REVSTAT – Statistical Journal
Volume 10, Number 1, March 2012, 1–31

AN OVERVIEW AND OPEN RESEARCH TOPICS IN 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 appropriate 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 open research topics.

Key-Words:

- *extreme value index; parameters of extreme events; 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 that we have 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 societal impact. The domains of application of SUE are thus quite diverse. We mention the fields of hydrology, meteorology, geology, insurance, finance, structural engineering, telecommunications and biostatistics (see, for instance, and among others, Coles, 2001; Reiss & Thomas, 2001, 2007; Beirlant *et al.*, 2004, §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 below.

1.1. Main limiting results in EVT

The main limiting results in EVT date back to the papers by Fréchet (1927), Fisher & 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,

$$\lim_{n\to\infty} \mathbb{P}\big\{(X_{n,n} - b_n)/a_n \leq x\big\} = G(x) \;, \tag{1.1}$$

we can redefine the constants in such a way that

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

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 often use 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, governs the behaviour of the right-tail of F. The EVD, in (1.2), is often separated into the three types:

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

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 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*, $F(x) = K \exp\bigl\{-c/(x^F - x)\bigr\}$, 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 & de Haan, 1974; Pickands, 1975), which can be written as

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

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)

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

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 constitutes an important initial 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 reference 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 supports the approximation

$$P\big(X_{n,n} \le x\big) \;=\; F^n(x) \;\approx\; G_\gamma\big\{(x-\lambda_n)/\delta_n\big\}\;. \tag{1.6}$$

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 make inference on the right-tail weight of the underlying model, it seems sensible to consider 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, contrary 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, $(\underline{X}_1,\ldots,\underline{X}_k)$, where $\underline{X}_j=(X_{1j},\ldots,X_{i_j j})$, $1\le j\le k$, 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 $1\le j\le k$ give the BM model. The choices $m'=n$ $(m=1)$ and $i_1=k$ give the LO model.

The peaks over threshold (POT) approaches. The *Paretian model* for the excesses, $X_j-u>0,\ 1\le j\le k$, 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 $P\big(X-u\le x\mid X>u\big)\approx P_\gamma(x/\sigma)$, 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 & Thomas (2001), the ones by Bermudez & Amaral-Turkman (2003), Bottolo *et al.* (2003), Stephenson & 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 in a central position, turns out to be an important statistical problem. From a parametric point of view, empirical tests of $H_0\colon\gamma=0$ 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 reference only one of the most recent (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 & Hadi (1997) and Chaouche & Bacro (2004). The problem of goodness-of-fit tests for the GPD has been studied by Choulakian & 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 1970s, there was a move from a *parametric approach* based on the limiting models in EVT, towards a *semi-parametric approach*, where tail estimation is done under a quite general framework. In §2 of this review paper we cover 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 §3. Finally, in §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), γ being the key parameter of extreme events to be estimated, using a few large observations, and with suitable 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 & 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 introducing first and second-order conditions in §2.1, §2.2 describes several classical semi-parametric EVI-estimators. In §2.3, we give results on the testing of the EV condition $F \in \mathcal{D}(G_\gamma)$, under a semi-parametric framework. Finally, in §2.4, we outline the semi-parametric estimation of other parameters of extreme events.

2.1. First, second (and higher) order conditions

As mentioned above, in §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 & Ferreira (2006). Indeed, with U standing for a (reciprocal) quantile type function associated with F and defined by $U(t) := \big(1/(1-F)\big)^{\leftarrow}(t) = F^{\leftarrow}(1-1/t) = \inf\big\{x\colon F(x) \geq 1-1/t\big\}$, the *extended regular variation* property,

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

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 many 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

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A^*(t)} = \frac{1}{\rho^*}\left(\frac{x^{\gamma+\rho^*}-1}{\gamma+\rho^*} - \frac{x^\gamma - 1}{\gamma}\right), \qquad x > 0\,, \tag{2.2}$$

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 $\lim_{t\to\infty} A^*(tx)/A^*(t) = x^{\rho^*}$, $x > 0$, i.e., $|A^*| \in RV_{\rho^*}$ (de Haan & Stadtmüller, 1996). For heavy tails, the second-order condition is usually written as

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

where $\rho \leq 0$ and $A(t) \to 0$ as $t \to \infty$. More precisely, $|A| \in RV_\rho$ according to Geluk & de Haan (1987). For the link between $(A^*(t), \rho^*)$ and $(A(t), \rho)$, see de Haan & Ferreira (2006) and Fraga Alves *et al.* (2007). Similarly, third-order conditions specify 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 the chosen CDFs in $\mathcal{D}_{\mathcal{M}}(G_\gamma)$ more strictly.

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 §2.2.1. Next, in §2.2.2, we briefly discuss 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 $\hat\gamma^{\rm H}_{n,k}$, is the average of the scaled log-spacings as well as of the log-excesses, given by

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

respectively. Its asymptotic properties have been thoroughly studied (see de Haan & 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

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

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

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

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

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

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 $GP_\gamma(\alpha x/\gamma)$, $\gamma, \alpha \in \mathbb{R}$, 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 $\hat\gamma^{\rm PORT-ML}_{n,k} := \frac{1}{k}\sum_{i=1}^{k} \ln(1 + \hat\alpha\, D_{ik})$, 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), this 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 $\hat{\gamma}^{\mathcal{K}}_{n,k} := \sum_{i=1}^n \mathcal{K}(i/k)\,(\ln X_{n-i+1,n} - \ln X_{n-k,n})/\sum_{i=1}^n \mathcal{K}(i/k)$, where $\mathcal{K}(\cdot)$ is some non-negative, non-increasing kernel defined on $(0,\infty)$ and with unit integral. 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; see Beirlant *et al.* (1996a,c), Schultze & Steinbach (1996), Kratz & Resnick (1996), Csörgő & 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 $\hat{\gamma}^{\mathrm{P}(\theta)}_{n,k} := -\ln\big\{\big(X_{n-[\theta^2 k]+1,n} - X_{n-[\theta k]+1,n}\big)\big/\big(X_{n-[\theta k]+1,n} - X_{n-k,n}\big)\big\}\big/\ln\theta, \quad 0 < \theta < 1.$ (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, $\hat{\gamma}^{\mathrm{GH}}_{n,k} = \hat{\gamma}^{\mathrm{H}}_{n,k} + \frac{1}{k}\sum_{i=1}^k\big(\ln\hat{\gamma}^{\mathrm{H}}_{n,i} - \ln\hat{\gamma}^{\mathrm{H}}_{n,k}\big)$. 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 $\hat{\varphi}_{n,k} := \big(M^{(1)}_{n,k} - L^{(1)}_{n,k}\big)\big/\big(L^{(1)}_{n,k}\big)^2$, with $L^{(1)}_{n,k} := \frac{1}{k}\sum_{i=1}^k\big(1 - X_{n-k,n}/X_{n-i+1,n}\big)$, and where $M^{(1)}_{n,k}$ 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 $\hat{\gamma}^{\mathrm{MM}}_{n,k} := \big\{\hat{\varphi}_n(k) - 1\big\}\big/\big[1 + 2\min\{\hat{\varphi}_n(k) - 1, 0\}\big]$. This seems 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 & 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 $\hat{\gamma}_{n,k}^{\mathrm{GPPWM}} := 1 - 2\,\hat{a}_1^\star(k)/\big(\hat{a}_0^\star(k) - 2\,\hat{a}_1^\star(k)\big)$, $1 \le k < n$, and $\hat{a}_s^\star(k) := \sum_{i=1}^{k} \big(\frac{i}{k}\big)^s \big(X_{n-i+1:n} - X_{n-k:n}\big)/k$, $s = 0, 1$. On the basis of the Pareto model, Caeiro & Gomes (2011) introduced the PPWM EVI-estimators, with PPWM standing for *Pareto PWM*, given by $\hat{\gamma}_{n,k}^{\mathrm{PPWM}} := 1 - \hat{a}_1(k)/\big\{\hat{a}_0(k) - \hat{a}_1(k)\big\}$, where $\hat{a}_s(k) := \frac{1}{k+1} \sum_{i=1}^{k+1} \big(\frac{i}{k+1}\big)^s X_{n-i+1:n}$, $s = 0, 1$ with $1 \le k < n$.

Other estimators. Falk (1995a) proposed the location-invariant estimator, $\hat{\gamma}_{n,k} := \frac{1}{k} \sum_{i=1}^{k-1} \ln\big(X_{n,n} - X_{n-i,n}\big)/\big(X_{n,n} - X_{n-k,n}\big)$, 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 & Müller (2005). The non-invariance for shifts of the H-estimator led Fraga Alves (2001) to the consideration for $k > k_0$, with k_0 appropriately chosen, of the location invariant Hill-type estimator $\hat{\gamma}_{n,k,k_0} := \frac{1}{k_0} \sum_{i=1}^{k} \ln\big(\big(X_{n-i+1,n} - X_{n-k,n}\big)/\big(X_{n-k_0+1,n} - X_{n-k,n}\big)\big)$. 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., §6.4 of Embrechts *et al.* (1997), Beirlant *et al.* (1996a;1998), Csörgő & Viharos (1998), §3 of de Haan & 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, with T denoting any of these EVI-estimators, and with $B(t)$ a bias function converging to zero as $t \to \infty$ and closely 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

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

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 $n \to \infty$,

$$\sqrt{k}\,\big(\hat{\gamma}_{n,k}^T - \gamma\big) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \mathrm{Normal}(\lambda\, b_T, \sigma_T^2)\ .$$

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 $H_0\colon F \in \mathcal{D}_{\mathcal{M}}(G_0)$ against $H_1\colon F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$, $\gamma \neq 0$, or the corresponding one-sided alternatives, under a semi-parametric framework is obviously natural and sensible. 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 & 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 & 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 & Peng (2008) and Neves & Fraga Alves (2008). See also Koning & Peng (2008) and Goegebeur & 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 $X_{n-k+1,n} \overset{p}{\sim} U(n/k)$ enables us to estimate χ_{1-p} on the basis of this approximation and appropriate 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 $\hat{\chi}_{1-p,k} := X_{n-k:n}\big\{k/(np)\big\}^{\hat{\gamma}_k}$, where $\hat{\gamma}_k$ is any consistent semi-parametric EVI-estimator. This estimator is of the type introduced by Weissman (1978). Details on semi-parametric estimation of extremely high quantiles for $\gamma \in \mathbb{R}$, can be found in Dekkers & de Haan (1989), de Haan & 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 & Beirlant (2003). None of the above mentioned quantile estimators is equivariant. Araújo Santos *et al.* (2006) provide a class of semi-parametric VaR_p estimators which enjoy equivariance, 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 & 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 & Mason (1989), and Aarssen & 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. See also de Haan & 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 & Hall (1999) and Gomes *et al.* (2000), where the classical bias-variance trade-off always appears. Such a trade-off was removed with an appropriate 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 that of the Hill EVI-estimator but an asymptotic bias of smaller order, and thus beat the classical estimators for all k. In § 3.1 we deal with SORB semiparametric EVI-estimation and in § 3.2, we briefly describe 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 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 σ^2/k and $b^2B^2(n/k)$ respectively. Consequently, these estimators exhibit the same 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 the adaptive choice of the threshold problematic on the basis of any sample path stability criterion; and a very peaked MSE, making the choice of the value $k_0 := \arg\min_k \mathrm{MSE}\,(\hat{\gamma}_{n,k})$ difficult. These peculiarities have led researchers to consider the possibility of dealing with the bias term in an appropriate manner, building new estimators $\hat{\gamma}^R_{n,k}$, here called SORB EVI-estimators. In particular, for heavy tails, i.e., $\gamma > 0$, bias reduction is very important for the estimation of γ or of the Pareto index, $\alpha = 1/\gamma$, when the slowly varying part of the Pareto type model disappears at a very slow rate. We consider the following definition (Reiss & Thomas, 2007, §6).

Definition 3.1. Under the second-order condition in (2.2) and for intermediate k, the statistic $\hat{\gamma}^R_{n,k}$, 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^R_k, such that for a large class of models in $\mathcal{D}_{\mathcal{M}}(EV_\gamma)$, and with $B(\cdot)$ the function in (2.6),

$$\hat{\gamma}^R_{n,k} \;\stackrel{d}{=}\; \gamma + \sigma_R P^R_k/\sqrt{k} + o_p\{B(n/k)\}\ . \tag{3.1}$$

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,

$$\sqrt{k}\big(\hat{\gamma}^R_{n,k} - \gamma\big) \;\underset{n\to\infty}{\stackrel{d}{\longrightarrow}}\; \mathrm{Normal}\big(0, \sigma^2_R\big)$$

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 to the optimal level.

Such an approach has been carried out for heavy tails in different manners. The key ideas are either to find ways of getting rid of the dominant component $b\,B(n/k)$ 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 date back to Gomes (1994b), Drees (1996) and Peng (1998), among others. Gomes uses the *generalized jackknife* (GJ) methodology in Gray & Schucany (1972), and Peng deals with linear combinations of appropriate EVI-estimators, in a spirit close to that associated to the GJ technique. Feuerverger & 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 §3.1.1, we provide details about the GJ EVI-estimation. In §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 §3.1.3. Finally in §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 & Schucany (1972). Whenever we are dealing with semi-parametric EVI-estimators, or even estimators of other parameters of extreme events, we usually have 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 was revisited by Gomes *et al.* (2011c). Indeed, if the second-order condition in (2.3) holds, we can easily find two statistics $\hat{\gamma}^{(j)}_{n,k}$, such that (2.6) holds for both. The ratio between the dominant components of bias of $\hat{\gamma}^{(1)}_{n,k}$ and $\hat{\gamma}^{(2)}_{n,k}$ is $q = b_1/b_2 = q(\rho)$, and we get the GJ random variable,

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

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

$$\sqrt{k}\big(\hat{\gamma}^{\mathrm{GJ}(\hat{\rho})}_{n,k} - \gamma\big) \ \underset{n\to\infty}{\overset{d}{\longrightarrow}}\ \mathrm{Normal}\big(0, \sigma^2_{_{GJ}}\big)\ . \tag{3.3}$$

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

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

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 the asymptotic variance of the EVI-estimator will change. Gomes *et al.* (2000) have suggested the misspecification of ρ at $\rho = -1$, essentially due not only to the high bias and variance of the existing 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 & 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 a 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),

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

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

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

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

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

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,

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

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

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

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

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

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,

$$\sqrt{k}\,\Big(\hat\gamma_{n,k}^{\mathrm{FH}(\hat\beta,\hat\rho)} - \gamma\Big) \;\underset{n\to\infty}{\overset{d}{\longrightarrow}}\; \mathrm{Normal}\!\left(0,\, \sigma^2_{\mathrm{FH}} = \gamma^2 \left(\frac{1-\rho}{\rho}\right)^{4}\right). \tag{3.10}$$

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 $\sigma^2_{\text{FH}} = \gamma^2\{(1-\rho)/\rho\}^4$. Note that the asymptotic variance is smaller, and given by $\gamma^2\{(1-\rho)/\rho\}^2$, 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 & Martins (2002) to an explicit estimator for β, given by

$$(3.11) \qquad \hat\beta^{\text{GM}(\hat\rho)}_{n,k} := \left(\frac{k}{n}\right)^{\hat\rho} \frac{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\hat\rho}\right)\hat C_0 - \hat C_1}{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\hat\rho}\right)\hat C_1 - \hat C_2}\,, \qquad \hat C_j = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-j\hat\rho} U_i\,,$$

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

$$(3.12) \qquad \hat\gamma^{\text{GM}(\hat\rho)}_{n,k} := \frac{1}{k}\sum_{i=1}^{k} U_i - \hat\beta^{\text{GM}(\hat\rho)}_{n,k}\left(\frac{n}{k}\right)^{\hat\rho}\hat C_1\,.$$

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 $\hat\beta^{\text{GM}(\hat\rho)}_{n,k}(n/k)^{\hat\rho}\,\hat C_1$, and directly removed from the H-estimator, which can also be written as $\gamma^{\text{H}}_{n,k} = \sum_{i=1}^{k} U_i/k$. Under the same conditions as before, the asymptotic variance of $\hat\gamma^{\text{GM}(\hat\rho)}_{n,k}$ is $\sigma^2_{\text{GM}} = \gamma^2(1-\rho)^2/\rho^2 < \sigma^2_{\text{FH}}$, but still greater than $\sigma^2_{\text{H}} = \gamma^2$.

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}, $1 \le i \le k < n$, also in (2.4), giving rise to the WH EVI-estimator in Gomes *et al.* (2008c),

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

where $(\hat\beta,\hat\rho)$ are suitable consistent estimators of second-order parameters (β,ρ). The key to the 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 (β,ρ) 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). Compared 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, in comparison with the Hill estimator, it keeps the same asymptotic variance $\sigma^2_{\text{WH}} = \sigma^2_{\text{H}} = \gamma^2$

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

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

where $\gamma_{n,k}^{\mathrm{H}}$ denotes the H-estimator, and $(\hat{\beta}, \hat{\rho})$ are appropriate consistent estimators of the second-order parameters (β, ρ). With the same objectives, but with a simpler expression, we also mention the estimator (Caeiro *et al.*, 2005).

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

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 appropriate 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. Reiss & Thomas (2007), §6, and Gomes *et al.* (2008a) review this topic.

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 & 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 & Welsh (1985), Beirlant *et al.* (1996c) and Drees & Kaufmann (1998). Another estimator of ρ appears in Gomes *et al.* (2002a), and more recently, we mention the classes of ρ-estimators in Goegebeur *et al.* (2008; 2010b) and Ciuperca & Mercadier (2010). Here we choose particular members of the class of estimators of the second-order parameter ρ proposed by Fraga Alves *et al.* (2003a). Under appropriate 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

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

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

Gomes & 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 & 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 can be found in Gomes & 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 $n \to \infty$, $\big(\hat{\rho}^{(\tau)}_{n,k_1} - \rho\big) \ln n = o_p(1)$, 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 approaches to bias reduction, in the estimation of a positive EVI can be found in Gomes & Martins (2001, 2004), Caeiro & Gomes (2002), Gomes *et al.* (2004a; 2005a; 2005b; 2007c; 2011a), Canto e Castro & 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 & Figueiredo (2006), who consider the classical SORB EVI-estimators. Gomes & 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 & 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 & 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 § 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 may or may not support use of the commonest models in SUE. As noted by Fisher & Tippett (1928), although 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 approximation 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 & de Haan (1999) derived, for all $\gamma \in \mathbb{R}$, exact penultimate approximation rates with respect to the variational distance, under appropriate differentiability assumptions. Kaufmann (2000) proved, under weaker conditions, a result related to that in Gomes & de Haan (1999). This penultimate or pre-asymptotic behaviour has further been studied by Raoult & Worms (2003) and Diebolt & 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

$$PG_\gamma(x; r) = \exp\Big[-(1 + \gamma x)^{-1/\gamma} \Big\{1 + r(1 + \gamma x)^{-1/\gamma}\Big\}\Big]. \tag{4.1}$$

These models surely deserve deeper statistical consideration. *Penultimate models* seem interesting alternatives to the classical models but have never been much used. Concomitantly, the convergence of the estimators can be very slow when $\rho = 0$ or $\rho^* = 0$, as happens with normal and loggamma distributions, important models in many areas, and alternative estimation procedures are still needed.

4.2. Max-semistable laws

We also mention 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 & 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

$$G_{\gamma,\nu}(x) = \begin{cases} \exp\Big[-\nu\big\{\ln(1+\gamma x)\big\}\,(1+\gamma x)^{-1/\gamma}\Big], \quad 1+\gamma x>0, & \text{if } \gamma\neq 0\,, \\ \exp\big\{-\nu(x)\exp(-x)\big\}, \quad x\in\mathbb{R}, & \text{if } \gamma=0\,, \end{cases}$$

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

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

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

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 & Welsh (1985), Hall (1990), Beirlant *et al.* (1996c), Drees & 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 & Oliveira (2001) also uses the bootstrap methodology to provide an adaptive choice of the threshold, alternative to that 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. Beirlant *et al.* (2002) consider the exponential regression model (ERM) introduced in Beirlant *et al.* (1999), discuss applications of the ERM to the selection of the optimal sample fraction in EV estimation, and derive a connection between the new choice strategy in the paper and the diagnostic of Guillou & Hall (2001). Csörgő & Viharos (1998) provide a data-driven choice of k for kernel estimators. Apart from the papers by Drees & 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 & 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 mentioned in §1.2 and 2.3. And what about testing second-order and third-order conditions? Change-point detection is also a challenging topic. And SUE for weakly dependent data, with all problems related with clustering of extreme values, merits further research. SUE for randomly censored data is another challenging topic. See Beirlant *et al.* (2007; 2010), Einmahl *et al.* (2008a) and Gomes & 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 & Magnus (2008), Li & Peng (2009), Einmahl & 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

Research partially supported by National Funds through FCT — *Fundação para a Ciência e a Tecnologia*, projects PEst-OE/MAT/UI0006/2011 (CEAUL) and PEst-OE/MAT/UI0297/2011 (CMA/UNL) and PTDC/FEDER.

REFERENCES

Aarssen, K. & de Haan, L. (1994). On the maximal life span of humans, *Math. Pop. Stud.*, **4**, 259–281.

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

Araújo Santos, P.; Fraga Alves, M.I. & Gomes, M.I. (2006). Peaks over random threshold methodology for tail index and quantile estimation, *Revstat*, **4**, 227–247.

Balkema, A.A. & de Haan, L. (1974). Residual life time at great age, *Ann. Prob.*, **2**, 792–804.

Beirlant, J.; Teugels, J. & Vynckier, P. (1996a). *Practical Analysis of Extreme Values*, Leuven University Press, Leuven, Belgium.

Beirlant, J.; Vynckier, P. & Teugels, J. (1996b). Excess functions and estimation of the extreme-value index, *Bernoulli*, **2**, 293–318.

Beirlant, J.; Vynckier, P. & Teugels, J. (1996c). Tail index estimation, Pareto quantile plots and regression, *J. Am. Statist. Assoc.*, **91**, 1659–1667.

Beirlant, J.; Teugels, J. & Vynckier, P. (1998). *Some thoughts on extreme values.* In “Probability Towards 2000” (L. Accardi and C.C. Heyde, Eds.), Lecture Notes in Statistics, **128**, 58–73, Springer, New York.

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

Beirlant, J.; Dierckx, G.; Guillou, A. & Starica, C. (2002). On exponential representations of log-spacings of extreme order statistics, *Extremes*, **5**, 157–180.

Beirlant, J.; Goegebeur, Y.; Segers, J. & Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*, Wiley, Chichester.

Beirlant, J.; Dierckx, G. & Guillou, A. (2005). Estimation of the extreme-value index and generalized quantile plots, *Bernoulli*, **11**, 949–970.

Beirlant, J.; de Wet, T. & Goegebeur, Y. (2006). A goodness-of-fit for Pareto-type behaviour, *J. Comput. Appl. Math.*, **186**, 99–116.

Beirlant, J.; Guillou, A.; Dierckx, G. & Fils-Viletard, A. (2007). Estimation of the extreme value index and extreme quantiles under random censoring, *Extremes*, **10**, 151–174.

Beirlant, J.; Figueiredo, F.; Gomes, M.I. & Vandewalle, B. (2008). Improved reduced bias tail index and quantile estimators, *J. Statist. Plann. Infer.*, **138**, 1851–1870.

Beirlant, J.; Joossens, E. & Segers, J. (2009). Second-order refined peaks-over-threshold modelling for heavy-tailed distributions, *J. Statist. Plann. Infer.*, **139**, 2800–2815.

Beirlant, J.; Guillou, A. & Toulemonde, G. (2010). Peaks-over-threshold modeling under random censoring, *Comm. Statist. Th. Meth.*, **39**, 1158–1179.

Beirlant, J.; Boniphace, E. & Dierckx, G. (2011). Generalized sum plots, *Revstat*, **9**, 181–198.

Bermudez, P. de Zea & Amaral-Turkman, M.A. (2003). Bayesian approach to parameter estimation of the generalized Pareto distribution, *Test*, **12**, 259–277.

Bingham, N.; Goldie, C.M. & Teugels, J.L. (1987). *Regular Variation*, Cambridge Univ. Press, Cambridge.

Bottolo, L.; Consonni, G.; Dellaportas, P. & Lijoi, A. (2003). Bayesian analysis of extreme values by mixture modeling, *Extremes*, **6**, 25–47.

Brilhante, M.F. (2004). Exponential versus generalized Pareto — a resistant and robust test, *Revstat*, **2**, 1–13.

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

Caeiro, C. & Gomes, M.I. (2006). A new class of estimators of a scale second order parameter, *Extremes*, **9**, 193–211.

Caeiro, F. & Gomes, M.I. (2009). Semi-parametric second-order reduced-bias high quantile estimation, *Test*, **18**, 392–413.

Caeiro, F. & Gomes, M.I. (2011). Semi-parametric tail inference through probability-weighted moments, *J. Statist. Plann. Infer.*, **141**, 937–950.

Caeiro, C.; Gomes, M.I. & Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator, *Revstat*, **3**, 113–136.

Caeiro, F.; Gomes, M.I. & Henriques-Rodrigues, L. (2009). Reduced-bias tail index estimators under a third order framework, *Comm. Statist. Th. Meth.*, **38**, 1019–1040.

Cai, J.; de Haan, L. & Zhou, C. (2011). *Bias correction in extreme value statistics with index around zero*, available at: `http://people.few.eur.nl/ldehaan/BiasCorrection11.pdf`

Canto e Castro, L. & Dias, S. (2011). Generalized Pickands estimators for the tail index parameter and max-semistability, *Extremes*, **14**, 429–449.

Canto e Castro, L. & de Haan, L. (2006). A class of distribution functions with less bias in extreme value estimation, *Statist. Prob. Lett.*, **76**, 1617–1624.

Canto e Castro, L.; de Haan, L. & Temido, M.G. (2000). Rarely observed maxima, *Th. Prob. Appl.*, **45**, 779–782.

Canto e Castro, L.; Dias, S. & Temido, M.G. (2011). Looking for max-semistability: a new test for the extreme value condition, *J. Statist. Plann. Infer.*, **141**, 3005–3020.

Castillo, E. & Hadi, A. (1997). Fitting the generalized Pareto distribution to data, *J. Am. Statist. Assoc.*, **92**, 1609–1620.

Castillo, E.; Hadi, A.; Balakrishnan, N. & Sarabia, J.M. (2005). *Extreme Value and Related Models with Applications in Engineering and Science*, Wiley, New Jersey.

Chaouche, A. & Bacro, J.-N. (2004). A statistical test procedure for the shape parameter of a generalized Pareto distribution, *Comput. Statist. Data Anal.*, **45**, 787–803.

Choulakian, V. & Stephens, M.A. (2001). Goodness-of-fit tests for the generalized Pareto distribution, *Technometrics*, **43**, 478–484.

Ciuperca, G. & Mercadier, C. (2010). Semi-parametric estimation for heavy tailed distributions, *Extremes*, **13**, 55–87.

Coles, S. (2001). *An Introduction to Statistical Modelling of Extreme Values*, Springer, New York.

Csörgő, S. & Mason, D.M. (1989). *Simple estimators of the endpoint of a distribution.* In "Extreme Value Theory, Proceedings Oberwolfach 1987" (J. Hüsler and R.-D. Reiss, Eds., 1989), 132–147, Springer, Berlin–Heidelberg.

Csörgő, S. & Viharos, L. (1998). *Estimating the tail index.* In "Asymptotic Methods in Probability and Statistics" (B. Szyszkowicz, Ed.), 833–881, North-Holland.

Csörgő, S.; Deheuvels, P. & Mason, D.M. (1985). Kernel estimates of the tail index of a distribution, *Ann. Statist.*, **13**, 1050–1077.

Danielsson, J.; de Haan, L.; Peng, L. & de Vries, C.G. (2001). Using a bootstrap method to choose the sample fraction in the tail index estimation, *J. Mult. Anal.*, **76**, 226–248.

Davison, A.C. (1984). *Modelling excesses over high thresholds.* In "Statistical Extremes and Applications" (J. Tiago de Oliveira, Ed.), 461–482, D. Reidel, Dordrecht.

Dekkers, A.L.M. & de Haan, L. (1989). On the estimation of the extreme-value index and large quantile estimation, *Ann. Statist.*, **17**, 1795–1832.

Dekkers, A.; Einmahl, J. & de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution, *Ann. Statist.*, **17**, 1833–1855.

Diebolt, J. & Guillou, A. (2005). Asymptotic behavior of regular estimators, *Revstat*, **3**, 19–44.

Diebolt, J.; El-Aroui, M.A.; Garrido, M. & Girard, S. (2005). Quasi-conjugate Bayes estimates for GPD parameters and application to heavy tails modelling, *Extremes*, **8**, 57–78.

Diebolt, J.; Guillou, A. & Rached, I. (2007). Approximation of the distribution of excesses through a generalized probability-weighted moments method, *J. Statist Plann. Infer.*, *137*, 841–857.

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

Diebolt, J.; Gardes, L.; Girard, S. & Guillou, A. (2008b). Bias-reduced extreme quantiles estimators of Weibull tail-distributions, *J. Statist. Plann. Infer.*, **138**, 1389–1401.

Diebolt, J.; Guillou, A.; Naveau, P. & Ribereau, P. (2008c). Improving probability-weighted moment methods for the generalized extreme value distribution, *Revstat*, **6**, 33–50.

Dietrich, D.; de Haan, L. & Hüsler, J. (2002). Testing extreme value conditions, *Extremes*, **5**, 71–85.

Dijk, V. & de Haan, L. (1992). *On the estimation of the exceedance probability of a high level.* In "Order statistics and nonparametrics: theory and applications" (P.K. Sen and I.A. Salama, Eds.), 79–92, Elsevier, Amsterdam.

Draisma, G.; de Haan, L.; Peng, L. & Pereira, T.T. (1999). A bootstrap-based method to achieve optimality in estimating the extreme-value index, *Extremes*, **2**, 367–404.

Drees, H. (1995). Refined Pickands estimator of the extreme value index, *Ann. Statist.*, **23**, 2059–2080.

Drees, H. (1996). Refined Pickands estimators with bias correction, *Comm. Statist. Th. Meth.*, **25**, 837–851.

Drees, H. (1998). A general class of estimators of the extreme value index, *J. Statist. Plann. Inf.*, **66**, 95–112.

Drees, H. & Kaufmann, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation, *Stoch. Proc. Application*, **75**, 149–172.

Drees, H.; Ferreira, A. & de Haan, L. (2004). On maximum likelihood estimation of the extreme value index, *Ann. Appl. Prob.*, **14**, 1179–1201.

Drees, H.; de Haan, L. & Li, D. (2006). Approximations to the tail empirical distribution function with application to testing extreme value conditions, *J. Stat. Plann. Infer.*, **136**, 3498–3538.

Dwass, M. (1964). Extremal processes, *Ann. Math. Statist.*, **35**, 1718–1725.

Einmahl, J. & Magnus, J.R. (2008). Records in athletics through extreme-value theory, *J. Am. Statist. Assoc.*, **103**, 1382–1391.

Einmahl, J. & Smeets, S.G.W.R. (2011). Ultimate 100-m world records through extreme-value theory, *Statist. Neerl.*, **65**, 32–42.

Einmahl, J.H.J.; Fils-Villetard, A. & Guillou, A. (2008a). Statistics of extremes under random censoring, *Bernoulli*, **14**, 207–227.

Embrechts, P.; Klüppelberg, C. & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin–Heidelberg.

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

Falk, M. (1995a). Some best parameter estimates for distributions with finite endpoint, *Statistics*, **27**, 115–125.

Falk, M.; Hüsler, J. & Reiss, R.-D. (1994; 2004). *Laws of Small Numbers: Extremes and Rare Events*, 1^{st} edition; 2^{nd} edition, Birkhäuser, Basel.

Falk, F.; Guillou, A. & Toulemonde, G. (2008). A LAN based Neyman smooth test for the generalized Pareto distribution, *J. Statist. Plann. Infer.*, **138**, 2867–2886.

Ferreira, A. (2002). Optimal asymptotic estimation of small exceedance probabilities, *J. Statist. Plann. Infer.*, **104**, 83–102.

Ferreira, A.; de Haan, L. & Peng, L. (2003). On optimizing the estimation of high quantiles of a probability distribution, *Statistics*, **37**, 401–434.

Ferreira, M.; Gomes, M.I. & Leiva, V. (2011). On an extreme value version of the Birnbaum–Saunders distribution, *Revstat*. Accepted.

Feuerverger, A. & Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution, *Ann. Statist.*, **27**, 760–781.

Fisher, R.A. & 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.

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. Infer.*, **45**(1–2), 143–173.

Fraga Alves, M.I. (2001). A location invariant Hill-type estimator, *Extremes*, **4**, 199–217.

Fraga Alves, M.I.; Gomes, M.I. & de Haan, L. (2003a). A new class of semi-parametric estimators of the second order parameter, *Portugaliae Math.*, **60**, 193–213.

Fraga Alves, M.I.; de Haan, L. & Lin, T. (2003b). Estimation of the parameter controlling the speed of convergence in extreme value theory, *Math. Meth. Statist.*, **12**, 155–176.

Fraga Alves, M.I.; de Haan, L. & Lin, T. (2006). Third order extended regular variation, *Pub. l'Institut Math.*, **80**, 109–120.

Fraga Alves, M.I.; Gomes, M.I.; de Haan, L. & Neves, C. (2007). A note on second order condition in extremes: linking general and heavy tails conditions, *Revstat*, **5**, 285–305.

Fraga Alves, M.I.; Gomes, M.I.; de Haan, L. & Neves, C. (2009). The mixed moment estimator and location invariant alternatives, *Extremes*, **12**, 149–185.

Fréchet, M. (1927). Sur le loi de probabilité de l'écart maximum, *Ann. Soc. Polonaise Math.*, **6**, 93–116.

Galambos, J. (1984). *Rates of convergence in extreme value theory.* In "Statistical Extremes and Applications" (J. Tiago de Oliveira, Ed.), 347–352, D. Reidel, Dordrecht.

Geluk, J. & de Haan, L. (1980). *Regular Variation, Extensions and Tauberian Theorems*, CWI Tract 40, Centre for Mathematics and Computer Science, Amsterdam, The Netherlands.

Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.*, **44**, 423–453.

Girard, S. (2004). A Hill type estimate of the Weibull tail-coefficient, *Comm. Statist. Th. Meth.*, **33**, 205–234.

Goegebeur, Y. & Guillou, A. (2010). Goodness-of-fit testing for Weibull-type behavior, *J. Statist. Plann. Infer.*, **140**, 1417–1436.

Goegebeur, Y.; Beirlant, J. & de Wet, T. (2008). Linking Pareto-tail kernel goodness-of-fit statistics with tail index at optimal threshold and second order estimation, *Revstat*, **6**, 51–69.

Goegebeur, Y.; Beirlant, J. & de Wet, T. (2010a). Generalized kernel estimators for the Weibull-tail coefficient, *Comm. Statist. Th. Meth.*, **39**, 3695–3716.

Goegebeur, Y.; Beirlant, J. & de Wet, T. (2010b). Kernel estimators for the second order parameter in extreme value statistics, *J. Statist. Plann. Infer.*, **140**, 2632–2652.

Gomes, M.I. (1989). Generalized Gumbel and likelihood ratio test statistics in the multivariate GEV model, *Comput. Statist. Data Anal.*, **7**, 259–267.

Gomes, M.I. (1994a). *Penultimate behaviour of the extremes.* In "Extreme Value Theory and Applications" (J. Galambos *et al.*, Eds.), 403–418, Kluwer, The Netherlands.

Gomes, M.I. (1994b). *Metodologias Jackknife e bootstrap em estatística de extremos.* In "Actas II Congresso SPE" (Mendes-Lopes *et al.*, Eds.), 31–46 (in Portuguese).

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

Gomes, M.I. & de Haan, L. (1999). Approximation by penultimate extreme value distributions, *Extremes*, **2**, 71–85.

Gomes, M.I. & Martins, M.J. (2001). Generalizations of the Hill estimator: asymptotic versus finite sample behaviour, *J. Statist. Planning Infer.*, **93**, 161–180.

Gomes, M.I. & Martins, M.J. (2002). Asymptotically unbiased estimators of the tail index based on external estimation of the second order parameter, *Extremes*, **5**, 5–31.

Gomes, M.I. & Martins, M.J. (2004). Bias reduction and explicit semi-parametric estimation of the tail index, *J. Statist. Plann. Infer.*, **124**, 361–378.

Gomes, M.I. & Neves, M.M. (2011). Estimation of the extreme value index for randomly censored data, *Biomet. Lett.*, **48**, 1–22.

Gomes, M.I. & Oliveira, O. (2001). The bootstrap methodology in Statistics of Extremes: choice of the optimal sample fraction, *Extremes*, **4**, 331–358.

Gomes, M.I. & Pestana, D. (2007a). A simple second order reduced bias' tail index estimator, *J. Statist. Comput. Simul.*, **77**, 487–504.

Gomes, M.I. & Pestana, D. (2007b). A sturdy reduced bias extreme quantile (VaR) estimator, *J. Am. Statist. Assoc.*, **102**, 280–292.

Gomes, M.I.; Martins, M.J. & Neves, M. (2000). Alternatives to a semi-parametric estimator of parameters of rare events — the Jackknife methodology, *Extremes*, **3**, 207–229.

Gomes, M.I.; de Haan, L. & Peng, L. (2002a). Semi-parametric estimation of the second order parameter — asymptotic and finite sample behaviour, *Extremes*, **5**, 387–414.

Gomes, M.I.; Martins, M.J. & Neves, M.M. (2002b). Generalized Jackknife semi-parametric estimators of the tail index, *Portugaliae Math.*, **59**, 393–408.

Gomes, M.I.; Caeiro, F. & Figueiredo, F. (2004a). Bias reduction of a extreme value index estimator trough an external estimation of the second order parameter, *Statistics*, **38**, 497–510.

Gomes, M.I.; de Haan, L. & Henriques-Rodrigues, L. (2004b). Tail index estimation through the accomodation of bias in the weighted log-excesses, *Notas e Comunicações CEAUL*, 14/2004.

Gomes, M.I.; Figueiredo, F. & Mendonça, S. (2005a). Asymptotically best linear unbiased tail estimators under a second order regular variation, *J. Statist. Plann. Infer.*, **134**, 409–433.

Gomes, M.I.; Miranda, C. & Pereira, H. (2005b). Revisiting the role of the Jackknife methodology in the estimation of a positive extreme value index, *Comm. Statist. Th. Meth.*, **34**, 1–20.

Gomes, M.I.; Canto e Castro, L.; Fraga Alves, M.I. & Pestana, D. (2007a). Statistics of extremes for iid data and breakthroughs in the estimation of the extreme value index: Laurens de Haan leading contributions, *Notas e Comunicações CEAUL*, 16/2007. Available at: http://www.ceaul.fc.ul.pt/notas.html?ano=2007

Gomes, M.I.; Martins, M.J. & Neves, M. (2007b). Improving second order reduced bias extreme value index estimation, *Revstat*, **5**, 177–207.

Gomes, M.I.; Miranda, C. & Viseu, C. (2007c). Reduced bias extreme value index estimation and the Jackknife methodology, *Statist. Neerl.*, **61**, 243–270.

Gomes, M.I.; Canto e Castro, L.; Fraga Alves, M.I. & 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**, 3–34.

Gomes, M.I.; de Haan, L. & Henriques-Rodrigues, L. (2008c). Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses, *J. R. Statist. Soc. B*, **70**, 31–52.

Gomes, M.I.; Henriques-Rodrigues, L.; Vandewalle, B. & Viseu, C. (2008e). A heuristic adaptive choice of the threshold for bias-corrected Hill estimators, *J. Statist. Comput. Simul.*, **78**, 133–150.

Gomes, M.I.; Henriques-Rodrigues, L.; Pereira, H. & Pestana, D. (2010). Tail index and second order parameters' semi-parametric estimation based on the log-excesses, *J. Statist. Comput. Simul.*, **80**, 653–666.

Gomes, M.I.; Figueiredo, F. & 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.

Gomes, M.I.; Henriques-Rodrigues, L. & Miranda, C. (2011b). Reduced-bias location-invariant extreme value index estimation: a simulation study, *Comm. Statist. Simul. Comput.*, **40**, 424–447.

Gomes, M.I.; Martins, M.J. & Neves, M.M. (2011c). Revisiting the role of the generalized JackKnife methodology in the field of extremes, *Notas e Comunicações CEAUL*, 20/2011. Submitted.

Gomes, M.I.; Mendonça, S. & Pestana, D. (2011d). Adaptive reduced-bias tail index and VaR estimation via the bootstrap methodology, *Comm. Statist. Th. Meth.*, **40**, 2946–2968.

Gray, H.L. & Schucany, W.R. (1972). *The Generalized Jackknife Statistic*, Dekker, New York.

Greenwood, J.A.; Landwehr, J.M.; Matalas, N.C. & Wallis, J.R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form, *Wat. Resour. Res.*, **15**, 1049–1054.

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

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

Groeneboom, P.; Lopuha, H.P. & de Wolf, P.P. (2003). Kernel-type estimators for the extreme value index, *Ann. Statist.*, **31**, 1956–1995.

Guillou, A. & Hall, P. (2001). A diagnostic for selecting the threshold in extreme-value analysis, *J. R. Statist. Soc. B*, **63**, 293–305.

Guillou, A.; Naveau, P. & You, A. (2010). A folding method for extreme quantiles estimation, *Revstat*, **8**, 21–35.

Gumbel, E.J. (1958). *Statistics of Extremes*, Columbia University Press, New York.

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

de Haan, L. (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*, Mathematical Centre Tract 32, Amsterdam.

de Haan, L. (1984). *Slow variation and characterization of domains of attraction.* In "Statistical Extremes and Applications" (J. Tiago de Oliveira, Ed.), 31–48, D. Reidel, Dordrecht, Holland.

de Haan, L. & Ferreira, A. (2006). *Extreme Value Theory: An Introduction*, Springer, New York.

de Haan, L. & Peng, L. (1998). Comparison of tail index estimators, *Statist. Neerl.*, **52**, 60–70.

de Haan, L. & Rootzén, H. (1993). On the estimation of high quantiles, *J. Statist. Plann. Infer.*, **35**, 1–13.

de Haan, L. & Stadtmüller, U. (1996). Generalized regular variation of second order, *J. Austral. Math. Soc. A*, **61**, 381–395.

Hall, P. (1982). On some simple estimates of an exponent of regular variation, *J. R. Statist. Soc. B*, **44**, 37–42.

Hall, P. (1990). Using the bootstrap to estimate mean-square error and selecting smoothing parameter in non-parametric problems, *J. Mult. Anal.*, **32**, 177–203.

Hall, P. & Welsh, A.H. (1985). Adaptative estimates of parameters of regular variation, *Ann. Statist.*, **13**, 331–341.

Henriques-Rodrigues, L. & Gomes, M.I. (2009). High quantile estimation and the PORT methodology, *Revstat*, **7**, 245–264.

Henriques-Rodrigues, L.; Gomes, M.I. & Pestana, D. (2011). Statistics of extremes in athletics, *Revstat*, **9**, 127-153.

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

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

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

Hosking, J.R.M.; Wallis, J.R. & Wood, E.F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, *Technometrics*, **27**, 251–261.

Hüsler, J. & Li, D. (2006). On testing extreme value conditions, *Extremes*, **9**, 69–86.

Hüsler, J. & Müller, S. (2005). Iterative estimation of the extreme value index, *Math. Comput. Appl. Probab.*, **7**, 139–148.

Hüsler, J. & Peng, L. (2008). Review of testing issues in extremes: in honor of Professor Laurens de Haan, *Extremes*, **11**, 99–111.

Jenkinson, A.F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements, *Quart. J. R. Meteorol. Soc.*, **81**, 158–171.

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

Jurečková, J. & Picek, J. (2001). A class of tests on the tail index, *Extremes*, **4**, 165–183.

Kaufmann, E. (2000). Penultimate approximations in extreme value theory, *Extremes*, **3**, 39–55.

Koning, A.J. & Peng, L. (2008). Goodness-of-fit tests for a heavy tailed distribution, *J. Statist. Plann. Infer.*, **138**, 3960–3981.

Kratz, M. & Resnick, S. (1996). The qq-estimator of the index of regular variation, *Stochastic Models*, **12**, 699–724.

Li, D. & Peng, L. (2009). Does bias reduction with external estimator of second order parameter work for endpoint? *J. Statist. Plann. Infer.*, **139**, 1937–1952.

Li, D.; Peng, L. & Yang, J. (2010). Bias reduction for high quantiles, *J. Statist. Plann. Infer.*, **140**, 2433–2441.

Li, D.; Peng, L. & Qi, Y. (2011). Empirical likelihood confidence intervals for the endpoint of a distribution function, *Test*, **20**, 353–366.

Ling, C.; Peng, Z. & Nadarajah, S. (2011). Location invariant Weiss–Hill estimator, *Extremes*, DOI 10.1007/s10687-011-0134-x.

Luceño, A. (2006). Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators, *Comput. Statist. Data Anal.*, **51**, 904–917.

Marohn, F. (2000). Testing extreme value models, *Extremes*, **3**, 363–384.

Matthys, G. & Beirlant, J. (2003). Estimating the extreme value index and high quantiles with exponential regression models, *Statist. Sinica*, **13**, 853–880.

Matthys, G.; Delafosse, M.; Guillou, A. & Beirlant, J. (2004). Estimating catastrophic quantile levels for heavy-tailed distributions, *Insur. Math. Econ.*, **34**, 517–537.

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.

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

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

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

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

Peng, L. (1998). Asymptotically unbiased estimator for the extreme-value index, *Statist. Prob. Lett.*, **38**(2), 107–115.

Peng, L. (2001). Estimating the mean of a heavy-tailed distribution, *Statist. Prob. Lett.*, **52**, 255–264.

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

Raoult, J.P. & Worms, R. (2003). Rate of convergence for the generalized Pareto approximation of the excesses, *Adv. Appl. Probab.*, **35**, 1007–1027.

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

Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer, New York.

Schultze, J. & Steinbach, J. (1996). On least squares estimators of an exponential tail coefficient, *Statist. Decis.*, **14**, 353–372.

Segers, J. (2005). Generalized Pickands estimators for the extreme value index, *J. Statist. Plann. Infer.*, **28**, 381–396.

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

Stephenson, A. & Tawn, J. (2004). Bayesian inference for extremes: accounting for the three extremal types, *Extremes*, **7**, 291–307.

Temido, M.G. & Canto e Castro, L. (2003). Max-semistable laws in extremes of stationary random sequences, *Th. Prob. Appl.*, **47**, 365–374.

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

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

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

Zhou, C. (2009). Existence and consistency of the maximum likelihood estimator for the extreme value index, *J. Mult. Anal.*, **100**, 794–815.

Zhou, C. (2010). The extent of the maximum likelihood estimator for the extreme value index, *J. Mult. Anal.*, **101**, 971–983.

You, A.; Schneider, U.; Guillou, A. & Naveau, P. (2010). Improving extreme quantile estimation by folding observations, *J. Statist. Plann. Infer.*, **140**, 1775–1787.

Yun, S. (2002). On a generalized Pickands estimator of the extreme value index, *J. Statist. Plann. Infer.*, **102**, 389–409.