HAL open science

Estimation of extreme quantiles from heavy and light tailed distributions

Jonathan El Methni, Laurent Gardes, Stéphane Girard, Armelle Guillou

To cite this version:

Jonathan El Methni, Laurent Gardes, Stéphane Girard, Armelle Guillou. Estimation of extreme quantiles from heavy and light tailed distributions. Journal of Statistical Planning and Inference, 2012, 142 (10), pp.2735-2747. 10.1016/j.jspi.2012.03.025 . hal-00627964v3

HAL Id: hal-00627964

https://hal.science/hal-00627964v3

Submitted on 19 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimation of extreme quantiles from heavy and light tailed distributions

Jonathan El Methni[(1)], Laurent Gardes[(2)], Stéphane Girard[(1)] & Armelle Guillou[(2)]

[(1)] Team Mistis, INRIA Rhône-Alpes & LJK, Inovallée, 655, av. de l'Europe, Montbonnot, 38334 Saint-Ismier cedex, France
[(2)] Université de Strasbourg & CNRS, IRMA, UMR 7501, 7 rue René Descartes, 67084 Strasbourg cedex, France

Abstract

In [18], a new family of distributions is introduced, depending on two parameters τ and θ, which encompasses Pareto-type distributions as well as Weibull tail-distributions. Estimators for θ and extreme quantiles are also proposed, but they both depend on the unknown parameter τ, making them useless in practical situations. In this paper, we propose an estimator of τ which is independent of θ. Plugging our estimator of τ in the two previous ones allows us to estimate extreme quantiles from Pareto-type and Weibull tail-distributions in an unified way. The asymptotic distributions of our three new estimators are established and their efficiency is illustrated on a small simulation study and on a real data set.

AMS Subject Classifications: 62G05, 62G20, 62G30.

Keywords: Weibull tail-distributions, Pareto-type distributions, extreme quantile, maximum domain of attraction, asymptotic normality.

1 Introduction

Let $X_1, \ldots, X_n$ be a sequence of independent and identically distributed random variables with a cumulative distribution function F and let $X_{1,n} \leq \cdots \leq X_{n,n}$ denote the order statistics associated to this sample. The Gnedenko theorem [20] insures that for a large class of cumulative distribution functions, the maximum $X_{n,n}$ (after proper renormalization) converges in distribution to an extreme-value distribution with shape parameter γ. Depending on its sign, three possible maximum domains of attraction for F are possible: Fréchet ($\gamma > 0$), Gumbel ($\gamma = 0$) and Weibull ($\gamma < 0$). Since distributions in the Weibull maximum domain of attraction have a finite right tail, in most applications this maximum domain of attraction is not relevant. In this paper, we focus on the Fréchet and Gumbel maximum domains of attraction.

Distributions in the Fréchet maximum domain of attraction can be characterized through their survival function $\overline{F} = 1 - F$ as $\overline{F}(x) = x^{-1/\gamma} L(x)$ where $\gamma > 0$ and L is a slowly varying function at infinity *i.e.* $L(\lambda x)/L(x) \to 1$ as $x \to \infty$ for all $\lambda \geq 1$. $\overline{F}$ is said to be regularly varying at infinity with index $-1/\gamma$. This property is denoted by $\overline{F} \in \mathcal{R}_{-1/\gamma}$, (see [6] for more details on regular variations theory) and F is called a Pareto-type distribution. To make inference on the distribution tail, most approaches consist of using the k_n upper order statistics $X_{n-k_n+1,n} \leq \cdots \leq X_{n,n}$ since the tail information is only contained in the extreme upper part of the sample. Here, (k_n) is an

intermediate sequence of integers *i.e.* such that

$$\lim_{n\to\infty} k_n = \infty \text{ and } \lim_{n\to\infty} k_n/n = 0.$$

A large part of the extreme-value literature is devoted to the estimation of the tail-index $\gamma > 0$, the most known estimator being the Hill estimator [24]. It can be equivalently defined in terms of log-excesses $\log(X_{n-i+1,n}/X_{n-k_n+1,n})$ or in terms of log-spacings $\log(X_{n-i+1,n}/X_{n-i,n})$:

$$H_n(k_n) = \frac{1}{k_n - 1}\sum_{i=1}^{k_n-1} \log\left(\frac{X_{n-i+1,n}}{X_{n-k_n+1,n}}\right) = \frac{1}{k_n - 1}\sum_{i=1}^{k_n-1} i\log\left(\frac{X_{n-i+1,n}}{X_{n-i,n}}\right). \tag{1}$$

At the opposite, there is no simple representation for distributions in the Gumbel maximum domain of attraction. However, an interesting subset of the Gumbel maximum domain of attraction is the Weibull tail-distributions family. It encompasses for instance Weibull, Gaussian and Gamma distributions. Let us recall that a cumulative distribution function F has a Weibull tail if $\overline{F}(x) = \exp\left(-H(x)\right)$, where $H^{\leftarrow}(t) := \inf\{x,\ H(x) \geq t\} \in \mathcal{R}_\beta$. The function $H^{\leftarrow}$ is the so-called generalized inverse of H. The tail of such distributions is driven by the shape parameter $\beta > 0$ called the Weibull tail-coefficient. Many papers are dedicated to the estimation of the Weibull tail-coefficient. A first family of approaches [2, 3, 7, 13] is based on the log-excesses while a second one relies on the log-spacings [5, 11, 16, 17, 19, 21, 22]. All these estimators are thus similar to the Hill statistic (1).

In order to understand the similarity between most estimators of the Weibull tail-coefficient and the Hill estimator, a new family of distributions has been proposed in [18]. These distributions depend on two parameters $\tau \in [0,1]$ and $\theta > 0$. More specifically, letting $K_x(y) = \int_1^y u^{x-1}du$ where $x \in \mathbb{R}$, the considered family of survival distribution functions is given by:

$(\mathbf{A}_1(\tau,\theta))$ $\overline{F}(x) = \exp(-K_\tau^{\leftarrow}(\log H(x)))$ for $x \geq x_* > 0$ with $\tau \in [0,1]$.

Here, H is an increasing function such that $H^{\leftarrow} \in \mathcal{R}_\theta$ where $\theta > 0$. The parameter τ allows us to represent a large panel of distribution tails ranging from Weibull-type tails (in this case $\tau = 0$ and θ coincides with the Weibull tail-coefficient β) to distributions belonging to the Fréchet maximum domain of attraction (in this case $\tau = 1$ and θ corresponds to the tail index γ).

In [18], an estimator of θ based on the Hill statistic is also introduced:

$$\widehat{\theta}_{n,\tau}(k_n) = \frac{H_n(k_n)}{\mu_\tau(\log(n/k_n))}, \tag{2}$$

with, for all $t > 0$,

$$\mu_\tau(t) = \int_0^\infty \left(K_\tau(x+t) - K_\tau(t)\right)\mathrm{e}^{-x}dx,$$

and an estimator of the extreme quantile $x_{p_n} = \overline{F}^{\leftarrow}(p_n)$ with $p_n \to 0$ as $n \to \infty$ is derived:

$$\widehat{x}_{p_n,\widehat{\theta}_{n,\tau}} = X_{n-k_n+1,n}\exp\left(\widehat{\theta}_{n,\tau}(k_n)\left[K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n))\right]\right). \tag{3}$$

We refer to [14, 26, 15] for applications of extreme quantiles respectively in reliability, hydrology and finance. The asymptotic distributions of $\widehat{\theta}_{n,\tau}$ and $\widehat{x}_{p_n,\widehat{\theta}_{n,\tau}}$ are established in [18]. Let us highlight that estimators (2) and (3) are only of theoretical interest since, in practical situations τ is unknown, and therefore, they cannot be used.

This paper builds on the work of [18]. We propose an estimator $\widehat{\tau}_n$ of τ, independent of θ. Replacing τ by $\widehat{\tau}_n$ in (2) and (3) yields two new estimators which can be computed in practical

situations. As a result, we are able to estimate extreme quantiles from Pareto-type and Weibull tail-distributions in an unified way. The asymptotic normality of our three new estimators is also established.

The paper is organized as follows. The estimators are defined in Section 2, their asymptotic properties are established in Section 3. The behavior of the extreme quantile estimator is illustrated on simulated data in Section 4 and on a real data set in Section 5. Proofs of the main results are presented in Section 6 while proofs of auxiliary results are postponed to the Appendix.

2 Definition of the estimators

Let us first describe the construction of an estimator of τ. Let (k_n) and (k'_n) with $k'_n > k_n$ be two intermediate sequences of integers such that $\widehat{\theta}_{n,\tau}(k_n) \xrightarrow{\mathbb{P}} \theta$ and $\widehat{\theta}_{n,\tau}(k'_n) \xrightarrow{\mathbb{P}} \theta$. It straightforwardly follows that

$$\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k'_n)} = \frac{H_n(k_n)}{H_n(k'_n)} \frac{\mu_\tau(\log(n/k'_n))}{\mu_\tau(\log(n/k_n))} \xrightarrow{\mathbb{P}} 1.$$

Introducing for all $t > t' > 0$ the function defined by $\psi(.;t;t') : \mathbb{R} \to (-\infty, \exp(t-t'))$ and $\psi(x;t,t') = \mu_x(t)/\mu_x(t')$, it follows that

$$\frac{H_n(k_n)}{H_n(k'_n)} \overset{\mathbb{P}}{\sim} \psi(\tau; \log(n/k_n), \log(n/k'_n)). \tag{4}$$

Moreover, it can be shown (see Lemma 3 in Section 6) that $\psi(.;\log(n/k_n), \log(n/k'_n))$ is a bijection from $\mathbb{R}$ to $(-\infty, k'_n/k_n)$. As a consequence, the following estimator of τ is considered:

$$\widehat{\tau}_n = \begin{cases} \psi^{-1}\left(\frac{H_n(k_n)}{H_n(k'_n)}; \log(n/k_n), \log(n/k'_n)\right) & \text{if} \quad \frac{H_n(k_n)}{H_n(k'_n)} < \frac{k'_n}{k_n} \\ u & \text{if} \quad \frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}, \end{cases} \tag{5}$$

where u is the realization of a standard uniform distribution. In practice, only the first situation has to be considered, since Lemma 5 in Section 6 shows that, for n large enough, $H_n(k_n)/H_n(k'_n)$ is almost surely smaller than k'_n/k_n. It is thus possible to plug $\widehat{\tau}_n$ in (2) to obtain a new estimator of θ:

$$\widehat{\theta}_{n,\widehat{\tau}_n}(k_n) = \frac{H_n(k_n)}{\mu_{\widehat{\tau}_n}(\log(n/k_n))}. \tag{6}$$

Similarly, replacing τ and $\widehat{\theta}_{n,\tau}$ by their estimates in (3) yields a new estimator of extreme quantiles:

$$\widehat{x}_{p_n,\widehat{\theta}_{n,\widehat{\tau}_n}} = X_{n-k_n+1,n} \exp\left(\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\left[K_{\widehat{\tau}_n}(\log(1/p_n)) - K_{\widehat{\tau}_n}(\log(n/k_n))\right]\right). \tag{7}$$

The asymptotic behavior of the three new estimators $\widehat{\tau}_n$, $\widehat{\theta}_{n,\widehat{\tau}_n}$ and $\widehat{x}_{p_n,\widehat{\theta}_{n,\widehat{\tau}_n}}$ is established in the next section.

3 Asymptotic properties

As a first result, we establish the consistency of $\widehat{\tau}_n$ under the following assumption:

$$H^{\leftarrow}(t) = t^\theta \ell(t) = c\, t^\theta \exp\left(\int_1^x \frac{\varepsilon(u)}{u} du\right) \tag{8}$$

with c a positive constant and $\varepsilon(s) \to 0$ as $s \to \infty$. Let us highlight that (8) amounts to supposing that $H^{\leftarrow} \in \mathcal{R}_\theta$ and that the slowly varying function at infinity ℓ is normalised, see [6], page 15.

Proposition 1 *Suppose that* $(\mathbf{A}_1(\tau,\theta))$ *holds with* $\ell(.)$ *a normalised slowly varying function. If* (k_n) *and* (k'_n) *are two intermediate sequences of integers such that* $k_n/k'_n \to 0$*, then* $\widehat{\tau}_n \xrightarrow{\mathbb{P}} \tau$.

Let us note that the consistency of $\widehat{\tau}_n$ is established for all $\theta > 0$ and $\tau \in [0,1]$ in an unified way. In this sense, the asymptotic behavior of this estimator is more a consequence of the log-spacings property than a tail behavior (which can be exponential for $\tau = 0$ as well as polynomial for $\tau = 1$). Next, to establish the asymptotic normality of the three estimators (5), (6) and (7), a second-order condition on ℓ is necessary:

$(\mathbf{A}_2(\rho))$ There exist $\rho < 0$ and $b(x) \to 0$ such that uniformly locally on $\lambda \geq \lambda_0 > 0$

$$\log\left(\frac{\ell(\lambda x)}{\ell(x)}\right) \sim b(x)K_\rho(\lambda), \text{ when } x \to \infty,$$

with $|b|$ asymptotically decreasing.

It can be shown that necessarily $|b| \in \mathcal{R}_\rho$. The second order parameter $\rho < 0$ tunes the rate of convergence of $\ell(\lambda x)/\ell(x)$ to 1. The closer is ρ to 0, the slower is the convergence. Condition $(\mathbf{A}_2(\rho))$ is the cornerstone in all the proofs of asymptotic normality for extreme value estimators. It is used in [4, 23, 24] to prove the asymptotic normality of several estimators of the extreme value index. Let $\log_2(.) = \log(\log(.))$. The first theorem establishes the asymptotic normality of $\widehat{\tau}_n$:

Theorem 1 *Suppose that* $(\mathbf{A}_1(\tau,\theta))$ *and* $(\mathbf{A}_2(\rho))$ *hold. If* (k_n) *and* (k'_n) *are two intermediate sequences of integers such that* $k_n/k'_n \to 0$ *and*

$$\sqrt{k'_n}b(\exp K_\tau(\log n/k'_n)) \to 0, \tag{9}$$

$$\sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right) \to \infty, \tag{10}$$

$$\log(n/k'_n)\left(\log_2(n/k_n) - \log_2(n/k'_n)\right) \to \infty, \tag{11}$$

then

$$\sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right)(\widehat{\tau}_n - \tau) \xrightarrow{d} \mathcal{N}(0,1).$$

Condition (9) is standard in extreme-value theory. It imposes that the bias induced by the slowly varying function is asymptotically negligible. Condition (10) forces the speed of convergence of $\widehat{\tau}_n$ to tend to infinity. Finally, (11) is of the same nature as (10): It imposes some minimal spacing between the two sequences (k_n) and (k'_n). Besides, if $\tau = 0$, conditions (9) and (10) imply that $xb(x) \to 0$ as $x \to \infty$. As a consequence, if $\tau = 0$ and $\rho > -1$, it is not possible to choose sequences (k_n) and (k'_n) satisfying the above assumptions. In such a case, only the consistency of $\widehat{\tau}_n$ can be guaranteed.
In our next result, it is established that $\widehat{\theta}_{n,\widehat{\tau}_n}$ inherits from the asymptotic normality of $\widehat{\tau}_n$.

Theorem 2 *Suppose the assumptions of Theorem 1 hold. If, moreover,*

$$\left(\log_2(n/k_n) - \log_2(n/k'_n)\right)/\log_2(n/k_n) \to 0, \tag{12}$$

$$\sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right)/\log_2(n/k_n) \to \infty, \tag{13}$$

then

$$\frac{\sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right)}{\log_2(n/k_n)}\left(\widehat{\theta}_{n,\widehat{\tau}_n}(k_n) - \theta\right) \xrightarrow{d} \mathcal{N}(0,\theta^2).$$

It appears that the estimation of τ has a cost in terms of rates of convergence. Condition (12) implies that $\widehat{\theta}_{n,\widehat{\tau}_n}$ converges slower than $\widehat{\theta}_{n,\tau}$, see Lemma 7 in Section 6. As previously, (13) forces the speed of convergence of $\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)$ to tend to infinity. Note that this condition implies (10) in Theorem 1. Similarly as for Theorem 1, it can be shown that, if $\tau = 0$, (9) and (13) imply $x \log(x)\, b(x) \to 0$ as $x \to \infty$. Thus, again no sequences (k_n) and (k'_n) exist in case $\tau = 0$ and $\rho > -1$. Now, if $\tau \in (0,1]$ or if $\tau = 0$ and $\rho < -1$, a possible choice for the two intermediate sequences is $\log(k_n) = aK_\tau(\log(n))$ and $\log(k'_n) = a'K_\tau(\log(n))$ with the following restrictions on $(a, a') \in \mathbb{R}$:

$$\begin{cases} 0 < a < a' < 2\rho/(2\rho - 1) & \text{if} \quad \tau = 1 \\ 0 < a < a' < -2\rho & \text{if} \quad 0 < \tau < 1 \\ 2 < a < a' < -2\rho & \text{if} \quad \tau = 0. \end{cases} \tag{14}$$

Finally, in the case where $\tau = 0$ and $\rho = -1$, the existence of sequences (k_n) and (k'_n) depends on the underlying distribution.
The last result is dedicated to the asymptotic distribution of the extreme quantile estimator.

Theorem 3 *Suppose the assumptions of Theorem 2 hold. If, moreover,*

$$\sqrt{k_n}(\log_2(n/k_n) - \log_2(n/k'_n))\,/\log_2(1/p_n) \to \infty, \tag{15}$$

$$(\log(n/k_n))^{1-\tau}[K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n))] \to \infty, \tag{16}$$

$$\log_2(n/k_n)[K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n))] \Big/ \int_{\log(n/k_n)}^{\log(1/p_n)} \log(u)u^{\tau-1}du \to 0, \tag{17}$$

then

$$\frac{\sqrt{k_n}\,(\log_2(n/k_n) - \log_2(n/k'_n))}{\int_{\log(n/k_n)}^{\log(1/p_n)} \log(u)u^{\tau-1}du} \left(\frac{\widehat{x}_{p_n,\widehat{\theta}_{n,\widehat{\tau}_n}}}{x_{p_n}} - 1 \right) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

A sufficient condition to verify (16) and (17) is $\log_2(1/p_n)/\log_2(n/k_n) \to \infty$. This imposes an upper bound on the order p_n of the extreme quantile. At the opposite, condition (15) provides a lower bound on p_n. A possible choice for the order p_n of the extreme quantile is given by $\log_2(1/p_n) = [\log_2(n)]^\alpha$ for all $\alpha > 1$. The finite sample performances of $\widehat{x}_{p_n,\widehat{\theta}_{n,\widehat{\tau}_n}}$ are illustrated in the next section.

4 A small simulation study

In this section, our extreme quantile estimator is compared to the Moment estimator of Dekkers et al. [9] and the Peaks Over Threshold (POT), see for instance [8], on simulated data. The POT approach relies on the approximation of the excesses distribution, over a high threshold, by a Generalized Pareto Distribution (GPD). Among the numerous methods available for estimating the GPD parameters, we focus on the moments method which yields the best results in our simulations. Let us emphasize that the Moment and POT estimators are designed to work in any domain of attraction. Let us recall that the extreme quantile estimator (7) requires the numerical inversion of the function ψ. This computation is achieved thanks to a dichotomy procedure since Lemma 3(i) ensures that ψ is increasing.

The comparison is achieved on twelve different distributions:

- five Pareto-type distributions ($\tau = 1$): the absolute value of the standard Cauchy distribution ($\theta = 1$ and $\rho = -2$), standard Pareto distribution ($\theta = 2$ and $\rho = -\infty$), the absolute value of the Student distribution with two degrees of freedom ($\theta = 1/2$ and $\rho = -2$) and two Burr distributions ($\theta = 1/2$ and $\rho \in \{-1, -1/2\}$) with quantile function $\overline{F}^{\leftarrow}(x) = (x^{\rho} - 1)^{-\theta\rho}$, $x \in (0, 1)$,
- five Weibull tail-distributions ($\tau = 0$): the absolute value of the standard Gaussian distribution ($\theta = 1/2$ and $\rho = -1$), a Weibull distribution with 2 as shape parameter ($\theta = 2$ and $\rho = -\infty$), a Gamma distribution with 2 as shape parameter ($\theta = 1$ and $\rho = -1$) and two distributions with quantile function $\overline{F}^{\leftarrow}(x) = (-\log x)^{\theta}(1+(\rho+\theta)(-\log x)^{\rho})$, $x \in (0, 1)$ with $\theta = 1/2$ and $\rho \in \{-1/2, -1/4\}$,
- two log-Weibull tail distributions ($\tau = 1/2$), see [18], paragraph 2.2: the standard lognormal distribution ($\theta = \sqrt{2}/2$ and $\rho = 0$) and the distribution with quantile function $\overline{F}^{\leftarrow}(x) = \exp\{\sqrt{2}[(-\log x)^{1/2} - 1]\}$, $x \in (0, 1)$ for which $\theta = \sqrt{2}/2$ and $\rho = -\infty$.

These distributions represent various situations ($\tau \in \{0, 1/2, 1\}$, $\theta \in \{1/2, \sqrt{2}/2, 1, 2\}$ and $\rho \in \{-\infty, -2, -1, -1/2, -1/4, 0\}$) in which the asymptotic normality of our estimator is not always established. In the following, we take $p_n = 10^{-3}$ and simulate $N = 100$ samples $(\mathcal{X}_{n,j})_{j=1,\ldots,N}$ of size $n = 500$. For each sample $\mathcal{X}_{n,j}$, the estimator $\widehat{x}_{p_n,\widehat{\theta}_{n,\widehat{\tau}_n}}$ is computed for $k'_n = 3, \ldots, 500$ and $k_n = \lfloor ck'_n \rfloor$ with $c = 0.1$. This value has been chosen on the basis of intensive Monte-Carlo simulations. With such a choice, our estimator and the Moment and POT estimators depend on only one intermediate sequence of integers (k'_n). For all the considered distributions, the mean-squared errors associated to the estimators are computed as functions of k'_n and are reported on Figures 1–3. It appears that our estimator outperforms the Moment and POT estimators for almost all the values of k'_n. Also for $k'_n \geq 50$, the mean-squared errors associated to our estimator are almost constant as a function of k'_n, whatever the distribution is.

5 A real data set

The performance of our estimator (7) is illustrated through the analysis of extreme events on the Nidd river data set, which is common in extreme values studies, see for instance [8]. It consists of 154 exceedances of the levels $65m^3s^{-1}$ by the river Nidd (Yorkshire, England) during the period 1934-1969 (35 years). There is no general agreement on a maximum domain of attraction for this data set. In [10], a Fréchet maximum domain of attraction is assumed and heavy tailed-distributions are considered as a possible model for such data. However, according to [25], the Nidd data may reasonably be assumed to come from a distribution in the Gumbel maximum domain of attraction. This result was in accordance with [12] where it has been shown that a Weibull tail-distribution could be considered for such a data. The estimation of τ is therefore of great interest. The estimated values of τ and θ are depicted on Figure 4 as functions of k'_n (recall that $k_n = \lfloor 0.1k'_n \rfloor$). It appears that the estimators become stable for $k'_n \geq 80$ with $\widehat{\tau}_n \simeq 1$ and $\widehat{\theta}_{n,\widehat{\tau}_n}(k_n) \simeq 0.3$. These results indicate that the data may be assumed to come from a distribution in the Fréchet maximum domain of attraction. They are in accordance with the ones obtained by Bayesian methods [10], Figure 7 where the tail index is also estimated at 0.3.

The standard quantity of interest in environmental studies is the N-year return level, defined as the level which is exceeded on average once in N years. Here, we focus on the estimation of the 50- and 100- year return levels. In Figure 4, our extreme quantile estimator is compared to the Moment and POT estimators, plotting the associated N-year return level as a function of k'_n for $N = 50$

and $N = 100$. It appears that our estimator and the Moment estimator yield similar curves. The POT sample path is different, but for all methods, choosing $k'_n \simeq 60$, we obtain an estimation of the 50-year return level which belongs approximately to the interval $[340m^3s^{-1}, 375m^3s^{-1}]$, and an estimation of the 100-year return in $[400m^3s^{-1}, 470m^3s^{-1}]$. Again, these results are in accordance with the the credibility intervals obtained in [10], Table 1.

6 Proofs

We first give some preliminary lemmas. Their proofs are postponed to the appendix.

6.1 Preliminary lemmas

In the following, C is a compact subset such that $[0,1] \subset C \subset (-\infty, 2)$. The first lemma is a standard result on the behavior at infinity of the Laplace transform.

Lemma 1 *Let $x \in C$ and $h_x \in C^\infty(\mathbb{R}^+)$. Set $i = \min\{j \in \mathbb{N}/h_x^{(j)}(0) \neq 0\}$. If*

$$\sup_{\substack{x\in C\\ y\geq 0}} \left| h_x^{(i+1)}(y) \right| < \infty,$$

then

$$\lim_{t\to\infty} \sup_{x\in C} \left| t^{i+1}\widetilde{h}_x(t) - h_x^{(i)}(0) \right| = 0,$$

with $\widetilde{h}_x(t) = \int_0^{+\infty} \exp(-tu)h_x(u)du$ is the Laplace transform of h_x.

Lemma 1 is the key tool for establishing the uniform asymptotic expansions of μ_x and $\partial\mu_x/\partial x$ given in Lemma 2:

Lemma 2 *For all $x \in C$ and $t > 0$, we have*

$$(i) \quad \lim_{t\to\infty} \sup_{x\in C} \left| \frac{\mu_x(t)}{t^{x-1}} - 1 \right| = 0,$$

$$(ii) \quad \lim_{t\to\infty} \sup_{x\in C} \left| \frac{\frac{\partial}{\partial x}\mu_x(t) - \log(t)\mu_x(t)}{t^{x-2}} - 1 \right| = 0.$$

As a consequence of the above expansions, some important properties of ψ can be derived in the two next lemmas:

Lemma 3 *For all $t > t' > 0$ and $x \in \mathbb{R}$, we have*

(i) $x \to \psi(x;t,t')$ is an increasing function,

(ii) $\lim_{x\to\infty} \psi(x;t,t') = \exp(t-t')$.

Lemma 4 *For all $x \in C$ and $t > t' > 0$, we have*

$$\frac{\partial}{\partial x}\psi(x;t,t') \quad = \quad \log(t/t')\psi(x;t,t')\left(1 + O\left(\frac{1}{t'\log(t/t')}\right)\right), \quad \text{as} \quad t' \to \infty.$$

The following lemma ensures that the estimator $\widehat{\tau}_n$ is well-defined. In (5), the situation where $H_n(k_n)/H_n(k'_n) \geq k'_n/k_n$ is eventually impossible.

Lemma 5 *Let (k_n) and (k'_n) be two intermediate sequences of integers such that $k_n/k'_n \to 0$. If $\widehat{\theta}_{n,\tau}(k_n)/\widehat{\theta}_{n,\tau}(k'_n) \xrightarrow{\mathbb{P}} 1$ then*

$$\mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right) \to 0 \text{ as } n \to \infty.$$

We now prove that $\widehat{\theta}_{n,\tau}(k_n)$ is a consistent estimator for θ when τ is known and under general assumptions.

Lemma 6 *Suppose that $(\mathbf{A}_1(\tau,\theta))$ holds with a normalised slowly varying function $\ell(.)$. If (k_n) is an intermediate sequence of integers, then $\widehat{\theta}_{n,\tau}(k_n) \xrightarrow{\mathbb{P}} \theta$.*

The asymptotic normality of $\widehat{\theta}_{n,\tau}$ and $\widehat{x}_{p_n,\widehat{\theta}_{n,\tau}}$ has already been established in the particular case where τ is known by [18]. Let us quote two results from this work.

Lemma 7 *(Theorem 1, [18]). Suppose that $(\mathbf{A}_1(\tau,\theta))$ and $(\mathbf{A}_2(\rho))$ hold. If (k_n) is an intermediate sequence such that $\sqrt{k_n}\, b(\exp K_\tau(\log(n/k_n))) \to 0$, then*

$$\sqrt{k_n}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\theta} - 1\right) \xrightarrow{d} \mathcal{N}(0,1). \tag{18}$$

Lemma 8 *(Theorem 2, [18]). Under the assumptions of Lemma 7 and if, moreover,*

$$(\log(n/k_n))^{1-\tau}(K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n))) \to \infty,$$

then,

$$\frac{\sqrt{k_n}}{K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n))} \log\left(\frac{\widehat{x}_{p_n,\widehat{\theta}_{n,\tau}}}{x_{p_n}}\right) \xrightarrow{d} \mathcal{N}(0,\theta^2).$$

The next lemma establishes that θ can be replaced by $\widehat{\theta}_{n,\tau}(k'_n)$ in (18) without changing the asymptotic distribution.

Lemma 9 *Suppose that $(\mathbf{A}_1(\tau,\theta))$ and $(\mathbf{A}_2(\rho))$ hold. Let (k_n) and (k'_n) be two intermediate sequences of integers such that $\sqrt{k'_n} b(\exp K_\tau(\log n/k'_n)) \to 0$ and $k_n/k'_n \to 0$. We have*

$$\sqrt{k_n}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k'_n)} - 1\right) \xrightarrow{d} \mathcal{N}(0,1).$$

The last lemma quantifies the effect of estimating τ in $\widehat{\theta}_{n,\tau}(k_n)$.

Lemma 10 *Suppose that $(\mathbf{A}_1(\tau,\theta))$ and $(\mathbf{A}_2(\rho))$ hold. Let (k_n) and (k'_n) be two intermediate sequences of integers such that*

$$k_n/k'_n \to 0, \quad \sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right)/\log_2(n/k_n) \to \infty,$$

$$\sqrt{k'_n} b(\exp K_\tau(\log n/k'_n)) \to 0 \quad \text{and} \quad \log(n/k'_n)\left(\log_2(n/k_n) - \log_2(n/k'_n)\right) \to \infty.$$

We have

$$\frac{\sqrt{k_n}(\log_2(n/k_n) - \log_2(n/k'_n))}{\log_2(n/k_n)}\left(\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)} - 1\right) \xrightarrow{d} \mathcal{N}(0,1).$$

6.2 Proofs of the main results

Proof of Proposition 1 – For all $\varepsilon > 0$, let us write

$$\mathbb{P}\left(|\widehat{\tau}_n - \tau| > \varepsilon\right) = \mathbb{P}\left(\widehat{\tau}_n > \tau + \varepsilon\right) + \mathbb{P}\left(\widehat{\tau}_n < \tau - \varepsilon\right)$$

and consider the two terms separately. Clearly, one has

$$\begin{aligned}
\mathbb{P}(\widehat{\tau}_n > \tau + \varepsilon) &= \mathbb{P}\left(\{\widehat{\tau}_n > \tau + \varepsilon\} \middle| \left\{\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right\}\right) \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right) \\
&+ \mathbb{P}\left(\{\widehat{\tau}_n > \tau + \varepsilon\} \cap \left\{\frac{H_n(k_n)}{H_n(k'_n)} < \frac{k'_n}{k_n}\right\}\right) \\
&\leq \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right) + \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \psi(\tau + \varepsilon; \log(n/k_n), \log(n/k'_n))\right). \quad (19)
\end{aligned}$$

Focusing on the first probability of (19), note that, $\widehat{\theta}_{n,\tau}(k_n)$ and $\widehat{\theta}_{n,\tau}(k'_n)$ are consistent estimators for θ in probability by Lemma 6. Thus $\widehat{\theta}_{n,\tau}(k_n)/\widehat{\theta}_{n,\tau}(k'_n) \xrightarrow{\mathbb{P}} 1$ and Lemma 5 entail

$$\mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right) \to 0.$$

The second probability in (19) tends to 0 by (4) and Lemma 3(i). Similarly, we can prove that $\mathbb{P}\left(\widehat{\tau}_n < \tau - \varepsilon\right) \to 0$ which achieves the proof. ■

Proof of Theorem 1 – Letting $v_n = \sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right)$, our goal is to prove that $\mathbb{P}\left(v_n(\widehat{\tau}_n - \tau) \leq s\right) \to \Phi(s)$, for all $s \in \mathbb{R}$ where Φ is the cumulative distribution function of the standard Gaussian distribution. To this aim, let us first remark that Lemma 9 yields $\widehat{\theta}_{n,\tau}(k_n)/\widehat{\theta}_{n,\tau}(k'_n) \xrightarrow{\mathbb{P}} 1$. Thus, introducing $E_n(s) = \{v_n(\widehat{\tau}_n - \tau) \leq s\}$, we have

$$\begin{aligned}
\mathbb{P}(E_n(s)) &= \mathbb{P}\left(E_n(s) \cap \left\{\frac{H_n(k_n)}{H_n(k'_n)} < \frac{k'_n}{k_n}\right\}\right) + \mathbb{P}\left(E_n(s) \middle| \left\{\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right\}\right) \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right) \\
&=: T_n^{(1)}(s) + o(1),
\end{aligned}$$

by Lemma 5. From the definition of $\widehat{\tau}_n$ given in (5) and recalling that, from Lemma 3(i), $\psi(.; \log(n/k_n), \log(n/k'_n))$ is an increasing function, we obtain

$$\begin{aligned}
T_n^{(1)}(s) &= \mathbb{P}\left(\left\{\frac{H_n(k_n)}{H_n(k'_n)} \leq \psi(\tau + s/v_n; \log(n/k_n), \log(n/k'_n))\right\} \cap \left\{\frac{H_n(k_n)}{H_n(k'_n)} < \frac{k'_n}{k_n}\right\}\right) \\
&= \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \leq \min\left(\psi(\tau + s/v_n; \log(n/k_n), \log(n/k'_n)), \frac{k'_n}{k_n}\right)\right).
\end{aligned}$$

Lemma 3(ii) thus yields

$$\begin{aligned}
T_n^{(1)}(s) &= \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \leq \psi(\tau + s/v_n; \log(n/k_n), \log(n/k'_n))\right) \\
&= \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \leq \frac{\mu_{\tau+s/v_n}(\log(n/k_n))}{\mu_{\tau+s/v_n}(\log(n/k'_n))}\right),
\end{aligned}$$

and, from (2) and the fact that

$$\zeta_n := \sqrt{k_n}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k'_n)} - 1\right) \xrightarrow{d} \mathcal{N}(0,1),$$

(see Lemma 9), we have

$$\begin{aligned} T_n^{(1)}(s) &= \mathbb{P}\left(\zeta_n \leq \sqrt{k_n}\left(\frac{\mu_\tau(\log(n/k'_n))}{\mu_\tau(\log(n/k_n))}\frac{\mu_{\tau+s/v_n}(\log(n/k_n))}{\mu_{\tau+s/v_n}(\log(n/k'_n))}-1\right)\right) \\ &= \mathbb{P}\left(\zeta_n \leq \sqrt{k_n}\frac{\mu_\tau(\log(n/k'_n))}{\mu_\tau(\log(n/k_n))}\left[\psi(\tau+s/v_n;\log(n/k_n),\log(n/k'_n))-\psi(\tau;\log(n/k_n),\log(n/k'_n))\right]\right). \end{aligned}$$

A first order Taylor expansion leads to

$$T_n^{(1)}(s) = \mathbb{P}\left(\zeta_n \leq \frac{s}{(\log_2(n/k_n)-\log_2(n/k'_n))}\frac{\mu_\tau(\log(n/k'_n))}{\mu_\tau(\log(n/k_n))}\frac{\partial}{\partial x}\psi(\tau_0;\log(n/k_n),\log(n/k'_n))\right),$$

where $\tau_0 = \tau + s\eta_0/v_n$ with $\eta_0 \in (0,1)$. Since $\tau_0 \to \tau$, for n large enough, $\tau_0 < 2$ and Lemma 4 entails

$$\begin{aligned} T_n^{(1)}(s) &= \mathbb{P}\left(\zeta_n \leq s\frac{\mu_{\tau_0}(\log(n/k_n))}{\mu_{\tau_0}(\log(n/k'_n))}\frac{\mu_\tau(\log(n/k'_n))}{\mu_\tau(\log(n/k_n))}\left(1+O\left(\frac{1}{\log(n/k'_n)\left(\log_2(n/k_n)-\log_2(n/k'_n)\right)}\right)\right)\right) \\ &= \mathbb{P}\left(\zeta_n \leq s\frac{\mu_{\tau_0}(\log(n/k_n))}{\mu_{\tau_0}(\log(n/k'_n))}\frac{\mu_\tau(\log(n/k'_n))}{\mu_\tau(\log(n/k_n))}(1+o(1))\right). \end{aligned}$$

Besides, Lemma 2(i) entails that

$$s\frac{\mu_{\tau_0}(\log(n/k_n))}{\mu_{\tau_0}(\log(n/k'_n))}\frac{\mu_\tau(\log(n/k'_n))}{\mu_\tau(\log(n/k_n))} = s\left(\frac{\log(n/k_n)}{\log(n/k'_n)}\right)^{\tau_0-\tau}(1+o(1)) = s\exp\left(\frac{s\eta_0}{\sqrt{k_n}}\right)(1+o(1)) \underset{n\to\infty}{\longrightarrow} s,$$

and thus

$$T_n^{(1)}(s) = \mathbb{P}(\zeta_n \leq s(1+o(1))) \leq \Phi(s) + \sup_{x\in\mathbb{R}}|\Phi(x)-\mathbb{P}(\zeta_n \leq x)| \to \Phi(s)$$

by [15], page 552. This achieves the proof of Theorem 1. ■

Proof of Theorem 2 – On the first hand, Lemma 7 shows that

$$\widehat{\theta}_{n,\tau}(k_n) - \theta = \frac{\zeta_n\theta}{\sqrt{k_n}} \quad \text{where} \quad \zeta_n \xrightarrow{d} \mathcal{N}(0,1),$$

and on the other hand, Lemma 10 entails

$$\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)} - 1 = \frac{\delta_n\log_2(n/k_n)}{\sqrt{k_n}(\log_2(n/k_n)-\log_2(n/k'_n))} \quad \text{where} \quad \delta_n \xrightarrow{d} \mathcal{N}(0,1). \tag{20}$$

Collecting these results, it follows that

$$\begin{aligned} \widehat{\theta}_{n,\widehat{\tau}_n}(k_n) - \theta &= \widehat{\theta}_{n,\tau}(k_n)\left(\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)}-1\right) + \left(\widehat{\theta}_{n,\tau}(k_n)-\theta\right) \\ &= \widehat{\theta}_{n,\tau}(k_n)\frac{\delta_n\log_2(n/k_n)}{\sqrt{k_n}(\log_2(n/k_n)-\log_2(n/k'_n))} + \frac{\zeta_n\theta}{\sqrt{k_n}}. \end{aligned}$$

Consequently, one immediately has

$$\frac{\sqrt{k_n}(\log_2(n/k_n)-\log_2(n/k'_n))}{\log_2(n/k_n)}\left(\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)-\theta\right) = \widehat{\theta}_{n,\tau}(k_n)\delta_n + \zeta_n\theta\frac{\log_2(n/k_n)-\log_2(n/k'_n)}{\log_2(n/k_n)}$$

and Theorem 2 follows. ■

Proof of Theorem 3 – Let us consider the following expansion:

$$\log\left(\frac{\widehat{x}_{p_n,\widehat{\theta}_{n,\widehat{\tau}_n}}}{x_{p_n}}\right) = \log\left(\frac{\widehat{x}_{p_n,\widehat{\theta}_{n,\widehat{\tau}_n}}}{\widehat{x}_{p_n,\widehat{\theta}_{n,\tau}}}\right) + \log\left(\frac{\widehat{x}_{p_n,\widehat{\theta}_{n,\tau}}}{x_{p_n}}\right) =: T_n^{(2)} + T_n^{(3)}.$$

By (3), we have

$$\begin{aligned}
T_n^{(2)} &= \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)[K_{\widehat{\tau}_n}(\log(1/p_n)) - K_{\widehat{\tau}_n}(\log(n/k_n))] - \widehat{\theta}_{n,\tau}(k_n)[K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n))] \\
&= [K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n))]\left(\widehat{\theta}_{n,\widehat{\tau}_n}(k_n) - \widehat{\theta}_{n,\tau}(k_n)\right) \\
&\quad + \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\left[(K_{\widehat{\tau}_n}(\log(1/p_n)) - K_{\widehat{\tau}_n}(\log(n/k_n))) - (K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n)))\right] \\
&=: T_n^{(2,1)} + T_n^{(2,2)}.
\end{aligned}$$

Focusing on the first term, (20) leads to

$$T_n^{(2,1)} = \frac{\log_2(n/k_n)[K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n))]}{\sqrt{k_n}\,(\log_2(n/k_n) - \log_2(n/k_n'))}\widehat{\theta}_{n,\tau}(k_n)\delta_n$$

which implies that

$$\frac{\sqrt{k_n}(\log_2(n/k_n) - \log_2(n/k_n'))}{\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}\log(u)du}T_n^{(2,1)} = \frac{\delta_n \log_2(n/k_n)[K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n))]}{\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}\log(u)du}\widehat{\theta}_{n,\tau}(k_n) = o_{\mathbb{P}}(1).$$

Let us now consider $T_n^{(2,2)}$. Theorem 1 states that

$$\widehat{\tau}_n = \tau + \frac{\xi_n}{\sqrt{k_n}\,(\log_2(n/k_n) - \log_2(n/k_n'))} =: \tau + \xi_n\sigma_n \quad \text{where} \quad \xi_n \xrightarrow{d} \mathcal{N}(0,1).$$

Replacing in $T_n^{(2,2)}$, we obtain

$$T_n^{(2,2)} = \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)[(K_{\tau+\sigma_n\xi_n}(\log(1/p_n)) - K_{\tau+\sigma_n\xi_n}(\log(n/k_n))) - (K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n)))].$$

By definition

$$K_\tau(\log(1/p_n)) - K_\tau(\log(n/k_n)) = \int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}du,$$

and therefore it immediately follows that

$$T_n^{(2,2)} = \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\int_{\log(n/k_n)}^{\log(1/p_n)}\left(u^{\tau+\sigma_n\xi_n-1} - u^{\tau-1}\right)du = \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}\left(u^{\sigma_n\xi_n} - 1\right)du.$$

Letting $\varphi(x) := \exp(x) - 1 - x$, we deduce that

$$T_n^{(2,2)} = \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\sigma_n\xi_n\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}\log(u)du + \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}\varphi(\sigma_n\xi_n\log(u))du.$$

Now, there exists $c > 0$ such that $x < \log(c)$ implies $|\varphi(x)| < \frac{c}{2}x^2$. As a consequence, since $\sigma_n\log_2(1/p_n) \to 0$ and $\sigma_n\log_2(n/k_n) \to 0$, for n large enough, we have

$$\begin{aligned}
\left|\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}\varphi(\sigma_n\xi_n\log(u))du\right| &\leq \int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}|\varphi(\sigma_n\xi_n\log(u))|du \\
&\leq \frac{c}{2}\sigma_n^2\xi_n^2\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}[\log(u)]^2du.
\end{aligned}$$

Thus,

$$\begin{aligned}\frac{\left|\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}\varphi(\sigma_n\xi_n\log(u))du\right|}{\sigma_n\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}\log(u)du} &\leq \frac{c\sigma_n\xi_n^2\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}[\log(u)]^2du}{2\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}\log(u)du}\\ &\leq \frac{c\xi_n^2\log_2(1/p_n)}{2\sqrt{k_n}(\log_2(n/k_n)-\log_2(n/k_n'))} = o_{\mathbb{P}}(1),\end{aligned}$$

and, replacing in $T_n^{(2,2)}$, we obtain

$$\frac{\sqrt{k_n}\,(\log_2(n/k_n)-\log_2(n/k_n'))}{\int_{\log(n/k_n)}^{\log(1/p_n)}\log(u)u^{\tau-1}du}T_n^{(2,2)} = \xi_n\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)+o_{\mathbb{P}}(1)\xrightarrow{d}\mathcal{N}(0,\theta^2).$$

Finally, Lemma 8 states that

$$T_n^{(3)} = \frac{\varrho_n[K_\tau(\log(1/p_n))-K_\tau(\log(n/k_n))]}{\sqrt{k_n}} \quad \text{where} \quad \varrho_n\xrightarrow{d}\mathcal{N}(0,\theta^2),$$

and thus, under our assumptions,

$$\frac{\sqrt{k_n}(\log_2(n/k_n)-\log_2(n/k_n'))}{\int_{\log(n/k_n)}^{\log(1/p_n)} u^{\tau-1}\log(u)du}T_n^{(3)} = o_{\mathbb{P}}(1).$$

Combining the above results and using the delta method, Theorem 3 follows. ■

Acknowledgments

The authors are grateful to the referees for their helpful comments.

References

[1] Artin, E., (1964), *The gamma function*, Holt, Rinehart and Winston.

[2] Beirlant, J., Bouquiaux, C., Werker, B., (2006), Semiparametric lower bounds for tail index estimation, *Journal of Statistical Planning and Inference*, **136**, 705–729.

[3] Beirlant, J., Broniatowski, M., Teugels, J.L., Vynckier, P., (1995), The mean residual life function at great age: Applications to tail estimation, *Journal of Statistical Planning and Inference*, **45**, 21–48.

[4] Beirlant, J., Dierckx, G., Goegebeur, Y., Matthys, G., (1999), Tail index estimation and an exponential regression model, *Extremes*, **2**, 177–200.

[5] Beirlant, J., Teugels, J., Vynckier, P., (1996), *Practical analysis of extreme values*, Leuven university press, Leuven.

[6] Bingham, N.H., Goldie, C.M., Teugels, J.L., (1987), *Regular Variation*, Cambridge University Press.

[7] Broniatowski, M., (1993), On the estimation of the Weibull tail coefficient, *Journal of Statistical Planning and Inference*, **35**, 349–366.

[8] Davison, A., Smith, R. (1990), Models for exceedances over high thresholds, *Journal of the Royal Statistical Society Series B*, **52**, 3, 393–442.

[9] Dekkers, A. L. M., Einmahl, J. H. J., de Haan, L., (1989), A moment estimator for the index of an extreme-value distribution, *The Annals of Statistics*, **17**, 1833–1855.

[10] Diebolt, J., El-Aroui, M., Garrido, M., Girard, S., (2005), Quasi-Conjugate Bayes Estimates for GPD Parameters and Application to Heavy Tails Modeling, *Extremes*, **8**, 57–78.

[11] Diebolt, J., Gardes, L., Girard, S., Guillou, A., (2008), Bias-reduced estimators of the Weibull tail-coefficient, *Test*, **17**, 311–331.

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

[13] Dierckx, G., Beirlant, J., De Waal, D., Guillou, A. (2009), A new estimation method for Weibull-type tails based on the mean excess function, *Journal of the Statistical Planning and Inference*, **139**, 1905–1920.

[14] Ditlevsen, O., (1994), Distribution Arbitrariness in Structural Reliability, *Structural Safety and Reliability*, 1241–1247, Balkema, Rotterdam.

[15] Embrechts, P., Klüppelberg, C., Mikosch, T., (1997), *Modelling extremal events*, Springer.

[16] Gardes, L., Girard, S., (2006), Comparison of Weibull tail-coefficient estimators, *REVSTAT - Statistical Journal*, **4**, 163–188.

[17] Gardes, L., Girard, S., (2008), Estimation of the Weibull tail-coefficient with linear combination of upper order statistics, *Journal of Statistical Planning and Inference*, **138**, 1416–1427.

[18] Gardes, L., Girard, S., Guillou, A., (2011), Weibull tail-distributions revisited: a new look at some tail estimators, *Journal of Statistical Planning and Inference*, **141**, 429–444.

[19] Girard, S., (2004), A Hill type estimate of the Weibull tail-coefficient, *Communication in Statistics - Theory and Methods*, **33**(2), 205–234.

[20] Gnedenko, B.V., (1943), Sur la distribution limite du terme maximum d'une série aléatoire, *Annals of Mathematics*, **44**, 423–453.

[21] Goegebeur,Y., Beirlant, J., de Wet, T. (2010), Generalized kernel estimators for the Weibull-tail coefficient, *Communications in Statistics - Theory and Methods*, **39**(20), 3695–3716.

[22] Goegebeur, Y., Guillou, A. (2010), Goodness-of-fit testing for Weibull-type behavior, *Journal of Statistical Planning and Inference*, **140**, 1417–1436.

[23] Häusler, E., Teugels, J.L. (1985), On asymptotic normality of Hill's estimator for the exponent of regular variation, *The Annals of Statistics*, **13**, 743–756.

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

[25] Hosking, J., Wallis, J., (1987), Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics*, **29**, 339–349.

[26] Smith, J., (1991), Estimating the upper tail of flood frequency distributions, *Water Resources Research*, **23**(8), 1657–1666.

Appendix: Proof of auxiliary results

Proof of Lemma 1 – A $(i+1)^{th}$ order Taylor expansion leads to

$$h_x(u) = \frac{u^i}{i!} h_x^{(i)}(0) + \frac{u^{i+1}}{(i+1)!} h_x^{(i+1)}(\eta u),$$

with $\eta \in (0,1)$ and consequently,

$$\widetilde{h}_x(t) = \int_0^\infty \exp(-tu) \frac{u^i}{i!} h_x^{(i)}(0) du + \int_0^\infty \exp(-tu) \frac{u^{i+1}}{(i+1)!} h_x^{(i+1)}(\eta u) du =: T_x^{(1)}(t) + T_x^{(2)}(t).$$

It follows that

$$T_x^{(1)}(t) = \frac{h_x^{(i)}(0)}{i!} \int_0^\infty u^i \exp(-tu) du = \frac{h_x^{(i)}(0)}{t^{i+1}},$$

and the change of variable $v = tu$ yields

$$T_x^{(2)}(t) = \frac{1}{(i+1)!} \frac{1}{t^{i+2}} \int_0^\infty \exp(-v) v^{i+1} h_x^{(i+1)} \left(\eta \frac{v}{t}\right) dv.$$

Therefore, we have, uniformly in $x \in C$:

$$|T_x^{(2)}(t)| \leq \sup_{\substack{z \in C \\ y \geq 0}} \left| h_z^{(i+1)}(y) \right| \frac{1}{(i+1)!} \frac{1}{t^{i+2}} \int_0^\infty \exp(-v) v^{i+1} dv = \frac{1}{t^{i+2}} \sup_{\substack{z \in C \\ y \geq 0}} \left| h_z^{(i+1)}(y) \right| = O\left(\frac{1}{t^{i+2}}\right)$$

and thus

$$\widetilde{h}_x(t) = \frac{h_x^{(i)}(0)}{t^{i+1}} + O\left(\frac{1}{t^{i+2}}\right) = \frac{1}{t^{i+1}} \left(h_x^{(i)}(0) + O\left(\frac{1}{t}\right) \right),$$

which achieves the proof. ■

Proof of Lemma 2 – (i) By definition, for all $x \in \mathbb{R}$ and $t > 0$, we have

$$\mu_x(t) = \int_0^\infty (K_x(u+t) - K_x(t)) \exp(-u) du = \int_0^\infty \left(\int_t^{u+t} y^{x-1} dy \right) \exp(-u) du.$$

Using Fubini's theorem, it follows

$$\mu_x(t) = \int_t^\infty y^{x-1} \left(\int_{y-t}^\infty \exp(-u) du \right) dy = \exp(t) \int_t^\infty y^{x-1} \exp(-y) dy, \tag{21}$$

and the change of variable $u = y/t - 1$ yields

$$\mu_x(t) = t^x \int_0^\infty \exp(-tu)(u+1)^{x-1} du = t^x \widetilde{h}_x(t)$$

with $h_x(t) = (t+1)^{x-1}$. Applying Lemma 1 with $i = 0$ concludes the proof of (i).
(ii) From (21) we obtain, for $x \in \mathbb{R}$ and $t > 0$,

$$\mu_x(t) = \exp(t) \int_t^\infty y^{x-1} \exp(-y) dy = \exp(t) \Gamma(x,t), \tag{22}$$

where $\Gamma(x,t)$ is the upper incomplete gamma function. We thus have (see for instance [1])

$$\frac{\partial}{\partial x} \mu_x(t) = \exp(t) \int_t^\infty \exp(-y) \log(y) y^{x-1} dy, \tag{23}$$

and the change of variable $u = y/t - 1$ yields

$$\begin{aligned}
\frac{\partial}{\partial x}\mu_x(t) &= \exp(t)\int_0^\infty \exp(-tu)\exp(-t)t^{x-1}(u+1)^{x-1}\log(t(u+1))tdu \\
&= t^x\int_0^\infty \exp(-tu)(u+1)^{x-1}\log(u+1)du + \log(t)t^x\int_0^\infty \exp(-tu)(u+1)^{x-1}du \\
&=: t^x\int_0^\infty \exp(-tu)g_x(u)du + \log(t)\mu_x(t),
\end{aligned}$$

with $g_x(u) := (u+1)^{x-1}\log(u+1)$. Applying Lemma 1 with $i = 1$, the conclusion follows. ■

Proof of Lemma 3 – (i) First of all, remark that $\psi(x;t,t') > 0$ for all $x \in \mathbb{R}$ and $t > t' > 0$. Besides, routine calculations show that

$$\begin{aligned}
\frac{\partial}{\partial x}\psi(x;t,t') &= \psi(x;t,t')\left(\frac{\partial/\partial x(\mu_x(t))}{\mu_x(t)} - \frac{\partial/\partial x(\mu_x(t'))}{\mu_x(t')}\right) \\
&=: \psi(x;t,t')\left(Q_x(t) - Q_x(t')\right),
\end{aligned} \tag{24}$$

where, by (21) and (23),

$$Q_x(z) = \frac{\int_z^\infty \log(y)y^{x-1}\exp(-y)dy}{\int_z^\infty y^{x-1}\exp(-y)dy}.$$

Let us remark that Q_x is an increasing function on $(0,\infty)$ since

$$Q_x'(z) = \frac{z^{x-1}\exp(-z)\int_z^\infty y^{x-1}\exp(-y)\log(y/z)dy}{\left(\int_z^\infty y^{x-1}\exp(-y)dy\right)^2} > 0.$$

As a consequence $t > t'$ implies $\partial/\partial x(\psi(x;t,t')) = \psi(x;t,t')(Q_x(t) - Q_x(t')) > 0$ and concludes the first part of the proof.
(ii) From (22), we have

$$\frac{1}{\psi(x;t,t')}\frac{\exp(-t')}{\exp(-t)} - 1 = \frac{\mu_x(t')\exp(-t')}{\mu_x(t)\exp(-t)} - 1 = \frac{\int_{t'}^t y^{x-1}\exp(-y)dy}{\int_t^\infty y^{x-1}\exp(-y)dy} =: M_x(t,t').$$

Besides, considering the following inequalities,

$$0 < \int_{t'}^t y^{x-1}\exp(-y)dy < t^{x-1}\left(\exp(-t') - \exp(-t)\right),$$

and

$$\int_t^\infty y^{x-1}\exp(-y)dy > \int_{2t}^\infty y^{x-1}\exp(-y)dy > (2t)^{x-1}\exp(-2t),$$

it follows that

$$M_x(t,t') < 2^{1-x}\frac{\exp(-t') - \exp(-t)}{\exp(-2t)},$$

and thus $M_x(t,t') \to 0$ as $x \to \infty$. The conclusion follows. ■

Proof of Lemma 4 – Recall that C is a compact subset such that $[0,1] \subset C \subset (-\infty, 2)$. From (24) and Lemma 2(ii), we have

$$\frac{\partial}{\partial x}\psi(x;t,t') = \log(t/t')\psi(x;t,t')\left(1+\frac{1}{\log(t/t')}\frac{t^{x-2}}{\mu_x(t)}(1+o(1)) - \frac{1}{\log(t/t')}\frac{t'^{x-2}}{\mu_x(t')}(1+o(1))\right),$$

for all $x \in C$ as $t' \to \infty$. Moreover, a direct application of Lemma 2(i) implies $t^{x-2}/\mu_x(t) = (1/t)(1+o(1))$. Replacing in the previous expression and recalling that $t > t'$, it follows

$$\frac{\partial}{\partial x}\psi(x;t,t') = \log(t/t')\psi(x;t,t')\left(1+O\left(\frac{1}{t'\log(t/t')}\right)\right).$$

The result is proved. ■

Proof of Lemma 5 – Let us remark that

$$\begin{aligned}\mathbb{P}\left(\frac{H_n(k_n)}{H_n(k_n')} \geq \frac{k_n'}{k_n}\right) &= \mathbb{P}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k_n')} \geq \frac{k_n'}{k_n}\frac{\mu_\tau(\log(n/k_n'))}{\mu_\tau(\log(n/k_n))}\right)\\ &= \mathbb{P}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k_n')} \geq \frac{k_n'}{k_n}\left(\frac{\log(n/k_n)}{\log(n/k_n')}\right)^{1-\tau}(1+o(1))\right)\end{aligned}$$

by Lemma 2(i). Moreover $k_n'/k_n \to \infty$ implies $(\log(n/k_n)/\log(n/k_n'))^{1-\tau} \geq 1$ eventually while $\widehat{\theta}_{n,\tau}(k_n)/\widehat{\theta}_{n,\tau}(k_n') \xrightarrow{\mathbb{P}} 1$ by assumption. The conclusion follows. ■

Proof of Lemma 6 – According to Lemma 3 in Gardes *et al.* (2011), we have

$$\begin{aligned}\widehat{\theta}_{n,\tau}(k_n) = \frac{1}{\mu_\tau(\log(n/k_n))}\Bigg\{&\theta\frac{1}{k_n-1}\sum_{i=1}^{k_n-1}\left[K_\tau(F_i+E_{n-k_n+1,n}) - K_\tau(E_{n-k_n+1,n})\right]\\ &+\frac{1}{k_n-1}\sum_{i=1}^{k_n-1}\log\frac{\ell(\exp K_\tau(F_i+E_{n-k_n+1,n}))}{\ell(\exp K_\tau(E_{n-k_n+1,n}))}\Bigg\}. \qquad (25)\end{aligned}$$

Now, Lemma 2(ii) and Lemma 5 in Gardes *et al.* (2011) yield

$$\frac{1}{\mu_\tau(E_{n-k_n+1,n})}\frac{1}{k_n-1}\sum_{i=1}^{k_n-1}\left[K_\tau(F_i+E_{n-k_n+1,n}) - K_\tau(E_{n-k_n+1,n})\right] \xrightarrow{\mathbb{P}} 1. \qquad (26)$$

Moreover, since the slowly varying function $\ell(.)$ is assumed to be normalised, it follows that

$$\log\frac{\ell(\exp(K_\tau(F_i+E_{n-k_n+1,n})))}{\ell(\exp(K_\tau(E_{n-k_n+1,n})))} = \int_{\exp(K_\tau(E_{n-k_n+1,n}))}^{\exp(K_\tau(F_i+E_{n-k_n+1,n}))}\varepsilon(u)\frac{du}{u},$$

with $\varepsilon(s) \to 0$ as $s \to \infty$. Thus for all $i = 1, \ldots, k_n - 1$, we have

$$\begin{aligned}\left|\int_{\exp(K_\tau(E_{n-k_n+1,n}))}^{\exp(K_\tau(F_i+E_{n-k_n+1,n}))}\varepsilon(u)\frac{du}{u}\right| &\leq \sup_{u\geq\exp(K_\tau(E_{n-k_n+1,n}))}|\varepsilon(u)|\int_{\exp(K_\tau(E_{n-k_n+1,n}))}^{\exp(K_\tau(F_i+E_{n-k_n+1,n}))}\frac{du}{u}\\ &= o_{\mathbb{P}}(1)\left[K_\tau(F_i+E_{n-k_n+1,n}) - K_\tau(E_{n-k_n+1,n})\right].\end{aligned}$$

Thus, (26) entails

$$\frac{1}{\mu_\tau(\log(n/k_n))}\frac{1}{k_n-1}\sum_{i=1}^{k_n-1}\log\frac{\ell(\exp K_\tau(F_i+E_{n-k_n+1,n}))}{\ell(\exp K_\tau(E_{n-k_n+1,n}))} = o_{\mathbb{P}}(1). \qquad (27)$$

Collecting (25)-(27), Lemma 6 follows. ■

Proof of Lemma 9 – First of all, let us note that the two conditions $\sqrt{k'_n}b(\exp K_\tau(\log n/k'_n)) \to 0$ and $k_n/k'_n \to 0$ imply $\sqrt{k_n}b(\exp K_\tau(\log n/k_n)) \to 0$ since $|b|$ is asymptotically decreasing. Lemma 7 states that $\widehat{\theta}_{n,\tau}(k_n) = \theta + \theta\zeta_n/\sqrt{k_n}$ where $\zeta_n \xrightarrow{d} \mathcal{N}(0,1)$ and $\widehat{\theta}_{n,\tau}(k'_n) = \theta + \theta\zeta'_n/\sqrt{k'_n}$ where $\zeta'_n \xrightarrow{d} \mathcal{N}(0,1)$. As a consequence, we have

$$\begin{aligned}
\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k'_n)} &= \frac{1+\zeta_n/\sqrt{k_n}}{1+\zeta'_n/\sqrt{k'_n}} \\
&= \left(1+\frac{\zeta_n}{\sqrt{k_n}}\right)\left(1-\frac{\zeta'_n}{\sqrt{k'_n}}+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k'_n}}\right)\right) \\
&= 1+\frac{\zeta_n}{\sqrt{k_n}}+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right),
\end{aligned}$$

and therefore

$$\sqrt{k_n}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k'_n)}-1\right) = \zeta_n + o_{\mathbb{P}}(1).$$

It concludes the proof. ■

Proof of Lemma 10 – A first order Taylor expansion yields

$$\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)}-1 = \frac{\mu_\tau(\log(n/k_n))}{\mu_{\widehat{\tau}_n}(\log(n/k_n))}-1 = (\tau-\widehat{\tau}_n)\frac{\partial/\partial x(\mu_x(\log(n/k_n)))\big|_{x=\tau_0}}{\mu_{\widehat{\tau}_n}(\log(n/k_n))},$$

with $\tau_0 = \widehat{\tau}_n + \eta(\tau-\widehat{\tau}_n)$ and $\eta \in (0,1)$. Let

$$w_n := \frac{\sqrt{k_n}\left(\log_2(n/k_n)-\log_2(n/k'_n)\right)}{\log_2(n/k_n)},\ F_n(s) := \left\{w_n\left(\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)}-1\right) \le s\right\}$$

and recall that Φ is the cumulative distribution function of the standard Gaussian distribution. Our aim is to prove that $\mathbb{P}(F_n(s)) \to \Phi(s)$, for all $s \in \mathbb{R}$. Keeping in mind that C is a compact subset such that $[0,1] \subset C \subset (-\infty, 2)$ and $A_n = \{\tau_0 \in C\} \cap \{\widehat{\tau}_n \in C\}$, we have

$$\mathbb{P}(F_n(s)) = \mathbb{P}(F_n(s)\cap A_n) + \mathbb{P}\left(F_n(s)\,|A_n^c\right)\mathbb{P}\left(A_n^c\right) =: T_n^{(4)}(s) + T_n^{(5)}(s),$$

where A_n^c is the complementary event of A_n. Let us first consider $T_n^{(5)}(s)$. Theorem 1 states that

$$\sqrt{k_n}\left(\log_2(n/k_n)-\log_2(n/k'_n)\right)(\widehat{\tau}_n-\tau) = \xi_n \xrightarrow{d} \mathcal{N}(0,1).$$

Consequently, under the assumptions of Theorem 1, $\widehat{\tau}_n \xrightarrow{\mathbb{P}} \tau \in [0,1]$ and $\tau_0 \xrightarrow{\mathbb{P}} \tau \in [0,1]$. Thus, $\mathbb{P}(A_n^c) \to 0$ which implies $T_n^{(5)}(s) \to 0$. Let us focus on $T_n^{(4)}(s)$. Under the event A_n, combining Lemma 2(ii) with Theorem 1, it follows

$$\begin{aligned}
w_n\left(\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)}-1\right) &= w_n(\tau-\widehat{\tau}_n)\left(\frac{\partial/\partial x(\mu_x(\log(n/k_n)))\big|_{x=\tau_0}}{\mu_{\widehat{\tau}_n}(\log(n/k_n))}\right) \\
&= \xi_n\frac{\log_2(n/k_n)\mu_{\tau_0}(\log(n/k_n))+\log(n/k_n)^{\tau_0-2}(1+o_{\mathbb{P}}(1))}{\log_2(n/k_n)\mu_{\widehat{\tau}_n}(\log(n/k_n))} \\
&= \xi_n\frac{\mu_{\tau_0}(\log(n/k_n))}{\mu_{\widehat{\tau}_n}(\log(n/k_n))}\left(1+\frac{1+o_{\mathbb{P}}(1)}{\log(n/k_n)\log_2(n/k_n)}\right),
\end{aligned}$$

from Lemma 2(i). Moreover, under the event A_n, using Lemma 2(i) and Theorem 1, we have

$$\begin{aligned}
\frac{\mu_{\tau_0}(\log(n/k_n))}{\mu_{\widehat{\tau}_n}(\log(n/k_n))} = (\log(n/k_n))^{\tau_0-\widehat{\tau}_n}(1+o_{\mathbb{P}}(1)) &= (\log(n/k_n))^{\eta(\tau-\widehat{\tau}_n)}(1+o_{\mathbb{P}}(1)) \\
&= \exp(\eta(\tau-\widehat{\tau}_n)\log_2(n/k_n))(1+o_{\mathbb{P}}(1)) \\
&= \exp\left(-\frac{\eta\xi_n\log_2(n/k_n)}{\sqrt{k_n}(\log_2(n/k_n)-\log_2(n/k_n'))}\right)(1+o_{\mathbb{P}}(1)) \\
&\xrightarrow{\mathbb{P}} 1.
\end{aligned}$$

We thus obtain

$$w_n\left(\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)}-1\right) = \xi_n(1+o_{\mathbb{P}}(1)),$$

and, replacing in $T_n^{(4)}(s)$, it finally follows that

$$\begin{aligned}
T_n^{(4)}(s) &= \mathbb{P}(\{\xi_n \le s(1+o_{\mathbb{P}}(1))\}\cap A_n) \\
&= \mathbb{P}(\xi_n \le s(1+o_{\mathbb{P}}(1))) - \mathbb{P}(\{\xi_n \le s(1+o_{\mathbb{P}}(1))\}\,|A_n^c)\mathbb{P}(A_n^c).
\end{aligned}$$

Using the same arguments as in the proof of Theorem 1, it is easily seen that $\mathbb{P}(\xi_n \le s(1+o_{\mathbb{P}}(1))) \to \Phi(s)$ and thus $T_n^{(4)}(s) \to \Phi(s)$ since $\mathbb{P}(A_n^c) \to 0$. Combining the above results the proof of Lemma 10 is achieved. ■

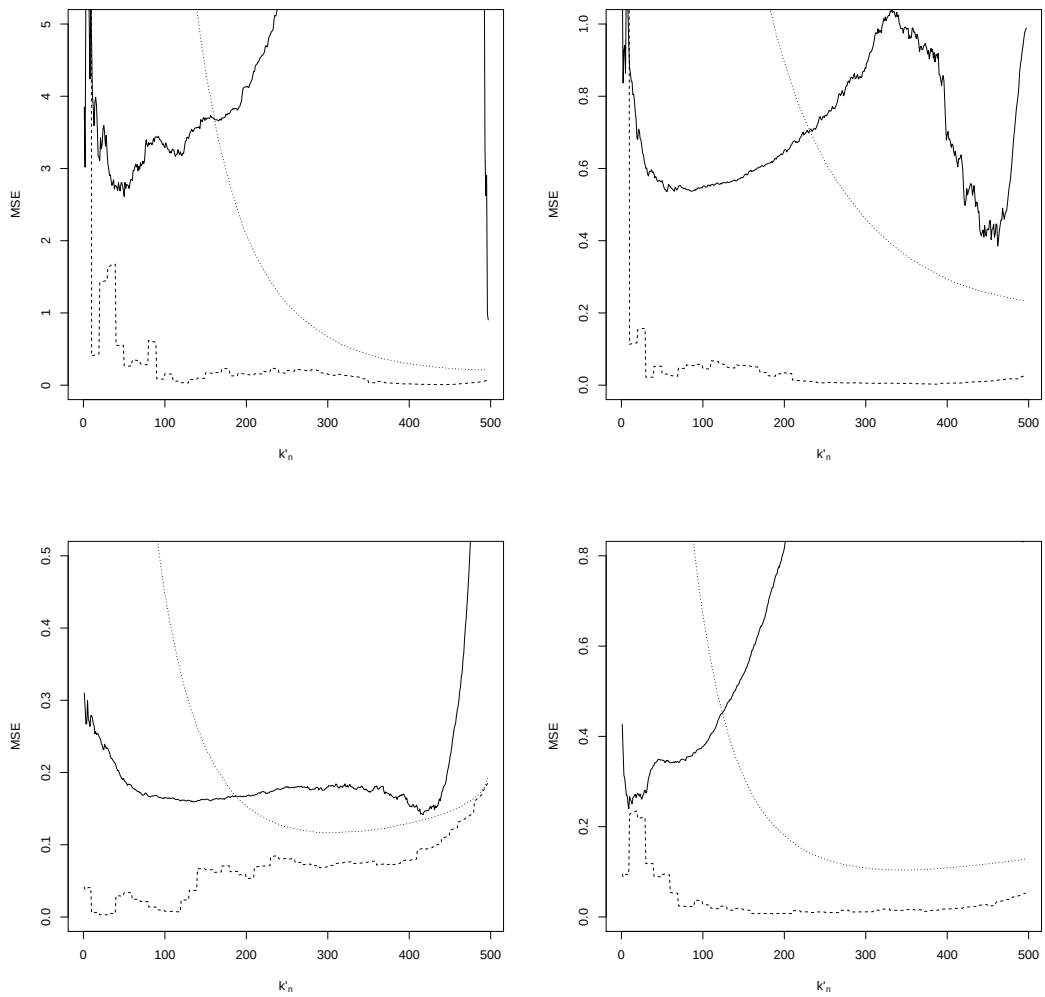

Figure 1: Mean-squared errors as a function of k'_n associated to $\widehat{x}_{p_n,\widehat{\theta}_n,\widehat{\tau}_n}$ (dotted line) and to the Moment estimator of Dekkers *et al.* (bold line) and the POT estimator (solid line). They are computed on 100 samples of size 500. Upper left: absolute value of Cauchy distribution, upper right: absolute value of Student distribution, bottom left: Burr distribution with $\rho = -1$, bottom right: Burr distribution with $\rho = -1/2$.

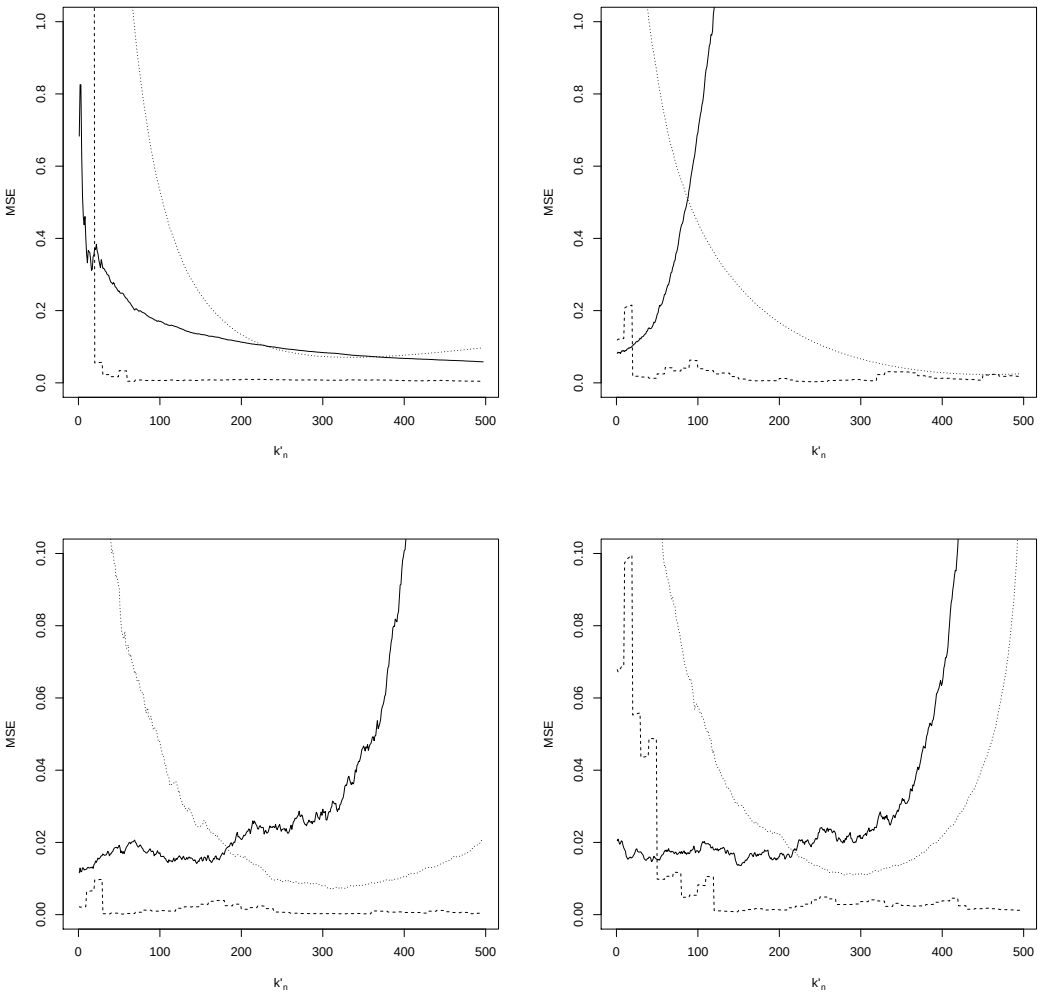

Figure 2: Mean-squared errors as a function of k'_n associated to $\widehat{x}_{p_n,\widehat{\theta}_{n,\widehat{\tau}_n}}$ (dotted line) and to the Moment estimator of Dekkers *et al.* (bold line) and the POT estimator (solid line). They are computed on 100 samples of size 500. Upper left: Pareto distribution, upper right: Weibull distribution, bottom left: absolute value of Gaussian distribution, bottom right: Gamma distribution.

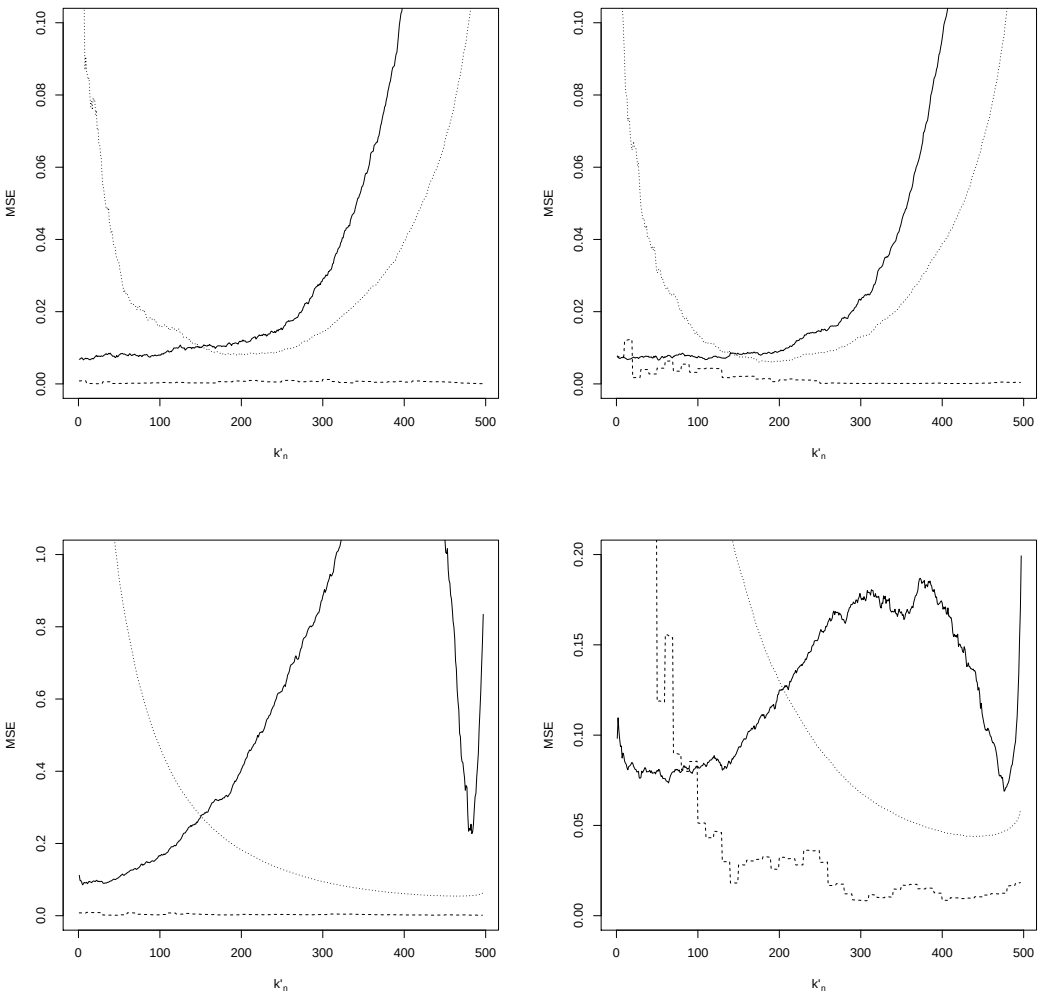

Figure 3: Mean-squared errors as a function of k'_n associated to $\widehat{x}_{p_n,\widehat{\theta}_{n,\widehat{\tau}_n}}$ (dotted line) and to the Moment estimator of Dekkers *et al.* (bold line) and the POT estimator (solid line). They are computed on 100 samples of size 500. Upper left: Weibull-tail distribution with $\rho = -1/2$, upper right: Weibull-tail distribution with $\rho = -1/4$, bottom left: lognormal distribution, bottom right: log-Weibull distribution.

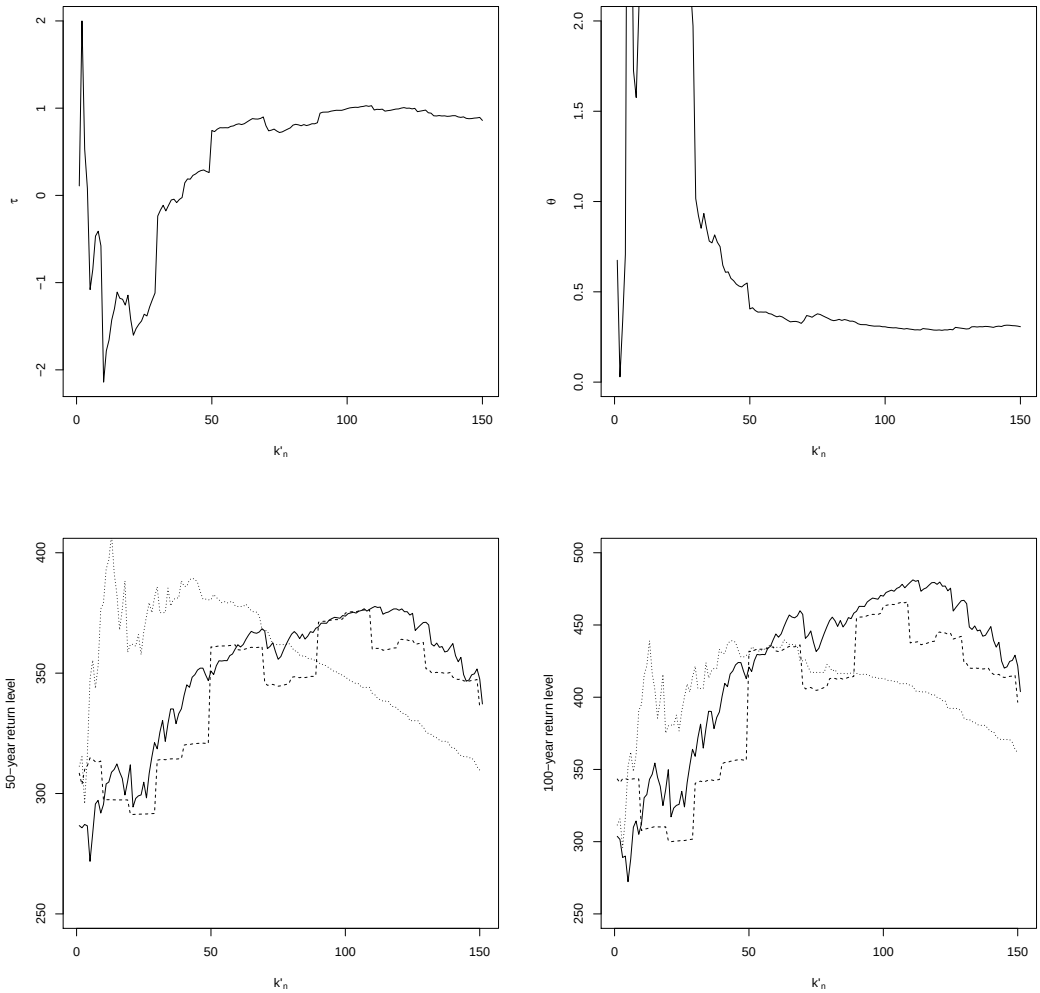

Figure 4: Results obtained on the Nidd river data set. Top: estimation of τ (left) and θ (right) as functions of k'_n. Bottom: N-year return levels as functions of k'_n obtained with $\widehat{x}_{p_n,\widehat{\theta}_{n,\widehat{\tau}_n}}$ (dotted line) and the Moment estimator of Dekkers *et al.* (bold line) and the POT estimator (solid line) (left: $N = 50$, right: $N = 100$).