AN OVERVIEW AND OPEN RESEARCH TOPICS IN THE FIELD OF STATISTICS OF UNIVARIATE EXTREMES

Authors: Jan Beirlant
– Catholic University of Leuven,
Belgium (Jan.Beirlant@wis.kuleuven.be)

Frederico Caeiro
– New University of Lisbon, F.C.T. and C.M.A.,
Portugal (fac@fct.unl.pt)

M. Ivette Gomes
– University of Lisbon, F.C.U.L. (D.E.I.O.) and C.E.A.U.L.,
Portugal (ivette-gomes@fc.ul.pt)

Abstract:

- This review paper focuses on statistical issues arising in modeling univariate extremes of a random sample. In the last three decades there has been a shift from the area of parametric statistics of extremes, based on probabilistic asymptotic results in extreme value theory, towards a semi-parametric approach, where the estimation of the right and/or left tail-weight is performed under a quite general framework. But new parametric models can still be of high interest for the analysis of extreme events, if associated with adequate statistical inference methodologies. After a brief reference to Gumbel's classical block methodology and later improvements in the parametric framework, we present an overview of the developments on the estimation of parameters of extreme events and testing of extreme value conditions under a semi-parametric framework, and discuss a few challenging topics of open research in the area.

Key-Words:

- *Parameters of extreme events; extreme value index; parametric and semi-parametric estimation and testing; statistics of univariate extremes.*

AMS Subject Classification:

- 62G32, 62E20.

1. INTRODUCTION, LIMITING RESULTS IN THE FIELD OF EXTREMES AND PARAMETRIC APPROACHES

We shall assume to have access to a sample $(X_1, \ldots, X_n)$ of n *independent, identically distributed* (IID) or possibly *stationary, weakly dependent* random variables from an underlying *cumulative distribution function* (CDF), F, and shall use the notation $(X_{1,n} \leq \cdots \leq X_{n,n})$ for the sample of associated ascending *order statistics* (OSs). *Statistics of univariate extremes* (SUE) helps us to learn from disastrous or almost disastrous events, of high relevance in society and with a high social impact. The domains of application of SUE are thus quite diversified. We mention the fields of hydrology, meteorology, geology, insurance, finance, structural engineering, telecommunications and biostatistics (see, for instance, and among others, Coles, 2001; Reiss and Thomas, 2001, 2007; Beirlant *et al.*, 2004, Section 1.3; Castillo *et al.*, 2005; Resnick, 2007). Although it is possible to find some historical papers with applications related to extreme events, the field dates back to Gumbel, in papers from 1935 on, summarized in his book (Gumbel, 1958). Gumbel develops statistical procedures essentially based on Gnedenko's (Gnedenko, 1943) *extremal types theorem* (ETT), one of the main limiting results in the field of *extreme value theory* (EVT), briefly summarized in the following.

1.1. Main limiting results in EVT

The main limiting results in EVT date back to the papers by Fréchet (1927), Fisher and Tippett (1928), von Mises (1936) and Gnedenko (1943). Gnedenko's ETT provides the possible limiting behaviour of the sequence of maximum or minimum values, linearly normalised, and an incomplete characterization, fully achieved in de Haan (1970), of the domains of attraction of the so-called *max-stable* (MS) or *min-stable* laws. Here, we shall always deal with the right-tail, $\overline{F}(x) := 1 - F(x)$, for large x, i.e. we shall deal with top OSs. But all results for maxima (top OSs) can be easily reformulated for minima (low OSs). Indeed, $X_{1,n} = -\max_{1 \leq i \leq n}(-X_i)$, and consequently, $\mathbb{P}(X_{1,n} \leq x) = 1 - (1 - F(x))^n$. MS laws are defined as laws S such that the functional equation $S^n(\alpha_n x + \beta_n) = S(x)$, $n \geq 1$, holds for some $\alpha_n > 0$, $\beta_n \in \mathbb{R}$. More specifically, all possible non-degenerate weak limit distributions of the normalized partial maxima $X_{n,n}$, of IID random variables $X_1, \ldots, X_n$, are (generalized) *extreme value distributions* (EVDs), i.e. if there are normalizing constants $a_n > 0$, $b_n \in \mathbb{R}$ and some non-

degenerate CDF G such that, for all x,

$$(1.1) \qquad \lim_{n \to \infty} \mathbb{P}\left\{ (X_{n,n} - b_n)/a_n \leq x \right\} = G(x),$$

we can redefine the constants in such a way that,

$$(1.2) \qquad G(x) \equiv G_\gamma(x) := \begin{cases} \exp\left(-(1 + \gamma x)^{-1/\gamma}\right), \ 1 + \gamma x > 0 \ \text{if} \ \gamma \neq 0 \\ \exp(-\exp(-x)), \ x \in \mathbb{R} \qquad\qquad \text{if} \ \gamma = 0, \end{cases}$$

given here in the von Mises-Jenkinson form (von Mises, 1936; Jenkinson, 1955). If (1.1) holds, we then say that the CDF F which is underlying $X_1, X_2, \ldots$, is in the *max-domain of attraction* (MDA) of G_γ, in (1.2), and use often the notation $F \in \mathcal{D}_\mathcal{M}(G_\gamma)$. The limiting CDFs, G, in (1.1), are then MS. They are indeed the unique MS laws. The real parameter γ, the primary parameter of interest in extreme value analysis (EVA), is called the *extreme value index* (EVI). The EVI, γ, rules the behaviour of the right-tail of F. The EVD, in (1.2), is often separated in the three following types:

$$(1.3) \qquad \begin{array}{lll} \text{Type I (Gumbel)}: & \Lambda(x) = \exp(-\exp(-x)), \ x \in \mathbb{R}, \\ \text{Type II (Fréchet)}: & \Phi_\alpha(x) = \exp(-x^{-\alpha}), \ x \geq 0, \\ \text{Type III (max-Weibull)}: & \Psi_\alpha(x) = \exp(-(-x)^\alpha), \ x \leq 0. \end{array}$$

Indeed, with $\gamma = 0$, $\gamma = 1/\alpha > 0$ and $\gamma = -1/\alpha < 0$, respectively, we have $\Lambda(x) = G_0(x)$, $\Phi_\alpha(x) = G_{1/\alpha}(\alpha(1-x))$ and $\Psi_\alpha(x) = G_{-1/\alpha}(\alpha(x+1))$, with G_γ the EVD in (1.2). The Fréchet domain of attraction ($\gamma > 0$) contains heavy-tailed CDFs like the Pareto and the Student's t-distributions, i.e. tails of a negative polynomial type and infinite right endpoint. Short-tailed CDFs, with finite right endpoint, like the beta CDFs, belong to the Weibull MDA ($\gamma < 0$). The Gumbel MDA ($\gamma = 0$), is relevant for many applied sciences, and contains a great variety of CDFs with an exponential tail, like the normal, the exponential and the gamma, but not necessarily with an infinite right endpoint. As an example of a CDF $F \in \mathcal{D}_\mathcal{M}(G_0)$, with a finite right endpoint x^F, we have the *exponential-type distribution*, $F(x) = K \exp(-c/\{x^F - x\})$, for $x < x^F$, $c > 0$, and $K > 0$.

Apart from the ETT and the already mentioned EVD, in (1.2), it is also worth mentioning the *generalized Pareto distribution* (GPD), the limit distribution of scaled excesses over high thresholds (see the pioneering papers by Balkema and de Haan, 1974; Pickands, 1975), which can be written as

$$(1.4) \quad P_\gamma(x) = 1 + \ln G_\gamma(x) = \begin{cases} 1 - (1 + \gamma x)^{-1/\gamma}, \ 1 + \gamma x > 0, \ x > 0 \ \text{if} \ \gamma \neq 0 \\ 1 - \exp(-x), \ x > 0, \qquad\qquad\qquad \text{if} \ \gamma = 0, \end{cases}$$

with G_γ given in (1.2), as well as the *multivariate* EVD, related with the limiting distribution of the k largest values $X_{n-i+1:n}, 1 \leq i \leq k$, also called the *extremal*

process (Dwass, 1964), with associated probability density function (PDF)

$$(1.5) \qquad h_\gamma(x_1, x_2, \cdots, x_k) = g_\gamma(x_k) \prod_{j=1}^{k-1} \frac{g_\gamma(x_j)}{G_\gamma(x_j)} \quad \text{if} \quad x_1 > x_2 > \cdots > x_k,$$

where $g_\gamma(x) = dG_\gamma(x)/dx$, with $G_\gamma(x)$ given in (1.2).

1.2. Parametric approaches to SUE

Deciding upon the right tail-weight for the distribution underlying the sample data, through a proper EVI-estimation, constitutes a very important starting task in EVA. On the other hand, statistical inference about rare events is clearly linked to observations which are extreme in some sense. There are different ways to define such observations, leading to different approaches to SUE. We next briefly refer the most common parametric approaches to SUE. For further details on the topic, and pioneering papers on the subject, see Gomes *et al.* (2008a).

Block maxima (BM) method. With $(\lambda_n, \delta_n) \in \mathbb{R} \times \mathbb{R}^+$, a vector of unknown location and scale parameters, the ETT validates the approximation

$$(1.6) \qquad\qquad P\big(X_{n,n} \leq x\big) = F^n(x) \approx G_\gamma\left((x - \lambda_n)/\delta_n\right).$$

Gumbel was pioneer in the use of approximations of the type of the one provided in (1.6), but for any of the models in (1.3), suggesting the first model in SUE, usually called the BM model or the *annual maxima* model or the *extreme value* (EV) *univariate model* or merely *Gumbel's model*. The sample of size n is divided into k sub-samples of size r (usually associated to k years, with $n = r \times k$, r reasonably large). Next, the maximum of the r observations in each of the k sub-samples is considered, and one of the extremal models in (1.3), obviously with extra unknown location and scale parameters, is fitted to such a sample. Nowadays, whenever using this approach, still quite popular in environmental sciences, it is more common to fit to the data a univariate EVD, $G_\gamma((x - \lambda_r)/\delta_r)$, with G_γ given in (1.2), $(\lambda_r, \delta_r, \gamma) \in (\mathbb{R}, \mathbb{R}^+, \mathbb{R})$ unknown location, scale and "shape" parameters. All statistical inference is then related to EVDs.

The method of largest observations (LO). Although the BM-method has proved to be fruitful in the most diversified situations, several criticisms have been made on Gumbel's technique, and one of them is the fact that we are wasting information when using only observed maxima and not further top OSs, if available, because they surely contain useful information about the right-tail of the CDF underlying the data. To infer on the right-tail weight of the underlying

model, it seems sensible to think on a small number k of top OSs from the original data, and when the sample size n is large and k fixed, it is sensible to consider the *multivariate* EVD, with a standardized PDF given in (1.5). Again, unknown location and scale parameters, λ_n and δ_n, respectively, are considered and estimated on the basis of the k top OSs, out of n. This approach to SUE is the so-called LO method or *multivariate* EV *model*. It is now easier to increase the number k of observations, contrarily to what happens in Gumbel's approach.

Multi-dimensional EV approaches.   It is obviously feasible to combine the two aforementioned approaches to SUE. In each of the sub-samples associated to Gumbel's classical approach, we can collect a few top OSs modelled through a *multivariate* EV *model*, and then consider the so-called *multidimensional* EV *model*. Under this approach, we have access to the multivariate sample, $(\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_k)$,   where $\underline{X}_j = (X_{1j}, \ldots, X_{i_j j})$, $1 \leq j \leq k$, are *multivariate* EV *vectors*. The *multi-dimensional* EV *model* is indeed the multivariate EV model for the i_j top observations, $j = 1, \ldots, m$, in sub-samples of size m', with $m \times m' = n$. The choices $m = k$ ($m' = r$) and $i_j = 1$ for $1 \leq j \leq k$ originate the BM model. The choices $m' = n$ ($m = 1$) and $i_1 = k$ originate the LO model.

The peaks over threshold (POT) approaches. The *Paretian model* for the excesses, $X_j - u > 0$, $1 \leq j \leq k$, over a high threshold u, suitably chosen, is considered under this approach, in a certain sense parallel to the *multivariate* EV *model*, but where we restrict our attention only to observations that exceed a certain high *threshold* u, fitting the appropriate statistical model to the *excesses* over u. On the basis of the approximation $P(X - u \leq x | X > u) \approx P_\gamma(x/\sigma)$, with $P_\gamma(x)$ given in (1.4), we come to the so-called *Paretian excesses model* or POT *model*. Statistical inference is then related to the GPD.

Bayesian approaches.   The use of Bayesian methodology, within EVA, has recently become quite common. We mention only some recent papers, written after the monographs by Coles (2001) and Reiss and Thomas (2001), the ones by Bermudez and Amaral-Turkman (2003), Bottolo *et al.* (2003), Stephenson and Tawn (2004), on the use of reversible jump MCMC techniques for inference for the EVD and the GPD and Diebolt *et al.* (2005), on a quasi-conjugate Bayesian inference approach for the GPD with $\gamma > 0$, through the representation of a heavy-tailed GPD as a mixture of an exponential and a gamma distribution.

Statistical choice of EV models under parametric frameworks.   The Gumbel type CDF, $\Lambda \equiv G_0$ or the exponential (E) type CDF, $E \equiv P_0$, with G_γ and P_γ given in (1.2) and (1.4), respectively, are favorites in SUE, essentially

because of the simplicity of associated inference. Additionally, $\gamma = 0$ can be regarded as a change-point, and any separation between EV models, with Λ or E playing a central and proeminent position, turns out to be an important statistical problem. From a parametric point of view, empirical tests of $H_0 : \gamma = 0$ versus a sensible one-sided or two-sided alternative, either for the EVD or the GPD, date back to Jenkinson (1955) and Gumbel (1965). Next, we can find in the literature, different heuristic tests, among which we refer only one of the most recent ones (Brilhante, 2004). We can also find locally asymptotically normal tests (see Marohn, 2000, and Falk *et al.*, 2008, among others). The fitting of the GPD to data has been worked out in Castillo and Hadi (1997) and Chaouche and Bacro (2004). The problem of goodness-of-fit tests for the GPD has been studied by Choulakian and Stephens (2001) and Luceño (2006), again among others. Tests from large sample theory, like the likelihood ratio test have been dealt with by Hosking (1984) and Gomes (1989). Further details on this topic can be found in Gomes *et al.* (2007a), an enlarged version of Gomes *et al.* (2008a).

1.3. Scope of the paper

In the late seventies, there has been an inflection from a *parametric approach* to SUE, based on the limiting models in EVT, towards a *semi-parametric approach*, where the estimation of the left and the right-tails is done under a quite general framework. Section 2 of this review paper is devoted to classical semi-parametric inference. Recently, essentially for heavy tails, i.e. for $\gamma > 0$, but also for a general $\gamma \in \mathbb{R}$, the accommodation of bias of the classical estimators of parameters of extreme events has been deeply considered in the literature. The topic of second-order reduced-bias (SORB) estimation still seems to open interesting perspectives in the field, and will be addressed in Section 3. Finally, in Section 4, we shall discuss some still challenging topics in SUE, providing some overall comments on the subject.

---

2.   CLASSICAL SEMI-PARAMETRIC INFERENCE

Under these semi-parametric approaches, we work with the k top OSs associated to the n available observations or with the excesses over a high random threshold, assuming only that, for a certain γ, the model F underlying the data is in $\mathcal{D}_{\mathcal{M}}(G_\gamma)$ or in specific sub-domains of $\mathcal{D}_{\mathcal{M}}(G_\gamma)$, with G_γ provided in (1.2), being γ the unique primary parameter of extreme events to be estimated, on the

basis of a few top observations, and according to adequate methodology. There is thus no fitting of a specific parametric model, dependent upon a location λ, a scale δ and a shape γ. We usually need to base the EVI-estimation on the k top OSs in the sample, with k intermediate, i.e. such that $k = k_n \to \infty$ and $k = o(n)$, i.e. $k/n \to 0$, as $n \to \infty$. Such estimators, together with semi-parametric estimators of location and scale (see, for instance, de Haan and Ferreira, 2006), can next be used to estimate extreme quantiles, return periods of high levels, upper tail probabilities and other parameters of extreme events. After a brief introduction to first and second-order conditions in Section 2.1, we shall briefly refer, in Section 2.2, several classical semi-parametric EVI-estimators. In Section 2.3, we mention a few results on the testing of the EV condition $F \in \mathcal{D}(G_\gamma)$, under a semi-parametric framework. Finally, in Section 2.4, we refer, also briefly, the semi-parametric estimation of other parameters of extreme events.

2.1. First, second (and higher) order conditions

As mentioned above, in Section 1, the full characterization of $\mathcal{D}_{\mathcal{M}}(G_\gamma)$ has been given in de Haan (1970), and can be also found in Falk *et al.* (2004) and de Haan and Ferreira (2006). Indeed, with U standing for a (reciprocal) quantile type function associated with F and defined by $U(t) := (1/(1-F))^{\leftarrow}(t)$ $= F^{\leftarrow}(1-1/t) = \inf\{x : F(x) \geq 1 - 1/t\}$, the *extended regular variation* property,

$$(2.1) \qquad F \in \mathcal{D}_{\mathcal{M}}(G_\gamma) \iff \lim_{t \to \infty} \frac{U(tx) - U(t)}{a(t)} = \begin{cases} \frac{x^\gamma - 1}{\gamma} & \text{if } \gamma \neq 0 \\ \ln x & \text{if } \gamma = 0, \end{cases}$$

for every $x > 0$ and some positive measurable function a, is a well-known necessary and sufficient condition for $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$ (de Haan, 1984). Heavy-tailed models, i.e. models $F \in \mathcal{D}_{\mathcal{M}}^+ := \mathcal{D}_{\mathcal{M}}(G_{\gamma>0})$, are quite important in the most diversified areas. We can then choose $a(t) = \gamma\, U(t)$ in (2.1), and $F \in \mathcal{D}_{\mathcal{M}}^+$ if and only if, for every $x > 0$, $\lim_{t\to\infty} U(tx)/U(t) = x^\gamma$, i.e. U is of regular variation with index γ, denoted $U \in RV_\gamma$. More generally, $F \in \mathcal{D}_{\mathcal{M}}^+ \iff \overline{F} := 1 - F \in RV_{-1/\gamma} \iff U \in RV_\gamma$. For full details on regular variation see Bingham *et al.* (1987).

Under a semi-parametric framework, apart from the first-order condition in (2.1), we often need to assume a second-order condition, specifying the rate of convergence in (2.1). It is then common to assume the existence of a function A^*, possibly not changing in sign and tending to zero as $t \to \infty$, such that $\forall x > 0$,

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

where $\rho^* \leq 0$ is a *second-order* parameter controlling the speed of convergence of maximum values, linearly normalized, towards the limit law in (1.2). Then $\lim_{t\to\infty} A^*(tx)/A^*(t) = x^{\rho^*}$, $\forall x > 0$, i.e. $|A^*| \in RV_{\rho^*}$ (de Haan and Stadtmüller, 1996). For heavy tails, the second-order condition is usually written as

$$(2.3) \qquad \lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho},$$

where $\rho \leq 0$ and $A(t) \to 0$ as $t \to \infty$. More precisely, $|A| \in RV_\rho$ according to Geluk and de Haan (1987). For the link between $(A^*(t), \rho^*)$ and $(A(t), \rho)$, see de Haan and Ferreira (2006) and Fraga Alves *et al.* (2007). Third-order conditions specify, in a parallel way, the rate of convergence either in (2.2) or in (2.3). For further details on the third-order condition for heavy tails, see Gomes *et al.* (2002a) and Fraga Alves *et al.* (2003a). For a general third-order framework, see Fraga Alves *et al.* (2003b, Appendix; 2006). Higher-order conditions can be similarly postulated, but restrict more and more the elected CDFs in $\mathcal{D}_\mathcal{M}(G_\gamma)$.

2.2. Classical semi-parametric EVI-estimation

The most basic EVI-estimators that have motivated several other refined estimators, i.e. the *Hill* (H), *Pickands* (P), *moment* (M) and *peaks over random threshold-maximum likelihood* (PORT-ML) estimators, are described in Section 2.2.1. Next, in Section 2.2.2, we briefly refer other classical EVI-estimators.

2.2.1. H, P, M and PORT-ML EVI-estimators

The H-estimator. For heavy tailed models, i.e. in $\mathcal{D}_\mathcal{M}^+$, a simple EVI-estimator has been proposed in Hill (1975). The H-estimator, denoted $\hat{\gamma}_{n,k}^{\mathrm{H}}$, is the average of the scaled log-spacings as well as of the log-excesses, given by

$$(2.4) \qquad U_i := i\left\{ \ln \frac{X_{n-i+1,n}}{X_{n-i,n}} \right\} \quad \text{and} \quad V_{ik} := \ln \frac{X_{n-i+1,n}}{X_{n-k,n}}, \quad 1 \leq i \leq k < n,$$

respectively. Its asymptotic properties have been thoroughly studied by several authors (see de Haan and Peng, 1998, and the review in Gomes *et al.*, 2008a).

The P-estimator. For a general EVI, $\gamma \in \mathbb{R}$, and considering as the basis of the estimation the k top OSs, we can write the P-estimator (Pickands, 1975) as

$$\hat{\gamma}_{n,k}^{\mathrm{P}} := \ln\left(\left(X_{n-[k/4]+1,n} - X_{n-[k/2]+1,n} \right) / \left(X_{n-[k/2]+1,n} - X_{n-k+1,n} \right) \right)/\ln 2,$$

where $[x]$ denotes the integer part of x. Asymptotic properties of this estimator are provided in Dekkers and de Haan (1989).

The M-estimator. Dekkers *et al.* (1989) proposed the M-estimator, based on

$$(2.5) \qquad M_{n,k}^{(j)} := \frac{1}{k} \sum_{i=1}^{k} \{\ln X_{n-i+1,n} - \ln X_{n-k,n}\}^j, \quad j > 0,$$

the j-moment of the log-excesses, being $M_{n,k}^{(1)} \equiv \hat{\gamma}_{n,k}^{H}$ the H-estimator. The M-estimator is given by $\hat{\gamma}_{n,k}^{M} := M_{n,k}^{(1)} + \frac{1}{2}\big(1 - \big(M_{n,k}^{(2)}/[M_{n,k}^{(1)}]^2 - 1\big)^{-1}\big)$.

The PORT-ML-estimator. Conditionally on $X_{n-k,n}$, with k intermediate, $D_{ik} := X_{n-i+1,n} - X_{n-k,n}$, $1 \le i \le k$, are approximately the k top OSs associated to a sample of size k from $GP_\gamma(\alpha x/\gamma)$, γ, $\alpha \in \mathbb{R}$, with $GP_\gamma(x)$ given in (1.4). The solution of the maximum-likelihood (ML) equations associated to the above mentioned set-up (Davison, 1984) gives rise to an explicit EVI-estimator, the PORT-ML EVI-estimator, named PORT after Araújo Santos *et al.* (2006), and given by $\hat{\gamma}_{n,k}^{PORT-ML} := \frac{1}{k} \sum_{i=1}^{k} \ln(1+\hat{\alpha}\, D_{ik})$, where $\hat{\alpha}$ is the implicit ML estimator of the unknown "scale" parameter α. A comprehensive study of the asymptotic properties of this ML estimator has been undertaken in Drees *et al.* (2004). As recently shown by Zhou (2009, 2010), such estimator is valid for $\gamma > -1$.

2.2.2. Other "classical" semi-parametric EVI-estimators

Kernel ($\mathcal{K}$) and QQ-estimators. A general class of estimators for a positive EVI are the $\mathcal{K}$-estimators proposed by Csörgő *et al.* (1985), given by $\hat{\gamma}_{n,k}^{\mathcal{K}} := \sum_{i=1}^{n} \mathcal{K}(i/k)\{\ln X_{n-i+1,n} - \ln X_{n-k,n}\}/\sum_{i=1}^{n} \mathcal{K}(i/k)$, where $\mathcal{K}(\cdot)$ is some non-negative, non-increasing kernel defined on $(0, \infty)$ and integrating to one. As an example, the H-estimator is a kernel estimator associated to the kernel $\mathcal{K}(t) = I_{]0,1]}(t)$, where $I_A(t)$ denotes the indicator function ($I_A(t) = 1$ if $t \in A$, and equal to 0 otherwise). Kernel estimators for a real EVI are considered in Groeneboom *et al.* (2003). The H-estimator can also be obtained from the Pareto QQ-plot, through the use of a naïve estimator of the slope in the ultimate right-end of the QQ-plot. More flexible regression methods can be applied to the highest k points of the Pareto QQ-plot. We refer Beirlant *et al.* (1996a,c), Schultze and Steinbach (1996), Kratz and Resnick (1996), Csörgő and Viharos (1998) and Oliveira *et al.* (2006). They are all $\mathcal{K}$-estimators.

Generalized P-estimators. The large asymptotic variance of the P-estimator has motivated different generalizations of the type $\hat{\gamma}_{n,k}^{P(\theta)} :=$

$-\ln\left(\left(X_{n-[\theta^2 k]+1,n}-X_{n-[\theta k]+1,n}\right)/\left(X_{n-[\theta k]+1,n}-X_{n-k,n}\right)\right)/\ln\theta,\ 0<\theta<1.$ (Fraga Alves, 1992, 1995; Themido Pereira, 1993; Yun, 2002). Drees (1995) establishes the asymptotic normality of linear combinations of P-estimators, obtaining optimal weights that can be adaptively estimated from the data. Related work appears in Falk (1994). In Segers (2005), the P-estimator is generalized in a way that includes all of its previously known variants.

The generalized Hill (GH) estimator. The slope of a generalized quantile plot led Beirlant *et al.* (1996b) to the GH-estimator, valid for all $\gamma\in\mathbb{R}$, with the functional form, $\hat{\gamma}_{n,k}^{\mathrm{GH}}=\hat{\gamma}_{n,k}^{\mathrm{H}}+\frac{1}{k}\sum_{i=1}^{k}\left\{\ln\hat{\gamma}_{n,i}^{\mathrm{H}}-\ln\hat{\gamma}_{n,k}^{\mathrm{H}}\right\}$. Further study of this estimator has been performed in Beirlant *et al.* (2005).

The Mixed Moment (MM) estimator. Fraga Alves *et al.* (2009) introduced the so-called MM-estimator, involving not only the log-excesses but also another type of moment-statistics given by $\hat{\varphi}_{n,k}:=\left(M_{n,k}^{(1)}-L_{n,k}^{(1)}\right)/\left(L_{n,k}^{(1)}\right)^2$, with $L_{n,k}^{(1)}:=\frac{1}{k}\sum_{i=1}^{k}\left(1-X_{n-k,n}/X_{n-i+1,n}\right)$, and where $M_{n,k}^{(1)}$ is defined in (2.5). The statistic $\hat{\varphi}_{n,k}$ can easily be transformed into what has been called the MM-estimator, valid for any $\gamma\in\mathbb{R}$, and given by $\hat{\gamma}_{n,k}^{\mathrm{MM}}:=\left(\hat{\varphi}_n(k)-1\right)/\left(1+2\min\left(\hat{\varphi}_n(k)-1,0\right)\right)$. The MM-estimator appears indeed as a promising alternative to the most popular EVI-estimators for $\gamma\in\mathbb{R}$.

Semi-parametric probability weighted moment (PWM) estimators. The PWM method is a generalization of the *method of moments*, introduced in Greenwood *et al.* (1979). For $\gamma<1$ and for CDFs like the EVD, $EV_\gamma((x-\lambda)/\delta)$, with $EV_\gamma(x)$ given in (1.2), the *Pareto* d.f., $P_\gamma(x;\delta)=1-(x/\delta)^{-1/\gamma}$, $x>\delta$, and the GPD, $GP_\gamma(x/\delta)$, with $GP_\gamma(x)$ defined in (1.4), the PWM have simple expressions, which allow a simple parametric estimation of the EVI (see Hosking *et al.*, 1985; Hosking and Wallis, 1987; Diebolt *et al.*, 2007, 2008c). On the basis of the GPD, de Haan and Ferreira (2006) considered, for $\gamma<1$, the semi-parametric GPPWM EVI-estimator, with GPPWM standing for *generalized Pareto PWM*, given by $\hat{\gamma}_{n,k}^{\mathrm{GPPWM}}:=1-2\hat{a}_1^\star(k)/\left(\hat{a}_0^\star(k)-2\hat{a}_1^\star(k)\right)$, $1\le k<n$, and $\hat{a}_s^\star(k):=\sum_{i=1}^{k}\left(\frac{i}{k}\right)^s\left(X_{n-i+1:n}-X_{n-k:n}\right)/k$, $s=0,\ 1$. On the basis of the Pareto model, Caeiro and Gomes (2011) introduced the PPWM EVI-estimators, with PPWM standing for *Pareto PWM*, given by $\hat{\gamma}_{n,k}^{\mathrm{PPWM}}:=1-\hat{a}_1(k)/\left(\hat{a}_0(k)-\hat{a}_1(k)\right)$, where $\hat{a}_s(k):=\frac{1}{k+1}\sum_{i=1}^{k+1}\left(\frac{i}{k+1}\right)^s X_{n-i+1:n}$, $s=0,1$ with $1\le k<n$.

Other estimators. Falk (1995a) proposed the location-invariant estimator, $\hat{\gamma}_{n,k}:=\frac{1}{k}\sum_{i=1}^{k-1}\ln\left(X_{n,n}-X_{n-i,n}\right)/\left(X_{n,n}-X_{n-k,n}\right)$, as a complement of the PORT-ML estimator for $\gamma<-1/2$. Such an estimator has been improved, on the basis of an iterative procedure, in Hüsler and Müller (2005). The non-

invariance for shifts of the H-estimator led Fraga Alves (2001) to the consideration for $k > k_0$, k_0 adequately chosen, of the location invariant Hill-type estimator $\hat{\gamma}_{n,k,k_0} := \frac{1}{k_0} \sum_{i=1}^{k} \ln \left((X_{n-i+1,n} - X_{n-k,n})/(X_{n-k_0+1,n} - X_{n-k,n}) \right)$. Beirlant *et al.* (1996b) consider a general class of estimators based on the mean, median and trimmed excess functions. Drees (1998) obtains asymptotic results for a general class of EVI-estimators, arbitrary smooth functionals of the empirical tail quantile function $Q_n(t) = X_{n-[k_n t],n}$, $t \in [0,1]$. Such a class includes H, P and $\mathcal{K}$-estimators, among others. For further references see, e.g., Section 6.4 of Embrechts *et al.* (1997), Beirlant *et al.* (1996a;1998), Csörgő and Viharos (1998), Chapter 3 of de Haan and Ferreira (2006), and Ling *et al.* (2011).

2.2.3. Consistency and asymptotic normal behaviour of the estimators

---

Weak consistency of any of the aforementioned EVI-estimators is achieved in the sub-domain of $\mathcal{D}_\mathcal{M}(EV_\gamma)$ where they are valid, whenever (2.1) holds and k is intermediate. Under the validity of the second-order condition in (2.2), it is possible to guarantee their asymptotic normality. More precisely, denoting T any of the aforementioned EVI-estimators, and with $B(t)$ a bias function converging to zero as $t \to \infty$ and strongly related with the $A^*(t)$ function in (2.3), it is possible to guarantee the existence of $\mathcal{C}_T \subset \mathbb{R}$ and $(b_T, \sigma_T) \in \mathbb{R} \times \mathbb{R}^+$, such that:

$$(2.6) \qquad \hat{\gamma}_{n,k}^T \stackrel{d}{=} \gamma + \sigma_T P_k^T / \sqrt{k} + b_T B(n/k) + o_p(B(n/k)),$$

with P_k^T an asymptotically standard normal random variable. Consequently, for values k such that $\sqrt{k}\, B(n/k) \to \lambda$, finite, as $n \to \infty$,

$$\sqrt{k} \left(\hat{\gamma}_{n,k}^T - \gamma \right) \xrightarrow[n\to\infty]{d} \text{Normal}(\lambda b_T, \sigma_T^2).$$

The values b_T and σ_T^2 are usually called the *asymptotic bias* and *asymptotic variance* of $\hat{\gamma}_{n,k}^T$ respectively. Details on the values of (b_T, σ_T) and the function B, in (2.6) are given in the aforementioned papers associated with the T-estimators.

2.3. Testing under a semi-parametric framework

Testing the hypothesis $H_0 : F \in \mathcal{D}_\mathcal{M}(G_0)$ against $H_1 : F \in \mathcal{D}_\mathcal{M}(G_\gamma)$, $\gamma \neq 0$, or the corresponding one-sided alternatives, under a semi-parametric framework, it is obviously a natural and sensible issue. In a broad sense, tests of this nature can already be found in papers prior to 2000 (see Gomes *et al.*, 2007a). Non-parametric tests appear in Jurečková and Picek (2001). But the testing of

extreme value conditions can be dated back to Dietrich *et al.* (2002), who propose a test statistic to test whether the hypothesis $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$ is supported by the data, together with a simpler version devised to test whether $F \in \mathcal{D}_{\mathcal{M}}(G_{\gamma \geq 0})$. Further results of this last nature can be found in Drees *et al.* (2006) for testing $F \in \mathcal{D}_{\mathcal{M}}(G_{\gamma > -1/2})$. Tables of associated critical points are provided in Hüsler and Li (2006). Beirlant *et al.* (2006) tackle the goodness-of-fit problem for the class of heavy-tailed or Pareto-type distributions. For overviews of the subject see Hüsler and Peng (2008) and Neves and Fraga Alves (2008). See also Koning and Peng (2008) and Goegebeur and Guillou (2010).

2.4. Estimation of other parameters of extreme events

High quantiles of probability $1 - p$, p small, or equivalently in financial frameworks the Value at Risk at a level p (VaR_p) are possibly the most important parameters of extreme events, functions of the EVI, as well as of location/scale parameters. In a semi-parametric context, the most usual estimators of a quantile $\chi_{1-p} := U(1/p)$, with p small, can be easily derived from (2.1), through the approximation $U(tx) \approx U(t) + a(t)(x^\gamma - 1)/\gamma$. The fact that $X_{n-k+1,n} \stackrel{p}{\sim} U(n/k)$ enables us to estimate χ_{1-p} on the basis of this approximation and adequate estimates of γ and $a(n/k)$. For the simpler case of heavy tails, the approximation is $U(tx) \approx U(t)x^\gamma$, and we get $\hat{\chi}_{1-p,k} := X_{n-k:n} \ (k/(np))^{\hat{\gamma}_k}$, where $\hat{\gamma}_k$ is any consistent semi-parametric EVI-estimator. This estimator is of the type of the one introduced by Weissman (1978). Details on semi-parametric estimation of extremely high quantiles for $\gamma \in \mathbb{R}$, can be found in Dekkers and de Haan (1989), de Haan and Rootzén (1993) and more recently in Ferreira *et al.* (2003). Fraga Alves *et al.* (2009) also provide, jointly with the MM-estimator, accompanying shift and scale estimators that make high quantile estimation almost straightforward. Other approaches to high quantile estimation can be found in Matthys and Beirlant (2003). None of the above mentioned quantile estimators reacts adequately to a shift of the data. Araújo Santos *et al.* (2006) provide a class of semi-parametric VaR_p estimators which enjoy such a feature, the empirical counterpart of the theoretical linearity of a quantile χ_p, $\chi_p(\delta X + \lambda) = \delta\chi_p(X) + \lambda$, for any real λ and positive δ. This class of estimators is based on the PORT methodology, providing exact properties for risk measures in finance: translation-equivariance and positive homogeneity. The estimation of the probability of exceedance of a fixed high level, has been dealt with by Dijk and de Haan (1992) and Ferreira (2002), among others. See also Guillou *et al.* (2010) and You *et al.* (2010). The estimation of the endpoint of an underlying CDF has been studied by Hall

(1982), Csörgő and Mason (1989), Aarssen and de Haan (1994), among others. Estimation of the mean of a heavy-tailed distribution has been undertaken by Peng (2001) and Johansson (2003). Estimation of the Weibull tail coefficient dates back to Girard (2004). See also Goegebeur *et al.* (2010a), among others. Further details on the topic can be found in de Haan and Ferreira (2006).

3.   SORB ESTIMATION

Most of the classical semi-parametric estimators of any parameter of extreme events have a strong bias for moderate up to large values of k, including the optimal k, in the sense of minimal *mean squared error* (MSE). Accommodation of bias of classical estimators of parameters of extreme events has been deeply considered in the recent literature. We mention the pioneering papers of Peng (1998), Beirlant *et al.* (1999), Feuerverger and Hall (1999) and Gomes *et al.* (2000), where appeared always the old trade-off between variance and bias. Such a trade-off was removed with an adequate estimation of the second-order parameters, as done in Caeiro *et al.* (2005) and Gomes *et al.* (2007b; 2008c), who introduced different types of minimum-variance reduced-bias (MVRB) EVI-estimators. Such estimators have an asymptotic variance equal to the one of the Hill EVI-estimator but an asymptotic bias of smaller order, overpassing the classical ones for all k. In Section 3.1 we deal with SORB semiparametric EVI-estimation and in Section 3.2, we briefly refer the recent literature on SORB semi-parametric estimation of other parameters of extreme events.

3.1. SORB semi-parametric EVI-estimation

Let us consider any "classical" semi-parametric EVI-estimator, $\hat{\gamma}_{n,k}$. Let us also assume that a distributional representation similar to the one in (2.6), with (b_T, σ_T) replaced by (b, σ), holds for $\hat{\gamma}_{n,k}$. For intermediate k, $\hat{\gamma}_{n,k}$ is consistent for the EVI-estimation, and it is asymptotically normal if we further assume that $\sqrt{k}B(n/k) \to \lambda$, finite. Approximations for the variance and the squared-bias of $\hat{\gamma}_{n,k}$ are then given by σ^2/k and $b^2B^2(n/k)$ respectively. Consequently, the pattern of these estimators exhibit the same type of peculiarities: a high variance for high thresholds $X_{n-k,n}$, i.e. for small k; a high bias for low thresholds, i.e. for large k; a small region of stability of the sample path (plot of the estimates versus k), making problematic the adaptive choice of the threshold, on the basis of any sample paths' stability criterion; a "very peaked" MSE, making difficult the choice

of the value $k_0 := \arg \min_k \mathrm{MSE}\,(\hat{\gamma}_{n,k})$. These peculiarities have led researchers to consider the possibility of dealing with the bias term in an appropriate manner, building new estimators $\hat{\gamma}_{n,k}^R$, here called SORB EVI-estimators. Particularly, for heavy tails, i.e. $\gamma > 0$, the reduction of bias is a very important problem for the estimation of γ or of the Pareto index, $\alpha = 1/\gamma$, in case the slowly varying part of the Pareto type model disappears at a very slow rate. We consider the following definition, already provided in Reiss and Thomas (2007), Chapter 6.

Definition 3.1.   Under the second-order condition in (2.2) and for intermediate k, the statistic $\hat{\gamma}_{n,k}^R$, a consistent EVI-estimator, based on the k top OSs in a sample from $F \in \mathcal{D}_\mathcal{M}(EV_\gamma)$, is said to be a SORB semi-parametric EVI-estimator, if there exist $\sigma_R > 0$ and an asymptotically standard normal random variable P_k^R, such that for a large class of models in $\mathcal{D}_\mathcal{M}(EV_\gamma)$, and with $B(.)$ the function in (2.6),

(3.1) $$\hat{\gamma}_{n,k}^R \;\overset{d}{=}\; \gamma + \sigma_R P_k^R/\sqrt{k} + o_p(B(n/k)).$$

Notice that for the SORB EVI-estimators, we no longer have a dominant component of bias of the order of $B(n/k)$, as in (2.6). Therefore,

$$\sqrt{k}\big(\hat{\gamma}_{n,k}^R - \gamma\big) \;\xrightarrow[n\to\infty]{d}\; \mathrm{Normal}\big(0, \sigma_R^2\big)$$

not only when $\sqrt{k}B(n/k) \to 0$ (as for classical estimators), but also when $\sqrt{k}B(n/k) \to \lambda$, finite and non-null. Such a bias reduction provides usually a stable sample path for a wider region of k-values, a "bath-shaped" MSE and a reduction of the MSE at the optimal level, in the sense of minimum MSE.

Such an approach has been carried out essentially for heavy tails in the most diversified manners. The key ideas are either to find ways of getting rid of the dominant component $bB(n/k)$ of bias, in (2.6), or to go further into the second-order behaviour of the basic statistics used for the estimation of γ, like the log-excesses or the scaled log-spacings, in (2.4). We first mention some pre-2000 results about bias-corrected estimators in EVT. Such estimators may be dated back to Gomes (1994b), Drees (1996) and Peng (1998), among others. Gomes uses the *generalized jackknife* (GJ) methodology in Gray and Schucany (1972), and Peng deals with linear combinations of adequate EVI-estimators, in a spirit close to the one associated to the GJ technique. Feuerverger and Hall (1999) discuss the question of the possible misspecification of the second-order parameter ρ at -1, a value that corresponds to many commonly used heavy-tailed models, like the Fréchet. Within the second-order framework, Beirlant *et*

al. (1999) investigate the accommodation of bias in the scaled log-spacings and derive approximate "ML" and "least squares" SORB EVI-estimators. In Section 3.1.1, we provide details about the GJ EVI-estimation. In Section 3.1.2 we briefly review an approximate ML approach, together with the introduction of simple SORB EVI-estimators based on the scaled log-spacings or the log-excesses, in (2.4). Second-order parameters are usually decisive for the bias reduction, and we deal with their estimation in Section 3.1.3. Finally in Section 3.1.4, we conclude with some remarks about further literature on SORB EVI-estimation, including the recent first steps on SORB EVI-estimation for a general $\gamma \in \mathbb{R}$.

3.1.1. A brief review of GJ estimators of a positive EVI

The pioneering SORB EVI-estimators are, in a certain sense, GJ estimators, i.e. affine combinations of well-known estimators of γ. For details on the GJ methodology, see Gray and Schucany (1972). Whenever we are dealing with semi-parametric EVI-estimators, or even estimators of other parameters of extreme events, we have usually information about their asymptotic bias. We can thus choose estimators with similar asymptotic properties, and build the associated GJ random variable or statistic. This methodology has been used in Gomes *et al.* (2000, 2002b), among others, and revisited by Gomes *et al.* (2011c). Indeed, if the second-order condition, in (2.3), holds, we easily find two statistics $\hat{\gamma}_{n,k}^{(j)}$, such that (2.6) holds for both statistics. The ratio between the dominant components of bias of $\hat{\gamma}_{n,k}^{(1)}$ and $\hat{\gamma}_{n,k}^{(2)}$ is $q = b_1/b_2 = q(\rho)$, and we get the GJ random variable,

$$(3.2) \qquad \hat{\gamma}_{n,k}^{\mathrm{GJ}(\rho)} := \big(\hat{\gamma}_{n,k}^{(1)} - q(\rho)\,\hat{\gamma}_{n,k}^{(2)}\big)/(1 - q(\rho)).$$

We can then say that, under the second-order condition, in (2.3), a distributional representation of the type of the one in (3.1) holds for $\hat{\gamma}_{n,k}^{\mathrm{GJ}(\rho)}$, with $\sigma_{GJ}^2 > \sigma_H^2 = \gamma^2$ and $(P_k^R, B(n/k))$ replaced by $(P_k^{\mathrm{GJ}}, A(n/k))$. The same result remains true for the GJ EVI-estimator, $\gamma_{n,k}^{\mathrm{GJ}(\hat{\rho})}$, provided that $\hat{\rho} - \rho = o_p(1)$ for all k on which we initially base the EVI-estimation. Then (Gomes and Martins, 2002), if $\sqrt{k}\,A(n/k) \to \lambda$, finite,

$$(3.3) \qquad \sqrt{k}\left(\hat{\gamma}_{n,k}^{\mathrm{GJ}(\hat{\rho})} - \gamma\right) \xrightarrow[n\to\infty]{d} \mathrm{Normal}\big(0, \sigma_{GJ}^2\big).$$

The result in (3.3), comes from the fact that, through the use of Taylor's expansion, we can write

$$(3.4) \qquad \hat{\gamma}_{n,k}^{\mathrm{GJ}(\hat{\rho})} \overset{d}{=} \hat{\gamma}_{n,k}^{\mathrm{GJ}(\rho)}(k) + \big(\hat{\rho} - \rho\big)\big(O_p\big(1/\sqrt{k}\big) + O_p\big(A(n/k)\big)\big)(1 + o_p(1)).$$

A closer look at (3.4) reveals that it does not seem convenient to compute $\hat{\rho}$ at the value k considered for the EVI-estimation. Indeed, if we do that, and since

we have $\hat\rho - \rho = \hat\rho_k - \rho = O_p\big(1/\big(\sqrt{k}\, A(n/k)\big)\big)$ (see Fraga Alves *et al.*, 2003a), $(\hat\rho - \rho)\, A(n/k)$ is a term of the order of $1/\sqrt{k}$, and we are going to have a change in the asymptotic variance of the EVI-estimator. Gomes *et al.* (2000) have indeed suggested the misspecification of ρ at $\rho = -1$, essentially due not only to the high bias and variance of the exhisting estimators of ρ at that time, but also to the idea of considering $\hat\rho = \hat\rho_k$. Nowadays, the use of any of the algorithms in Gomes and Pestana (2007a,b), among others, enables us to get the limiting result in (3.3), for k-values such that $\sqrt{k}\, A(n/k) \to \infty$, as $n \to \infty$.

3.1.2. Accommodation of bias in the scaled log-spacings and in the log-excesses: alternative SORB EVI-estimators

---

The ML EVI-estimation based on the scaled log-spacings. The accommodation of bias in the scaled log-spacings U_i, in (2.4), has also been another source of inspiration for the building of SORB EVI-estimators. Under the second-order condition in (2.3), but for $\rho < 0$, i.e. working in Hall's class of Pareto-type models (Hall, 1982), with a right-tail function $\overline{F}(x) = Cx^{-1/\gamma}\big(1 + Dx^{\rho/\gamma} + o\big(x^{\rho/\gamma}\big)\big)$, as $x \to \infty$, $C > 0$, D real, $\rho < 0$, we can choose in (2.3),

$$(3.5) \qquad\qquad A(t) = \alpha\, t^\rho =: \gamma\, \beta\, t^\rho, \quad \beta \in \mathbb{R}, \quad \rho < 0,$$

where β can be regarded as a slowly varying function. Beirlant *et al.* (1999) provide the approximation

$$(3.6) \qquad\qquad U_i \sim \big(\gamma + A(n/k)\,(i/k)^{-\rho}\big)\, E_i, \quad 1 \le i \le k,$$

where E_i, $i \ge 1$, denotes a sequence of IID standard exponential random variables. Feuerverger and Hall (1999) consider the approximation,

$$(3.7) \quad U_i \sim \gamma \exp\big(A(n/k)\,(i/k)^{-\rho}\,/\gamma\big)\, E_i = \gamma \exp\big(A(n/i)/\gamma\big)\, E_i, \quad 1 \le i \le k.$$

The approximation (3.6), or equivalently (3.7), has been made more precise in the asymptotic sense, in Beirlant *et al.* (2002). The use of the approximation in (3.7) and the joint maximization, in γ, β and ρ, of the approximate log-likelihood of the scaled log-spacings,

$$\log L(\gamma, \beta, \rho; U_i, 1 \le i \le k) = -k \log \gamma - \beta \sum_{i=1}^{k} (i/n)^{-\rho} - \frac{1}{\gamma} \sum_{i=1}^{k} e^{-\beta(i/n)^{-\rho}}\, U_i,$$

led Feuerverger and Hall to an explicit expression for $\hat\gamma$, given by

$$(3.8) \qquad\qquad \hat\gamma = \hat\gamma_{n,k}^{\mathrm{FH}(\hat\beta,\hat\rho)} := \frac{1}{k} \sum_{i=1}^{k} e^{-\hat\beta(i/n)^{-\hat\rho}}\, U_i,$$

as a function of $\hat{\beta}$ and $\hat{\rho}$, where $\hat{\beta} = \hat{\beta}_{n,k}^{FH(\hat{\rho})}$ and $\hat{\rho} = \hat{\rho}_{n,k}^{FH}$ are both computed at the same k used for the EVI-estimation, and are numerically obtained through

$$(3.9) \qquad (\hat{\beta}, \hat{\rho}) := \underset{(\beta,\rho)}{\arg\min} \left\{ \log\left(\frac{1}{k} \sum_{i=1}^{k} e^{-\beta(i/n)^{-\rho}} U_i \right) + \beta\left(\frac{1}{k} \sum_{i=1}^{k} (i/n)^{-\rho} \right) \right\}.$$

If k is intermediate and the second-order condition (2.3) hold, it is possible to state that if ρ is unknown as well as β, as usually happens, and they are both estimated through the above mentioned ML technique,

$$(3.10) \qquad \sqrt{k}\left(\hat{\gamma}_{n,k}^{\mathrm{FH}(\hat{\beta},\hat{\rho})} - \gamma \right) \xrightarrow[n\to\infty]{d} \mathrm{Normal}\left(0, \sigma_{FH}^2 = \gamma^2\left(\frac{1-\rho}{\rho} \right)^4 \right).$$

Again, even when $\sqrt{k}\, A(n/k) \to \lambda$, non-null, we have a null asymptotic bias for the reduced-bias EVI-estimator, but at the expenses of a larger asymptotic variance, ruled by $\sigma_{FH}^2 = \gamma^2\left((1-\rho)/\rho\right)^4$. Note that the asymptotic variance is smaller, and given by $\gamma^2\left((1-\rho)/\rho\right)^2$, if we assume ρ to be known.

A simplified maximum likelihood EVI-estimator based on the external estimation of ρ. The use of the first-order approximation, $e^x = 1+x$, as $x \to 0$, in the two ML equations that provided before $(\hat{\beta}, \hat{\rho})$, led Gomes and Martins (2002) to an explicit estimator for β, given by

$$(3.11) \qquad \hat{\beta}_{n,k}^{\mathrm{GM}(\hat{\rho})} := \left(\frac{k}{n} \right)^{\hat{\rho}} \frac{\left(\frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{k} \right)^{-\hat{\rho}} \right)\hat{C}_0 - \hat{C}_1}{\left(\frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{k} \right)^{-\hat{\rho}} \right)\hat{C}_1 - \hat{C}_2}, \quad \hat{C}_j = \frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{k} \right)^{-j\hat{\rho}} U_i.$$

and, on the basis of an adequate consistent estimator $\hat{\rho}$ of ρ, they suggest the following approximate ML estimator for the EVI, γ,

$$(3.12) \qquad \hat{\gamma}_{n,k}^{\mathrm{GM}(\hat{\rho})} := \frac{1}{k} \sum_{i=1}^{k} U_i - \hat{\beta}_{n,k}^{\mathrm{GM}(\hat{\rho})} \left(\frac{n}{k} \right)^{\hat{\rho}} \hat{C}_1.$$

The estimator in (3.12) is clearly a bias-corrected Hill estimator, i.e. the dominant component of the bias of the H-estimator, equal to $A(n/k)/(1-\rho) = \gamma\beta(n/k)^\rho/(1-\rho)$ is estimated through $\hat{\beta}_{n,k}^{\mathrm{GM}(\hat{\rho})} (n/k)^{\hat{\rho}} \hat{C}_1$, and directly removed from the H-estimator, which can also be written as $\gamma_{n,k}^{\mathrm{H}} = \sum_{i=1}^{k} U_i/k$. Under the same conditions as before, the asymptotic variance of $\hat{\gamma}_{n,k}^{\mathrm{GM}(\hat{\rho})}$ is $\sigma_{GM}^2 = \gamma^2(1-\rho)^2/\rho^2 < \sigma_{FH}^2$, but still greater than $\sigma_H^2 = \gamma^2$.

External estimation of second-order parameters and the weighted Hill (WH) EVI-estimator. In a trial to accommodate bias in the excesses over a high random threshold, Gomes *et al.* (2004b) were led, for heavy tails, to a weighted combination of the log-excesses V_{ik}, $1 \leq i \leq k < n$, also in (2.4), giving

rise to the WH EVI-estimator in Gomes *et al.* (2008c), given by

$$(3.13) \quad \hat{\gamma}_{n,k,\hat{\beta},\hat{\rho}}^{WH} := \frac{1}{k} \sum_{i=1}^{k} p_{ik}(\hat{\beta}, \hat{\rho}) \, V_{ik}, \quad p_{ik}(\hat{\beta}, \hat{\rho}) := e^{\hat{\beta} \, (n/k)^{\hat{\rho}} \left((i/k)^{-\hat{\rho}} - 1 \right) / (\hat{\rho} \, \ln(i/k))},$$

where $(\hat{\beta}, \hat{\rho})$ are suitable consistent estimators of second-order parameters (β, ρ). The key of success of the WH-estimator lies in the estimation of β and ρ at a level k_1, such that $k = o(k_1)$, with k the number of top OSs used for the EVI-estimation. The level k_1 needs to be such that $(\hat{\beta}, \, \hat{\rho})$ is consistent for the estimation of $(\beta, \, \rho)$ and $\hat{\rho} - \rho = o_p(1/\ln n)$. For more details on the choice of k_1, see Gomes *et al.* (2008c), and more recently Caeiro *et al.* (2009). Comparatively to the SORB EVI-estimators available in the literature and published prior to 2005, this EVI-estimator is a MVRB EVI-estimator, in the sense that, comparatively to the Hill estimator, it keeps the same asymptotic variance $\sigma_{WH}^2 = \sigma_H^2 = \gamma^2$ and a smaller order asymptotic bias, outperforming the H-estimator for all k. Related work appears in Caeiro *et al.* (2005) and Gomes *et al.* (2007b). Gomes *et al.* (2007b) suggest the computation of the β-estimator $\hat{\beta}_{n,k}^{\mathrm{GM}(\hat{\rho})}$, used at (3.12), at the level k_1 used for the estimation of ρ. With the notation $\hat{\beta} := \hat{\beta}_{n,k_1}^{\mathrm{GM}(\hat{\rho})}$, they suggest thus the replacement of the estimator in (3.12) by

$$(3.14) \qquad\qquad \hat{\gamma}_{n,k}^{\overline{\mathrm{M}}(\hat{\beta},\hat{\rho})} := \gamma_{n,k}^{\mathrm{H}} - \hat{\beta} \left(\frac{n}{k} \right)^{\hat{\rho}} \hat{C}_1,$$

where $\gamma_{n,k}^{\mathrm{H}}$ denotes the H-estimator, and $(\hat{\beta}, \hat{\rho})$ are adequate consistent estimators of the second-order parameters (β, ρ). With the same objectives, but with a slightly simpler analytic expression, we also refer the estimator

$$(3.15) \qquad\qquad \hat{\gamma}_{n,k}^{\overline{\mathrm{H}}(\hat{\beta},\hat{\rho})} := \gamma_{n,k}^{\mathrm{H}} \left(1 - \hat{\beta} \, (n/k)^{\hat{\rho}} / (1 - \hat{\rho}) \right),$$

studied in Caeiro *et al.* (2005). Notice that the dominant component of the bias of the H-estimator is estimated in (3.15) through $\gamma_{n,k}^{\mathrm{H}} \hat{\beta}(n/k)^{\hat{\rho}} / (1 - \hat{\rho})$, and directly removed from Hill's classical EVI-estimator. The adequate estimation of β and ρ at a level k_1 of a higher order than the level k used for the EVI-estimation, enables, for a large diversity of heavy-tailed models, the reduction of bias without increasing the asymptotic variance, which is kept at the value γ^2, the asymptotic variance of Hill's estimator. For overviews of this subject see Reiss and Thomas (2007), Chapter 6, as well as Gomes *et al.* (2008a).

3.1.3. Second-order parameters estimation for heavy tails

The first estimator of the parameter ρ, in (2.3), with $A(\cdot)$ given in (3.5), but where β can possibly be any slowly varying function, appears in Hall and

Welsh (1985). Peng (1998) claims that no good estimator for the second-order parameter ρ was then available in the literature, and considers a new ρ-estimator, alternative to the ones in Hall and Welsh (1985), Beirlant *et al.* (1996c) and Drees and Kaufmann (1998). Another estimator of ρ appears in Gomes *et al.* (2002a), and more recently, we refer the classes of ρ-estimators in Goegebeur *et al.* (2008; 2010b) and Ciuperca and Mercadier (2010). We elect here particular members of the class of estimators of the second-order parameter ρ proposed by Fraga Alves *et al.* (2003a). Under adequate general conditions, they are asymptotically normal estimators of ρ, if $\rho < 0$, which show highly stable sample paths as functions of k, the number of top OSs used, for a wide range of large k-values. Such a class of estimators, parameterised in a tuning real parameter $\tau \in \mathbb{R}$, is defined as,

$$(3.16) \quad \hat{\rho}_{n,k}^{(\tau)} := - \left| 3(T_{n,k}^{(\tau)} - 1)/(T_{n,k}^{(\tau)} - 3) \right|, \quad T_{n,k}^{(\tau)} := \frac{\left(M_{n,k}^{(1)}\right)^{\tau} - \left(M_{n,k}^{(2)}/2\right)^{\tau/2}}{\left(M_{n,k}^{(2)}/2\right)^{\tau/2} - \left(M_{n,k}^{(3)}/6\right)^{\tau/3}},$$

with $M_{n,k}^{(j)}$ given in (2.5) and with the notation $a^{b\tau} = b \ln a$ whenever $\tau = 0$.

Gomes and Martins (2002) provide an explicit estimator for β, based on the scale log-spacings U_i, in (2.4), and already given in (3.11). An additional estimator of β, is provided in Caeiro and Gomes (2006). See also Gomes *et al.* (2010), for a β-estimator based on the log-excesses.

Algorithms for the estimation of second-order parameters, in the above mentioned lines, can be found in Gomes and Pestana (2007a,b). The use of such algorithms, where the ρ-estimator is computed at $k_1 = [n^{1-\epsilon}]$, with ϵ small, say $\epsilon = 0.001$, enables us to guarantee that, for a large class of heavy-tailed models, as $n \to \infty$, $(\hat{\rho}_{n,k_1}^{(\tau)} - \rho) \ln n = o_p(1)$, a crucial property of the ρ-estimator, if we do not want to increase the asymptotic variance of the random variable, function of (β, ρ), underlying the SORB EVI-estimator. Such a crucial property can potentially be achieved if we compute $\hat{\rho}$ at its optimal level (see Caeiro *et al.*, 2009), but the adaptive choice of such a level is still an open research topic.

3.1.4. Additional Literature on SORB EVI-estimation

Other approches to bias reduction, in the estimation of a positive EVI can be found in Gomes and Martins (2001, 2004), Caeiro and Gomes (2002), Gomes *et al.* (2004a; 2005a; 2005b; 2007c; 2011a), Canto e Castro and de Haan (2006) and Willems *et al.* (2007), among others. Recently, Cai *et al.* (2011) introduced the first SORB estimators for $\gamma \in \mathbb{R}$, based on the PWM methodology.

3.2.   SORB semi-parametric estimation of other parameters of extreme events

Reduced bias quantile estimators have been studied in Matthys *et al.* (2004) and Gomes and Figueiredo (2006), who consider the classical SORB EVI-estimators. Gomes and Pestana (2007b) and Beirlant *et al.* (2008) incorporate the MVRB EVI-estimators in Caeiro *et al.* (2005) and Gomes *et al.* (2007b) in high quantile semi-parametric estimation. See also Diebolt *et al.* (2008b), Beirlant *et al.* (2009), Caeiro and Gomes (2009), Li *et al.* (2010). For a SORB estimation of the Weibull-tail coefficient, we mention Diebolt *et al.* (2008a). Finally, for SORB endpoint estimation, we mention Li and Peng (2009).

4.   OVERALL COMMENTS AND FURTHER RESEARCH

We shall next discuss a few areas where a lot has been already done but further research is still welcome. In our opinion, SUE is still a lively topic of research. Important developments have appeared recently in the area of *spatial extremes*, where *parametric models* seem again to be quite relevant. In this case, and now that we have access to highly sophisticated computational techniques, a great variety of *parametric models* can further be considered. And in a semi-parametric framework, topics like *threshold selection, trends and change points* in the *tail behaviour*, and *clustering*, among others, are still challenging.

4.1.   Rates of convergence and penultimate approximations.

An important problem in EVT concerns the rate of convergence of $F^n(a_n x + b_n)$ towards $G_\gamma(x)$, in (1.2), or, equivalently, the search for estimates of the difference $d_n(F, G_\gamma, x) := F^n(a_n x + b_n) - G_\gamma(x)$. Indeed, as detailed in Section 1, parametric inference on the right-tail of F, usually unknown, is done on the basis of the identification of $F^n(a_n x + b_n)$ and of $G_\gamma(x)$. And the rate of convergence can validate or not the most usual models in SUE. As noted by Fisher and Tippett (1928), despite of the fact that the normal CDF, $\Phi \in \mathcal{D}_\mathcal{M}(G_0)$, the convergence of $\Phi^n(a_n x + b_n)$ towards $G_0(x)$ is extremely slow. They then show, through the use of skewness and kurtosis coefficients as indicators of closeness, that $\Phi^n(x)$ is "closer" to a suitable penultimate $G_{-1/\gamma_n}((x - \lambda_n)/\delta_n)$, for $\gamma_n > 0$, $\lambda_n \in \mathbb{R}$, $\delta_n > 0$, than to the ultimate $G_0((x - b_n)/a_n)$. Such an approxima-

tion is the so-called *penultimate approximation* and several penultimate models have been advanced by several authors. Dated overviews of the modern theory of rates of convergence in EVT, introduced in Anderson (1971), can be seen in Galambos (1984) and Gomes (1994a). More recently, Gomes and de Haan (1999) derived, for all $\gamma \in \mathbb{R}$, exact penultimate approximation rates with respect to the variational distance, under adequate differentiability assumptions. Kaufmann (2000) proved, under weaker conditions, a result related to the one in Gomes and de Haan (1999). This penultimate or pre-asymptotic behaviour has further been studied by Raoult and Worms (2003) and Diebolt and Guillou (2005), among others. Other type of penultimate approximations have been considered in Smith (1987b). Among them, we mention a penultimate parametric model of the type,

$$(4.1) \qquad PG_\gamma(x;r) = \exp\left(-(1+\gamma x)^{-1/\gamma}\left(1 + r(1+\gamma x)^{-1/\gamma}\right)\right).$$

This type of models surely deserves a deeper consideration under statistical backgrounds. *Penultimate models* seem to be possible and interesting alternatives to the classical models but have never been deeply used in the literature. Concomitantly, the convergence of the estimators can be really slow when $\rho = 0$ or $\rho^* = 0$, like happens with normal and loggamma distributions, important models in the most diversified areas, and alternative estimation procedures are still needed.

4.2. Max-semistable laws

We also refer the class of *max-semistable* (MSS) laws, introduced by Grienvich (1992a, 1992b), Pancheva (1992), and further studied in Canto e Castro *et al.* (2000) and in Temido and Canto e Castro (2003). Such a class is more general than the class of MS laws, given in (1.2). Indeed, the possible MSS laws are

$$G_{\gamma,\nu}(x) = \begin{cases} \exp\left(-\nu(\ln(1+\gamma x))(1+\gamma x)^{-1/\gamma}\right), \ 1+\gamma x > 0 \ \text{if} \ \gamma \neq 0 \\ \exp\left(-\nu(x)\exp(-x)\right), \ x \in \mathbb{R} \qquad\qquad\quad \text{if} \ \gamma = 0, \end{cases}$$

where $\nu(\cdot)$ is a positive, limited and periodic function. A unit ν-function enables us to get the MS laws in (1.2). Discrete models like the geometric and negative binomial, and some multimodal continuous models, are in $\mathcal{D}_{\mathcal{MSS}}$ but not in $\mathcal{D}_{\mathcal{M}}$. A recent survey of the topic can be found in Pancheva (2010). Generalized P-statistics have been used in Canto e Castro and Dias (2011), to develop methods of estimation in the MSS context. See also Canto e Castro *et al.* (2011). Such a diversity of models, if duly exploited from a statistical point of view, can surely provide fruitful topics of research, both in parametric and semi-parametric setups.

4.3.   Invariance versus non-invariance

In *statistics of extremes* most of the methods of estimation are dependent on the log-excesses, and consequently, are *non-invariant* with respect to *shifts of the data*. But the invariance not only to changes in scale but also to changes in location of any EVI estimator is statistically appealing. Wouldn't be sensible to use the PORT methodology in Araújo Santos *et al.* (2006), and consider PORT EVI-estimators based on the transformed sample

$$(4.2) \qquad X_i^* := X_i - X_{[np]+1,n}, \ 0 < p < 1, \ 1 \le i \le n?$$

A similar procedure has been used in Fraga Alves *et al.* (2009), who also propose a class of EVI-estimators alternative to the MM-estimator, invariant for changes in location, and dependent on a similar *tuning parameter p*, $0 < p < 1$. Such estimators have the same functional expression of the original estimator T, say, but the original observation X_i is replaced everywhere by X_i^*, in (4.2), $1 \le i \le n$. Note that a similar procedure has already been used for the H and M EVI-estimators, as well as for quantile estimation in Araújo Santos *et al.* (2006). For PORT quantile estimation, see also Henriques-Rodrigues and Gomes (2009). The shift invariant versions, dependent on the tuning parameter p have properties similar to the ones of the original estimator T, provided we keep to adequate k-values and choose an adequate *tuning parameter p*. Recent research on this topic can be seen in Gomes *et al.* (2011b), but further research is still welcome.

4.4.   Adaptive selection of sample fraction or threshold

A *threshold* is often set "almost arbitrarily" (for instance at the 90% or the 95% sample quantile). However, the choice of the threshold, or equivalently of the number k of top OSs to be used is crucial for a reliable estimation of any parameter of extreme events. The topic has already been extensively studied for classical EVI-estimators, for which (2.6) holds. In Hall and Welsh (1985), Hall (1990), Beirlant *et al.* (1996c), Drees and Kaufmann (1998) and Danielsson *et al.* (2001), methods for the adaptive choice of k are proposed for the H-estimator, some of them involving the bootstrap technique. Gomes and Oliveira (2001) also uses the bootstrap methodology to provide an adaptive choice of the threshold, alternative to the one in Danielsson *et al.* (2001), and easy to generalise to other semi-parametric estimators of parameters of extreme events. For a general γ and for the M-estimator and a generalized P-estimator, see Draisma *et al.* (1999).

These authors also use the bootstrap methodology. Beirlant *et al.* (2002) consider the exponential regression model (ERM) introduced in Beirlant *et al.* (1999), and discuss applications of the ERM to the selection of the optimal sample fraction in EV estimation. They also derive a connection between the new choice strategy in the paper and the diagnostic proposed in Guillou and Hall (2001). Csörgő and Viharos (1998) provide a data-driven choice of k for the kernel class of estimators. Apart from the papers by Drees and Kaufmann (1998) and Guillou and Hall (2001), where choice of the optimal sample fraction is based on bias stability, the other papers make the optimal choice minimizing the estimated MSE. Possible heuristic choices are provided in Gomes and Pestana (2007b), Gomes *et al.* (2008e) and Beirlant *et al.* (2011). The adaptive SORB estimation is still giving its first steps. We can however mention the recent papers by Gomes *et al.* (2011a,d). Is it sensible to use bootstrap computational intensive procedures for threshold selection or there will be simpler techniques possibly related with bias pattern? Is it possible to apply a similar methodology for the estimation of other parameters of extreme events?

4.5. Other possible topics of research in SUE

Testing whether $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$, for a certain γ, is a crucial topic, already dealt with in several articles referred both in Section 1.2 as well as in Section 2.3. And what about testing second-order and third-order conditions? Change-points detection is also a challenging topic of research. And SUE for weakly dependent data, with all problems related with clustering of extreme values still deserves further research. SUE for randomly censored data is another challenging topic. See the recent papers Beirlant *et al.* (2007; 2010), Einmahl *et al.* (2008a) and Gomes and Neves (2011). Statistics of extremes in athletics and estimation of the endpoint is another of the relevant topics in SUE. We mention the recent papers by Einmahl and Magnus (2008), Li and Peng (2009), Einmahl and Smets (2011), Henriques-Rodrigues *et al.* (2011) and Li *et al.* (2011). Recent models, like the *extreme value Birnbaum-Saunders* model in Ferreira *et al.* (2011), can also become relevant in the area of SUE. Moreover, the estimation of second and higher order parameters still deserves further attention, particularly due to the importance of such estimation in SORB estimators of parameters of extreme events.

ACKNOWLEDGMENTS

REFERENCES

[1]   AARSSEN, K.; and HAAN, L. de (1994). On the maximal life span of humans, *Mathematical Population Studies* **4**(4), 259–281.

[2]   ANDERSON, C.W. (1971). *Contributions to the Asymptotic Theory of Extreme Values.* Ph.D. Thesis, University of London.

[3]   ARAÚJO SANTOS, P.; FRAGA ALVES, M.I.; and GOMES, M.I. (2006). Peaks over random threshold methodology for tail index and quantile estimation, *Revstat* **4**(3), 227–247.

[4]   BALKEMA, A.A.; and HAAN, L. DE (1974). Residual life time at great age, *Annals of Probability* **2**, 792–804.

[5]   BEIRLANT, J.; TEUGELS, J.; and VYNCKIER, P. (1996a). *Practical Analysis of Extreme Values.* Leuven University Press, Leuven, Belgium.

[6]   BEIRLANT, J.; VYNCKIER, P.; and TEUGELS, J. (1996b). Excess functions and estimation of the extreme-value index, *Bernoulli* **2**, 293–318.

[7]   BEIRLANT, J.; VYNCKIER, P.; and TEUGELS, J. (1996c). Tail index estimation, Pareto quantile plots and regression, *J. Amer. Statist. Assoc.* **91**, 1659–1667.

[8]   BEIRLANT, J.; TEUGELS, J.; and VYNCKIER, P. (1998). Some thoughts on extreme values. In Accardi, L., and Heyde, C. C. (eds.). *Probability Towards 2000*, Lecture Notes in Statistics **128**, 58-73, Springer, New York.

[9]   BEIRLANT, J., DIERCKX, G.; GOEGEBEUR, Y.; and MATTHYS, G. (1999). Tail index estimation and an exponential regression model, *Extremes* **2**, 177–200.

[10]   BEIRLANT, J., DIERCKX, G.; GUILLOU, A.; and STARICA, C. (2002). On exponential representations of log-spacings of extreme order statistics, *Extremes* **5**(2), 157–180.

[11]   BEIRLANT, J.; GOEGEBEUR, Y.; SEGERS, J.; and TEUGELS, J. (2004). *Statistics of Extremes: Theory and Applications.* Wiley, England.

[12]   BEIRLANT, J.; DIERCKX, G.; and GUILLOU, A. (2005). Estimation of the extreme-value index and generalized quantile plots, *Bernoulli* **11**(6), 949–970.

[13]   BEIRLANT, J.; DE WET, T.; and GOEGEBEUR, Y. (2006). A goodness-of-fit for Pareto-type behaviour, *J. Computational and Applied Mathematics* **186**, 99–116.

[14] BEIRLANT, J.; GUILLOU, A.; DIERCKX, G.; and FILS-VILETARD, A. (2007). Estimation of the extreme value index and extreme quantiles under random censoring. *Extremes* **10**(3), 151-174.

[15] BEIRLANT, J.; FIGUEIREDO, F.; GOMES, M.I.; and VANDEWALLE, B. (2008). Improved reduced bias tail index and quantile estimators, *J. Statistical Planning and Inference* **138**(6), 1851–1870.

[16] BEIRLANT, J.; JOOSSENS, E.; and SEGERS, J. (2009) Second-order refined peaks-over-threshold modelling for heavy-tailed distributions, *J. Statist. Plann. Infer.* **139**, 2800-2815.

[17] BEIRLANT, J.; GUILLOU, A.; and TOULEMONDE, G. (2010). Peaks-Over-Threshold modeling under random censoring, *Comm. Statist. Theory and Methods* **39**, 1158–1179.

[18] BEIRLANT, J.; BONIPHACE, E.; and DIERCKX, G. (2011). Generalized sum plots, *Revstat* **9**(2), 181–198.

[19] BERMUDEZ, P. DE ZEA; and AMARAL-TURKMAN, M.A. (2003). Bayesian approach to parameter estimation of the generalized pareto distribution, *Test* **12**, 259–277.

[20] BINGHAM, N., GOLDIE, C.M.; and TEUGELS, J.L. (1987). *Regular Variation.* Cambridge Univ. Press, Cambridge.

[21] BOTTOLO, L., CONSONNI, G.; DELLAPORTAS, P.; and LIJOI, A. (2003). Bayesian analysis of extreme values by mixture modeling, *Extremes* **6**(1), 25–47.

[22] BRILHANTE, M.F. (2004). Exponential versus generalized Pareto – a resistant and robust test, *Revstat* **2**(1), 1–13.

[23] CAEIRO, F.; and GOMES, M.I. (2002). A class of asymptotically unbiased semi-parametric estimators of the tail index. *Test* **11**(2), 345–364.

[24] CAEIRO, C.; and GOMES, M.I. (2006). A new class of estimators of a scale second order parameter, *Extremes* **9**, 193–211.

[25] CAEIRO, F.; and GOMES, M.I. (2009). Semi-parametric second-order reduced-bias high quantile estimation, *Test* **18**(2), 392–413.

[26] CAEIRO, F.; and GOMES, M.I. (2011). Semi-parametric tail inference through probability-weighted moments, *J. Statist. Plann. Inference* **141**(2), 937–950.

[27] CAEIRO, C.; GOMES, M.I.; and PESTANA, D. (2005). Direct reduction of bias of the classical Hill estimator, *Revstat* **3**(2), 113–136.

[28] CAEIRO, F.; GOMES, M.I.; up HENRIQUES-RODRIGUES, L. (2009). Reduced-bias tail index estimators under a third order framework, *Comm. Statist. – Theory and Methods* **38**(7), 1019–1040.

[29] CAI, J.; DE HAAN, L.; and ZHOU, C. (2011). *Bias correction in extreme value statistics with index around zero.* Available at: http://people.few.eur.nl/ldehaan/BiasCorrection11.pdf

[30] CANTO E CASTRO, L.; and DIAS, S. (2011). Generalized Pickands estimators for the tail index parameter and max-semistability, *Extremes* **14**(4), 429–449.

[31]   CANTO E CASTRO, L.; and HAAN. L. de (2006). A class of distribution functions with less bias in extreme value estimation, *Statist., and Probab. Letters* **76**, 1617–1624.

[32]   CANTO E CASTRO, L.; HAAN, L. DE; and TEMIDO, M.G. (2000). Rarely observed maxima, *Th. Prob. Appl.* **45**, 779–782.

[33]   CANTO E CASTRO, L. DIAS, SANDRA AND TEMIDO, M.G. (2011). Looking for max-semistability: a new test for the extreme value condition, *Journal of Statistical Planning and Inference* **141**, 3005-3020.

[34]   CASTILLO, E.; and HADI, A. (1997). Fitting the generalized Pareto distribution to data, *J. Amer. Statist. Assoc.* **92**, 1609–1620.

[35]   CASTILLO, E.; HADI, A.; BALAKRISHNAN, N.; and SARABIA, J.M. (2005). *Extreme Value and Related Models with Applications in Engineering and Science.* Wiley, Hoboke, New Jersey.

[36]   CHAOUCHE, A.; and BACRO, J.-N. (2004). Fitting the generalized Pareto distribution to data, *Comput. Statist. and Data Analysis* **45**, 787–803.

[37]   CHOULAKIAN, V.; and STEPHENS, M.A. (2001). Goodness-of-fit tests for the generalized Pareto distribution, *Technometrics* **43**(4), 478–484.

[38]   CIUPERCA, and MERCADIER, C. (2010). Semi-parametric estimation for heavy tailed distributions, *Extremes* **13**(1), 55–87.

[39]   COLES, S. (2001). *An Introduction to Statistical Modelling of Extreme Values.* Springer Series in Statistics.

[40]   CSÖRGŐ, S.; and MASON, D.M. (1989). Simple estimators of the endpoint of a distribution. In Hüsler, J., and Reiss, R.-D. (1989), *Extreme Value Theory, Proceedings Oberwolfach* 1987, 132–147, Springer-Verlag, Berlin, Heidelberg.

[41]   CSÖRGŐ, S.; and VIHAROS, L. (1998). Estimating the tail index. In Szyszkowicz, B. ed., *Asymptotic Methods in Probability and Statistics*, 833–881, North-Holland.

[42]   CSÖRGŐ, S., DEHEUVELS, P.; and MASON, D.M. (1985). Kernel estimates of the tail index of a distribution, *Ann. Statist.* **13**, 1050–1077.

[43]   DANIELSSON, J.; HAAN, L. DE; PENG, L.; and DE VRIES, C.G. (2001). Using a bootstrap method to choose the sample fraction in the tail index estimation, *J. Multivariate Analysis* **76**, 226–248.

[44]   DAVISON, A.C. (1984). Modelling excesses over high thresholds. In *Statistical Extremes and Applications*, ed. J. Tiago de Oliveira, Dordrecht: D. Reidel, 461–482.

[45]   DEKKERS, A.L.M.; and HAAN, L. DE (1989). On the estimation of the extreme-value index and large quantile estimation, *Ann. Statist.* **17**, 1795–1832.

[46]   DEKKERS, A.; EINMAHL, J.; and DE HAAN, L. (1989). A moment estimator for the index of an extreme-value distribution, *Annals of Statistics* **17**, 1833–1855.

[47]   DIEBOLT, J.; and GUILLOU, A. (2005). Asymptotic behavior of regular estimators, *Revstat* **3**(1), 19–44.

[48] DIEBOLT, J., EL-AROUI, M.A.; GARRIDO, M.; and GIRARD, S. (2005). Quasi-conjugate Bayes estimates for GPD parameters and application to heavy tails modelling, *Extremes* **8**, 57–78.

[49] DIEBOLT, J.; GUILLOU, A.; and RACHED, I. (2007). Approximation of the distribution of excesses through a generalized probability-weighted moments method, *J. Statist Plann. Inference 137*, 841–857.

[50] DIEBOLT, J.; GARDES, L.: GIRARD, S.; and GUILLOU, A. (2008a). Bias-reduced estimators of the Weibull tail-coefficient, *Test* **17**, 311–331.

[51] DIEBOLT, J.; GARDES, L.: GIRARD, S.; and GUILLOU, A. (2008b). Bias-reduced extreme quantiles estimators of Weibull tail-distributions, *J. Statist. Plann. Infer.* **138**, 1389-1401.

[52] DIEBOLT, J.; GUILLOU, A.; NAVEAU, P.; and RIBEREAU, P. (2008c). Improving probability-weighted moment methods for the generalized extreme value distribution. *Revstat* **6**(1), 33–50.

[53] DIETRICH, D.; HAAN, L. DE; and HÜSLER, J. (2002). Testing extreme value conditions, *Extremes* **5**, 71–85.

[54] DIJK, V.; and HAAN, L. de (1992). On the estimation of the exceedance probability of a high level. Order statistics and nonparametrics: theory and applications. In Sen, P. K., and Salama, I. A. (eds.), 79–92, Elsevier, Amsterdam.

[55] DRAISMA, G.; HAAN, L. DE; PENG, L.; and PEREIRA, T.T. (1999). A bootstrap-based method to achieve optimality in estimating the extreme-value index, *Extremes* **2**, 367–404.

[56] DREES, H. (1995). Refined Pickands estimator of the extreme value index, *Ann. Statist.* **23**, 2059-2080.

[57] DREES, H. (1996). Refined Pickands estimators with bias correction, *Comm. Statist. Theory and Meth.* **25**, 837-851.

[58] DREES, H. (1998). A general class of estimators of the extreme value index, *J. Statist. Plann. Inf.* **66**, 95-112.

[59] DREES, H.; and KAUFMANN, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation, *Stoch. Proc. Application* **75**, 149-172.

[60] DREES, H.; FERREIRA, A.; and DE HAAN, L. (2004). On maximum likelihood estimation of the extreme value index, *Annals of Applied Probability* **14**, 1179–1201.

[61] DREES, H.; HAAN, L. DE; and LI, D. (2006). Approximations to the tail empirical distribution function with application to testing extreme value conditions, *J. Stat. Planning and Inference* **136**, 3498-3538.

[62] DWASS, M. (1964). Extremal processes, *Ann. Math. Statist.* **35**, 1718-1725.

[63] EINMAHL, J.; and MAGNUS, J.R. (2008). Records in athletics through extreme-value theory, *J. American Statistical Association* **103**, 1382–1391.

[64] EINMAHL, J.; and SMEETS, S.G.W.R. (2011). Ultimate 100-m world records through extreme-value theory, *Statistica Neerlandica* **65**(1), 32–42.

[65]   EINMAHL, J.H.J.; FILS-VILLETARD, A.; and GUILLOU, A. (2008a). Statistics of extremes under random censoring, *Bernoulli* **14**(1), 207-227.

[66]   EMBRECHTS, P.; KLÜPPELBERG, C.; and MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, Heidelberg.

[67]   FALK, M. (1994). Efficiency of a convex combination of Pickands estimator of the extreme value index, *J. Nonparam. Statist.* **4**, 133–147.

[68]   FALK, M. (1995a). Some best parameter estimates for distributions with finite endpoint, *Statistics* **27**(1-2), 115–125.

[69]   FALK, M.; HÜSLER, J.; and REISS, R.-D. (1994; 2004). *Laws of Small Numbers: Extremes and Rare Events*, 1st edition; 2nd edition, Birkhäuser, Basel.

[70]   FALK, F.; GUILLOU, A.; and TOULEMONDE, G. (2008). A LAN based Neyman smooth test for the generalized Pareto distribution, *J. Statist. Plann. Inference* **138**, 2867–2886.

[71]   FERREIRA, A. (2002). Optimal asymptotic estimation of small exceedance probabilities, *J. Statist. Plannning and Inference* **104**, 83–102.

[72]   FERREIRA, A.; HAAN, L. DE; and PENG, L. (2003). On optimizing the estimation of high quantiles of a probability distribution, *Statistics* **37**(5), 401–434.

[73]   FERREIRA, M.; GOMES, M.I. and LEIVA, V. (2011). On an extreme value version of the Birnbaum-Saunders distribution. *Revstat.* Accepted.

[74]   FEUERVERGER, A.; and HALL, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution, *Ann. Statist.* **27**, 760–781.

[75]   FISHER, R.A.; and TIPPETT, L.H.C. (1928). Limiting forms of the frequency of the largest or smallest member of a sample. *Proc. Cambridge Phil. Soc.* **24**, 180–190.

[76]   FRAGA ALVES, M.I. (1992; 1995). Estimation of the tail parameter in the domain of attraction of an extremal distribution. *Extreme Value Theory and Applications* (Villeneuve d'Ascq, 1992) and *J. Statist. Plann. Inference* **45**(1-2), 143–173.

[77]   FRAGA ALVES, M.I. (2001). A location invariant Hill-type estimator, *Extremes* **4**(3), 199–217.

[78]   FRAGA ALVES, M.I.; GOMES, M.I.; and HAAN, L. DE (2003a). A new class of semi-parametric estimators of the second order parameter, *Portugaliae Mathematica* **60**(2), 193–213.

[79]   FRAGA ALVES, M.I.; DE HAAN, L.; and LIN, T. (2003b). Estimation of the parameter controlling the speed of convergence in extreme value theory, *Mathematical Methods of Statistics* **12**(2), 155–176.

[80]   FRAGA ALVES, M. I.; HAAN, L. DE; and LIN, T. (2006). Third order extended regular variation, *Publications de l'Institut Mathématique* **80**(94), 109–120.

[81]   FRAGA ALVES, M.I., GOMES, M.I., HAAN, L. DE; and NEVES, C. (2007). A note on second order condition in extremes: linking general and heavy tails conditions, *Revstat* **5**(3), 285–305.

[82]   FRAGA ALVES, M.I.; GOMES, M.I.; DE HAAN, L.; and NEVES, C. (2009). The mixed moment estimator and location invariant alternatives, *Extremes* **12**, 149–185.

[83]   FRÉCHET, M. (1927). Sur le loi de probabilité de l'écart maximum, *Ann. Société Polonaise de Mathématique* **6**, 93–116.

[84]   GALAMBOS, J. (1984). Rates of convergence in extreme value theory. In Tiago de Oliveira, J. (ed.) *Statistical Extremes and Applications*, 347–352, D. Reidel, Dordrecht.

[85]   GELUK, J.; and HAAN, L. DE (1980). *Regular Variation, Extensions and Tauberian Theorems*, CWI Tract 40, Centre for Mathematics and Computer Science, Amsterdam, The Netherlands.

[86]   GNEDENKO, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire, *Annals of Mathematics* **44**(6), 423–453.

[87]   GIRARD, S. (2004). A Hill type estimate of the Weibull tail-coefficient, *Comm.in Statist. – Theory and Methods* **33**(2), 205–234.

[88]   GOEGEBEUR, Y; and GUILLOU, A. (2010). Goodness-of-fit testing for Weibull-type behavior, *J. Statist. Plann. Inference* **140**, 1417–1436.

[89]   GOEGEBEUR, Y; BEIRLANT, J.; and DE WET, T. (2008). Linking Pareto-tail kernel goodness-of-fit statistics with tail index at optimal threshold and second order estimation, *Revstat* **6**(1), 51–69.

[90]   GOEGEBEUR, Y.; BEIRLANT, J.; and DE WET, T. (2010a). Generalized kernel estimators for the Weibull-tail coefficient, *Comm. in Statist. – Theory and Methods* **39**, 3695–3716.

[91]   GOEGEBEUR, Y.; BEIRLANT, J.; and DE WET, T. (2010b). Kernel estimators for the second order parameter in extreme value statistics, *J. Statist. Plann. Infer.* **140**, 2632–2652.

[92]   GOMES, M.I. (1989). Generalized Gumbel and likelihood ratio test statistics in the multivariate GEV model, *Computational Statistics and Data Analysis* **7**, 259–267.

[93]   GOMES, M.I. (1994a). Penultimate behaviour of the extremes. In Galambos, J. *et al.* (eds.) *Extreme Value Theory and Applications*, 403–418, Kluwer, The Netherlands.

[94]   GOMES, M.I. (1994b). Metodologias Jackknife e Bootstrap em Estatística de Extremos. In Mendes-Lopes et al. (eds.), *Actas II Congresso S.P.E.*, 31–46 (in Portuguese).

[95]   GOMES, M.I.; and FIGUEIREDO, F. (2006). Bias reduction in risk modelling: semi-parametric quantile estimation, *Test* **15**(2), 375–396.

[96]   GOMES, M.I.; and HAAN, L. DE (1999). Approximation by penultimate extreme value distributions, *Extremes* **2**(1), 71–85.

[97]   GOMES, M.I.; and MARTINS, M.J. (2001). Generalizations of the Hill estimator: asymptotic versus finite sample behaviour, *J. Statist. Planning and Inference* **93**, 161–180.

[98]   GOMES, M.I.; and MARTINS, M.J. (2002). "Asymptotically unbiased estima-
        tors of the tail index based on external estimation of the second order parameter,
        Extremes **5**(1), 5–31.

[99]   GOMES, M.I.; and MARTINS, M.J. (2004). Bias reduction and explicit semi-
        parametric estimation of the tail index, *J. Statist. Planning and Inference* **124**,
        361–378.

[100]   GOMES, M.I.; and NEVES, M.M. (2011). Estimation of the extreme value index
        for randomly censored data. *Biometrical Letters* **48**(1), 1–22.

[101]   GOMES, M.I.; and OLIVEIRA, O. (2001). The bootstrap methodology in
        Statistics of Extremes: choice of the optimal sample fraction, *Extremes* **4**(4),
        331–358.

[102]   GOMES, M.I.; and PESTANA, D. (2007a). A simple second order reduced bias'
        tail index estimator, *J. Statist. Comput. and Simulation* **77**(6), 487–504.

[103]   GOMES, M.I.; and PESTANA, D. (2007b). A sturdy reduced bias extreme quan-
        tile (VaR) estimator, *J. American Statistical Association* **102**(477), 280–292.

[104]   GOMES, M.I.; MARTINS, M.J.; and NEVES, M. (2000). Alternatives to a semi-
        parametric estimator of parameters of rare events — the Jackknife methodology,
        Extremes **3**(3), 207–229.

[105]   GOMES, M.I.; HAAN, L. DE; and PENG, L. (2002a). Semi-parametric estima-
        tion of the second order parameter — asymptotic and finite sample behaviour,
        Extremes **5**(4), 387–414.

[106]   GOMES, M.I., MARTINS, M.J., AND NEVES, M.M. (2002b). Generalized Jack-
        knife semi-parametric estimators of the tail index, *Portugaliae Mathematica*
        59(4), 393–408.

[107]   GOMES, M. I.; CAEIRO, F.; and FIGUEIREDO, F. (2004a). Bias reduction of a
        extreme value index estimator trough an external estimation of the second order
        parameter, *Statistics* **38**(6): 497–510.

[108]   GOMES, M.I.; DE HAAN, L.; and HENRIQUES-RODRIGUES, L. (2004b). *Tail In-
        dex Estimation through the Accomodation of Bias in the Weighted Log-Excesses*,
        Notas e Comunicações CEAUL 14/2004.

[109]   GOMES, M.I.; FIGUEIREDO, F.; and MENDONÇA, S. (2005a). Asymptotically
        best linear unbiased tail estimators under a second order regular variation, *J.
        Statist. Planning and Inference* **134**(2), 409–433.

[110]   GOMES, M.I.; MIRANDA, C.; and PEREIRA, H. (2005b). Revisiting the role of
        the Jackknife methodology in the estimation of a positive extreme value index,
        Comm. in Statistics: Theory and Methods **34**, 1–20.

[111]   GOMES, M.I.; CANTO E CASTRO, L.; FRAGA ALVES, M.I.; and PES-
        TANA, D. (2007a). Statistics of extremes for iid data and breakthroughs
        in the estimation of the extreme value index: Laurens de Haan lead-
        ing contributions, Notas e Comunicações CEAUL 16/2007, available at
        http://www.ceaul.fc.ul.pt/notas.html?ano=2007

[112] GOMES, M.I.; MARTINS, M.J.; and NEVES, M. (2007b). Improving second order reduced bias extreme value index estimation, *Revstat* **5**(2), 177–207.

[113] GOMES, M.I.; MIRANDA, C.; and VISEU, C. (2007c). Reduced bias extreme value index estimation and the Jackknife methodology, *Statistica Neerlandica* **61**(2), 243–270.

[114] GOMES, M.I.; CANTO E CASTRO, L.; FRAGA ALVES, M.I.; and PESTANA, D. (2008a). Statistics of extremes for iid data and breakthroughs in the estimation of the extreme value index: Laurens de Haan leading contributions, *Extremes* **11**(1), 3–34.

[115] GOMES, M.I.; DE HAAN, L.; and HENRIQUES-RODRIGUES, L. (2008c). Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses, *J. Royal Statistical Society* **B70**(1), 31–52.

[116] GOMES, M.I.; HENRIQUES-RODRIGUES, L.; VANDEWALLE, B.; and VISEU, C. (2008e). A heuristic adaptive choice of the threshold for bias-corrected Hill estimators, *J. Statist. Comput. and Simulation* **78**(2), 133–150.

[117] GOMES, M.I.; HENRIQUES-RODRIGUES, L.; PEREIRA, H.; and PESTANA, D. (2010). Tail index and second order parameters' semi-parametric estimation based on the log-excesses, *J. Statist. Comput. and Simul.* **80**(6), 653–666.

[118] GOMES, M.I.; FIGUEIREDO, F.; and NEVES, M.M. (2011a). Adaptive estimation of heavy right tails: the bootstrap methodology in action, *Extremes*, in press. DOI: 10.1007/s10687-011-0146-6

[119] GOMES, M.I.; HENRIQUES-RODRIGUES, L.; and MIRANDA, C. (2011b). Reduced-bias location-invariant extreme value index estimation: a simulation study, *Comm. Statist. Simul. and Comput.* **40**(3), 424–447.

[120] GOMES, M.I.; MARTINS, M.J.; and NEVES, M.M. (2011c). *Revisiting the Role of the Generalized JackKnife Methodology in the Field of Extremes,* Notas e Comunicações 20/2011. Submitted.

[121] GOMES, M.I.; MENDONÇA, S.; and PESTANA, D. (2011d). Adaptive reduced-bias tail index and VaR estimation via the bootstrap methodology, *Comm. in Statistics – Theory and Methods* **40**(16), 2946-2968.

[122] GRAY, H.L.; and SCHUCANY, W.R. (1972). *The Generalized Jackknife Statistic.* Marcel Dekker.

[123] GREENWOOD, J.A.; LANDWEHR, J.M.; MATALAS, N.C.; and WALLIS, J.R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research 15*, 1049–1054.

[124] GRIENVICH, I.V. (1992a). Max-semistable laws corresponding to linear and power normalizations, *Th. Probab. Appl.* **37**, 720–721.

[125] GRIENVICH, I.V. (1992b). Domains of attraction of max-semistable laws under linear and power normalizations, *Th. Probab. Appl.* **38**, 640–650.

[126] GROENEBOOM, P.; LOPUHA, H.P.; and DE WOLF, P.P. (2003). Kernel-type estimators for the extreme value index, *Ann. Statist.* **31**, 1956–1995.

[127]   GUILLOU, A.; and HALL, P. (2001). A diagnostic for selecting the threshold in extreme-value analysis, *J. R. Statist. Soc.* **B 63**, 293–305.

[128]   GUILLOU, A.; NAVEAU, P.; and YOU, A. (2010). A folding method for extreme quantiles estimation, *Revstat* **8**, 21–35.

[129]   GUMBEL, E.J. (1958). *Statistics of Extremes*. Columbia University Press, New York.

[130]   GUMBEL, E.J. (1965). A quick estimation of the parameters in Fréchet's distribution, *Review Intern. Statist. Inst.* **33**, 349–363.

[131]   HAAN, L. DE (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*, Mathematical Centre Tract 32, Amsterdam.

[132]   HAAN, L. DE (1984). Slow variation and characterization of domains of attraction, In Tiago de Oliveira, ed., *Statistical Extremes and Applications*, 31–48, D. Reidel, Dordrecht, Holland.

[133]   HAAN, L. DE; and FERREIRA, A. (2006). *Extreme Value Theory: an Introduction*, Springer Science+Business Media, LLC, New York, USA.

[134]   HAAN, L. DE; and PENG, L. (1998). Comparison of tail index estimators, *Statistica Neerlandica* **52**, 60–70.

[135]   HAAN, L. DE; and ROOTZÉN, H. (1993). On the estimation of high quantiles, *J. Statist. Planning and Inference* **35**, 1–13.

[136]   HAAN, L. DE; and STADTMÜLLER, U. (1996). Generalized regular variation of second order, *J. Austral. Math. Soc.* **A61**, 381–395.

[137]   HALL, P. (1982). On some simple estimates of an exponent of regular variation, *J. R. Statist. Soc.* **B44**(1), 37–42.

[138]   HALL, P. (1990). Using the bootstrap to estimate mean-square error and selecting smoothing parameter in non-parametric problems, *J. Multivariate Analysis* **32**, 177–203.

[139]   HALL, P.; and WELSH, A.H. (1985). Adaptative estimates of parameters of regular variation, *Ann. Statist.* **13**, 331–341.

[140]   HENRIQUES-RODRIGUES, L.; and GOMES, M.I. (2009). High quantile estimation and the PORT methodology, *Revstat* **7**(3), 245–264.

[141]   HENRIQUES-RODRIGUES, L.; GOMES, M.I.; and PESTANA, D. (2011). Statistics of extremes in athletics, *Revstat* **9**(2), 127-153.

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

[143]   HOSKING, J.R.M. (1984). Testing whether the shape parameter is zero in the generalized extreme value distribution, *Biometrika* **71**, 367–374.

[144]   HOSKING, J.R.M.; and WALLIS, J.R. (1987). Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics* **29**, 339–349.

[145]   HOSKING, J.R.M.; WALLIS, J.R.; and WOOD, E.F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics* **27**, 251–261.

[146] HÜSLER, J.; and LI, D. (2006). On testing extreme value conditions, *Extremes* **9**, 69–86.

[147] HÜSLER, J.; and MÜLLER, S. (2005). Iterative estimation of the extreme value index. *Math. Comput. Appl. Probab.* **7**, 139–148.

[148] HÜSLER, J.; and PENG, L. (2008). Review of testing issues in extremes: in honor of Professor Laurens de Haan, *Extremes* **11**, 99–111.

[149] JENKINSON, A.F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements, *Quart. J. Royal Meteorol. Society* **81**, 158–171.

[150] JOHANSSON, J. (2003). Estimating the mean of heavy-tailed distributions, *Extremes* **6**, 91–109.

[151] JUREČKOVÁ, J.; and PICEK, J. (2001). A class of tests on the tail index, *Extremes* **4**(2), 165–183.

[152] KAUFMANN, E. (2000). Penultimate approximations in extreme value theory, *Extremes* **3**(1), 39–55.

[153] KONING, A.J.; and PENG, L. (2008). Goodness-of-fit tests for a heavy tailed distribution, *J. Statist. Plann. Infer.* **138**, 3960–3981.

[154] KRATZ, M.; and RESNICK, S. (1996). The qq-estimator of the index of regular variation, *Comm. Statist. — Stochastic Models* **12**, 699–724.

[155] LI, D.; and PENG, L. (2009). Does Bias Reduction with External Estimator of Second Order Parameter Work for Endpoint? *J. Statist. Plann. Infer.* **139**, 1937–1952.

[156] LI, D.; PENG, L.; and YANG, J. (2010). Bias reduction for high quantiles, *J. Statist. Plann. Infer.* **140**, 2433–2441.

[157] LI, D.; PENG, L.; and QI, Y. (2011). Empirical likelihood confidence intervals for the endpoint of a distribution function, *Test* **20**, 353–366.

[158] LING, C.; PENG, Z.; and NADARAHJA, S. (2011). Location invariant Weiss-Hill estimator, *Extremes*, DOI 10.1007/s10687-011-0134-x

[159] LUCEÑO, A. (2006). Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators, *Comput. Statist. and Data Analysis* **51**, 904–917.

[160] MAROHN, F. (2000). Testing extreme value models, *Extremes* **3**(4), 363–384.

[161] MATTHYS, G.; and BEIRLANT, J. (2003). Estimating the extreme value index and high quantiles with exponential regression models, *Statistica Sinica* **13**, 853–880.

[162] MATTHYS, G.; DELAFOSSE, M.; GUILLOU, A.; and BEIRLANT, J. (2004). Estimating catastrophic quantile levels for heavy-tailed distributions, *Insurance: Mathematics and Economics* **34**, 517–537.

[163] MISES, R. VON (1936). La distribution de la plus grande de n valeurs, *Revue Math. Union Interbalcanique* **1**, 141-160. Reprinted in *Selected Papers of Richard von Mises*, Amer. Math. Soc. **2** (1964), 271–294.

[164] Neves, C.; and Fraga Alves, M.I. (2008). Testing extreme value conditions
       — an overview and recent approaches, *Revstat* **6**(1), 83–100.

[165] Oliveira, O., Gomes, M.I.; and Fraga Alves, M.I. (2006). Improvements
       in the estimation of a heavy tail, *Revstat* **4**(2), 81–109.

[166] Pancheva, E. (1992). Multivariate max-semistable distributions, *Th. Probab.
       and Appl.* **37**, 731–732.

[167] Pancheva, E. (2010). Max-semistability: a survey, *ProbStat Forum* **3**, 11–24.

[168] Peng, L. (1998). Asymptotically unbiased estimator for the extreme-value in-
       dex, *Statist. Probab. Letters* **38**(2), 107-115.

[169] Peng, L. (2001). Estimating the mean of a heavy-tailed distribution, *Statist.
       Probab. Letters* **52**, 255–264.

[170] Pickands III, J. (1975). Statistical inference using extreme order statistics,
       Ann. Statist. **3**, 119-131.

[171] Raoult, J.P.; and Worms, R. (2003). Rate of convergence for the generalized
       Pareto approximation of the excesses, *Adv. Appl. Probab.* **35**(4), 1007–1027.

[172] Reiss, R.-D.; and Thomas, M. (2001; 2007). *Statistical Analysis of Extreme
       Values, with Application to Insurance, Finance, Hydrology and Other Fields*,
       2nd edition; 3rd edition, Birkhäuser Verlag.

[173] Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Mod-
       eling*, Springer.

[174] Schultze, J.; and Steinbach, J. (1996). On least squares estimators of an
       exponential tail coefficient, *Statistics and Decisions* **14**, 353–372.

[175] Segers, J. (2005). Generalized Pickands estimators for the extreme value index,
       J. Statistical Planning and Inference **28**, 381–396.

[176] Smith, R.L. (1987b). *Approximations in extreme value theory*, Preprint, Univ.
       North-Carolina.

[177] Stephenson, A.; and Tawn, J. (2004). Bayesian inference for extremes: ac-
       counting for the three extremal types, *Extremes* **7**(4), 291–307.

[178] Temido, M.G. and Canto e Castro, L. (2003). Max-semistable laws in
       extremes of stationary random sequences, *Theory Probab. Appl.* **47**(2), 365–374.

[179] Themido Pereira, T. (1993). Second order behavior of domains of attraction
       and the bias of generalized Pickands' estimator. In Lechner, J., Galambos, J., and
       Simiu E. (eds.) *Extreme Value Theory and Applications* III, Proc. Gaithersburg
       Conference (NIST special publ.).

[180] Weissman, I. (1978). Estimation of parameters and large quantiles based on
       the k largest observations, *J. Amer. Statist. Assoc.* **73**, 812–815.

[181] Willems, P.; Guillou, A.; and Beirlant, J. (2007). Bias correction in
       hydrologic GPD based extreme value analysis by means of a slowly varying, *J.
       Hydrology* **338** 221–236.

[182] Zhou, C. (2009). Existence and consistency of the maximum likelihood estima-
       tor for the extreme value index, *J. Multivariate Analysis* **100**(4), 794–815.

[183] ZHOU, C. (2010). The extent of the maximum likelihood estimator for the extreme value index, *J. Multivariate Analysis* **101**(4), 971–983.

[184] YOU, A.; SCHNEIDER, U.; GUILLOU, A.; and NAVEAU, P. (2010). Improving extreme quantile estimation by folding observations, *J. Statist. Plann. Infer.* **140**, 1775–1787.

[185] YUN, S. (2002). On a generalized Pickands estimator of the extreme value index, *J. Statist. Plann. Inf.* **102**, 389–409.