\left(\frac{i}{k_t+1}\right)^{-\rho(t)},$$
$$T_{5,n} = \sum_{i=1}^{k_t} p(i/k_t, a, \lambda), \quad T_{6,n} = \left(\sum_{i=1}^{k_t} p^2(i/k_t, a, \lambda)\right)^{1/2},$$

Proposition 1 entails the following expansion:

$$\begin{aligned}\frac{T_{5,n}}{T_{6,n}}\left(\sum_{i=1}^{k_t} \tilde{p}(i/k_t, a, \lambda)C_{i,n,t}^{(1)} - \gamma(t) - \frac{T_{4,n}}{T_{5,n}}\right) &\stackrel{d}{=} \gamma(t)\frac{T_{1,n}}{T_{6,n}} + \Delta_t\frac{T_{2,n}}{T_{6,n}} + \frac{T_{3,n}}{T_{6,n}} \\ &+ \frac{T_{5,n}}{T_{6,n}} o_{\mathrm{P}}(\Delta_t). \end{aligned} \tag{10}$$

From Lindeberg theorem, a sufficient condition for $T_{1,n}/T_{6,n} \xrightarrow{d} \mathcal{N}(0,1)$ is

$$\sum_{i=1}^{k_t} p^3(i/k_t, a, \lambda)/T_{6,n}^3 \to 0. \tag{11}$$

Since for any integrable function φ, the following convergence of Riemann sum holds,

$$\frac{1}{k_t}\sum_{i=1}^{k_t} \varphi\left(\frac{i}{k_t}\right) \to \int_0^1 \varphi(s)ds \tag{12}$$

it follows that $T_{6,n} = k_t^{1/2}\mathcal{AV}(a,\lambda)^{1/2}(1+o(1))$. Thus, using (12), we have

$$\frac{1}{k_t}\sum_{i=1}^{k_t} p^3(i/k_t, a, \lambda) = \frac{\Gamma(3a)}{\Gamma^3(a)}(3-2\lambda)^{-3a}(1+o(1)),$$

and therefore $\sum_{i=1}^{k_t} p^3(i/k_t, a, \lambda)/T_{6,n}^3 = O(k_t^{-1/2})$, showing that condition (11) is satisfied and

$$T_{1,n}/T_{6,n} \xrightarrow{d} \mathcal{N}(0,1). \tag{13}$$

Next, let us focus on $T_{2,n}/T_{6,n}$. Remarking that this term is centered with finite variance, it follows that

$$T_{2,n}/T_{6,n} = O_{\mathrm{P}}(1). \tag{14}$$

Equation (9) in Proposition 1 yields

$$T_{3,n}/T_{6,n} = o_{\mathrm{P}}(k_t^{1/2}\Delta_t) = o_{\mathrm{P}}(1). \tag{15}$$

Finally, from (12) we obtain,

$$T_{4,n}/T_{5,n} = \Delta_t \mathcal{AB}(a, \lambda, \rho(t))(1+o(1)), \tag{16}$$
$$T_{5,n}/T_{6,n} = k_t^{1/2}\mathcal{AV}(a,\lambda)^{-1/2}(1+o(1)). \tag{17}$$

Replacing (13)–(17) in (10) shows that

$$k_t^{1/2}\left(\sum_{i=1}^{k_t}\tilde{p}(i/k_t,a,\lambda)C_{i,n,t}^{(1)}-\gamma(t)-\Delta_t\mathcal{AB}(a,\lambda,\rho(t))\right)$$

converges in distribution to a $\mathcal{N}(0,\gamma^2(t)\mathcal{AV}(a,\lambda))$ random variable. Taking account of

$$k_t^{1/2}\sum_{i=1}^{k_t}\tilde{p}(i/k_t,a,\lambda)C_{i,n,t}^{(2)}=O_{\mathrm{P}}(k_t^{3/2}\omega_n(m_t^{-(1+\delta)}))=o_{\mathrm{P}}(1)$$

concludes the proof. ∎

Proof of Theorem 2 – Let us consider the following expansion:

$$\begin{aligned}
\log\left(\frac{\hat{q}(\alpha_t,t)}{q(\alpha_t,t)}\right) &- \log\left(\frac{k_t}{m_t\alpha_t}\right)\Delta_t\mathcal{AB}(a,\lambda,\rho(t))\\
&= \log\left(\frac{Z_{m_t-k_t+1,m_t}}{q(k_t/m_t,t)}\right)\\
&+ \{\hat{\gamma}_n(t,a,\lambda)-\gamma(t)-\Delta_t\mathcal{AB}(a,\lambda,\rho(t))\}\log\left(\frac{k_t}{m_t\alpha_t}\right)\\
&- \left\{\log\left(\frac{q(\alpha_t,t)}{q(k_t/m_t,t)}\right)-\gamma(t)\log\left(\frac{k_t}{m_t\alpha_t}\right)\right\}\\
&= \zeta_{1,n}+\zeta_{2,n}+\zeta_{3,n}.
\end{aligned}$$

First, from Lemma 1, $\zeta_{1,n}\stackrel{d}{=}\log q(V_{k_t,m_t},t)-\log q(k_t/m_t,t)+O_{\mathrm{P}}(\omega_n(m_t^{-(1+\delta)}))$, and assumptions **(A.1)**, **(A.3)** imply that

$$\left|\log\left(\frac{q(V_{k_t,m_t},t)}{q(k_t/m_t,t)}\right)\right|=O_{\mathrm{P}}(\Delta_t)\left|\log\left(\frac{m_t}{k_t}V_{k_t,m_t}\right)\right|.$$

Moreover, under **(B)**, it is well-known that $k_t^{1/2}((m_t/k_t)V_{k_t,m_t}-1)\xrightarrow{d}\mathcal{N}(0,1)$, (see for instance (Girard 2004)) and thus

$$\zeta_{1,n}=O_{\mathrm{P}}(\Delta_t k_t^{-1/2})+O_{\mathrm{P}}(\omega_n(m_t^{-(1+\delta)})). \tag{18}$$

Besides, from Theorem 1, we have

$$k_t^{1/2}/\log(k_t/(m_t\alpha_t))\zeta_{2,n}\xrightarrow{d}\mathcal{N}(0,\gamma^2(t)\mathcal{AV}(a,\lambda)), \tag{19}$$

and, finally, (5) entails that

$$\log\left(\frac{q(\alpha_t,t)}{q(k_t/m_t,t)}\right)=\gamma(t)\log\left(\frac{k_t}{m_t\alpha_t}\right)+\int_{m_t/k_t}^{1/\alpha_t}\frac{\Delta(u,t)}{u}du.$$

Since $1/\alpha_t>m_t/k_t$, **(A.3)** yields

$$|\zeta_{3,n}|\le\int_{m_t/k_t}^{1/\alpha_t}\frac{|\Delta(u,t)|}{u}du\le\Delta_t\log\left(\frac{k_t}{m_t\alpha_t}\right) \tag{20}$$

and collecting (18)-(20) concludes the proof. ∎

7.3 Proofs of auxiliary results

Proof of Lemma 1 – Under **(A1)** the function $q(.,t)$ is continuous. Since the random variables $\{Z_i, i \in J_{m_t}\}$ are independent, we have:

$$\{\log Z_i, i \in J_{m_t}\} \stackrel{d}{=} \{\log q(V_i, x_i^*),\ i \in J_{m_t}\},$$

where x_i^* is the covariate associated to Z_i. Denoting by $\psi(i)$ the random index of the covariate associated to the observation Z_{m_t-i+1,m_t}, we obtain

$$\{\log Z_{m_t-i+1,m_t},\ i \in J_{m_t}\} \stackrel{d}{=} \{\log q(V_{\psi(i)}, x_{\psi(i)}^*),\ i \in J_{m_t}\}.$$

Let us consider the event $\mathcal{A}_n = \mathcal{A}_{1,n} \cap \mathcal{A}_{2,n}$ where

$$\begin{aligned}
\mathcal{A}_{1,n} &= \left\{ \min_{i \in J_{k_t} \setminus \{k_t\}} \log\left(\frac{q(V_{i,m_t}, u_i)}{q(V_{i+1,m_t}, u_{i+1})} \right) > 0, \forall (u_1, \ldots, u_{k_t}) \in \mathcal{V}_{n,t} \right\} \text{ and} \\
\mathcal{A}_{2,n} &= \left\{ \min_{i \in J_{m_t} \setminus J_{k_t}} \log\left(\frac{q(V_{k_t,m_t}, u_{k_t})}{q(V_{i,m_t}, u_i)} \right) > 0, \forall (u_{k_t+1}, \ldots, u_{m_t}) \in \mathcal{V}_{n,t} \right\}.
\end{aligned}$$

Conditionally to $\mathcal{A}_{1,n}$, the random variables $q(V_{i,m_t}, u_i)$, $i \in J_{k_t}$ are ordered as

$$q(V_{k_t,m_t}, u_{k_t}) \le q(V_{k_t-1,m_t}, u_{k_t-1}) \le \cdots \le q(V_{1,m_t}, u_1),$$

and, conditionally to $\mathcal{A}_{2,n}$, the remaining random variables $q(V_{i,m_t}, u_i)$, $i \in J_{m_t} \setminus J_{k_t}$ are smaller since

$$\max_{i \in J_{m_t} \setminus J_{k_t}} q(V_{i,m_t}, u_i) \le q(V_{k_t,m_t}, u_{k_t}).$$

Thus, conditionally to $\mathcal{A}_n$, the k_t largest random values taken from the set $\{\log q(V_{\psi(i)}, x_{\psi(i)}^*),\ i \in J_{m_t}\}$ are $\{\log q(V_{i,m_t}, x_{\psi(i)}^*),\ i \in J_{k_t}\}$. Consequently, letting $T_i \stackrel{def}{=} x_{\psi(i)}^*$, we have:

$$\{\log Z_{m_t-i+1,m_t}, i \in J_{k_t} | \mathcal{A}_n\} \stackrel{d}{=} \{\log q(V_{i,m_t}, T_i), i \in J_{k_t} | \mathcal{A}_n\}.$$

To conclude the proof, it remains to show that

$$\log q(V_{i,m_t}, T_i) - \log q(V_{i,m_t}, t) = O_{\mathrm{P}}(\omega_n(m_t^{-(1+\delta)})), \tag{21}$$

uniformly in $i \in J_{k_t}$ and that

$$\mathbb{P}(\mathcal{A}_n) \to 1, \tag{22}$$

as $n \to \infty$. Let us consider the event $\mathcal{A}_{3,n} = \{V_{1,m_t} > \delta_{m_t}\} \cap \{V_{m_t,m_t} < 1 - \delta_{m_t}\}$, where $\delta_{m_t} = m_t^{-(1+\delta)}$. Under $\mathcal{A}_{3,n}$, we have $\delta_{m_t} < V_{i,m_t} < 1 - \delta_{m_t}$ for all $i \in J_{m_t}$. Hence,

$$\begin{aligned}
&\mathbb{P}\left(|\log q(V_{i,m_t}, T_i) - \log q(V_{i,m_t}, t)| > \omega_n(m_t^{-(1+\delta)}) \right) \\
\le\ &\mathbb{P}\left(|\log q(V_{i,m_t}, T_i) - \log q(V_{i,m_t}, t)| > \omega_n(m_t^{-(1+\delta)}) \,\middle|\, \mathcal{A}_{3,n} \right) \mathbb{P}(\mathcal{A}_{3,n}) + \mathbb{P}(\mathcal{A}_{3,n}^C),
\end{aligned}$$

where $\mathcal{A}_{3,n}^C$ is the complementary event associated to $\mathcal{A}_{3,n}$. Since under $\mathcal{A}_{3,n}$,

$$|\log q(V_{i,m_t}, T_i) - \log q(V_{i,m_t}, t)| \le \omega_n(m_t^{-(1+\delta)}),$$

for all $i \in J_{k_t}$, it is clear that

$$\mathbb{P}\left(|\log q(V_{i,m_t}, T_i) - \log q(V_{i,m_t}, t)| > \omega_n(m_t^{-(1+\delta)}) \middle| \mathcal{A}_{3,n} \right) = 0.$$

Remarking that

$$\mathbb{P}(\mathcal{A}_{3,n}) \geq \mathbb{P}(V_{1,m_t} > \delta_{m_t}) + \mathbb{P}(V_{m_t,m_t} < 1 - \delta_{m_t}) - 1 = 2\mathbb{P}(V_{1,m_t} > \delta_{m_t}) - 1 \to 1,$$

since $V_{m_t,m_t} \overset{d}{=} 1 - V_{1,m_t}$ and $\mathbb{P}(V_{1,m_t} > \delta_{m_t}) = (1 - \delta_{m_t})^{m_t} \to 1$ concludes the proof of (21). Furthermore, for all $(u_i, u_j) \in \mathcal{V}_{n,t}^2$, we have, on the one hand

$$\begin{aligned} \log\left(\frac{q(V_{j,m_t}, u_j)}{q(V_{i,m_t}, u_i)}\right) &= \log\left(\frac{q(V_{j,m_t}, t)}{q(V_{i,m_t}, t)}\right) + \log\left(\frac{q(V_{j,m_t}, u_j)}{q(V_{j,m_t}, t)}\right) + \log\left(\frac{q(V_{i,m_t}, t)}{q(V_{i,m_t}, u_i)}\right) \\ &\geq \log\left(\frac{q(V_{j,m_t}, t)}{q(V_{i,m_t}, t)}\right) - 2\omega_n(\delta_{m_t}), \end{aligned}$$

and on the other hand,

$$\begin{aligned} \min_{i \in J_{m_t} \setminus J_{k_t}} \log\left(\frac{q(V_{k_t,m_t}, u_{k_t})}{q(V_{i,m_t}, u_i)}\right) &\geq \min_{i \in J_{m_t} \setminus J_{k_t}} \log\left(\frac{q(V_{k_t,m_t}, t)}{q(V_{i,m_t}, t)}\right) - 2\omega_n(\delta_{m_t}) \\ &\geq \log\left(\frac{q(V_{k_t,m_t}, t)}{q(V_{k_t+1,m_t}, t)}\right) - 2\omega_n(\delta_{m_t}). \end{aligned}$$

Consequently, considering the event

$$\mathcal{A}_{4,n} = \left\{ \min_{i \in J_{k_t}} \log\left(\frac{q(V_{i,m_t}, t)}{q(V_{i+1,m_t}, t)}\right) > 2\omega_n(\delta_{m_t}) \right\},$$

it is clear that $\mathcal{A}_{3,n} \cap \mathcal{A}_{4,n} \subset \mathcal{A}_n$. It thus remains to prove that $\mathbb{P}(\mathcal{A}_{4,n}) \to 1$ to show (22). From **(A.1)**, for all $\alpha \in (0,1)$,

$$q(\alpha, t) = c(t) \exp\left\{ \int_1^{1/\alpha} \frac{\gamma(t) + \Delta(u,t)}{u} du \right\}.$$

Hence, for all $i \in J_{k_t}$,

$$\log\left(\frac{q(V_{i,m_t}, t)}{q(V_{i+1,m_t}, t)}\right) = \int_{1/V_{i+1,m_t}}^{1/V_{i,m_t}} \frac{\gamma(t) + \Delta(u,t)}{u} du.$$

Since $V_{1,m_t}^{-1} \geq \ldots \geq V_{k_t+1,m_t}^{-1} = (m_t/k_t)(1 + o_{\mathrm{P}}(1)) \xrightarrow{\mathrm{P}} \infty$, it follows from **(A.3)** that

$$\log\left(\frac{q(V_{i,m_t}, t)}{q(V_{i+1,m_t}, t)}\right) \geq (\gamma(t) - |\Delta(V_{k_t+1,m_t}^{-1}, t)|) \log\left(\frac{V_{i+1,m_t}}{V_{i,m_t}}\right),$$

leading to

$$\begin{aligned} \mathbb{P}(\mathcal{A}_{4,n}) &\geq \mathbb{P}\left((\gamma(t) - |\Delta(V_{k_t+1,m_t}^{-1}, t)|) \min_{i \in J_{k_t}} \log\left(\frac{V_{i+1,m_t}}{V_{i,m_t}}\right) > 2\omega_n(\delta_{m_t}) \right) \\ &\geq \mathbb{P}\left(\left\{ \min_{i \in J_{k_t}} \log\left(\frac{V_{i+1,m_t}}{V_{i,m_t}}\right) \geq \frac{4\omega_n(\delta_{m_t})}{\gamma(t)} \right\} \cap \left\{ |\Delta(V_{k_t+1,m_t}^{-1}, t)| < \gamma(t)/2 \right\} \right) \\ &\geq \mathbb{P}\left(\min_{i \in J_{k_t}} \log\left(\frac{V_{i+1,m_t}}{V_{i,m_t}}\right) \geq \frac{4\omega_n(\delta_{m_t})}{\gamma(t)} \right) + \mathbb{P}\left(|\Delta(V_{k_t+1,m_t}^{-1}, t)| < \gamma(t)/2 \right) - 1 \\ &\overset{def}{=} P_{1,m_t} + P_{2,m_t} - 1. \end{aligned}$$

In view of Rényi representation $\{i \log(V_{i+1,m_t}/V_{i,m_t}),\ i \in J_{k_t}\} \stackrel{d}{=} \{F_i,\ i \in J_{k_t}\}$, we have

$$\begin{aligned} P_{1,m_t} &= \mathbb{P}\left(\min_{i \in J_{k_t}} \frac{F_i}{i} \geq \frac{4\omega_n(\delta_{m_t})}{\gamma(t)}\right) = \prod_{i=1}^{k_t} \exp\left(-\frac{4i\omega_n(\delta_{m_t})}{\gamma(t)}\right) \\ &= \exp\left(-\frac{2}{\gamma(t)} k_t(k_t+1)\omega_n(\delta_{m_t})\right) \to 1, \end{aligned}$$

since $k_t^2\omega_n(\delta_{m_t}) \to 0$. Furthermore, $V_{k_t+1,m_t} \xrightarrow{\mathrm{P}} 0$ and $\Delta(1/\alpha, t) \to 0$ as $\alpha \to 0$ entail $P_{2,m_t} \to 1$. The conclusion follows. ■

Proof of Lemma 3 – Letting

$$u(s) = \frac{d}{ds}(sp(s,a,\lambda)) = \frac{\lambda^{-a}}{\Gamma(a)} s^{1/\lambda - 1}\left(\frac{(-\log s)^{a-1}}{\lambda} - (a-1)(-\log s)^{a-2}\right),$$

it is easily seen that $p(j/k_t, a, \lambda) = \frac{1}{j}\int_0^{j/k_t} u(\nu)d\nu$. Note that, if $a = 1$, then $|u(s)|$ is a bounded function, (7) and (8) are thus satisfied. Let us now consider the case $a \neq 1$. Introducing $\tau = (2-a)\mathbb{I}\{a \in (1,2]\}$, we have:

$$\begin{aligned} |u(s)| &\leq \frac{\lambda^{-a}}{\Gamma(a)}(-\log s)^{-\tau}\left(\frac{(-\log s)^{a-1+\tau}}{\lambda} + (a-1)(-\log s)^{a-2+\tau}\right) \\ &\leq \tilde{g}(s) \stackrel{def}{=} \begin{cases} \lambda^{-a}/\Gamma(a)(\lambda^{-1}+a-1)(-\log s)^{a-1} & \text{if } s \in [0, 1/\mathrm{e}], \\ \lambda^{-a}/\Gamma(a)(\lambda^{-1}+a-1)(-\log s)^{-\tau} & \text{if } s \in [1/\mathrm{e}, 1]. \end{cases} \end{aligned}$$

Three situations are considered:

Situation a) – If $j < k_t/\mathrm{e}$, we have:

$$\left| k_t \int_{(j-1)/k_t}^{j/k_t} u(\nu)d\nu \right| \leq \begin{cases} \lambda^{-a}/\Gamma(a)(\lambda^{-1}+a-1)(\log(k_t/(j-1)))^{a-1} & \text{for } j \neq 1, \\ \lambda^{-a}/\Gamma(a)(\log k_t)^{a-1} & \text{for } j = 1. \end{cases}$$

Besides, straightforward calculations lead to

$$\left(\log\left(\frac{k_t}{j-1}\right)\right)^{a-1} \leq 2^{a-1}\left(\log\left(\frac{k_t+1}{j}\right)\right)^{a-1} \text{ and } (\log k_t)^{a-1} < (\log(k_t+1))^{a-1}.$$

Hence,

$$\left| k \int_{(j-1)/k_t}^{j/k_t} u(\nu)d\nu \right| \leq c_1(a,\lambda)\tilde{g}\left(\frac{j}{k_t+1}\right),$$

where $c_1(a,\lambda)$ is a positive constant.

Situation b) – If $j > (k_t+1)/\mathrm{e}$, then

$$\left| k_t \int_{(j-1)/k_t}^{j/k_t} u(\nu)d\nu \right| \leq \begin{cases} \frac{\lambda^{-a}}{\Gamma(a)}\left(\frac{1}{\lambda}+a-1\right)\left(\log\left(\frac{k_t}{j}\right)\right)^{-\tau} & \text{for } j \neq k_t, \\ \frac{\lambda^{-a}}{\Gamma(a)}(k_t-1)\left(\log\left(\frac{k_t}{k_t-1}\right)\right)^{a-1+\tau}\left(\log\left(\frac{k_t}{k_t-1}\right)\right)^{-\tau} & \text{for } j = k_t. \end{cases}$$

Since $\log(k_t/(k_t-1)) < 1/(k_t-1)$ and $a-2+\tau \geq 0$, it follows that

$$(k_t-1)\left(\log\left(\frac{k_t}{k_t-1}\right)\right)^{a-1+\tau} < \left(\frac{1}{k_t-1}\right)^{a-2+\tau} \leq 1,$$

and thus

$$\left| k_t \int_{(j-1)/k_t}^{j/k_t} u(\nu) d\nu \right| \leq \begin{cases} \lambda^{-a}/\Gamma(a)(\lambda^{-1}+a-1)(\log(k_t/j))^{-\tau} & \text{for } j \neq k_t, \\ \lambda^{-a}/\Gamma(a)(\log(k_t/(k_t-1)))^{-\tau} & \text{for } j = k_t. \end{cases}$$

Remarking that

$$\left(\log\left(\frac{k_t}{j}\right)\right)^{-\tau} \leq 3^\tau \left(\log\left(\frac{k_t+1}{j}\right)\right)^{-\tau} \text{ and } \left(\log\left(\frac{k_t}{k_t-1}\right)\right)^{-\tau} < \left(\log\left(\frac{k_t+1}{k_t}\right)\right)^{-\tau},$$

we have

$$\left| k \int_{(j-1)/k_t}^{j/k_t} u(\nu) d\nu \right| \leq c_2(a,\lambda)\tilde{g}\left(\frac{j}{k_t+1}\right),$$

where $c_2(a,\lambda)$ is a positive constant.

Situation c) – If $k_t/\mathrm{e} < j < (k_t+1)/\mathrm{e}$, then

$$\left| k \int_{(j-1)/k_t}^{j/k_t} u(\nu) d\nu \right| \leq \max\left(\tilde{g}\left(\frac{j}{k_t}\right); \tilde{g}\left(\frac{j-1}{k_t}\right)\right)$$

and one can show that

$$\max\left(\tilde{g}\left(\frac{j}{k_t}\right); \tilde{g}\left(\frac{j-1}{k_t}\right)\right) \leq c_3(a,\lambda)\tilde{g}\left(\frac{j}{k_t+1}\right),$$

where $c_3(a,\lambda)$ is a positive constant.

As a conclusion, for all $j \in J_{k_t}$, (7) is satisfied with

$$g(s) = \begin{cases} c_4(a,\lambda)(-\log s)^{a-1} & \text{if } s \in [0, 1/\mathrm{e}], \\ c_4(a,\lambda)(-\log s)^{-\tau} & \text{if } s \in [1/\mathrm{e}, 1], \end{cases}$$

where $c_4(a,\lambda) = \lambda^{-a}/\Gamma(a)(\lambda^{-1}+a-1)\max(c_1(a,\lambda); c_2(a,\lambda); c_3(a,\lambda))$. Finally, we have:

$$\begin{aligned} \int_0^1 \max(1, -\log s) g(s) ds &= c_4(a,\lambda)\left\{\int_0^{1/\mathrm{e}} (-\log s)^{a-1} ds \int_{1/\mathrm{e}}^1 (-\log s)^{-\tau} ds\right\} \\ &\leq c_4(a,\lambda)(\Gamma(a+1) + \Gamma(1-\tau)) < \infty, \end{aligned}$$

since $\tau < 1$, and (8) is proved. ■

References

Beirlant, J., Dierckx, G., Guillou, A. and Stărică, C.: On exponential representations of log-spacings of extreme order statistics. Extremes, 5, 157–180 (2002)

Beirlant, J. and Goegebeur, Y.: Regression with response distributions of Pareto-type. Computational Statistics and Data Analysis, 42, 595–619 (2003)

Beirlant, J. and Goegebeur, Y.: Local polynomial maximum likelihood estimation for Pareto-type distributions. Journal of Multivariate Analysis, 89, 97–118 (2004)

Bingham, N.H., Goldie, C.M. and Teugels, J.L.: Regular variation. Encyclopedia of Mathematics and its Applications, 27, Cambridge University Press (1987)

Bois, P., Obled, C., de Saintignon, M.F. and Mailloux, H.: Atlas expérimental des risques de pluies intenses Cévennes - Vivarais (2ème édition). Pôle grenoblois d'études et de recherche pour la prévention des risques naturels, Grenoble (1997)

Buishand, T.A., de Haan, L. and Zhou, C.: On spatial extremes: with application to a rainfall problem. Annals of applied statistics, 2, 624–642 (2008)

Chavez-Demoulin, V. and Davison, A.C.: Generalized additive modelling of sample extremes. Journal of the Royal Statistical Society, series C., 54, 207–222 (2005)

Coles, S. and Pericchi, L.R.: Anticipating catastrophes through extreme value modelling. Applied Statistics, 52, 405–416 (2003)

Coles, S., Pericchi, L.R. and Sisson, S.: A fully probabilistic approach to extreme rainfall modeling. Journal of Hydrology, 273, 35–50 (2003)

Coles, S. and Tawn, J.: A Bayesian analysis of extreme rainfall data. Applied Statistics, 45, 463–478 (1996)

Consul, P.C. and Jain, G.C.: On the log-gamma distribution and its properties. Statistische Hefe, 12(2), 100–106 (1971)

Cooley, D., Nychka, D. and Naveau, P.: Bayesian spatial modeling of extreme precipitation return levels. Journal of the American Statistical Association, 102, 824–840 (2007)

Davison, A.C. and Ramesh, N.I.: Local likelihood smoothing of sample extremes. Journal of the Royal Statistical Society, series B, 62, 191–208 (2000)

Davison, A.C. and Smith, R.L.: Models for exceedances over high thresholds. Journal of the Royal Statistical Society, series B, 52, 393–442 (1990)

Dekkers, A. and de Haan, L.: On the estimation of the extreme-value index and large quantile estimation. Annals of Statistics, 17, 1795–1832 (1989)

Diebolt, J., Gardes, L., Girard, S. and Guillou, A.: Bias-reduced extreme quantiles estimators of Weibull distributions. Journal of Statistical Planning and Inference, 138, 1389–1401 (2008)

Fawcett, L. and Walshaw, D.: Improved estimation for temporally clustered extremes. Environmetrics, 18, 173–188 (2007)

Falk, M., Hüsler, J. and Reiss, R.D.: Laws of small numbers: Extremes and rare events. 2nd edition, Birkhäuser (2004)

Gangopadhyay, A.K.: A note on the asymptotic behavior of conditional extremes. Statistics and Probability Letters, 25, 163–170 (1995)

Gardes, L. and Girard, S.: Estimating extreme quantiles of Weibull tail-distributions. Communication in Statistics - Theory and Methods, 34, 1065–1080 (2005)

Gardes L. and Girard, S.: A moving window approach for nonparametric estimation of the conditional tail index. Journal of Multivariate Analysis, 99, 2368–2388 (2008)

Gardes L., Girard, S. and Lekina, A.: Functional nonparametric estimation of conditional extreme quantiles. Journal of Multivariate Analysis, 101, 419–433 (2010)

Geluk, J.L. and de Haan, L.: Regular variation, extensions and Tauberian theorems. Math Centre tracts, 40, Centre for Mathematics and Computer Science, Amsterdam (1987)

Girard, S.: A Hill type estimate of the Weibull tail-coefficient. Communication in Statistics - Theory and Methods, 33, 205–234 (2004)

Gomes, M.I., de Haan, L., Peng, L.: Semi-parametric estimation of the second order parameter in statistics of extremes. Extremes, 5, 387–414 (2003)

Gomes, M.I., Caeiro, F., Figueiredo, F.: Bias reduction of a tail index estimator through an external estimation of the second-order parameter. Statistics, 38, 497–510 (2004)

Hall, P. and Tajvidi, N.: Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. Statistical Science, 15, 153–167 (2000)

Hill, B.M.: A simple general approach to inference about the tail of a distribution. Annals of Statistics, 3, 1163–1174 (1975)

Kratz, M. and Resnick, S.: The QQ-estimator and heavy tails. Stochastic Models, 12, 699–724 (1996)

Loftsgaarden, D. and Quesenberry, C.: A nonparametric estimate of a multivariate density function. Ann. Math. Statist., 36, 1049–1051 (1965)

Molinié, G., Yates, E., Ceresetti, D., Anquetin, S., Boudevillain, B., Creutin, J.D. and Bois, P.: Rainfall regimes in a mountainous Mediterranean region: Statistical analysis at short time steps. submitted (2008)

Padoan, S., Ribatet, M. and Sisson, S.: Likelihood-based inference for max-stable processes. Journal of the American Statistical Association, to appear (2009)

Schultze, J. and Steinebach, J.: On least squares estimates of an exponential tail coefficient, Statistics and Decisions, 14, 353–372 (1996)

Smith, R. L.: Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone (with discussion). Statistical Science, 4, 367–393 (1989)

Stone, C.: Consistent nonparametric regression. Annals of Statistics, 5, 595–645 (1977)

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

Acknowledgment

This work was supported by the Agence Nationale de la Recherche (French Research Agency) through its VMC program (Vulnérabilité: Milieux, Climats). The authors are grateful to Caroline Bernard-Michel (`http://carobm.free.fr`) and Gilles Molinié (`http://ltheln21.hmg.inpg.fr/PagePerso/molinie`) for their help in the application to rainfall data.

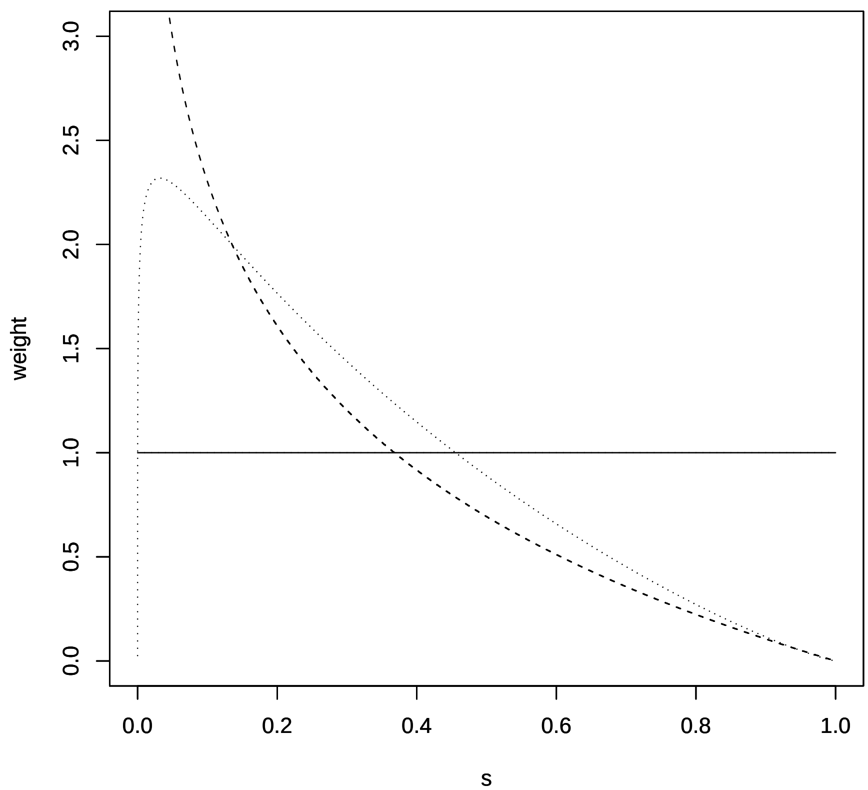

Figure 1: Log-gamma densities associated to $\hat{\gamma}_n^{\text{H}}$ (full line), $\hat{\gamma}_n^{\text{Z}}$ (dashed line) and $\hat{\gamma}_n^{\pi}$ (dotted line).

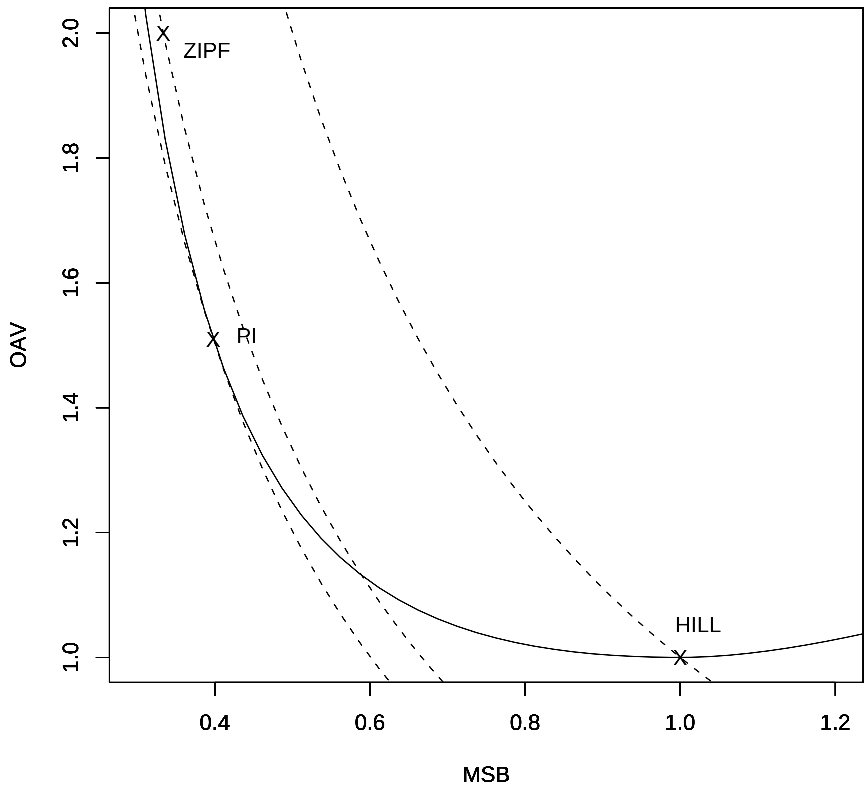

Figure 2: Optimal asymptotic variance ($\mathcal{OAV}$) as a function of the mean-squared bias ($\mathcal{MSB}$) (full line). Dashed lines represent some level curves of $\pi = \mathcal{AV} \times \mathcal{MSB}$. The estimators $\hat{\gamma}_n^{\text{H}}$ (HILL), $\hat{\gamma}_n^{\text{Z}}$ (ZIPF) and $\hat{\gamma}_n^{\pi}$ (PI) are depicted with a $\times$.

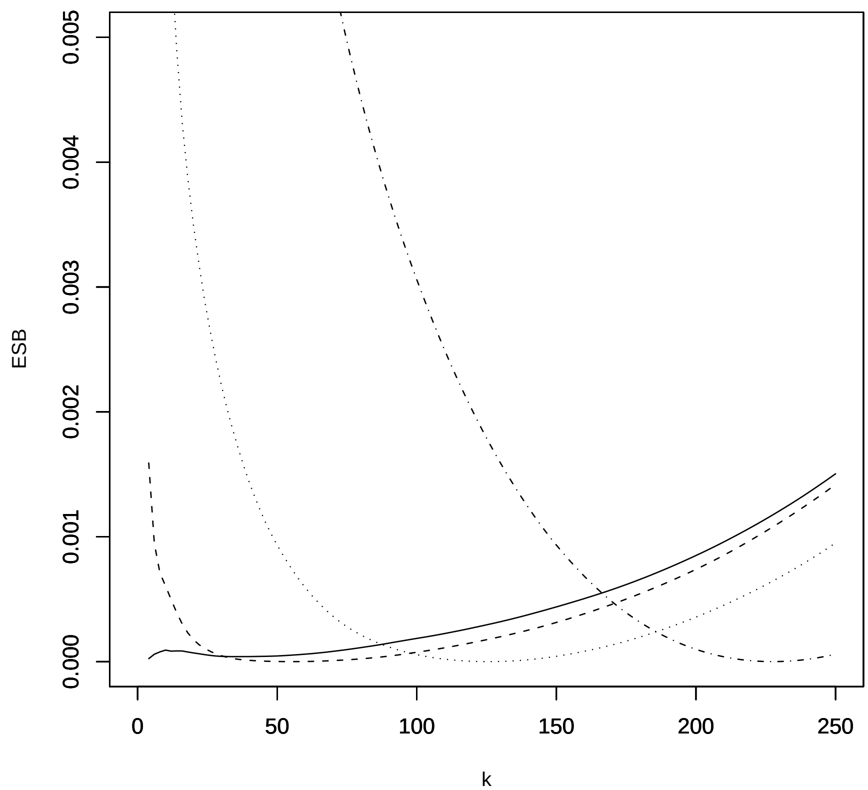

Figure 3: Empirical squared bias ($\mathcal{ESB}$) of $\hat{\gamma}_n^\pi$ as a function of k for a series with temporal dependence coefficient $\alpha = 1$ (full line), $\alpha = 0.8$ (dashed line), $\alpha = 0.5$ (dotted line) and $\alpha = 0.2$ (dashed-dotted line).

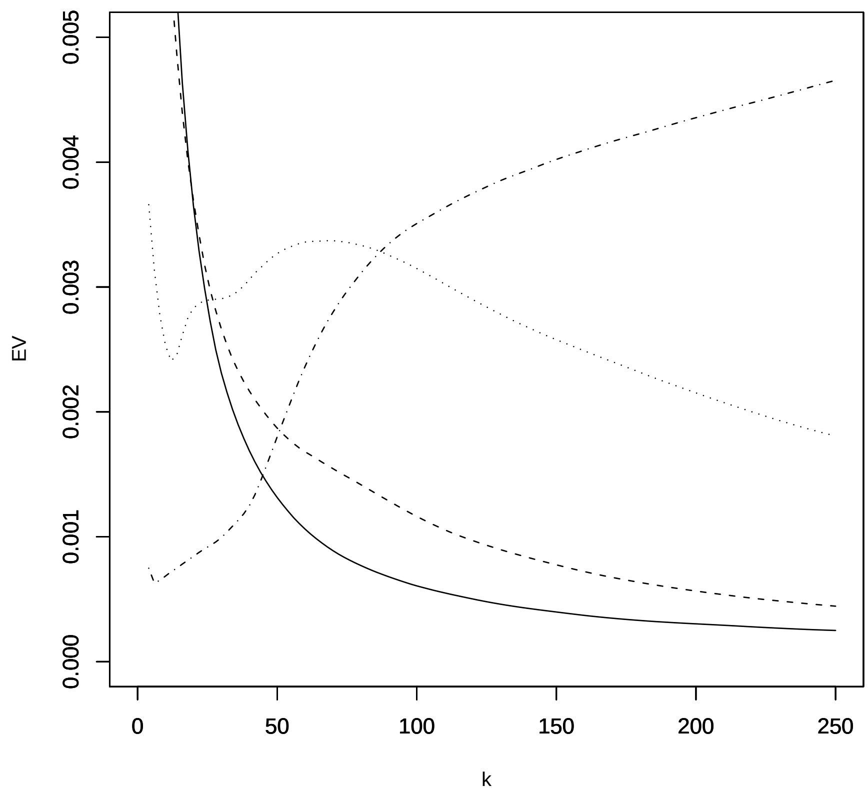

Figure 4: Empirical variance ($\mathcal{EV}$) of $\hat{\gamma}_n^{\pi}$ as a function of k for a series with temporal dependence coefficient $\alpha = 1$ (full line), $\alpha = 0.8$ (dashed line), $\alpha = 0.5$ (dotted line) and $\alpha = 0.2$ (dashed-dotted line).

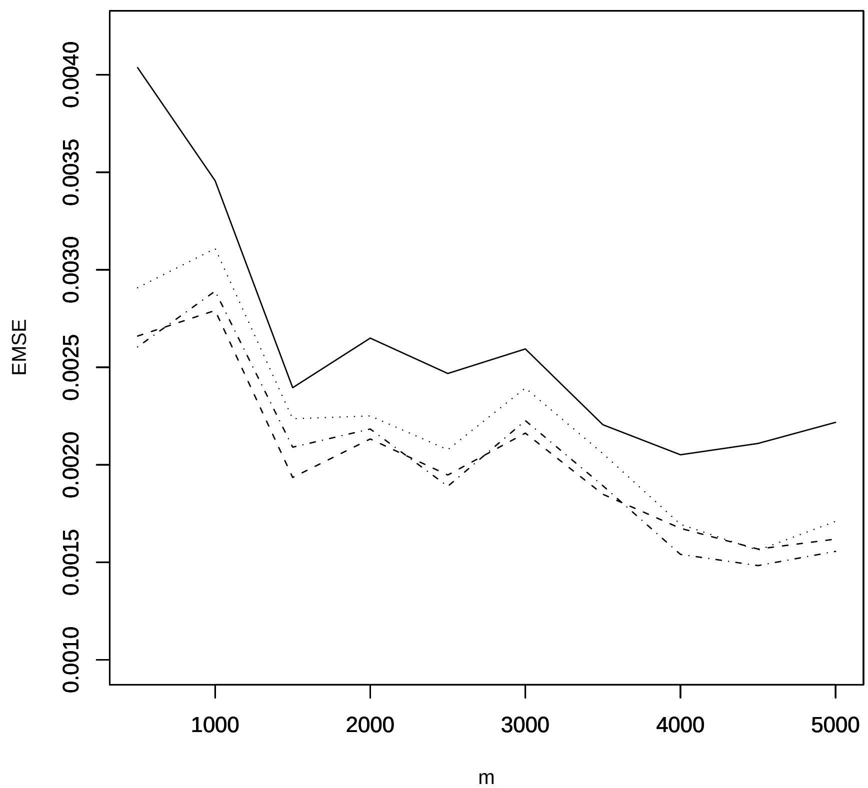

Figure 5: Empirical mean squared error ($\mathcal{EMSE}$) of $\hat{\gamma}_n^{\pi}$ as a function of the number of nearest neighbors m for $n_s = 10$ series with temporal dependence coefficient $\alpha = 0.5$ and with spatial dependence coefficient $\theta = 1$ (full line), $\theta = 0.8$ (dashed line), $\theta = 0.5$ (dotted line) and $\theta = 0.2$ (dashed-dotted line).

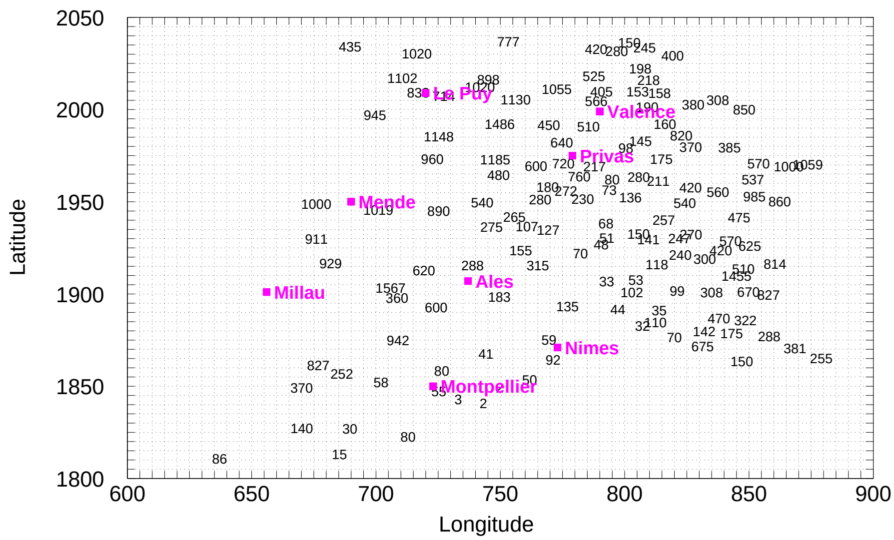

Figure 6: Geographical coordinates of the 142 raingauges stations. Horizontally: longitude (in kilometers), vertically: latitude (in kilometers), on the map: altitude (in meters).

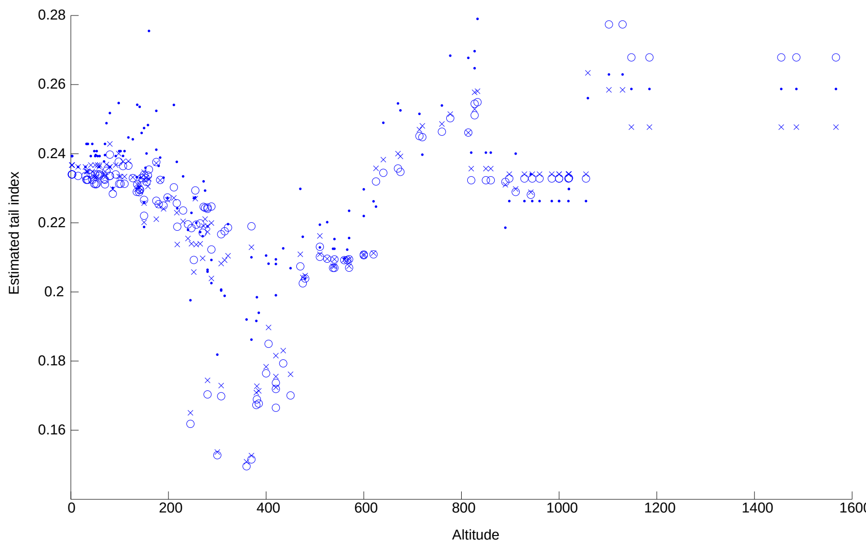

Figure 7: Estimated tail-index as a function of the altitude: $\hat{\gamma}_n^{\mathrm{H}}(\cdots)$, $\hat{\gamma}_n^{\pi}(\times\times\times)$ and $\hat{\gamma}_n^{\mathrm{Z}}(\bullet\bullet\bullet)$.

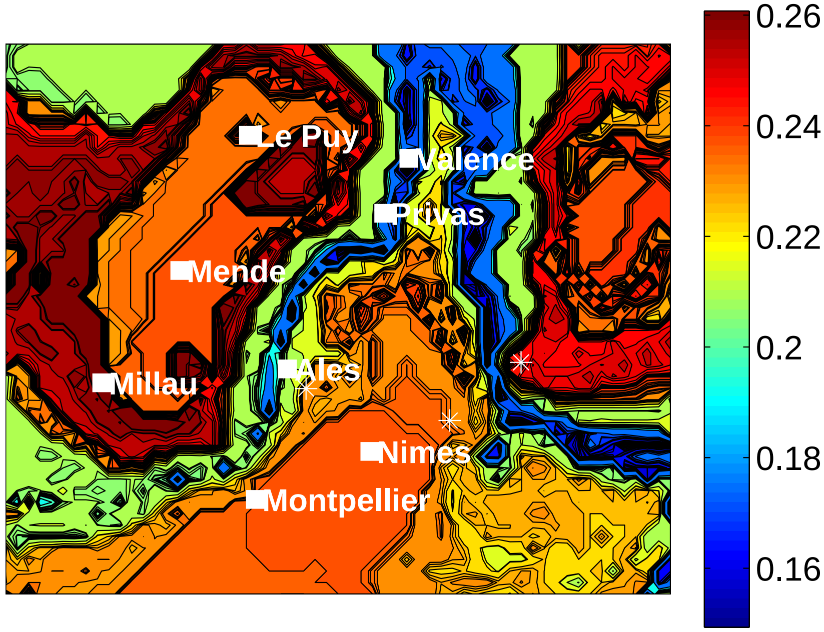

Figure 8: Estimated tail-index $\hat{\gamma}_n^{\pi}$ as a function of the longitude and latitude. Three stations (localized with a $*$ on the map) measured hourly rainfalls larger than 100 millimeters.

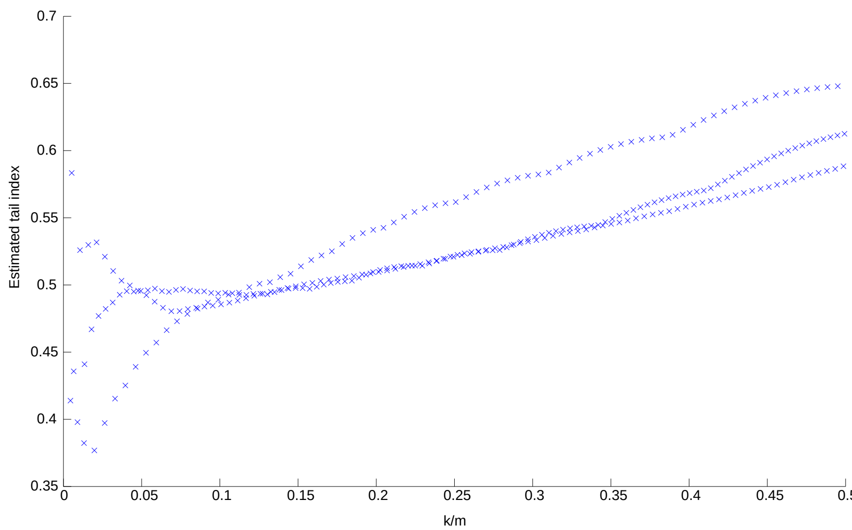

Figure 9: Estimated tail-index $\hat{\gamma}_n^\pi$ represented as a function of $k_{n,t}/m_{n,t}$ at three stations (localized with a $*$ on Figure 8). The estimation is performed without using neighbor stations.

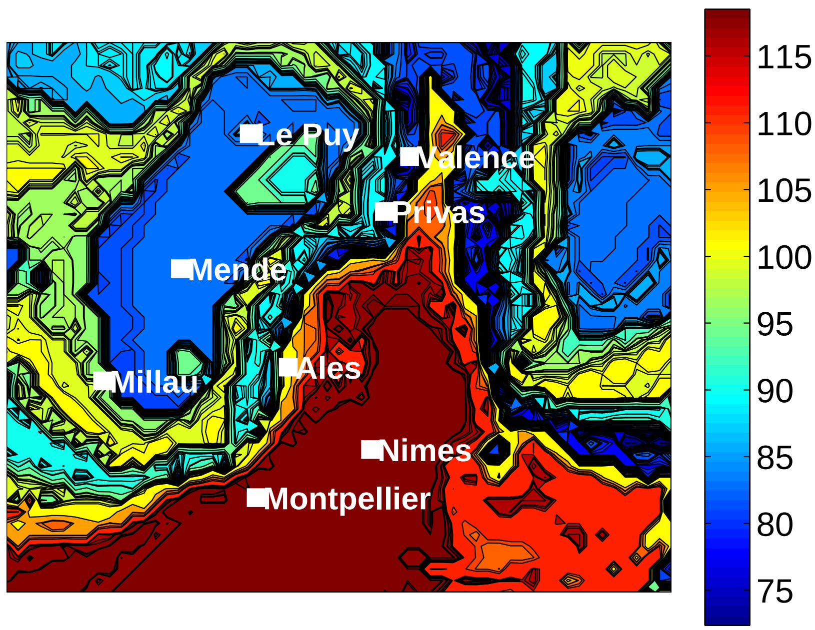

Figure 10: Representation of the pointwise 10 years- return level (in millimeters), estimated with $\hat{q}^{\pi}$, as a function of the longitude and latitude.