Bias Correction in Extreme Value Statistics with Index around Zero*

Juan-Juan Cai†

Tilburg University

Laurens de Haan‡

University of Lisbon and Erasmus University Rotterdam

Chen Zhou§

De Nederlandsche Bank and Erasmus University Rotterdam

July 28, 2011

Abstract Applying extreme value statistics in meteorology and environmental science requires accurate estimators on extreme value indices that are around zero. Without having prior knowledge on the sign of the extreme value indices, the probability weighted moment (PWM) estimator is a favorable candidate. As most other estimators on the extreme value index, the PWM estimator bears an asymptotic bias. In this paper, we develop a bias correction procedure for the PWM estimator. Moreover, we provide bias-corrected PWM estimators for high quantiles and, when the extreme value index is negative, the endpoint of a distribution. The choice of k, the number of high order statistics used for estimation, is crucial in applications. The asymptotically unbiased PWM estimators allows the choice of higher level k, which results in a lower asymptotic variance.

*We are grateful to Rijkswaterstaat for providing the still water level data at Hoek van Holland, The Netherlands.

†Postal address: Department of Econometrics & OR, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Email: j.cai@uvt.nl.

‡Postal address: Department of Economics, Erasmus university, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. Email: ldehaan@ese.eur.nl. Research is partially supported by ENES-Project PTDC/MAT/112770/2009.

§Postal address: Economic and Research Division, De Nederlandsche Bank, P.O. Box 98, 1000AB Amsterdam, The Netherlands. Email: c.zhou@dnb.nl; zhou@ese.eur.nl.

Moreover, since the bias-corrected PWM estimators can be applied for a wider range of k, one gets more flexibility in choosing k for finite sample applications. All advantages become apparent in simulations and an environmental application on estimating "once per 10,000 years" still water level at Hoek van Holland, The Netherlands.

Keywords: The probability weighted moment estimator; extreme value index; bias correction; high quantile estimation; endpoint estimation

AMS Classification Numbers: 62G32

1 Introduction

Extreme value statistics has been widely applied in modeling and analyzing rare events in meteorology, environmental science, finance, among other fields. An example in application is estimating high quantiles with extremely low tail probabilities. The major difficulty in application is to produce accurate estimates of the so-called *extreme value index*, denoted as γ which characterizes the shape of a distribution function in the tail region. The case $\gamma > 0$ corresponds to the "heavy-tailed" distributions, while the case $\gamma < 0$ corresponds to distributions with finite right endpoint. Empirical literature has documented that random variables investigated in meteorology and environmental science exhibit generally extreme value indices around zero [1]. In other words, when analyzing meteorological and environmental variables, scientists tend not to make assumption on the distribution function as either heavy-tailed ($\gamma > 0$) or with finite endpoint ($\gamma < 0$). Estimators of a general γ without having prior knowledge on its sign are necessary for such a situation.

Following the usual peak-over-threshold (POT) approach in extreme value analysis, with observations from a distribution that is in the max-domain of attraction, the excesses above a high threshold follows approximately the generalized Pareto distribution (GPD); see, e.g. de Haan and Ferreira (2006, Chap 3). The extreme value index is the shape parameter in the limiting GPD. Hosking and Wallis (1987) introduced the probability weighted moment (PWM) estimator for estimating the shape parameter in the GPD without having prior knowledge on the sign of the shape parameter. A modified version of the PWM estimator is then suitable for handling estimation of the extreme value index that is around zero, when the observations are from the max-domain of attraction. The estimator is widely used in meteorology and environmental applications. With independent and identically distributed (i.i.d.) observations $X_1, X_2, \cdots, X_n$ ranked as $X_{n,1} \leq \cdots \leq X_{n,n}$, we define the PWM estimator as

$$\hat{\gamma}_{pwm} := \frac{I_1 - 4I_2}{I_1 - 2I_2}, \tag{1.1}$$

[1]For example, for hourly surge level on the English east coast, Coles and Tawn (1991) find no significant evidence against $\gamma = 0$; for hourly maximum wind speed in Sheffield, UK, Coles and Walshaw (1994) estimate its γ at -0.12; for daily rainfall in the south-west of England, Coles and Tawn (1996) get the γ estimate at 0.066; for wave height and still water level on the Dutch coast, de Haan and de Ronde (1998) provide γ estimates of the two at -0.0074 and -0.12 respectively; for daily rainfall in North Holland, The Netherlands, Buishand et al. (2008) obtain a γ estimate at 0.1082.

where the probability weighted moments are given by

$$I_q = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{q-1}(X_{n,n-i+1} - X_{n,n-k}), \tag{1.2}$$

for $q = 1, 2, \cdots$. Here $k := k(n)$ is an intermediate sequence such that $k \to \infty$ and $k/n \to 0$ as $n \to \infty$.

Similar to other estimators of the extreme value index, with a suitable second order condition (see section 2.1), $\hat{\gamma}_{pwm}$ has the following asymptotic expansion,

$$\hat{\gamma}_{pwm} - \gamma \approx \frac{N_\gamma}{\sqrt{k}} + A(n/k)\frac{(1-\gamma)(2-\gamma)}{(1-\gamma-\rho)(2-\gamma-\rho)}, \tag{1.3}$$

where N_γ is a mean zero normal random variable, A is a positive or negative function with $A(t) \to 0$ as $t \to \infty$, and ρ is a parameter governing the second order behavior of the tail. From (1.3) we see that $\sqrt{k}(\hat{\gamma}_{pwm} - \gamma) \stackrel{d}{\to} N_\gamma$ provided that $\sqrt{k}A(n/k) \to 0$ as $n \to \infty$. The imposed condition requires a rather slow speed of convergence for the PWM estimator. When $\sqrt{k}A(n/k) \to \lambda$, $\lambda \neq 0$, a bias appears as $\lambda\frac{(1-\gamma)(2-\gamma)}{(1-\gamma-\rho)(2-\gamma-\rho)}$. Since the bias is an explicit function, the aim of this paper is to estimate the bias and subtract it from the original PWM estimator, thus creating a more efficient estimator. Our bias correction procedure has the following advantages.

Firstly, for the application of the bias-corrected PWM estimator, larger values of k can be used than that for the original PWM estimator. This results in a lower asymptotic variance.

Secondly, in applications the choice of k becomes less crucial since for a much wider range of $k-$ values, the estimation stays at a stable level.

Thirdly, the bias-corrected PWM estimator is asymptotically unbiased in the sense that the asymptotic normal distribution has zero mean. All these features become apparent in the simulations and application.

Bias correction has been studied extensively in the literature of estimating the extreme value index. Most studies focus on the γ positive case, with a few papers focusing on the γ negative case. (i) The first generation of bias correction methods is based on an indirect approach which uses weighted combination of different estimators in order to cancel out the bias terms;[2] see Peng

[2]This approach is sometimes called the "generalized jackknife estimators" of extreme value index.

(1998), Caeiro and Gomes (2002), Gomes and Martins (2002, 2004), Gomes et al. (2000, 2002, 2004, 2005, 2007). (ii) The second generation of bias correction methods considers subtracting the bias term in the asymptotic distribution directly; see Caeiro et al. (2005), Gomes and Pestana (2007b); Gomes et al. (2008) and Caeiro and Gomes (2008). (iii) Another category of methods comes from the maximum likelihood procedure: by applying a maximum likelihood procedure to a second order expansion of an explicit tail distribution function, the bias term in the extreme value index estimator is reduced to a lower level. This has been done by Feuerverger and Hall (1999) in the γ positive case. We remark that for all aforementioned bias correction methods, γ is assumed to be positive. Some recent studies focus on the γ negative case following the maximum likelihood method; see Li and Peng (2009) and Li et al. (2010a).

To our best knowledge, no bias correction method has been studied for the general case when γ is not restricted to be positive or negative. The present paper fills this gap because the PWM estimator does not require prior knowledge on the sign of γ.

We further provide a bias correction procedure for the PWM estimators of high quantiles and, when γ is negative, of the endpoint of a distribution. Simulations confirm that the bias-corrected estimators exhibit a superior performance compared to most existing estimators, except in the endpoint estimation. We conduct a real data application on the still water level at Hoek van Holland, The Netherlands.

2 Tail expansion and the choice of k

2.1 Tail expansion: the first, second and third order conditions

The fundamental probability setup in extreme value statistics, the so-called *domain of attraction condition*, models the tail above a high threshold. Suppose that a random variable X follows a distribution function F with right endpoint $x^* := \sup\{x|F(x) < 1\}$. The domain of attraction condition can be expressed as follows (Balkema and De Haan (1974)): there exists a scaling function

$f(t)$ and $\gamma \in \mathbb{R}$, such that for any x with $1 + \gamma x > 0$,

$$\lim_{t \to x^*} P\left(\frac{X - t}{f(t)} > x | X > t\right) = (1 + \gamma x)^{-1/\gamma}.$$

Based on such an approximation, the POT approach considers that the observed excesses above a high threshold follow approximately a GPD. An equivalent condition is given on the corresponding quantile function $U := (1/(1 - F))^{\leftarrow}$, where $^{\leftarrow}$ denotes the left-continuous inverse function. The domain of attraction condition is equivalent to

$$\lim_{t \to \infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \tag{2.1}$$

for $x > 1$, where $a(t)$ is a positive function linked to the function $f(t)$ by $a(t) = f(U(t))$, called the *(first order) scale function*, and the parameter γ is the so-called *extreme value index*.

Existing estimators of the extreme value index, such as the Hill estimator (Hill (1975)), the moment estimator (Dekkers et al. (1989)) and the maximum likelihood estimator (Smith (1987)), possess asymptotic normality under suitable second order conditions. We use the generalized second order condition which characterizes the speed of convergence in (2.1) as follows (de Haan and Stadtmüller (1996)): there exists a *second order scale function* $A(t)$ with constant sign near infinity satisfying $A(t) \to 0$ as $t \to \infty$ and a *second order index* $\rho < 0$[3], such that for all $x > 0$,

$$\lim_{t \to \infty} \frac{\frac{U(tx) - U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} = \frac{x^{\gamma + \rho} - 1}{\rho(\gamma + \rho)}. \tag{2.2}$$

As a consequence, $|A|$ is a regularly varying function with index ρ.

The direct approach on bias correction employs estimators on $A(n/k)$ as well as ρ. Similar to the logic that the second order condition ensures asymptotic normalities of estimators for first order parameters, to obtain the asymptotic properties of the estimators for the second order parameters, we need to impose a third order condition which characterizes the speed of convergence in the second order condition as follows. Suppose there exists a *third order scale function* $B(t)$ with

[3]In a general setup of the second order condition, it is possible to have a second order index ρ zero. Nevertheless, for bias correction studies, it is usually assumed that $\rho < 0$. We follow such an assumption. Similar argument holds for the third order index below.

constant sign near infinity satisfying $B(t) \to 0$ as $t \to \infty$ and a *third order index* $\rho' < 0$, such that for all $x > 0$,

$$\lim_{t \to \infty} \frac{\frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} - \frac{x^{\gamma+\rho}-1}{\rho(\gamma+\rho)}}{B(t)} = \frac{x^{\gamma+\rho+\rho'} - 1}{\gamma+\rho+\rho'}. \tag{2.3}$$

As a consequence, $|B|$ is a regularly varying function with index ρ'. The third order condition has been developed in Fraga Alves et al. (2003). In the literature on bias correction for positive γ, a similar but simpler third order condition applies.

2.2 The choice of k

With a bias-corrected estimator of the extreme value index, it is possible to choose a high value of k, but the choice can not be too high. As pointed out in Caeiro and Gomes (2008), under the third order condition, for an intermediate series k such that $\sqrt{k}\,|A(n/k)| \to \infty$ and $\sqrt{k}\left(A^2(n/k) + |A(n/k)B(n/k)|\right) \to \lambda'$ at a finite level as $n \to \infty$, the "bias-corrected estimators" in Caeiro et al. (2005) again have a bias. This applies in general to bias subtraction procedures for any extreme value index estimator. Therefore, to obtain a fully bias-corrected estimator, the highest k one can choose in the original γ estimator, denoted as k_γ, should satisfy $\sqrt{k_\gamma}\left(A^2(n/k_\gamma) + |A(n/k_\gamma)B(n/k_\gamma)|\right) \to 0$. We do choose such a k level in the bias correction procedure of the PWM estimator, i.e. k_γ satisfies the condition that, as $n \to \infty$,

$$\begin{cases} \sqrt{k_\gamma}\,|A(n/k_\gamma)| \to \infty, \\ \sqrt{k_\gamma}\left(A^2(n/k_\gamma) + |A(n/k_\gamma)B(n/k_\gamma)|\right) \to 0. \end{cases} \tag{2.4}$$

We remark that this condition is not too restrictive: considering the regular variation property of $|A|$ and $|B|$, condition (2.4) is equivalent to choosing a k of a order n^κ, with $\kappa \in \left(\frac{2\rho}{2\rho-1}, \frac{2(\rho+\max(\rho,\rho'))}{2(\rho+\max(\rho,\rho'))-1}\right)$. In the literature of the optimal choice of k for the original estimators of the extreme value index, the optimal choice k_{opt} is of order n^κ with $\kappa = \frac{2\rho}{2\rho-1}$.

When estimating the bias term, estimators of the second order index ρ are employed, which also involves a (different) choice of k. An accurate estimate of the second order index ρ is in general difficult to achieve according to extensive simulations in the literature. The recommendation in the literature is that for the ρ estimator, the choice of k, k_ρ, should be at a higher level compared to

k_γ [4]. More precisely, we choose a k_ρ such that as $n \to \infty$,

$$\begin{cases} \sqrt{k_\rho}\,|A(n/k_\rho)| \to \infty, \\ \sqrt{k_\rho}\left(A^2(n/k_\rho) + |A(n/k_\rho)B(n/k_\rho)|\right) = O(1), \\ \frac{k_\gamma}{k_\rho} \to 0. \end{cases} \tag{2.5}$$

3 Bias correction for the PWM estimators on the extreme value index, high quantiles and endpoint

3.1 Bias correction for the PWM estimator

Recall that we have i.i.d. observations $X_1, X_2, \cdots, X_n$ with a common distribution function F. The following lemma gives the asymptotic behavior of the PWMs, defined in (1.2). It can be obtained by applying the expansion of excesses above a high threshold as in Drees (1998, Theorem 2.1); see the proof in Appendix.

Lemma 3.1 *Suppose the third order condition* (2.3) *holds with* $\gamma < 1/2$ *and* $\rho, \rho' < 0$. *Let* k *be an intermediate sequence such that* $k \to \infty$, $k/n \to 0$ *and* $\sqrt{k}\left(A^2(n/k) + |A(n/k)B(n/k)|\right) = O(1)$ *as* $n \to \infty$. *Then for* $q = 1, 2, \cdots$, *we have that*

$$\frac{q(q-\gamma)I_q}{a(n/k)} = 1 + \frac{1}{\sqrt{k}}L_q + A(n/k)B_q + \varepsilon_{n,k}. \tag{3.1}$$

Here the main stochastic term is

$$L_q = q(q-\gamma)\int_0^1 s^{q-1}(s^{-\gamma-1}W(s) - W(1))ds$$

with W *a standard Brownian motion; the main bias term is*

$$B_q = q(q-\gamma)\int_0^1 s^{q-1}\frac{s^{-\gamma-\rho}-1}{\rho(\gamma+\rho)}ds = \frac{q-\gamma}{\rho(q-\gamma-\rho)},$$

[4] We remark that in the maximum likelihood approach on bias reduction, the maximum likelihood procedure is applied to all parameters including the γ and ρ parameter simultaneously. The choice of k for γ thus follows the optimal choice for the ρ estimator. This leads to less flexibility in obtaining the bias-reduced estimator of γ as pointed out in Li and Peng (2009).

and the residual term is $\varepsilon_{n,k} = O_p(A^2(n/k) + |A(n/k)B(n/k)|)$. *Moreover, for any positive integers* q *and* r,

$$Cov(W(1), L_q) = \gamma, \tag{3.2}$$

and

$$Cov(L_q, L_r) = \frac{qr}{q+r-1-2\gamma} + \gamma^2. \tag{3.3}$$

Under the condition (2.5), the $\varepsilon_{n,k}$ term in (3.1) is of a lower order compared to the other two terms. Throughout, $\varepsilon_{n,k}$ denotes any term at the level of $O_p(A^2(n/k) + |A(n/k)B(n/k)|)$. They are not necessarily equal.

The PWM estimator of the extreme value index γ is constructed by combining two different PWMs. With two different positive integers q and r, from Lemma 3.1, we get that, as $n \to \infty$,

$$\frac{qI_q}{rI_r} \xrightarrow{P} \frac{r-\gamma}{q-\gamma}.$$

This leads to a consistent PWM estimator of γ as

$$\hat{\gamma}_{q,r} = \frac{q^2 I_q - r^2 I_r}{qI_q - rI_r}.$$

The estimator $\hat{\gamma}_{1,2}$ coincides with $\hat{\gamma}_{pwm}$ defined in (1.1), which is a modified version of the classic PWM estimator initially proposed by Hosking and Wallis (1987). The original estimator, designed for observations from the GPD has no threshold, or one can say it has threshold zero, whereas we have threshold $X_{n,n-k}$ due to the fact that we deal with observations in the domain of attraction. We are going to correct for the bias of the estimator $\hat{\gamma}_{1,2}$.

The asymptotic expansion of the PWM estimator can be derived from (3.1) as

$$\hat{\gamma}_{q,r} - \gamma = -\frac{1}{\sqrt{k}} \frac{(q-\gamma)(r-\gamma)}{q-r}(L_q - L_r) + A(n/k)\frac{(q-\gamma)(r-\gamma)}{(q-\gamma-\rho)(r-\gamma-\rho)} + \varepsilon_{n,k}, \tag{3.4}$$

by applying Cramér's delta method. Thus, we omit the details of the derivation.

We see from the asymptotic expansion that in order to correct the bias term in the PWM estimator, $\hat{\gamma}_{1,2}$, it is necessary to construct estimators for the second order index ρ as well as the second order scale function $A(n/k)$. The estimation starts from comparing two PWM estimators. For simplicity, we compare $\hat{\gamma}_{q,r}$ and $\hat{\gamma}_{p,r}$, where p, q and r are three different integers. From the asymptotic property in (3.4), we get that

$$\hat{\gamma}_{q,r} - \hat{\gamma}_{p,r} = \frac{1}{\sqrt{k}} N_{p,q,r} + A(n/k) \frac{\rho(r-\gamma)(p-q)}{(p-\gamma-\rho)(q-\gamma-\rho)(r-\gamma-\rho)} + \varepsilon_{n,k}. \tag{3.5}$$

Here the stochastic term

$$N_{p,q,r} := (r-\gamma)\left(\frac{p-\gamma}{p-r} L_p - \frac{q-\gamma}{q-r} L_q + \frac{(p-q)(r-\gamma)}{(p-r)(q-r)} L_r\right)$$

is a normally distributed random variable with mean zero. We take specific values for p, q, r in (3.5): $p = 4$, $q = 3$, where r is first set to 1 and then 2. When choosing k_ρ as in (2.5), relation (3.5) implies

$$\frac{\hat{\gamma}_{3,1}(k_\rho) - \hat{\gamma}_{4,1}(k_\rho)}{\hat{\gamma}_{3,2}(k_\rho) - \hat{\gamma}_{4,2}(k_\rho)} \xrightarrow{P} \frac{1-\gamma}{2-\gamma} \cdot \frac{2-\gamma-\rho}{1-\gamma-\rho},$$

as $n \to \infty$. Thus, a consistent estimator of ρ can be given as

$$\hat{\rho}(k_\rho) := 1 - \hat{\gamma}_{1,2}(k_\rho) - \frac{1}{\frac{2-\hat{\gamma}_{1,2}(k_\rho)}{1-\hat{\gamma}_{1,2}(k_\rho)} \cdot \frac{\hat{\gamma}_{3,1}(k_\rho) - \hat{\gamma}_{4,1}(k_\rho)}{\hat{\gamma}_{3,2}(k_\rho) - \hat{\gamma}_{4,2}(k_\rho)} - 1}, \tag{3.6}$$

As discussed in Section 2.2, the intermediate sequence k we will use in the bias-corrected PWM estimator, k_γ, satisfies the condition (2.4), while k_ρ is at a higher level, in the sense that $k_\rho/k_\gamma \to \infty$ as $n \to \infty$. It ensures that the asymptotic distribution of ρ will not contaminate the bias subtraction procedure later: from Lemma 6.2, we get that

$$\sqrt{k_\gamma} A(n/k_\gamma)(\hat{\rho}(k_\rho) - \rho) = o_p(1). \tag{3.7}$$

From now on, we just write k for k_γ satisfying (2.4).

The estimator of $A(n/k)$ is based on the expansion of $\hat{\gamma}_{2,1} - \hat{\gamma}_{3,1}$ as in (3.5),

$$\hat{\gamma}_{2,1}(k) - \hat{\gamma}_{3,1}(k) = \frac{1}{\sqrt{k}} N_{3,2,1} + A(n/k) \frac{\rho(1-\gamma)}{(1-\gamma-\rho)(2-\gamma-\rho)(3-\gamma-\rho)} + \varepsilon_{n,k}. \tag{3.8}$$

This suggests the following estimator

$$\hat{A}(n/k) := (\hat{\gamma}_{2,1}(k) - \hat{\gamma}_{3,1}(k)) \frac{(1-\hat{\gamma}_{2,1}(k) - \hat{\rho}(k_\rho))(2-\hat{\gamma}_{2,1}(k) - \hat{\rho}(k_\rho))(3-\hat{\gamma}_{2,1}(k) - \hat{\rho}(k_\rho))}{\hat{\rho}(k_\rho)(1-\hat{\gamma}_{2,1}(k))}. \quad (3.9)$$

Using these estimators of the second order index and scale function, we can now directly subtract the bias term from the PWM estimator $\hat{\gamma}_{1,2}$ to obtain the bias-corrected estimator $\hat{\gamma}_{ub}$ given as

$$\hat{\gamma}_{ub} := \hat{\gamma}_{1,2}(k) - \hat{A}(n/k) \frac{(1-\hat{\gamma}_{1,2}(k))(2-\hat{\gamma}_{1,2}(k))}{(1-\hat{\gamma}_{1,2}(k) - \hat{\rho}(k_\rho))(2-\hat{\gamma}_{1,2}(k) - \hat{\rho}(k_\rho))}. \quad (3.10)$$

The following theorem gives the asymptotic normality of $\hat{\gamma}_{ub}$.

Theorem 3.2 *Under the conditions* (2.3), (2.4) *and* (2.5) *with* $\gamma < 1/2$ *and* ρ, $\rho' < 0$, *as* $n \to \infty$,

$$\sqrt{k}(\hat{\gamma}_{ub} - \gamma) \xrightarrow{d} \frac{(1-\gamma)(2-\gamma)(3-\gamma)}{2\rho} \left((\gamma+\rho-1)L_1 - 2(\gamma+\rho-2)L_2 + (\gamma+\rho-3)L_3\right), \quad (3.11)$$

where the covariance matrix of L_1, L_2 *and* L_3 *can be obtained from* (3.3).

Remark 3.3 *The bias-corrected PWM estimator* $\hat{\gamma}_{ub}$ *is asymptotically unbiased. It is still interesting to study the asymptotic variance. Denote the variance of the limit distribution in* (3.11) *as* $\sigma^2_{ub}(\gamma,\rho)$ *and the asymptotic variance of* $\sqrt{k}(\hat{\gamma}_{1,2} - \gamma)$ *in* (3.4) *as* $\sigma^2_{1,2}(\gamma)$. *Table 1 shows the value of* $\sigma^2_{ub}(\gamma,\rho)/\sigma^2_{1,2}(\gamma)$ *corresponding to different* ρ *and* γ. *Notice that the bias-corrected estimation allows a higher choice of* k, *hence ensures a lower mean square error; see simulation results in Section 4.*

Remark 3.4 *Similar to the PWM estimator* $\hat{\gamma}_{1,2}$, *this bias-corrected PWM estimator can be applied to any* γ *that is below* $1/2$ *and is shift and scale invariant. Therefore, it is particularly useful in dealing with the case that the extreme value index is around zero.*

3.2 Bias correction for the quantile estimator

Estimating high quantiles is one of the major interests in applications of extreme value statistics. The high quantile estimator has been introduced in Weissman (1978). Under the second order condition (2.2), the estimator has a bias term that can be corrected under the framework of the

Table 1: Comparison on the asymptotic variance of $\hat{\gamma}_{ub}$ and $\hat{\gamma}_{1,2}$

		γ								
		-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4
ρ	-2	2.5	2.5	2.4	2.3	2.3	2.2	2.2	2.3	2.4
	-1.5	3.9	3.7	3.5	3.3	3.1	2.9	2.8	2.8	2.7
	-1	7.6	7.1	6.6	6.0	5.4	4.8	4.3	3.9	3.6
	-0.5	27.3	25.1	22.7	20.0	17.1	14.2	11.4	9.1	7.4

Note: The value is $\sigma^2_{ub}(\gamma,\rho)/\sigma^2_{1,2}(\gamma)$, where $\sigma^2_{ub}(\gamma,\rho)$ indicates the asymptotic variance of $\sqrt{k}(\hat{\gamma}_{ub}-\gamma)$ in (3.11) and $\sigma^2_{1,2}(\gamma)$ indicates the asymptotic variance of $\sqrt{k}(\hat{\gamma}_{1,2}-\gamma)$ in (3.4). Notice that the convergence speed, $\sqrt{k}$ might be different for two estimations. The bias-corrected estimator allows a higher level of k.

third order condition. To our best knowledge, bias correction for high quantile estimators has been studied for the γ positive case only; see e.g. Matthys et al. (2004), Gomes and Figueiredo (2006), Gomes and Pestana (2007a), Beirlant et al. (2008) and Li et al. (2010b).

For a low probability level $p = p(n)$ depending on n such that $\lim_{n\to\infty} p(n) = 0$, we are interested in estimating the high quantile x_p defined by $x_p := \sup\{x|F(x) < 1-p\}$, or equivalently, $x_p = U(1/p)$. The high quantile estimator in de Haan and Rootzén (1993) is based on the first order expansion of U in (2.1) with taking $tx = 1/p$ and $t = n/k$. That is

$$\hat{x}_p = \hat{U}(n/k) + \hat{a}(n/k)\frac{d_n^{\hat{\gamma}} - 1}{\hat{\gamma}},$$

where $\hat{U}(n/k) := X_{n,n-k}$, $d_n = \frac{k}{np}$ and $\hat{a}(n/k)$ an estimator of the scale function. The second order expansion of U in (2.2) suggests a bias-corrected high quantile estimator as

$$\hat{x}_{p,ub} := \hat{U}(n/k) + \hat{a}_{ub}(n/k)\frac{d_n^{\hat{\gamma}_{ub}(k)} - 1}{\hat{\gamma}_{ub}(k)} + \hat{A}(n/k)\hat{a}_{ub}(n/k)\frac{d_n^{\hat{\rho}(k_\rho)+\hat{\gamma}_{ub}(k)} - 1}{\hat{\rho}(k_\rho)(\hat{\gamma}_{ub}(k) + \hat{\rho}(k_\rho))}. \tag{3.12}$$

Here we choose to use $\hat{\gamma}_{ub}$, $\hat{\rho}$ and $\hat{A}(n/k)$ as in Section 3.1 and a bias-corrected scale estimator as follows,

$$\hat{a}_{ub}(n/k) := \hat{a}(n/k)\exp\left(-\hat{A}(n/k)\frac{(1-\hat{\gamma}_{ub}(k))(2-\hat{\gamma}_{ub}(k)) - \hat{\rho}(k_\rho)(3-2\hat{\gamma}_{ub}(k))}{\hat{\rho}(k_\rho)(1-\hat{\gamma}_{ub}(k)-\hat{\rho}(k_\rho))(2-\hat{\gamma}_{ub}(k)-\hat{\rho}(k_\rho))}\right), \tag{3.13}$$

where $\hat{a}(n/k)$ is the corresponding PWM estimator of the scale function: $\hat{a}(n/k) = \frac{2I_1I_2}{I_1-2I_2}$, see de Haan and Ferreira (2006).

The asymptotic property of the bias-corrected quantile estimator is given in the following theorem.

Theorem 3.5 *Suppose the third order condition* (2.3) *holds with* $\gamma < 1/2$ *and* $\rho, \rho' < 0$. *Assume the conditions* (2.4) *and* (2.5) *hold and the probability sequence* $p(n)$ *satisfies* $np = o(k)$ *and* $\log(np) = o(\sqrt{k})$ *as* $n \to \infty$. *Then, as* $n \to \infty$, *with* $\gamma_- := \min(0, \gamma)$ *and* $q_\gamma(t) := \int_1^t s^{\gamma-1} \log s ds$ *for* $t > 1$,

$$\sqrt{k}\frac{\hat{x}_{p,ub} - x_p}{a(n/k)q_\gamma(d_n)} \xrightarrow{d} \Gamma_1 + \Gamma_2 + \Gamma_3 + \Gamma_4.$$

Γ_1, Γ_2, Γ_3, *and* Γ_4 *are centered normally distributed random variables stemming from the asymptotic expansions of* $\hat{U}(n/k)$, $\hat{\gamma}_{ub}$, $\hat{a}_{ub}(n/k)$ *and the bias correction procedure, respectively. More specifically,*

$$\begin{aligned}
\Gamma_1 =& (\gamma_-)^2 W(1),\\
\Gamma_2 =& \frac{(1-\gamma)(2-\gamma)(3-\gamma)}{2\rho}\left((\gamma+\rho-1)L_1 - 2(\gamma+\rho-2)L_2 + (\gamma+\rho-3)L_3\right),\\
\Gamma_3 =& -\frac{\gamma_-(1-\gamma)(2-\gamma)(3-\gamma)}{2\rho}\Bigg((\gamma+\rho-1)\left(\frac{1}{\rho} - \frac{1}{2-\gamma} - \frac{1}{3-\gamma}\right)L_1\\
&- 2(\gamma+\rho-2)\left(\frac{1}{\rho} - \frac{1}{1-\gamma} - \frac{1}{3-\gamma}\right)L_2\\
&+ (\gamma+\rho-3)\left(\frac{1}{\rho} - \frac{1}{1-\gamma} - \frac{1}{2-\gamma}\right)L_3\Bigg),\\
\Gamma_4 =& -\frac{(\gamma_-)^2(1-\gamma-\rho)(2-\gamma-\rho)(3-\gamma-\rho)}{2\rho^2(\gamma_-+\rho)}((1-\gamma)L_1) - 2(2-\gamma)L_2 + (3-\gamma)L_3).
\end{aligned}$$

Here W *is the standard Brownian motion in Lemma 3.1. The covariance matrix of* $(\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4)$ *can be obtained from* (3.2) *and* (3.3).

We remark that the asymptotic relation remains valid when replacing $a(n/k)$, $q_\gamma(n/k)$ by their consistent estimators. This is useful in constructing confidence interval of the high quantile estimates.

Notice that in the bias-corrected quantile estimator, we do not assume the sign of the extreme value index γ ex ante. Nevertheless, once the sign is known, the asymptotic property can be significantly simplified as in the following corollary.

Corollary 3.6 *If $\gamma \geq 0$, as $n \to \infty$,*

$$\sqrt{k}\frac{\hat{x}_{p,ub} - x_p}{a(n/k)q_\gamma(d_n)} \xrightarrow{d} \Gamma_2.$$

If $\gamma < 0$, as $n \to \infty$,

$$\begin{aligned}\sqrt{k}\frac{\hat{x}_{p,ub} - x_p}{a(n/k)q_\gamma(d_n)} \xrightarrow{d} & \gamma^2 W(1) + 3(1-\gamma)\left(1 - \frac{1}{\gamma+\rho}\right) L_1 \\ & - 3(2-\gamma)\left(1 - \frac{2}{\gamma+\rho}\right) L_2 + (3-\gamma)\left(1 - \frac{3}{\gamma+\rho}\right) L_3.\end{aligned}$$

3.3 Bias correction for the endpoint estimator

The endpoint of a distribution function F is defined as $x^* = \sup\{x : F(x) < 1\}$. If F belongs to the domain of attraction with $\gamma < 0$, the endpoint of F is finite. The estimation of the endpoint has been studied in literature; see Hall (1982) and Dekkers et al. (1989). The statistical procedure in such situations helps to answer many interesting questions. How long can we live (Aarssen and de Haan (1994))? What is the ultimate world record in a specific athletic event (Einmahl and Magnus (2008))? Similar to other estimation procedure in extreme value statistics, the estimation of the endpoint also bears a potential bias. Li and Peng (2009) propose a bias reduced maximum likelihood estimator for the endpoint.

An bias-corrected endpoint estimator can be derived from the bias-corrected quantile estimator defined in (3.12). By taking $p \to 0$ in (3.12), we get

$$\hat{x}^*_{ub} := X_{n,n-k} - \frac{\hat{a}_{ub}(n/k)}{\hat{\gamma}_{ub}(k)} - \frac{\hat{A}(n/k)\hat{a}_{ub}(n/k)}{\hat{\rho}(k_\rho)(\hat{\gamma}_{ub}(k) + \hat{\rho}(k_\rho))}. \tag{3.14}$$

The asymptotic property of the bias-corrected endpoint estimator is analogous to that of the bias-corrected high quantile estimator as shown in the following theorem.

Theorem 3.7 *Under the conditions* (2.4)*,* (2.5) *and* (2.3) *with* $\gamma < 0$ *and* $\rho, \rho' < 0$*, as* $n \to \infty$*,*

$$\frac{\sqrt{k}}{a(n/k)}(\hat{x}^*_{ub} - x^*) \xrightarrow{d} W(1) + \frac{1}{\gamma^2(\gamma+\rho)}\big(3(\gamma+\rho-1)(1-\gamma)L_1 - 3(\gamma+\rho-2)(2-\gamma)L_2 + (\gamma+\rho-3)(3-\gamma)L_3\big), \tag{3.15}$$

where the covariance matrix of $(W(1), L_1, L_2, L_3)$ *is given by* (3.2) *and* (3.3)*.*

Remark 3.8 *Notice that the rate of convergence in* (3.15) *is approximately* $\frac{k^{1/2+\gamma}}{n}$*. When* $\gamma < -1/2$ *it is a decreasing function of* k*. Thus choosing a high* k *will not reduce the estimation error. In this case, the sample maximum* $X_{n,n}$ *converges to* x^* *faster than* $\hat{x}^*_{ub}$ *(see de Haan and Ferreira (2006, Remark 4.5.5)). When* $\gamma > -1/2$*, theoretically, the bias-corrected estimator can benefit from choosing a higher level of* k*. Nevertheless, the high asymptotic variance imposed by the bias correction procedure dilutes the benefit in finite sample exercise. Such an effect is particularly severe when* γ *is close to zero. See the simulation results below.*

4 Simulations

In this section, we perform finite sample simulations on the proposed bias-corrected estimators. Data are simulated from four distributions: the standard Student-t distribution with degree of freedom 3, the Gumbel distribution ($F(x) = e^{-e^{-x}}$, $x \in R$), the reversed Burr distribution ($F(x) = 1 - (1 + (4-x)^{-4})^{-5/4}$, $x \leq 4$) and the Weibull distribution ($F(x) = 1 - e^{-x^2}$, $x \in R$). The first, second and third order indices of the four distributions are listed in Table 2. Notice that the Weibull distribution is not under our framework as we assume ρ negative. We draw 100 samples with sample size $n = 5000$ from each distribution.

As discussed in section 2.2, a higher level k_ρ is chosen for estimating the second order index ρ. With some pre-simulation, we choose a fixed k_ρ at $k_\rho = [n^{0.98}] = 4216$.

4.1 Simulations on γ estimation

Recall that the bias-corrected PWM estimator of γ is defined in (3.10). We compare our estimator $\hat{\gamma}_{ub}$ with the Hill estimator $\hat{\gamma}_{Hill}$ in Hill (1975), the moment estimator $\hat{\gamma}_M$ in Dekkers

Table 2: Distributions for simulation

parameters	Student-t(3)	Gumbel	reversed Burr	Weibull
γ	1/3	0	-0.2	0
ρ	-2/3	-1	-0.8	0
ρ'	-4/3	-1	-0.8	0

Note: The Student-t(3) distribution is the standard Student-t distribution with degree of freedom 3. The other three distribution functions are: $F(x) = e^{-e^{-x}}$, $x \in R$ (Gumbel); $F(x) = 1 - (1 + (4 - x)^{-4})^{-5/4}$, $x \leq 4$ (reversed Burr); and $F(x) = 1 - e^{-x^2}$, $x \in R$ (Weibull).

et al. (1989) and the original PWM estimator $\hat{\gamma}_{pwm} = \hat{\gamma}_{1,2}$. Moreover, to eliminate the impact of ρ estimation, we construct a pseudo bias-corrected estimator with the real value of ρ, defined as

$$\hat{\gamma}_\rho := \hat{\gamma}_{1,2}(k) - \hat{A}(n/k) \frac{(1 - \hat{\gamma}_{1,2}(k))(2 - \hat{\gamma}_{1,2}(k))}{(1 - \hat{\gamma}_{1,2}(k) - \rho)(2 - \hat{\gamma}_{1,2}(k) - \rho)}. \tag{4.1}$$

We do not apply the Hill estimator to data simulated from distributions with non-positive γ; neither $\hat{\gamma}_\rho$ to distribution with $\rho = 0$.

In Figure 1, we plot the average of γ estimates against k from 100 samples for each distribution. The plots show that the bias-corrected estimates remain stable for a large range of k, which is close to the real value. Other estimates bear dramatically increasing biases as the choice of k increases. It is remarkable that $\hat{\gamma}_\rho$ does not perform better than $\hat{\gamma}_{ub}$. Hence the inaccuracy in estimating ρ does not impose a significant error when estimating γ. Although for the Weibull distribution, the bias-corrected estimator also has a bias, it is still the most stable estimator with respect to the variation of k. Moreover, it is closest to the real value.

In our procedure, bias is corrected at the cost of imposing extra asymptotic variance . With the choice of a higher level k, the asymptotic root mean square error (RMSE) of the bias-corrected estimator is theoretically of a lower order compared to that of the original estimator. In Figure 2, we plot the RMSE against k and confirm this theoretical property. In general, the bias-corrected estimators are much less sensitive with respect to the choice of k.

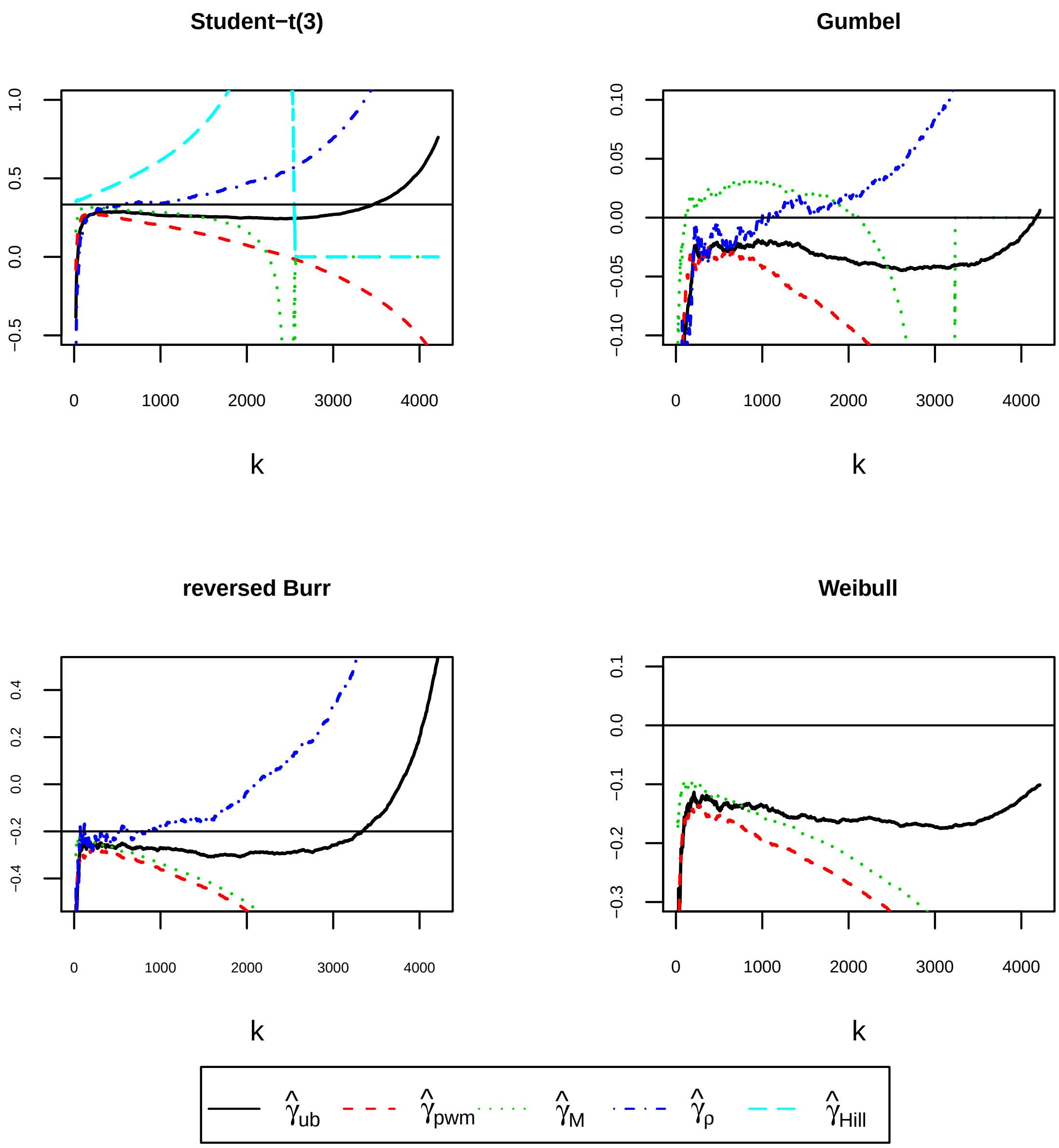

Figure 1: Average of γ estimates for 100 samples.

Note: Each sample consists of 5000 observations. $\hat{\gamma}_{ub}$ is the bias-corrected PWM estimator in (3.10); $\hat{\gamma}_{pwm}$ is the PWM estimator in (1.1); $\hat{\gamma}_{M}$ is the moment estimator in Dekkers et al. (1989); $\hat{\gamma}_{\rho}$ is a pseudo bias-corrected PWM estimator in (4.1); $\hat{\gamma}_{Hill}$ is the Hill estimator in Hill (1975). Horizontal lines indicate the real values of γ.

4.2 Simulations on the high quantile and endpoint estimation

The bias-corrected quantile estimator $\hat{x}_{p,ub}$ is given in (3.12). The competitors are the moment quantile estimator $\hat{x}_{p,M}$ introduced in Dekkers et al. (1989) and the PWM quantile estimator

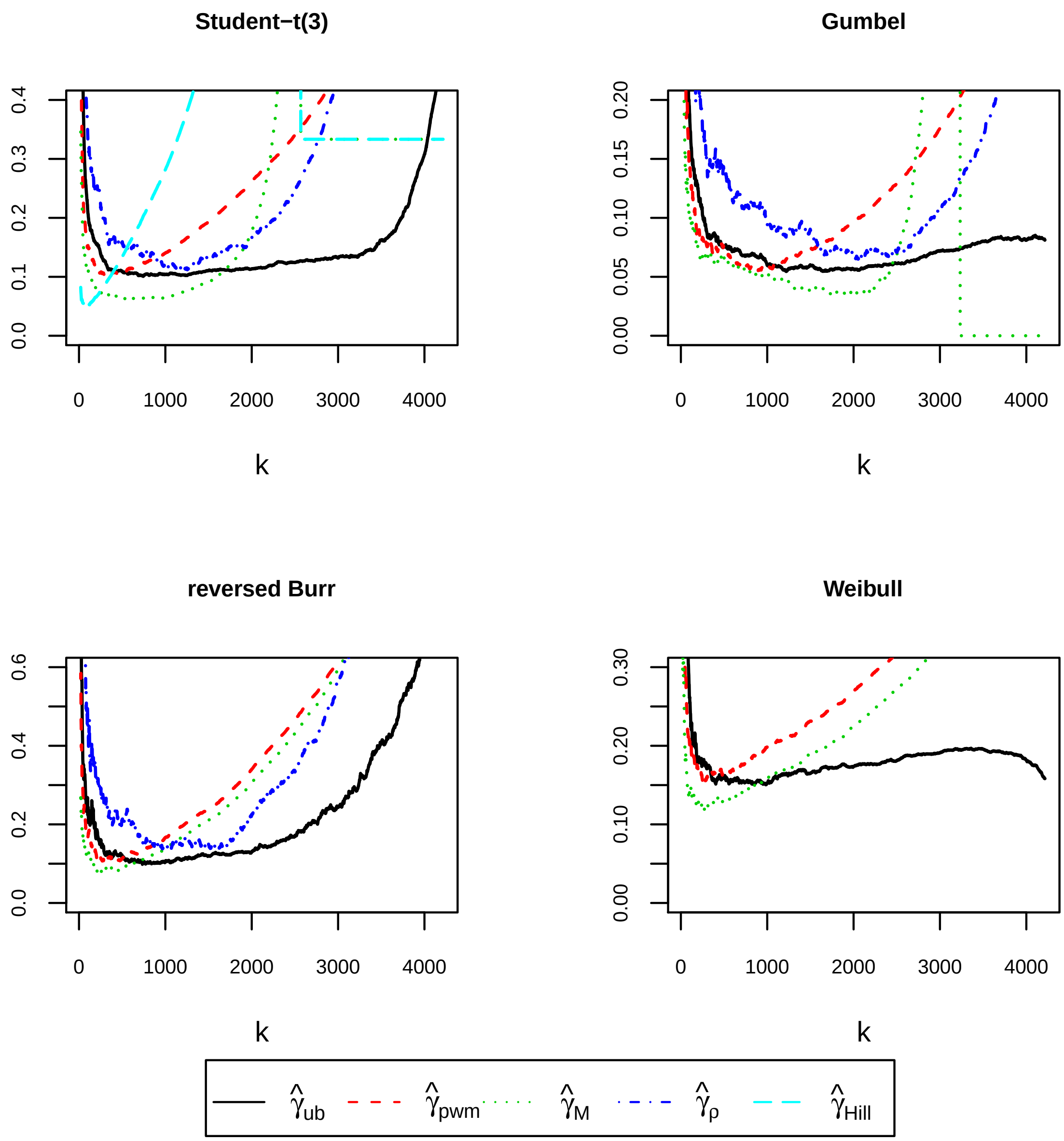

Figure 2: Root mean square error of γ estimates for 100 samples.

Note: Each sample consists of 5000 observations. $\hat{\gamma}_{ub}$ is the bias-corrected PWM estimator in (3.10); $\hat{\gamma}_{pwm}$ is the PWM estimator in (1.1); $\hat{\gamma}_{M}$ is the moment estimator in Dekkers et al. (1989); $\hat{\gamma}_{\rho}$ is a pseudo bias-corrected PWM estimator in (4.1); $\hat{\gamma}_{Hill}$ is the Hill estimator in Hill (1975).

without bias correction $\hat{x}_{p,pwm}$; see Exercise 4.7 in de Haan and Ferreira (2006) . We study the estimators of the $1-p$ quantile with $p = 1/2000$.

Figures 3 and 4 demonstrate the performance of the high quantile estimators in terms of the

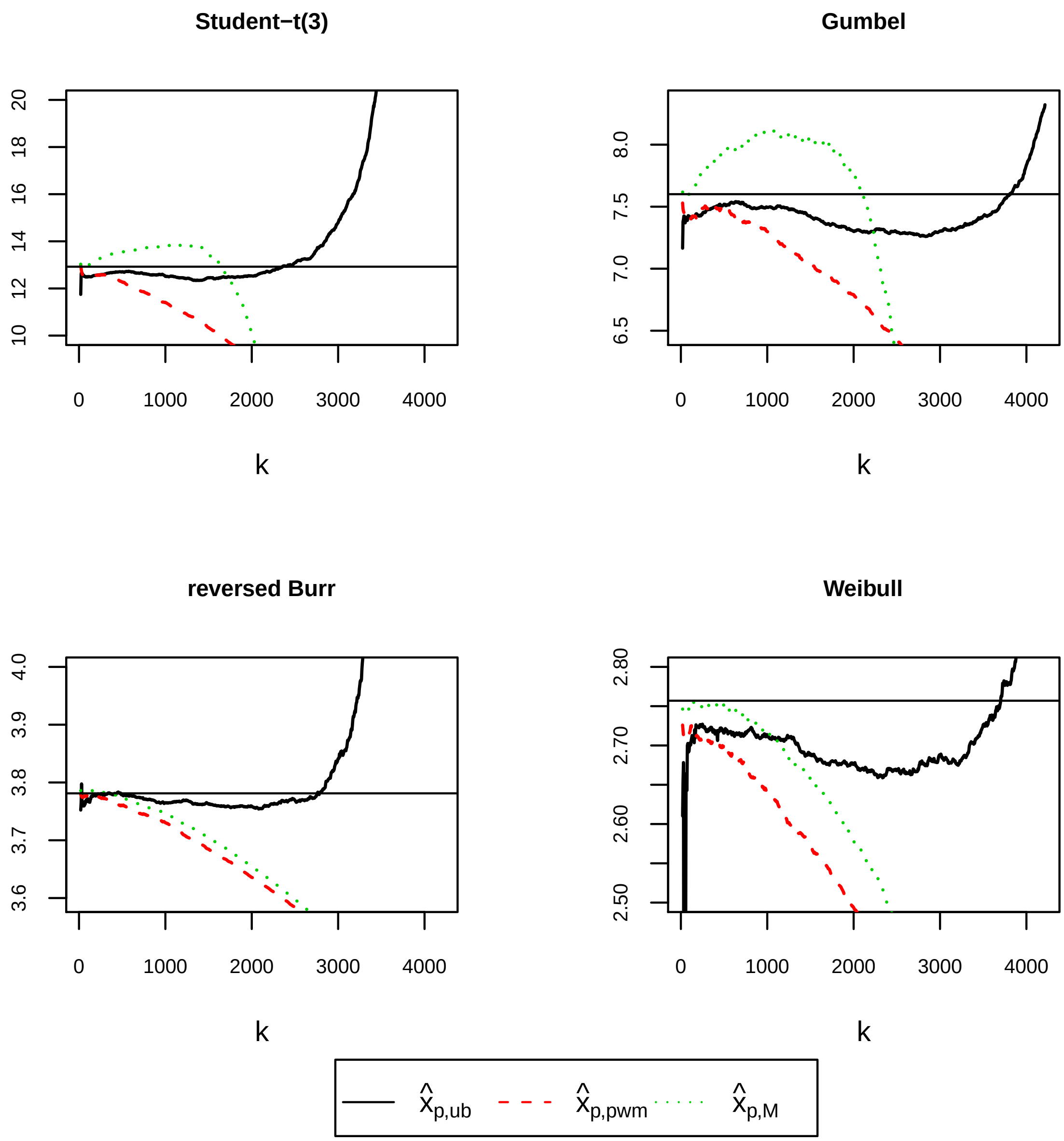

Figure 3: Average of the high quantile estimates for 100 samples.

Note: Each sample consists of 5000 observations. $\hat{x}_{p,ub}$ is the bias-corrected high quantile estimator in (3.12); $\hat{x}_{p,pwm}$ is the PWM quantile estimator in de Haan and Ferreira (2006); $\hat{x}_{p,M}$ is the moment quantile estimator in Dekkers et al. (1989). Horizontal lines indicate the value of x_p with $p = 1/2000$.

average of the estimates and the RMSE. The bias-corrected quantile estimator $\hat{x}_{p,ub}$ performs better than the other two estimators.

We further apply the bias-corrected endpoint estimator, $\hat{x}^*_{ub}$ defined in (3.14) to the random

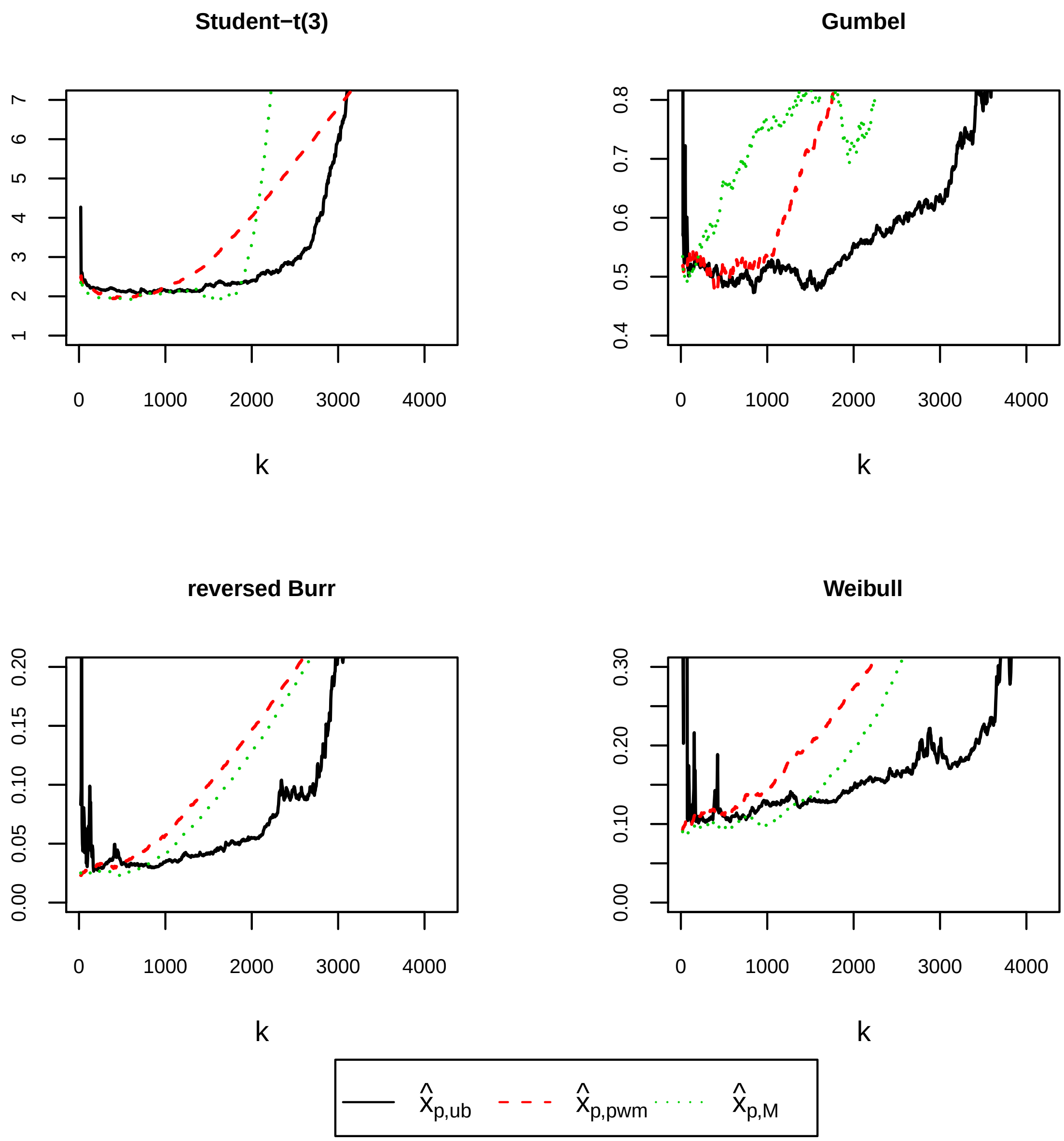

Figure 4: Root mean square error of the high quantile estimates for 100 samples.

Note: Each sample consists of 5000 observations. $\hat{x}_{p,ub}$ is the bias-corrected high quantile estimator in (3.12); $\hat{x}_{p,pwm}$ is the PWM quantile estimator in de Haan and Ferreira (2006); $\hat{x}_{p,M}$ is the moment quantile estimator in Dekkers et al. (1989).

samples generated from the reversed Burr distribution. We again employ the moment and the original PWM methods to produce two competitors. The results are shown in Figure 5. As explained in Remark 3.8, the $\hat{x}^*_{ub}$ does not have an obviously better performance in terms of

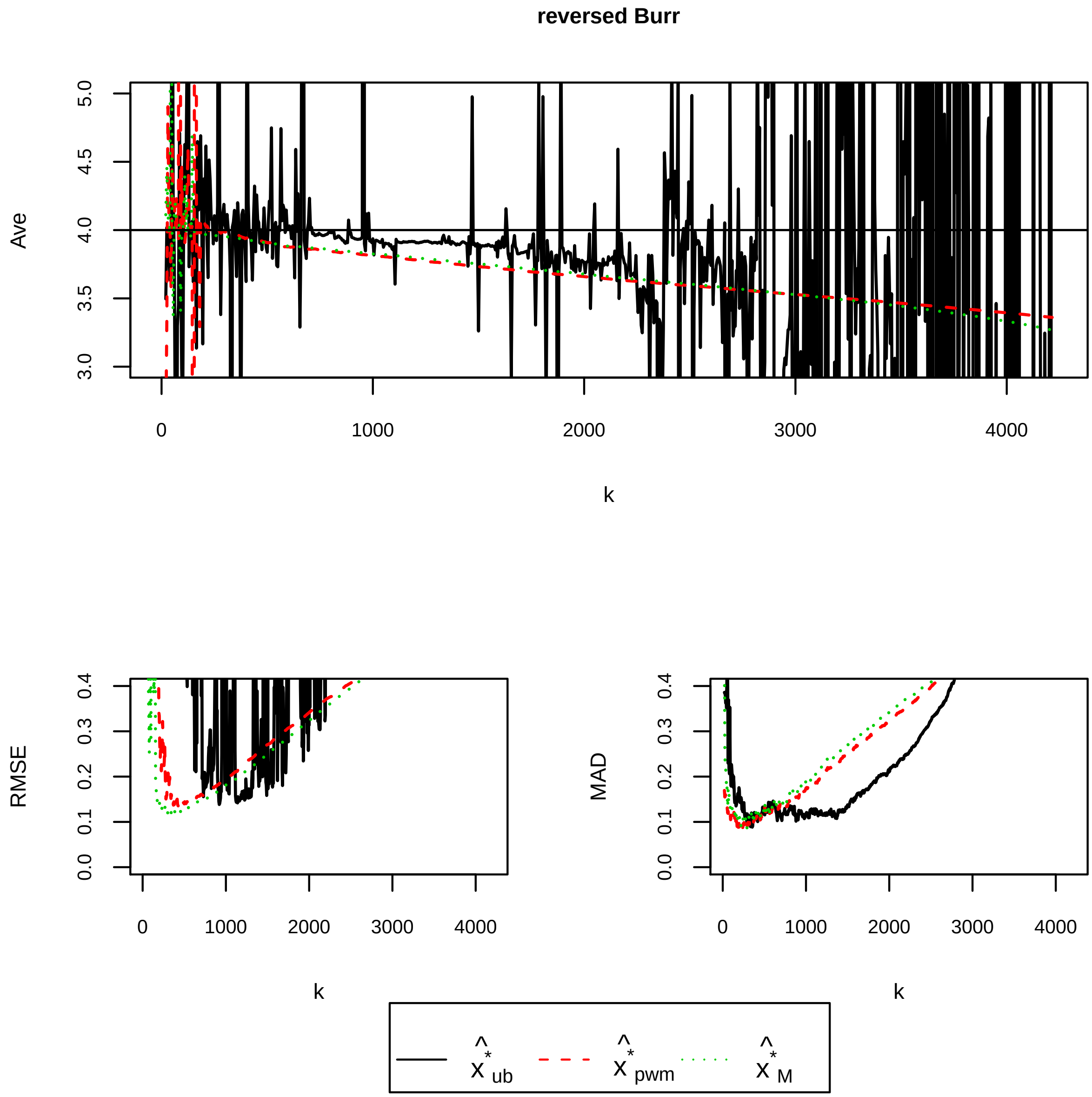

Figure 5: The endpoint estimates for 100 samples.

Note: Each sample consists of 5000 observations generated from reversed Burr distribution with distribution function $F(x) = 1 - (1 + (4 - x)^{-4})^{-5/4}$. The endpoint of the distribution is $x_0 = 4$. $\hat{x}^*_{ub}$ is the bias-corrected endpoint estimator in (3.14); $\hat{x}^*_{pwm}$ is the PWM endpoint estimator; $\hat{x}^*_M$ is the moment endpoint estimator in Dekkers et al. (1989).

RMSE. With an alternative evaluation criteria, the median absolute deviation (MAD)(see Hampel (1974) and Rousseeuw and Croux (1993)), we focus on comparing the estimators in terms of the estimation bias. The MAD is defined as

$$MAD := \text{median}\left(|\hat{x}^*(k) - x^*|\right).$$

As shown in the right bottom plot in Figure 5, the MAD of $\hat{x}^*_{ub}$ remains low and stable for a larger range of k compared to those of $\hat{x}^*_{pwm}$ and $\hat{x}^*_M$.

5 Application to the still water level

As part of the defence against extreme flood in the Netherlands, it is crucial to have accurate estimate of the extreme still water level on various stations along the Dutch coast. The question is to estimate the "once per $10,000$ years" level of the still water. For that reason, Center for Water Management (in Dutch: "Waterdienst") monitored the still water levels in the storm seasons.

We employ the data collected at the station Hoek van Holland. In 122 years, $n = 1965$ severe wind storms have been identified. We treat the still water levels during those storms as i.i.d. observations from a distribution function F. Thus, the "once per $10,000$ years" still water level corresponds to a high quantile of F with tail probability $p = \frac{122}{1965} \times 10^{-4} \approx 6.2 \times 10^{-6}$.

Assuming that F satisfies the third order condition, we start with estimating the extreme value index, γ. We plot the point estimates of γ using our bias-corrected PWM estimator[5] as well as the moment estimator and the original PWM estimator in the upper panel of Figure 6. Similar to the simulation exercises, the bias-corrected PWM method yields a relatively wider range of k, for which the corresponding γ estimates remain at a stable level (k varies from 150 to 280).

To construct a confidence interval, we estimate the asymptotic variance as follows. Denote the standard deviation of the limit distribution in (3.11) as $\sigma_1 := \sigma_1(\gamma, \rho)$, which can be consistently estimated by $\hat{\sigma}_1 := \sigma_1(\hat{\gamma}, \hat{\rho})$. Then the approximate $1 - \alpha$ confidence interval of γ is given by

$$\hat{\gamma}_{ub}(k) - z_{\alpha/2}\frac{\sigma_1(\hat{\gamma}_{ub}(k), \hat{\rho}(k_\rho))}{\sqrt{k}} \leq \gamma \leq \hat{\gamma}_{ub}(k) + z_{\alpha/2}\frac{\sigma_1(\hat{\gamma}_{ub}(k), \hat{\rho}(k_\rho))}{\sqrt{k}},$$

[5] Throughout the application, we choose $k_\rho = [n^{0.98}] = 1688$ for the estimation of the second order index, ρ.

Table 3: Application to the still water level data

k	$\hat{\gamma}_{ub}$	95% confidence interval of γ	$\hat{x}_{p,ub}$	95% confidence interval of x_p
100	0.03	$[-0.23, 0.30]$	526	$[170, 881]$
150	0.07	$[-0.14, 0.29]$	559	$[183, 935]$
250	0.13	$[-0.05, 0.30]$	611	$[206, 1016]$
280	0.10	$[-0.06, 0.27]$	577	$[235, 918]$
300	0.04	$[-0.11, 0.19]$	510	$[269, 752]$
400	-0.03	$[-0.16, 0.10]$	456	$[301, 610]$

Note: There are 1965 observations in the dataset. We choose $k_\rho = 1688$ in the estimation of the second order parameter ρ. $\hat{\gamma}_{ub}$ is defined in (3.10) and $\hat{x}_{p,ub}$ in (3.12) with $p = \frac{122}{1965} \times 10^{-4} \approx 6.2 \times 10^{-6}$.

where $z_{\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution. We plot the lower and upper bounds of the 95% asymptotic confidence intervals in the bottom panel of Figure 6. Since the value zero belongs to most of the confidence intervals with respect to different k, we can not rule out the possibility that $\gamma = 0$. Hence it is necessary to consider the PWM method for this dataset. We give the point estimates and the confidence intervals of γ for several choices of k in the second and third columns of Table 3.

We apply the bias-corrected high quantile estimator to estimate x_p with $p = \frac{122}{1965} \times 10^{-4} \approx 6.2 \times 10^{-6}$. The approximate $1 - \alpha$ confidence interval of x_p is given by

$$\hat{x}_p - z_{\alpha/2} \frac{\sigma_2(\hat{\gamma}_{ub}(k), \hat{\rho}(k_\rho))\hat{a}_{ub}(n/k) q_{\hat{\gamma}_{ub}(k)}(d_n)}{\sqrt{k}} \leq x_p \leq \hat{x}_p + z_{\alpha/2} \frac{\sigma_2(\hat{\gamma}_{ub}(k), \hat{\rho}(k_\rho))\hat{a}_{ub}(n/k) q_{\hat{\gamma}_{ub}(k)}(d_n)}{\sqrt{k}},$$

where $\sigma_2(\gamma, \rho)$ denotes the standard deviation of the limit distribution in Theorem 3.5. The high quantile estimators by the PWM and moment methods are also employed. The results are shown in Figure 7 and the fourth and fifth columns of Table 3. With a choice of $k = 250$, we conclude that the "once per 10000 years" still water level is estimated at 611 cm.

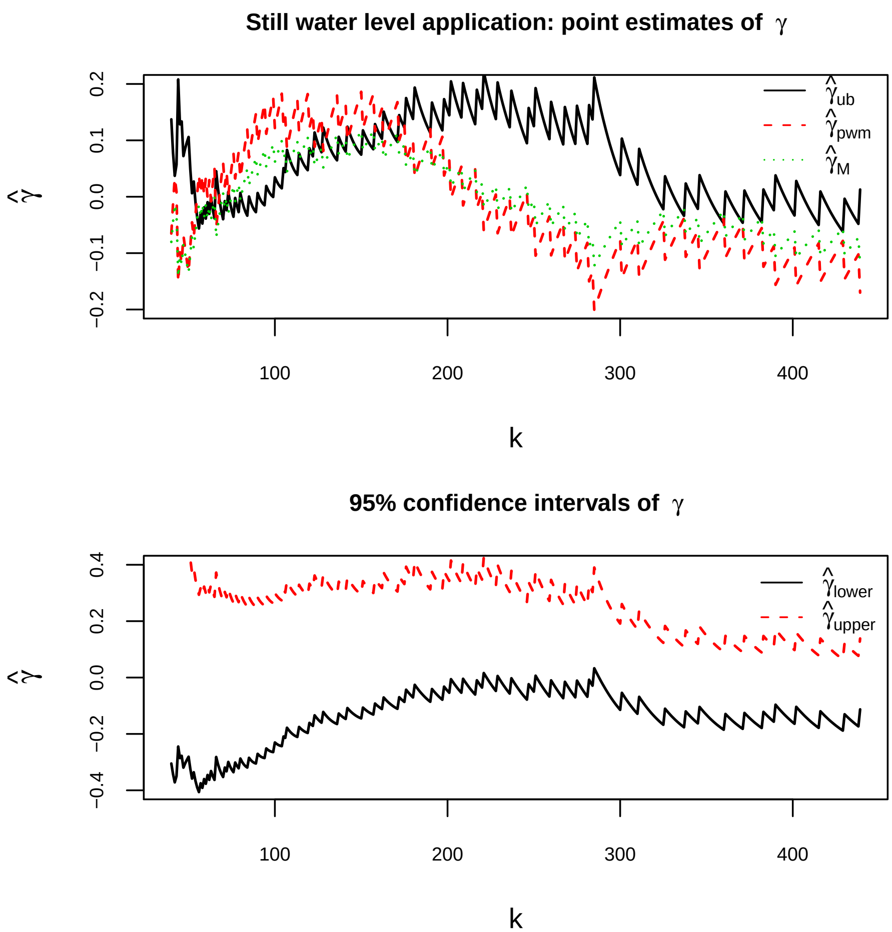

Figure 6: Still water level data: estimation of γ.

Note: The number of observations is 1965. The upper panel shows the point estimates of γ using three estimators: the bias-corrected PWM estimator $\hat{\gamma}_{ub}$ in (3.10), the PWM estimator $\hat{\gamma}_{pwm}$ in (1.1), and the moment estimator $\hat{\gamma}_M$ in Dekkers et al. (1989). The lower panel gives the 95% confidence interval from the bias-corrected estimator.

6 Appendix

Proof of Lemma 3.1 Under the third order condition, applying the expansion of the excesses above a high threshold as in the proof of Drees (1998, Theorem 2.1), one can prove that for each

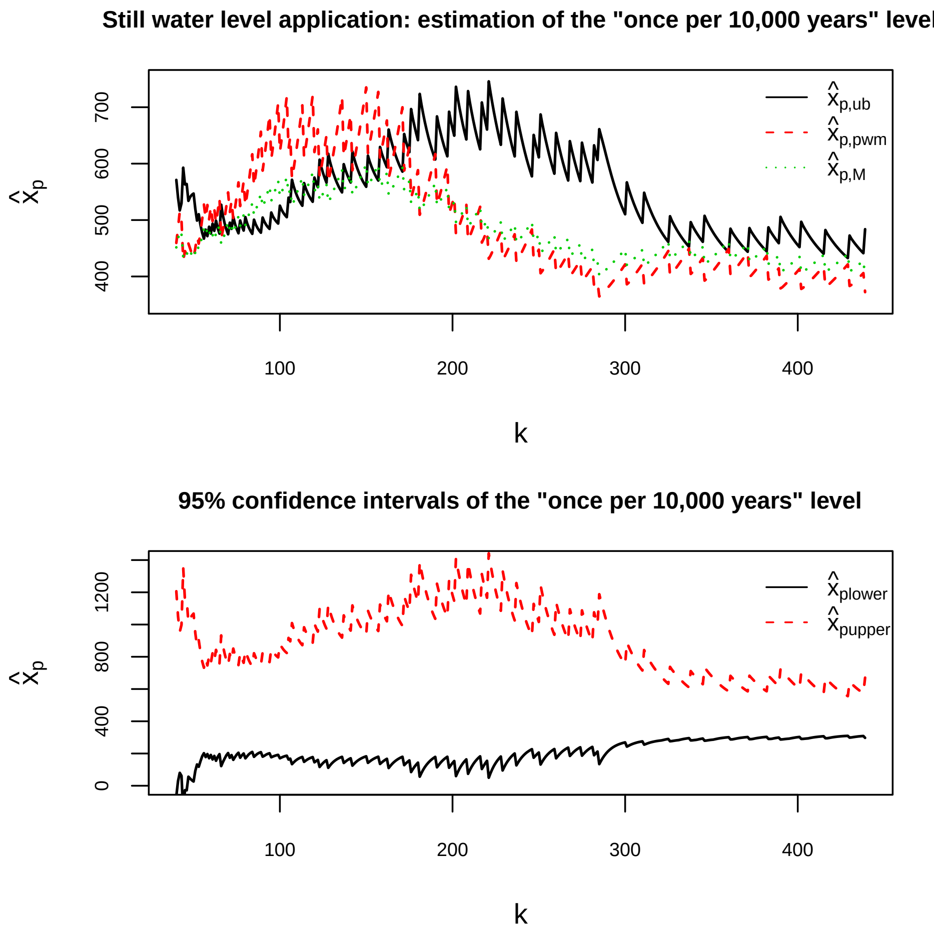

Figure 7: Still water level data: estimation of the "once per $10,000$ years" quantile.

Note: The number of observations is 1965. The "once per $10,000$" level corresponds to a high quantile with tail probability $p = \frac{122}{1965} \times 10^{-4} \approx 6.2 \times 10^{-6}$. The upper panel shows the point estimates of the "once per $10,000$" level using three estimators: the bias-corrected high quantile estimator $\hat{x}_{p,ub}$ in (3.12), the PWM quantile estimator $\hat{x}_{p,pwm}$ in de Haan and Ferreira (2006), and the moment quantile estimator $\hat{x}_{p,M}$ in Dekkers et al. (1989). The lower panel gives the 95% confidence intervals of the "once per $10,000$" level. The scale of the vertical axis is per centimeter.

$\epsilon > 0$,

$$\sup_{1/k \le s \le 1} s^{\gamma+1/2+\epsilon} \Bigg| \sqrt{k} \left(\frac{X_{n,n-[ks]} - U(n/k)}{a(n/k)} - \frac{s^{-\gamma} - 1}{\gamma} \right) - s^{-\gamma-1} W_n(s) \\ - \sqrt{k} A(n/k) \frac{s^{-(\gamma+\rho)} - 1}{\rho(\gamma+\rho)} - \sqrt{k} A(n/k) B(n/k) \frac{s^{-(\gamma+\rho+\rho')} - 1}{\gamma+\rho+\rho'} \Bigg| \xrightarrow{P} 0. \tag{6.1}$$

Similar to the proof of Corollary 2.4.6 in de Haan and Ferreira (2006), one can further show that

$$\sup_{0<s\le 1} \min\left(1, s^{\gamma+1/2+\epsilon}\right) \left| \sqrt{k}\left(\frac{X_{n,n-[ks]}-X_{n,n-k}}{a(n/k)} - \frac{s^{-\gamma}-1}{\gamma}\right) - s^{-\gamma-1}W_n(s) + W_n(1) \right.$$
$$\left. -\sqrt{k}A(n/k)\frac{s^{-(\gamma+\rho)}-1}{\rho(\gamma+\rho)} - \sqrt{k}A(n/k)B(n/k)\frac{s^{-(\gamma+\rho+\rho')}-1}{\gamma+\rho+\rho'} \right| \xrightarrow{P} 0.$$

Hence it follows that as $n \to \infty$,

$$\begin{aligned}
\frac{I_q}{a(n/k)} =& \frac{1}{k}\sum_{i=0}^{k-1}\left(\frac{i}{k}\right)^{q-1}\frac{X_{n,n-i}-X_{n,n-k}}{a(n/k)} \\
=& \int_0^1 \left(\frac{[ks]}{k}\right)^{q-1}\frac{X_{n,n-[ks]}-X_{n,n-k}}{a(n/k)}ds \\
=& \int_0^1 s^{q-1}\left(\frac{s^{-\gamma}-1}{\gamma} + \frac{1}{\sqrt{k}}(s^{-\gamma-1}W_n(s) - W_n(1)) + A(n/k)\frac{s^{-(\gamma+\rho)}-1}{\rho(\gamma+\rho)}\right)ds \\
&+ A(n/k)B(n/k)O_p(1) + O_p\left(\frac{1}{k}\right) \\
=& \frac{1}{q(q-\gamma)} + \frac{1}{\sqrt{k}}L_q + A(n/k)B_q + \varepsilon_{n,k}.
\end{aligned}$$

Here we apply $\left|\frac{[ks]}{k} - s\right| \le \frac{1}{k}$ for the third equality.

For the proof of the Theorems 3.2, 3.5 and 3.7, we need several lemmas. Firstly, we give a lemma on the uniform convergence of the third order condition. The proof is analogue to that in Fraga Alves et al. (2003), thus we omit it here.

Lemma 6.1 *Suppose F satisfies the third order condition* (2.3) *with $\rho, \rho' < 0$. There exists a particular choice of the first, second and third order scale functions $\tilde{a}(t)$, $\tilde{A}(t)$ and $\tilde{B}(t)$ such that for any $\delta > 0$, there exists a threshold $T(\delta)$, the inequality*

$$\left| \frac{\frac{\frac{U(tx)-U(t)}{\tilde{a}(t)} - \frac{x^{\gamma}-1}{\gamma}}{\tilde{A}(t)} - \frac{x^{\gamma+\rho}-1}{\rho(\gamma+\rho)}}{\tilde{B}(t)} - \frac{x^{\gamma+\rho+\rho'}-1}{\gamma+\rho+\rho'} \right| < \delta x^{\gamma+\rho+\rho'\pm\delta} \tag{6.2}$$

holds for all t, x such that $t, tx > T(\delta)$.

Without loss of generality, we still use the notations a, A and B for $\tilde{a}$, $\tilde{A}$ and $\tilde{B}$.

Lemma 6.2 and 6.3 give the asymptotic properties of the estimators of ρ and $A(n/k)$.

Lemma 6.2 *With an intermediate sequence k_ρ satisfying the condition (2.5), we have that as $n \to \infty$,*

$$\sqrt{k_\rho} A(n/k_\rho)(\hat{\rho}(k_\rho) - \rho) = O_p(1).$$

Lemma 6.3 *With an intermediate sequence k satisfying the condition* (2.4)*, we have that as $n \to \infty$*

$$\begin{aligned}\hat{A}(n/k) - A(n/k) = & \frac{1}{\sqrt{k}} \frac{(1-\gamma-\rho)(2-\gamma-\rho)(3-\gamma-\rho)}{2\rho}((1-\gamma)L_1 \\ & - 2(2-\gamma)L_2 + (3-\gamma)L_3) + A(n/k)A(n/k_\rho)O_p(1). \end{aligned} \tag{6.3}$$

Corollary 6.4 *Under the condition in Lemma 6.3, $\frac{\hat{A}(n/k)}{A(n/k)} \xrightarrow{p} 1$ as $n \to \infty$.*

The following lemma gives the asymptotic property of an intermediate order statistics, which can be regarded as an estimator of $U(n/k)$, see de Haan and Ferreira (2006).

Lemma 6.5 *Suppose the third order condition (2.3) holds with $\rho, \rho' < 0$. Let $k = k(n)$ satisfy $\sqrt{k}\,|A(n/k)B(n/k)| = O(1)$, as $n \to \infty$ Then*

$$\frac{X_{n,n-k} - U(n/k)}{a(n/k)} = \frac{1}{\sqrt{k}} W(1) + A(n/k)B(n/k)o_p(1),$$

where W is the Brownian motion defined in Lemma 3.1.

Lemma 6.6 states the asymptotic property of the bias-corrected estimator of the scale function a.

Lemma 6.6 *Suppose the third order condition (2.3) holds with $\gamma < 1/2$, $\rho, \rho' < 0$ and the intermediate sequence k satisfies condition* (2.4). *For $\hat{a}_{ub}(n/k)$ defined in (3.13), we have that, as $n \to \infty$,*

$$\begin{aligned}\frac{\hat{a}_{ub}(n/k)}{a(n/k)} - 1 = & \frac{1}{\sqrt{k}} \frac{(1-\gamma)(2-\gamma)(3-\gamma)}{2\rho} \left((\gamma+\rho-1)\left(\frac{1}{\rho} - \frac{1}{2-\gamma} - \frac{1}{3-\gamma}\right) L_1 \right. \\ & - 2(\gamma+\rho-2)\left(\frac{1}{\rho} - \frac{1}{1-\gamma} - \frac{1}{3-\gamma}\right) L_2 \\ & \left. + (\gamma+\rho-3)\left(\frac{1}{\rho} - \frac{1}{1-\gamma} - \frac{1}{2-\gamma}\right) L_3 \right) + A(n/k)A(n/k_\rho)O_p(1). \end{aligned} \tag{6.4}$$

Proof of Lemma 6.2 Since $\rho = 1-\gamma-\frac{1}{\frac{2-\gamma-\rho}{1-\gamma-\rho}-1}$, it is sufficient to notice that (3.5) implies—using condition (2.5)—that

$$\sqrt{k_\rho}A(n/k_\rho)\left(\frac{\hat\gamma_{3,1}(k_\rho)-\hat\gamma_{4,1}(k_\rho)}{\hat\gamma_{3,2}(k_\rho)-\hat\gamma_{4,2}(k_\rho)} - \frac{1-\gamma}{2-\gamma}\cdot\frac{2-\gamma-\rho}{1-\gamma-\rho}\right) = O_p(1)$$

and that (3.4) implies (with condition (2.5)) that

$$\sqrt{k_\rho}A(n/k_\rho)(\hat\gamma_{1,2}(k_\rho)-\gamma) = O_p(1).$$

Proof of Lemma 6.3 With $g(\gamma,\rho) := \frac{(1-\gamma-\rho)(2-\gamma-\rho)(3-\gamma-\rho)}{\rho(1-\gamma)}$,

$$\frac{\hat A(n/k)}{A(n/k)} - 1 = \frac{(\hat\gamma_{2,1}(k)-\hat\gamma_{3,1}(k))g(\gamma,\rho)}{A(n/k)}\cdot\frac{g(\hat\gamma_{2,1}(k),\hat\rho(k_\rho))}{g(\gamma,\rho)} - 1.$$

By (3.8), the right hand side is approximately

$$\frac{g(\gamma,\rho)N_{3,2,1}}{\sqrt{k}A(n/k)} + \left(\frac{g(\hat\gamma_{2,1}(k),\hat\rho(k_\rho))}{g(\gamma,\rho)} - 1\right).$$

Thus it is sufficient to show that

$$g(\hat\gamma_{2,1}(k),\hat\rho(k_\rho)) = g(\gamma,\rho) + A(n/k_\rho)O_p(1),$$

which follows from $\hat\gamma_{2,1}(k)-\gamma = A(n/k)O_p(1)$ (from (3.4)) and $\hat\rho-\rho = \frac{1}{\sqrt{k_\rho}A(n/k_\rho)}O_p(1) = A(n/k_\rho)O_p(1)$ (from Lemma 6.2).

Proof of Lemma 6.5 Write $X_{n,n-k} = U(Y_{n,n-k})$ with $Y_{n,n-k}$ the corresponding intermediate order statistics from the distribution function $1-1/x$, $x>1$. By taking $t = n/k$ and $x = \frac{k}{n}Y_{n,n-k}$ in (6.2), we get

$$\begin{aligned}\sqrt{k}\frac{X_{n,n-k}-U(n/k)}{a(n/k)} =& \sqrt{k}\frac{(\frac{k}{n}Y_{n,n-k})^\gamma - 1}{\gamma} + A(n/k)\sqrt{k}\frac{(\frac{k}{n}Y_{n,n-k})^{\gamma+\rho}-1}{\rho(\gamma+\rho)} \\ &+ \sqrt{k}A(n/k)B(n/k)\left(\frac{(\frac{k}{n}Y_{n,n-k})^{\gamma+\rho+\rho'}-1}{\gamma+\rho+\rho'} + \left(\frac{k}{n}Y_{n,n-k}\right)^{\gamma+\rho+\rho'}o(1)\right)\end{aligned}$$

Since $\sqrt{k}\left(\frac{k}{n}Y_{n,n-k}-1\right)\xrightarrow{d}W(1)$, as $n\to\infty$, we get that

$$\sqrt{k}\frac{X_{n,n-k}-U(n/k)}{a(n/k)}=W(1)+A(n/k)W(1)+o_p(1)+A(n/k)B(n/k)O_p\left(1\right)+\sqrt{k}A(n/k)B(n/k)o_p\left(1\right).$$

Hence the result follows from $\sqrt{k}\left|A(n/k)B(n/k)\right|=O(1)$.

Proof of Lemma 6.6 Write $h(\gamma,\rho):=\frac{(1-\gamma)(2-\gamma)-\rho(3-2\gamma)}{\rho(1-\gamma-\rho)(2-\gamma-\rho)}$. Then $\hat{a}_{ub}(n/k)=\hat{a}(n/k)\exp(-\hat{A}(n/k)h(\hat{\gamma},\hat{\rho}))$. From Lemma 3.1, one can obtain that

$$\frac{\hat{a}(n/k)}{a(n/k)}=\frac{\frac{I_1}{a(n/k)}}{\frac{I_1/a(n/k)}{2I_2/a(n/k)}-1}=1+\frac{1}{\sqrt{k}}((\gamma-1)L_1-(\gamma-2)L_2)+A(n/k)h(\gamma,\rho)+\varepsilon_{n,k}. \tag{6.5}$$

Similar to the proof of Lemma 6.3, we get

$$h(\hat{\gamma},\hat{\rho})=h(\gamma,\rho)+A(n/k_\rho)O_p\left(1\right) \tag{6.6}$$

Moreover, from Lemma 6.3, we get

$$\hat{A}(n/k)=A(n/k)+\frac{1}{\sqrt{k}}N_A+A(n/k)A(n/k_\rho)O_p(1) \tag{6.7}$$

with $N_A=\frac{(1-\gamma-\rho)(2-\gamma-\rho)(3-\gamma-\rho)}{2\rho}((1-\gamma)L_1-2(2-\gamma)L_2+(3-\gamma)L_3)$.

By combining the expansions (6.5), (6.6) and (6.7), we have that

$$\frac{\hat{a}_{ub}(n/k)}{a(n/k)}=1+\frac{1}{\sqrt{k}}((\gamma-1)L_1-(\gamma-2)L_2-h(\gamma,\rho)N_A)+A(n/k)A(n/k_\rho)O_p(1),$$

which can be verified as equivalent to (6.4).

Proof of Theorem 3.2 The result follows from (3.4), Lemma 6.2 and Lemma 6.3.

Proof of Theorem 3.5 Notice that $x_p=U(1/p)$. Thus, we can write

$$\begin{aligned}\frac{\sqrt{k}}{a(n/k)q_\gamma(d_n)}(\hat{x}_p-U(1/p))=&\frac{\sqrt{k}}{q_\gamma(d_n)}\frac{X_{n,n-k}-U(n/k)}{a(n/k)}\\&+\frac{\hat{a}_{ub}(n/k)}{a(n/k)}\cdot\frac{\sqrt{k}}{q_\gamma(d_n)}\left(\frac{d_n^{\hat{\gamma}}-1}{\hat{\gamma}}+\frac{d_n^{\gamma}-1}{\gamma}\right)\\&+\frac{d_n^{\gamma}-1}{\gamma q_\gamma(d_n)}\sqrt{k}\left(\frac{\hat{a}_{ub}(n/k)}{a(n/k)}-1\right)\end{aligned}$$

$$
\begin{aligned}
&+\frac{\sqrt{k}A(n/k)}{q_\gamma(d_n)}\left(\frac{\hat{A}(n/k)\hat{a}(n/k)}{A(n/k)a(n/k)}\frac{d_n^{\hat{\gamma}+\hat{\rho}}-1}{\hat{\rho}(\hat{\gamma}+\hat{\rho})}-\frac{d_n^{\gamma+\rho}}{\rho(\gamma+\rho)}\right)\\
&-\frac{\sqrt{k}A(n/k)}{q_\gamma(d_n)}\left(\frac{\frac{U(1/p)-U(n/k)}{a(n/k)}-\frac{d_n^{\gamma}-1}{\gamma}}{A(n/k)}-\frac{d_n^{\gamma+\rho}}{\rho(\gamma+\rho)}\right)\\
&=: I_1+I_2+I_3+I_4+I_5.
\end{aligned}
$$

By Lemma 6.5, $I_1 \xrightarrow{d} \Gamma_1$. As in the proof of Theorem 4.3.1 in de Haan and Ferreira (2006, page 136-137), $I_2 \xrightarrow{d} \Gamma_2$. By Lemma 6.6, $I_3 \xrightarrow{d} \Gamma_3$.

For the term I_4, similar to the term I_2, we have that

$$
\frac{d_n^{\hat{\gamma}+\hat{\rho}}-1}{(\hat{\gamma}+\hat{\rho})}-\frac{d_n^{\gamma+\rho}}{(\gamma+\rho)}=q_{\gamma+\rho}(d_n)(\hat{\gamma}+\hat{\rho}-\gamma-\rho)(1+o_p(1))=A(n/k_\rho)q_{\gamma+\rho}(d_n)O_p(1).
$$

Since $\frac{1}{\hat{\rho}}=\frac{1}{\rho}(1+A(n/k_\rho)O_p(1))$, we get

$$
\frac{d_n^{\hat{\gamma}+\hat{\rho}}-1}{\hat{\rho}(\hat{\gamma}+\hat{\rho})}=\frac{d_n^{\gamma+\rho}-1}{\rho(\gamma+\rho)}\left(1+\frac{(\gamma+\rho)A(n/k_\rho)q_{\gamma+\rho}(d_n)}{d_n^{\gamma+\rho}-1}O_p\left(1\right)\right). \tag{6.8}
$$

Lemma 6.6 implies that

$$
\frac{\hat{a}_{ub}(n/k)}{a(n/k)}=1+\frac{1}{\sqrt{k}}O_p(1) \tag{6.9}
$$

and by Lemma 6.3

$$
\frac{\hat{A}(n/k)}{A(n/k)}=1+\frac{N_A}{\sqrt{k}A(n/k)}(1+o_p(1)). \tag{6.10}
$$

Combing (6.8), (6.9) and (6.10) we get the expansion of I_4 as

$$
\begin{aligned}
I_4 =&\frac{\sqrt{k}A(n/k)}{q_\gamma(d_n)}\frac{d_n^{\gamma+\rho}-1}{\rho(\gamma+\rho)}\left(\left(1+\frac{N_A}{\sqrt{k}A(n/k)}(1+o_p(1))\right)\left(1+\frac{1}{\sqrt{k}}O_p(1)\right)\right.\\
&\left.\left(1+\frac{(\gamma+\rho)A(n/k_\rho)q_{\gamma+\rho}(d_n)}{d_n^{\gamma+\rho}-1}O_p(1)\right)-1\right)\\
=&\frac{d_n^{\gamma+\rho}-1}{\rho(\gamma+\rho)q_\gamma(d_n)}N_A+\frac{d_n^{\gamma+\rho}-1}{\rho(\gamma+\rho)q_\gamma(d_n)}o_p(1)+\frac{\sqrt{k}A(n/k)A(n/k_\rho)q_{\gamma+\rho}(d_n)}{q_\gamma(d_n)}O_p(1).
\end{aligned}
$$

Hence, together with the relation $\lim_{n\to\infty}\frac{d_n^{\gamma+\rho}-1}{\rho(\gamma+\rho)q_\gamma(d_n)}=-\frac{(\gamma_-)^2}{\rho(\gamma_-+\rho)}$, we get $I_4 \xrightarrow{d} \Gamma_4$.

For part I_5, we use the inequality (6.2) and obtain that

$$I_5 = -\frac{\sqrt{k}A(n/k)}{q_\gamma(d_n)} B(n/k) \left(\frac{d_n^{\gamma+\rho+\rho'} - 1}{\gamma+\rho+\rho'} + d_n^{\gamma+\rho+\rho'} o(1) \right) = \sqrt{k}A(n/k)B(n/k)O(1) \to 0.$$

Proof of Theorem 3.7 The proof is similar to that of Theorem 3.5.

References

K. Aarssen and L. de Haan. On the maximal life span of humans. *Mathematical Population Studies*, 4(4):259–281, 1994.

A.A. Balkema and L. De Haan. Residual life time at great age. *The Annals of Probability*, 2(5): 792–804, 1974.

J. Beirlant, F. Figueiredo, M.I. Gomes, and B. Vandewalle. Improved reduced-bias tail index and quantile estimators. *Journal of Statistical Planning and Inference*, 138(6):1851–1870, 2008.

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

F. Caeiro and M.I. Gomes. A class of asymptotically unbiased semi-parametric estimators of the tail index. *Test*, 11(2):345–364, 2002.

F. Caeiro and M.I. Gomes. Minimum-variance reduced-bias tail index and high quantile estimation. *REVSTAT–Statistical Journal*, 6(1):1–20, 2008.

F. Caeiro, M.I. Gomes, and D. Pestana. Direct reduction of bias of the classical Hill estimator. *REVSTAT–Statistical Journal*, 3(2):113–136, 2005.

S.G. Coles and J.A. Tawn. Modelling extreme multivariate events. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2):377–392, 1991.

S.G. Coles and J.A. Tawn. Modelling extremes of the areal rainfall process. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(2):329–347, 1996.

S.G. Coles and D. Walshaw. Directional modelling of extreme wind speeds. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 43(1):139–157, 1994.

L. de Haan and J. de Ronde. Sea and wind: multivariate extremes at work. *Extremes*, 1(1):7–45, 1998.

L. de Haan and A. Ferreira. *Extreme value theory: an introduction*. Springer Verlag, 2006.

L. de Haan and H. Rootzén. On the estimation of high quantiles. *Journal of Statistical Planning and Inference*, 35(1):1–13, 1993.

L. de Haan and U. Stadtmüller. Generalized regular variation of second order. *Journal of the Australian Mathematical Society (Series A)*, 61(03):381–395, 1996.

A.L.M. Dekkers, J.H.J. Einmahl, and L. de Haan. A moment estimator for the index of an extreme-value distribution. *Annals of Statistics*, 17(4):1833–1855, 1989.

H. Drees. On smooth statistical tail functionals. *Scandinavian Journal of Statistics*, 25(1):187–210, 1998.

J.H.J. Einmahl and J.R. Magnus. Records in athletics through extreme-value theory. *Journal of the American Statistical Association*, 103(484):1382–1391, 2008.

A. Feuerverger and P. Hall. Estimating a tail exponent by modelling departure from a Pareto distribution. *Annals of Statistics*, 27(2):760–781, 1999.

M.I. Fraga Alves, L. de Haan, and T. Lin. Estimation of the parameter controlling the speed of convergence in extreme value theory. *Mathematical Methods of Statistics*, 12(2):155–176, 2003.

M.I. Gomes and F. Figueiredo. Bias reduction in risk modelling: semi-parametric quantile estimation. *Test*, 15(2):375–396, 2006.

M.I. Gomes and M.J. Martins. Asymptotically Unbiased Estimators of the Tail Index Based on External Estimation of the Second Order Parameter. *Extremes*, 5(1):5–31, 2002.

M.I. Gomes and M.J. Martins. Bias reduction and explicit semi-parametric estimation of the tail index. *Journal of statistical planning and inference*, 124(2):361–378, 2004.

M.I. Gomes and D. Pestana. A sturdy reduced-bias extreme quantile (VaR) estimator. *Journal of the American Statistical Association*, 102(477):280–292, 2007a.

M.I. Gomes and D. Pestana. A simple second-order reduced-bias tail index estimator. *Journal of Statistical Computation and Simulation*, 77(6):487–502, 2007b.

M.I. Gomes, M.J. Martins, and M. Neves. Alternatives to a Semi-Parametric Estimator of Parameters of Rare EventsThe Jackknife Methodology. *Extremes*, 3(3):207–229, 2000.

M.I. Gomes, M.J. Martins, and M. Neves. Generalized Jackknife semi-parametric estimators of the tail index. *Portugaliae Mathematica*, 59(4):393–408, 2002.

M.I. Gomes, F. Caeiro, and F. Figueiredo. Bias reduction of a tail index estimator through an external estimation of the second-order parameter. *Statistics*, 38(6):497–510, 2004.

M.I. Gomes, H. Pereira, and M.C. Miranda. Revisiting the role of the Jackknife methodology in the estimation of a positive tail index. *Communications in statistics. Theory and methods*, 34 (2):319–335, 2005.

M.I. Gomes, C. Miranda, and C. Viseu. Reduced-bias tail index estimation and the jackknife methodology. *Statistica Neerlandica*, 61(2):243–270, 2007.

M.I. Gomes, L. de Haan, and L.H. Rodrigues. Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1):31–52, 2008.

P. Hall. On estimating the endpoint of a distribution. *Annals of Statistics*, 10(2):556–568, 1982.

Frank R. Hampel. The Influence Curve and its Role in Robust Estimation. *Journal of the American Statistical Association*, 69(346):383–393, 1974.

B.M. Hill. A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3(5):1163–1174, 1975.

J.R.M. Hosking and J.R. Wallis. Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, 29(3):339–349, 1987.

D. Li and L. Peng. Does bias reduction with external estimator of second order parameter work for endpoint? *Journal of Statistical Planning and Inference*, 139(6):1937–1952, 2009.

D. Li, L. Peng, and X. Xu. Bias reduction for endpoint estimation. *Extremes*, To appear, 2010a.

D. Li, L. Peng, and J. Yang. Bias reduction for high quantiles. *Journal of Statistical Planning and Inference*, To appear, 2010b.

G. Matthys, E. Delafosse, A. Guillou, and J. Beirlant. Estimating catastrophic quantile levels for heavy-tailed distributions. *Insurance: Mathematics and Economics*, 34(3):517–537, 2004.

L. Peng. Asymptotically unbiased estimators for the extreme-value index. *Statistics & Probability Letters*, 38(2):107–115, 1998.

Peter J. Rousseeuw and Christophe Croux. Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association*, 88(424):1273–1283, 1993.

R.L. Smith. Estimating tails of probability distributions. *Annals of Statistics*, 15(3):1174–1207, 1987.

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