Statistics of Extremes for IID Data: Laurens de Haan leading contributions*

M. IVETTE GOMES LUÍSA CANTO E CASTRO
M. ISABEL FRAGA ALVES
DINIS PESTANA

CEAUL, DEIO, Faculty of Science, University of Lisbon, Portugal

To Laurens de Haan, a token of friendship.

Abstract. In the last decades there has been a shift from the parametric statistics of extremes for IID random variables, based on the probabilistic asymptotic results in extreme value theory, towards a semi-parametric approach, where the estimation of the right tail-weight, under a quite general framework, is of major importance. After a brief presentation of classical Gumbel's block methodology and of later improvements in the parametric framework (multivariate and multi-dimensional extreme value models for largest observations and peaks over threshold approaches), we present a coordinated overview, over the last three decades, of the developments on the estimation of the extreme value index and testing of extreme value conditions under a semiparametric framework. Laurens de Haan has been one of the leading scientists in the field, (co-)author of many seminal ideas, that he generously shared with dozens (literally) of colleagues and students, thus achieving one of the main goals in a scientist's life: he gathered around him a bunch of colleagues united in the endeavour of building knowledge. The last section is a personal tribute to Laurens, who fully lives his ideal that "*co-operation is the heart of Science*".

AMS 2000 subject classification. 62G32, 62E20.
Keywords and phrases. *Extreme value index; parametric and semi-parametric estimation and testing; statistics of extremes.*

*Research partially supported by FCT / POCTI and POCI / FEDER.

Contents

1 Introduction

Statistics of extremes, either univariate, multivariate, multi-dimensional or infinite dimensional, helps us "to learn from almost disastrous events", a quotation of the title of the seminal Gulbenkian lecture of Laurens de Haan (de Haan, 2006). *The domains of application of statistics of extremes are thus quite diversified. We mention the fields of hydrology, meteorology, geology, insurance, finance, structural engineering, telecommunications and biostatistics, among others* (*see, for instance,* Reiss and Thomas, 2001*;* Beirlant, Goegebeur, Segers and Teugels, 2004, Section 1.3*;* Castillo, Hadi, Balakrishnan and Sarabia, 2005). 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).

It is common knowledge that *statistics of extremes* had initially a parametric nature only, working with models associated to limiting results in the field of extreme value theory (EVT). In the late seventies there was an inflection towards semi-parametric models and the estimation of parameters of extreme events has been essentially developed under such a framework. Laurens de Haan is indeed one of the main responsible for the harmonious development of this area. Following pioneering work of Bruce Hill (Hill, 1975) and James Pickands (Pickands, 1975), among others, the establishment of a rigourous limiting behaviour of semi-parametric estimators of parameters of extreme events, as well as the introduction of the moment estimator, a general estimator of the extreme value index (EVI) $\gamma \in \mathbb{R}$, which is a landmark in the field of statistics of extremes, are due to Laurens de Haan and some of his co-workers (Dekkers, Einmahl and de Haan, 1989).

Another landmark in the field of EVT is the result about the limiting Generalized Pareto (GP) behaviour of the scaled excesses (Balkema and de Haan, 1974; Pickands, 1975), which enabled the development of the so-called "maximum likelihood" (ML) EVI estimators. We here refer the *peaks over threshold* (POT) methodology of estimation (Smith, 1987) as well as the methodology used by Drees, Ferreira and de Haan (2004), named PORT (of *peaks over random threshold*) in Araújo Santos, Fraga Alves and Gomes (2006).

Recently, and for heavy tails, the accommodation of bias in the log-excesses led Gomes, de Haan and Henriques Rodrigues (2004) to investigate a second-order reduced-bias estimator, a weighted combination of the log-excesses denoted *weighted Hill* (WH) estimator. It is a minimum-variance reduced-bias (MVRB) estimator, in the sense that it has an asymptotic variance equal to the one of the Hill estimator and an asymptotic bias of smaller order, and therefore seems to open interesting new perspectives in the field.

Even more recently, the revisitation of Laurens de Haan Ph.D. thesis (de Haan, 1970), still a leading reference in most of the papers in the area of EVT, led to a new estimator, the so-called *mixed moment* (MM) estimator (Fraga Alves, Gomes, de Haan and Neves, 2007b), involving not only the log-excesses but also another type of moment-statistics. The MM estimator is an interesting alternative to the most popular EVI estimators for a general $\gamma \in \mathbb{R}$, broadening the scope of research in the field of semi-parametric estimation of EVT parameters.

This review paper, honoring Laurens de Haan for his many breakthroughs in EVT, focuses on statistical issues arising in modeling univariate extremes of a random sample, i.e., of a vector of n independent, identically distributed (IID) random variables from an underlying distribution function F. As usual, we shall denote $(X_1, \cdots, X_n)$ the (at least partially) available random sample, and $(X_{1,n} \leq \cdots \leq X_{n,n})$ the sample of associated ascending order statistics (OSs).

In Section 2 of this paper, we shall briefly review the main limiting results in EVT, important for the topics to be dealt with in this overview paper. Next, in Section 3, we deal also briefly with the most relevant parametric models in the field of extremes. Section 4 is devoted to semi-parametric inference, giving particular emphasis to the two most recent above mentioned estimation methods due to Laurens de Haan and some of his co-workers.

Finally, in Section 5, of a more personal nature, we try to express that we admire Laurens de Haan for his many achievements, and that we love him for his kindness, friendliness, modesty and generousity. In fact, Laurens has always shared his wealth of ideas with colleagues and students, and in his speech at the University of Lisbon, when he has been conferred an *Honoris Causa* doctorate, he said:

> "*I have written papers with 43 people.* [...] *Why do I mention it? Well, I am not only a statistician, but also a mathematician so I want to formulate a theorem. The Theorem is as follows: Co-operation is the heart of Science. I am tempted to add a corollary to this theorem, namely: this honorary doctorate has been earned mainly by my co-authors.*"

In our view, Laurens capacity to bring out the best in others is another major achievement in his highly successful research career, and we cannot but express how lucky we feel for the benefit of having had Laurens among us, at our research unit in Lisbon University, for most of the time in the last decade.

2 Main limiting results in extreme value theory

The main limiting results in EVT date back to the papers by Fréchet (1927), Fisher and Tippett (1928), von Mises (1936) and Gnedenko (1943), who fully characterizes the possible non-degenerate limit laws of the sequence of maximum values $X_{n,n}$, suitably normalized, as $n \to \infty$, and partially characterizes the domains of attraction of those possible limiting laws. The full characterization of domains of attraction is due to Laurens de Haan (de Haan, 1970), who gives the most complete and rigourous discussion of the problem, and can also be found in Galambos (1978; 1987), Falk, Hüsler and Reiss (1994; 2004) and de Haan and Ferreira (2006).

2.1 Max-stable laws: the univariare extreme value model

All possible non-degenerate weak limit distributions of the normalized partial maxima $X_{n,n}$, of i.i.d. random variables $X_1, \cdots, X_n$, are (generalized) *extreme value* (EV) distributions, i.e., if there are normalizing constants $a_n > 0$, $b_n \in \mathbb{R}$ and some non-degenerate distribution function G such that, for all x,

$$\lim_{n\to\infty} P\left\{\frac{X_{n,n}-b_n}{a_n} \le x\right\} = G(x), \tag{2.1}$$

we can redefine the constants in such a way that the limit $G(x)$ is, for some $\gamma \in \mathbb{R}$,

$$G(x) \equiv G_\gamma(x) := \begin{cases} \exp\left(-(1+\gamma x)^{-1/\gamma}\right), \ 1+\gamma x > 0 & \text{if} \quad \gamma \neq 0 \\ \exp(-\exp(-x)), \ x \in \mathbb{R} & \text{if} \quad \gamma = 0, \end{cases} \tag{2.2}$$

the EV distribution (EVD), given here in the von Mises-Jenkinson form (von Mises, 1936; Jenkinson, 1955). The real parameter γ is called the *extreme value index* and it is the primary parameter of interest in the whole of extreme value analysis. We say that the distribution function F underlying the random variables $X_1, X_2, \cdots$ is in the *domain of attraction* of G_γ in (2.2) if (2.1) holds with $G = G_\gamma$, and use the notation $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$. The limiting distribution functions G in (2.1) are then *max-stable*, i.e., the functional equation $G^n(\alpha_n x + \beta_n) = G(x)$, $n \ge 1$, holds for some $\alpha_n > 0$, $\beta_n \in \mathbb{R}$.

The following *extended regular variation* property (de Haan, 1984) is a well-known necessary and sufficient condition for $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$:

$$F \in \mathcal{D}_{\mathcal{M}}(G_\gamma) \iff \lim_{t\to\infty} \frac{U(tx)-U(t)}{a(t)} = \begin{cases} \frac{x^\gamma - 1}{\gamma} & \text{if} \quad \gamma \neq 0 \\ \ln x & \text{if} \quad \gamma = 0, \end{cases} \tag{2.3}$$

for every $x > 0$ and some positive measurable function a, with U standing for a quantile type function associated to F and defined by

$$U(t) := (1/(1-F))^{\leftarrow}(t) = F^{\leftarrow}(1-1/t) = \inf\left\{x : F(x) \ge 1 - \tfrac{1}{t}\right\}.$$

Heavy-tailed models, i.e., models F in $\mathcal{D}_{\mathcal{M}}(G_\gamma)$ with $\gamma > 0$, are quite important in areas like biostatistics, finance and insurance. For heavy tails, we may choose $a(t) = \gamma\, U(t)$ in (2.3), and we can say that a distribution function F is in the max-domain of attraction of G_γ, $\gamma > 0$, 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 $U \in RV_\gamma$, or equivalently, the tail function $\overline{F} := 1 - F \in RV_{-1/\gamma}$. For full details on regular variation see Bingham, Goldie and Teugels (1987).

Other limiting results, parallel to the one for the maximum $X_{n,n}$, have been obtained for the k-th top OS, denoted $X_{n-k+1,n}$, $k \geq 1$, with k fixed. Results in this direction appeared in the pioneering papers by Gumbel (1935) and Smirnov (1949; 1952), and surveys can be found in Galambos (1978; 1987), David (1970; 1981) and David and Nagaraja (2003). For a fixed k, if (2.1) holds, then, with $G_\gamma(x) \equiv G_{\gamma,1}(x)$ the EVD in (2.2),

$$\lim_{n\to\infty} P\left\{\frac{X_{n-k+1,n} - b_n}{a_n} \leq x\right\} = G_{\gamma,k}(x) = G_\gamma(x) \sum_{j=0}^{k-1} \frac{(-\ln G_\gamma(x))^j}{j!}. \tag{2.4}$$

If we allow $k \to \infty$, as the sample size $n \to \infty$, we no longer have the limiting extreme value distribution function in (2.4). We have instead an asymptotic normal behaviour, but these type of OSs, even if *central order statistics*, still have a strong connection with EVT. Let F be a distribution function, with second derivative and probability density function $f(x) = F'(x) > 0$ in a neighbourhood of the right endpoint $x^F = U(\infty)$. Let us further assume the validity of von Mises condition (von Mises, 1936),

$$\lim_{x\to x^F} ((1 - F(x))/f(x))' = \gamma, \tag{2.5}$$

and consider an *intermediate order statistic*, i.e., an OS $X_{n-k+1,n}$ such that

$$k = k_n \to \infty \quad \text{and} \quad k = o(n),\ \text{i.e.,}\ k/n \to 0,\ \text{as}\ n \to \infty. \tag{2.6}$$

Then (Smirnov, 1949, 1967; Falk, 1989)

$$\sqrt{k}\left(\frac{X_{n-k+1,n} - U(n/k)}{U'(n/k)}\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \text{Normal}(0,1). \tag{2.7}$$

This asymptotic normality can be obtained under a more general framework. For a detailed study, see Sections 2.2 and 2.4 of de Haan and Ferreira (2006); in particular, it is therein provided, in Lemma 2.2.3, a new elegant and simple proof of the result of Smirnov on asymptotic normality of uniform OSs. Details on *central order statistics* $X_{n-k+1,n}$, with $k/n \to p \in (0,1)$, as $n \to \infty$, can be seen in, e.g., Arnold, Balakrishnan and Nagaraja (1992).

2.2 The multivariate EVD

Under the same initial assumption (2.1), with $k \in \mathbb{N}$, fixed, and with $E_i,\ i \geq 1$, IID standard exponential random variables,

$$\left(\frac{X_{n,n} - b_n}{a_n}, \frac{X_{n-1,n} - b_n}{a_n}, \cdots, \frac{X_{n-k+1,n} - b_n}{a_n}\right)$$

converges in distribution to the random vector

$$\left(\frac{E_1^{-\gamma} - 1}{\gamma}, \frac{(E_1 + E_2)^{-\gamma} - 1}{\gamma}, \cdots, \frac{(E_1 + E_2 + \cdots + E_k)^{-\gamma} - 1}{\gamma}\right). \tag{2.8}$$

Let $H_\gamma(x_1, \cdots, x_k)$ be the k-variate distribution function of the random vector in (2.8). Then, denoting $g_\gamma(x) = \partial G_\gamma(x)/\partial x$ and $h_\gamma(x_1, \cdots, x_k) = \partial^k H_\gamma(x_1, \cdots, x_k)/(\partial x_1 \cdots \partial x_k)$, we have

$$h_\gamma(x_1, x_2, \cdots, x_k) = g_\gamma(x_k) \prod_{j=1}^{k-1} \frac{g_\gamma(x_j)}{G_\gamma(x_j)} \qquad \text{if} \qquad x_1 > x_2 > \cdots > x_k, \tag{2.9}$$

the so-called *multivariate* EV *model* or *extremal process* (Dwass, 1964). This result has been obtained by Lamperti (1964) for $k = 2$, and by Weissman (1975) in general, together with distributional representations of the extremal process.

2.3 Excesses over a high level

Another seminal result in the field of EVT is the one due to Balkema and de Haan (1974) and Pickands (1975): independently, they proved that, under adequate conditions, the generalized Pareto distribution (GPD),

$$P_\gamma(x) = \begin{cases} 1 - (1 + \gamma x)^{-1/\gamma},\ 1 + \gamma x > 0,\ x \geq 0 & \text{if} \quad \gamma \neq 0 \\ 1 - \exp(-x),\ x \geq 0 & \text{if} \quad \gamma = 0, \end{cases} \tag{2.10}$$

is the limit distribution of scaled excesses over high thresholds. More precisely, consider the *excess function*, $F_u(x) := P\left[X - u \leq x | X > u\right]$. Denoting $x^F := U(\infty)$, the right endpoint of F, $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$ if and only if there exists a positive real function $\sigma(u)$ such that

$$\lim_{u \to x^F} \left| F_u\big(\sigma(u) x\big) - P_\gamma(x) \right| = 0$$

(see, for instance, Embrechts, Klüppelberg and Mikosch, 1997, Section 3.4, and Reiss and Thomas (2007), Section 1.4, for more details).

2.4 Max-semistable laws

The class of *max-semistable* laws has been introduced by Grienvich (1992a, 1992b) and Pancheva (1992) and is more general than the class of *max-stable* laws (2.2), given in subsection 2.1. A distribution function F is in the domain of attraction of a max-semistable law, a fact that we denote by $F \in \mathcal{D}_{\mathcal{MSS}}(G_{\gamma,\nu})$, if there exist a sequence m_n, with

$$m_{n+1} \geq m_n > 0, \qquad \lim_{n\to\infty} \frac{m_{n+1}}{m_n} = r \geq 1$$

real numbers $a_n > 0$ and $b_n \in \mathbb{R}$ and a non-degenerate distribution function $G_{\gamma,\nu}(x)$ such that, for all continuity points of $G_{\gamma,\nu}(x)$,

$$\lim_{n\to\infty} P\left\{\frac{X_{m_n,m_n} - b_n}{a_n}\right\} = G_{\gamma,\nu}(x). \tag{2.11}$$

The possible max-semistable laws can be written as

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

where $\nu(\cdot)$ is a positive, limited and periodic function, with period p, related with γ and r through the equation

$$p = \begin{cases} \gamma \ln r & \text{if} \quad \gamma \neq 0 \\ \ln r & \text{if} \quad \gamma = 0. \end{cases}$$

If $m_n = n$, we get $r = 1$ and consequently $p = 0$. A unit function $\nu(\cdot)$ enables us to get the max-stable laws in (2.2). Max-semistable laws have an underlying repetitive pattern which is well captured by the non decreasing right continuous function $y(\cdot)$ in the following representation (Canto e Castro, de Haan and Temido, 2000):

$$-\ln(-\ln G_{\gamma,\nu}(r^{m\gamma}x + s_m)) = y(x) + m\ln r\,, \quad x \in [0, b], \tag{2.12}$$

where $s_m = b\,(1 + r^{\gamma} + r^{2\gamma} + \cdots + r^{(m-1)\gamma})$ and b is such that $-\ln(-\ln G_{\gamma,\nu}(b)) + \ln(-\ln G_{\gamma,\nu}(0)) = \ln r$. Discrete models like the geometric and the negative binomial, as well as some multimodal continuous models, are in $\mathcal{D}_{\mathcal{MSS}}$ but not in $\mathcal{D}_{\mathcal{M}}$.

2.5 Rates of convergence and penultimate approximations

Another important problem in EVT concerns the rate of convergence of $F^n(a_n x + b_n)$ towards $G(x)$ in (2.1) or, equivalently, the finding of estimates of the difference $d_n(F, G, x) := F^n(a_n x + b_n) - G(x)$. Note that, as detailed later on, in Section 3, 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 $G_\gamma(x)$ in (2.2), replacing $F^n(x)$ by $G_\gamma((x-b_n)/a_n)$, b_n and $a_n > 0$ being unknown parameters to be estimated from an adequate sample. The rate of convergence is thus important because it may validate or not the most usual models in *statistics of extremes*.

Note that in EVT there exists no analogue of the Berry-Esséen theorem that, under broad conditions, gives a rate of convergence of the order of $1/\sqrt{n}$ in the Central Limit Theorem (CLT). The rate of convergence depends here strongly on the right tail of F and on the choice of the attraction coefficients. Moreover, the rate of convergence of $F^n(a_n x+b_n)$ towards the ultimate limiting max-stable distribution function can be rather slow. Fisher and Tippett (1928) observed that, despite of the fact that the normal distribution function $\Phi \in \mathcal{D}_{\mathcal{M}}(G_0)$, the convergence of $\Phi^n(a_n x + b_n)$ towards $G_0(x)$ is extremely slow. They then concluded their paper by showing, 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.

The modern theory of rates of convergence in EVT began with Anderson (1971). Developments have followed several directions. For papers on the subject prior to 1992 see 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 von Mises-type conditions and some additional 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.

3 Statistics of univariate extremes: parametric approaches

Although the EVD constitutes a unified version of all possible non-degenerate weak limits of maxima of sufficiently long sequences of IID random variables, as stated in Section 2.1, the EVD reduces to the Fréchet ($\gamma > 0$), the Weibull ($\gamma < 0$) and the Gumbel ($\gamma = 0$) distribution functions, respectively.

The *extreme value index* γ rules the behaviour of the right tail of F. The Fréchet domain of attraction ($\gamma > 0$) contains heavy-tailed distribution functions with polynomially decaying tail, like the Pareto, the Cauchy, and the Student's-t distributions. Short-tailed distribution functions, with finite right endpoint (uniform and beta in general, e.g.), belong to the Weibull domain of attraction for maxima ($\gamma < 0$). The intermediate case, i.e., the Gumbel domain of attraction ($\gamma = 0$) is relevant for many applied sciences where extremes are important, and contains a great variety of distribution functions with an exponential tail, like the normal, the exponential and the gamma, but not necessarily with an infinite right endpoint. So, deciding the right tail-weight for the distribution underlying the sample data, through a proper estimation of γ, constitutes a very important starting task in statistical inference for extreme values. 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 the corresponding alternative approaches to statistics of extremes.

3.1 Gumbel's approach or block maxima method

When the sample size $n \to \infty$, and due to the limiting result given before for the normalized sequence of maximum values, often called the *extremal types theorem*, we may write

$$P[X_{n,n} \leq x] = F^n(x) \approx G_\gamma\left((x - \lambda_n)/\delta_n\right), \tag{3.1}$$

with $G_\gamma(x)$ given in (2.2) and $(\lambda_n, \delta_n) \in (\mathbb{R}, \mathbb{R}^+)$ an unknown vector of location and scale parameters, that replaces the attraction coefficients (b_n, a_n) in (2.1). The EVD in (2.2) is sometimes separated in the three following types:

$$\begin{array}{lll} \text{Type I (Gumbel)}: & \Lambda(x) = \exp(-\exp(-x)),\ x \in \mathbb{R} & (\gamma = 0) \\ \text{Type II (Fréchet)}: & \Phi_\alpha(x) = \exp(-x^{-\alpha}),\ x \geq 0 & (\gamma = \frac{1}{\alpha} > 0) \\ \text{Type III (max-Weibull)}: & \Psi_\alpha(x) = \exp(-(-x)^\alpha),\ x \leq 0 & (\gamma = -\frac{1}{\alpha} < 0). \end{array} \tag{3.2}$$

We have $\Lambda(x) = G_0(x)$, $\Phi_\alpha(x) = G_{1/\alpha}(\alpha(1-x))$ $(\gamma = 1/\alpha > 0)$ and $\Psi_\alpha(x) = G_{-1/\alpha}(\alpha(x+1))$ $(\gamma = -1/\alpha < 0)$, with G_γ the EVD in (2.2).

Remark 3.1. *Any result for maximum values has its counterpart for minimum values. Indeed,* $\min_{1\leq i\leq n} X_i = -\max_{1\leq i\leq n}(-X_i)$, *and consequently, when* $n \to \infty$, $P[X_{1,n} \leq x] = 1 - (1 - F(x))^n \approx G^*_\xi\left((x - \lambda')/\delta'\right)$, *with* $G^*_\xi(x) = 1 - G_\xi(-x) = 1 - \exp\left\{-(1 - \xi x)^{-1/\xi}\right\}$, *for* $1 - \xi x > 0$. *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. All results for maxima* (*top OSs*) *can be easily reformulated for minima* (*low OSs*).

The above mentioned *extremal types theorem* was used by Gumbel in several papers which culminated in his 1958 book, to give approximations of the type of the one provided in (3.1) but for any of the models in (3.2). He suggested the first model in *statistics of extremes*, usually called the *annual maxima* (AM) or *block maxima* (BM) *model* or the EV *univariate model* or merely *Gumbel's model*. Under *Gumbel's model*, the sample of size n is divided into k sub-samples of size r (usually associated to k years, with $n = r \times k$, and 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 (3.2), obviously with extra unknown location and scale parameters, is fitted to the sample of those k maximum values. Nowadays, whenever using this approach, still quite popular in environmental sciences, it is more common to fit to the data an extreme value distribution function $G_\gamma((x-\lambda_r)/\delta_r)$, with G_γ given in (2.2), $(\lambda_r, \delta_r, \gamma) \in (\mathbb{R}, \mathbb{R}^+, \mathbb{R})$ unknown location, scale and "shape" parameters. All statistical inference is then related to the above mentioned models.

Computational details on maximum likelihood estimation of $(\lambda, \delta, \gamma)$ *in the extreme value model* $G_\gamma((x-\lambda)/\delta)$ *can be found in* Prescott and Walden (1980, 1983)*,* Hosking (1985)*,* Smith (1985) *and* Macleod (1989)*. Probability weighted moments* (PWM) *parameter estimation for the* EVD *has been extensively studied in* Hosking, Wallis and Wood (1985)*. Several other estimation methods for the* EVD *can be found in the literature. Among them we mention: best linear unbiased estimation* (BLUE) (Balakrishnan and Chan, 1992)*; method of moments* (Christopeit, 1994)*; minimum distance estimation* (Dietrich and Hüsler, 1996)*. Robust methods for the* EVD *have been studied in* Dupuis and Field (1998)*, who derived B-optimal robust M-estimators for the case that the observations follow an* EVD*. Modifications of the* ML *estimator are presented by* Coles and Dixon (1999)*, who suggest penalised maximum likelihood* (PML) *estimators, showing that* PML *estimation improves the small-sample properties of a likelihood-based analysis.*

3.2 Multivariate and multi-dimensional EV approaches: the method of largest observations

Although Gumbel's statistical procedure 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 OSs, if available, because they certainly contain useful information about the right tail of the distribution function underlying the data. On the other hand, in the most diversified areas of application there is no natural seasonality of the data, and in such framework the method of sub-samples is subjective and artificial.

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. Indeed, if we have daily data, some years may have several values among those top OSs (that are for sure relevant to make inference upon the right tail of F), whereas other years may contain none of those top values. We may thus say that such an approach would provide additional information, that has been disregarded in the traditional Gumbel's methodology.

This approach depends on the joint limiting distributional behaviour of those top OSs.

When the sample size n is large and for fixed k, it is sensible to consider, on the basis of (2.9), the approximation,

$$P[X_{n,n} \le x_1, \cdots, X_{n-k+1,n} \le x_k] \approx H_\gamma\Big(\frac{x_1-\lambda_n}{\delta_n}, \cdots, \frac{x_k-\lambda_n}{\delta_n}\Big), \tag{3.3}$$

where H_γ is the *multivariate* EVD in subsection 2.2, with λ_n and δ_n unknown location and scale parameters, respectively, to be estimated on the basis of the k top OSs in the sample of size n. This approach to *statistics of extremes* lies then on the multivariate structure in (2.9), and it is the so-called *multivariate* EV *model* or *extremal process*. Under this approach it is easier to increase the number k of observations, contrarily to what happens in Gumbel's approach, where a larger number n of original observations is usually needed. Such an approach has been introduced first, in a slightly different context, by Pickands (1975) and used by Weissman (1978) and Gomes (1978, 1981). See also Weissman (1984), Smith (1986), Tawn (1988) and references therein.

Note finally that it is easy to combine both approaches. 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,

$$(\underline{X}_1, \underline{X}_2, \cdots, \underline{X}_k), \quad \text{where } \underline{X}_j = (X_{1j}, \cdots, X_{i_j j}),\ 1 \le j \le k,$$

are *multivariate* EV *vectors. Maximum likelihood estimators of the unknown parameters in this model have been studied in* Gomes (1981). *See also* Smith (1984) *and references therein. The use of concomitants of OSs, to deal with statistical inference techniques in this model, appears in* Gomes (1984, 1985a).

A comparison of the *multivariate* EV *model* and the *multi-dimensional* EV *model* is performed in Gomes (1985b, 1989a). Discrimination among *multivariate* EV *models* can also be found in Fraga Alves (1992).

3.3 The peaks over threshold approach

Another approach to *statistics of extremes*, in a certain sense parallel to the *multivariate* EV *model*, is the one at which we restrict our attention only to the observations that exceed a certain high *threshold* u, fitting the appropriate statistical model to the *excesses* over a high level u. From the results in subsection 2.3, we get the approximation $P\left[X-u \le x | X > u\right] \approx P_\gamma(x/\sigma)$, with $P_\gamma(x)$ given in (2.10). We are then led to consider a deterministic high level u and work with the excesses. The adequate model is then the *generalized Pareto* (GP).

Such a model is the so-called *Paretian excesses model* or POT *model*, with POT standing for *peaks over thresholds*, and was introduced in Smith (1987). Here, all statistical inference is related to the GPD.

ML *estimates of* γ *and* σ *in a* GPD*, non-regular for* $\gamma < -1/2$*, have been studied in* Smith (1987). A survey of the POT methodology, together with several applications can be found in Davison and Smith (1990). *As the* ML *estimators can be numerically hardly tractable* (see Grimshaw, 1993)*, there has been several methods, aside* ML*, proposed for the estimation of* γ *and* σ*.* Hosking and Wallis (1987) *suggest the use of* PWM *estimators* (*see also the comparative study of* Singh and Guo, 1997). Castillo and Hadi (1997) *have proposed estimators based on the elemental percentile method* (EPM). PML *estimators, containing a penalty function for the shape parameter, are presented by* Coles and Dixon (1999) *and* Martins and Stedinger (2000). *The* PML *estimator combines the flexibility of the* ML *estimator and the robustness of the* PWM *estimator. See also* Resnick (1997)*,* Crovella and Taqqu (1999) *and references therein. The estimates depend significantly on the choice of the threshold and several authors, among whom we mention* McNeill (1997) *and* Rootzén and Tajvidi (1997)*, explicitly state that the selection of an appropriate threshold u, above which the* GPD *assumption is appropriate, is a difficult task in practice. Robust estimation for the* GPD *was first addressed by* Dupuis (1998)*, who provides the optimally-biased robust estimator for the* GPD *and suggests a validation mechanism to guide the threshold selection.* Peng and Welsh (2001) *use the method of medians introduced by* He and Fung, 1999*, and get estimators of the unknow parameters of the* GPD *with bounded influence functions.* Juárez and Schucany (2004) *implement the minimum density power divergence estimator* (MDPDE) *for the shape and scale parameters of the* GPD*. The* MDPDE *is indexed by a non-negative constant that controls the trade-off between robustness and efficiency.* Frigessi, Haug and Rue (2002) suggest an unsupervised alternative to the classical POT model, where a GPD is fitted beyond a threshold which is selected in a supervised way. They suggest to model the data with a dynamical mixture: one term of the mixture is a GPD and the other is a light-tailed density distribution. The weight of the GPD component is predominant for large values, and takes the role of threshold selection. A recent comparison between Gumbel's approach of block maxima, i.e., the BM approach (with several block-sizes, not only annual maximum values) and the POT approach has been performed by Engeland, Hisdal and Frigessi (2004).

3.4 Bayesian approaches

It has recently become more and more common the use of Bayesian methodology within extreme value analysis. We shall give some examples of its most direct use: Smith and Naylor (1987) *compare Bayesian and* ML *estimators for the Weibull distribution;* Ashour and El-Adl (1980)*,* Lingappaiah

(1984), Achcar, Bolfarine and Pericchi (1987) *and* Engeland and Rackwitz (1992) *consider estimation for specific extremal types for maxima;* Lye, Hapuarachchi and Ryan (1993) *consider estimation of the* EVD*;* Pickands (1994) *and* Bermudez and Amaral-Turkman (2003) *consider estimation of the* GPD. Smith (1999) *discusses predictive inference aspects of Bayesian and frequentist approaches, and* Bermudez, Amaral-Turkman and Turkman (2001) *propose the use of a Bayesian predictive approach for the choice of the threshold, through a hierarchical Bayesian model.* Stephenson and Tawn (2004) *claim that in practice the appropriate asymptotically motivated extremal model, either* EVD *or* GPD, *fitted to the data that can be regarded as maxima or exceedances of a high threshold, reduces the Gumbel* (*exponential*) *type to a single point in the parameter space, and consequently the Gumbel* (*exponential*) *model is never selected. They then decide to incorporate knowledge of the structure of the extremal types theorem into inference for the* EVD *and the* GPD. *To do this, they associate the probability* p_γ *to the paramater subspace corresponding to the Gumbel* (*exponential*) *type. This approach requires an inference scheme which allows switching between the full* EVD (GPD) *and the Gumbel* (*exponential*) *sub-model. They then perform inference using reversible jump* MCMC *techniques. This Bayesian approach recognises the possibility that the data can come from any of the three extremal types. As a by-product of the analysis, posterior probabilities for* $P(\gamma > 0)$, $P(\gamma = 0)$ *and* $P(\gamma < 0)$ *are obtained.* Diebolt, El-Aroui, Garrido and Girard (2005) *propose 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. For other papers on Bayesian approaches to, for instance, high quantile estimation, see, e.g.,* Coles and Powell (1996) *and* Coles and Tawn (1996) , *who provide a detailed review of Bayesian methods in extreme value modelling up to this date. See also* Reiss and Thomas (1999), Walshaw (2000), Smith and Goodman (2000), Bottolo, Consonni, Dellaportas and Lijoi (2003), *and the monographs by* Reiss and Thomas (2001; 2007) *and* Coles (2001), *together with the references therein.*

3.5 Statistical choice of extreme value models

The Gumbel type d.f. $\Lambda = G_0$ *or the exponential type d.f.* P_0, *with* G_γ *and* P_γ *given in* (2.2) *and* (2.10), *respectively, are favorites in statistics of extremes, essentially because of the simplicity of inference associated to these populations. Additionally,* $\gamma = 0$ *can be regarded as a change-point, because for* $\gamma < 0$ *the data come from a d.f. with a finite right endpoint and for* $\gamma > 0$ *the right endpoint is infinite. Thus, any separation between extreme value models, with* Λ *playing a central and proeminent position* (*called "trillemma" by Tiago de Oliveira*), *turns out to be an important statistical problem, that has been recently considered under a semi-parametric framework. From a parametric point of view, empirical tests of the hypothesis* $H_0 : \gamma = 0$ *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:*

1. *Quick tests, suggested by heuristic reasons* (van Montfort, 1970, 1973; Bardsley, 1977; van

Montfort and Ottten, 1978; Otten and van Montfort, 1978; Galambos, 1982; Gomes, 1982, 1984; Tiago de Oliveira and Gomes, 1984; van Montfort and Gomes, 1985; van Monfort and Witter, 1985; Gomes and van Montfort, 1987; Brilhante, 2004).

2. *Modified locally most powerful tests* (MLMP) (Tiago de Oliveira, 1981, 1984; Tiago de Oliveira and Gomes (1984); Gomes and van Montfort, 1987).

3. *Locally asymptotically normal* (LAN) *tests* (Falk, 1995b,c; Marohn, 1994, 1998a,b, 2000).

4. *Goodness-of-fit tests for the* Gumbel *model* (Stephens, 1976, 1977, 1986; Kinnison, 1989). *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), *among others. Further non-parametric tests appear in* Jurečková and Picek (2001).

5. *Tests from large sample theory, like the likelihood ratio test and Wald test, among others* (Hosking, 1984; Gomes, 1989b).

Statistical choice in the multivariate EV model is developed in Gomes (1984), Gomes and Alpuim (1986), Gomes (1987, 1989b), Hasofer and Wang (1992), Wang (1995), Fraga Alves and Gomes (1996) *and* Wang, Cooke and Li (1996). *As we shall refer in* Section 4, *some of these authors already go beyond the extremal process, working under a semi-parametric framework.*

3.6 Summary of parametric approaches and a link to semi-parametric frameworks

More recently, the *method of largest observations*, the POT methodology and testing procedures have been considered under a semi-parametric framework. There is then no fitting of a specific parametric model, dependent upon a location parameter λ, a scale parameter δ and a shape parameter γ. It is merely assumed that F is in the domain of attraction for maxima of G_γ in (2.2), 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, to be dealt with in Section 4. We now summarise the different approaches to *statistics of univariate extremes* here discussed:

1. **Parametric approaches**:

   I The *univariate* EV *model* (for the k maximum values of sub-samples of size r, $n = r \times k$.) (Gumbel's classical approach or Block Maxima method).

   II The *multivariate* EV *model* or *method of largest observations* (for the k top OSs associated to the original sample of size n).

   III The *multi-dimensional* EV *model* (multivariate EV model for the i_j top observations, $j = 1, 2, \cdots, m$, in sub-samples of size m', $m \times m' = n$).

$m = k$ $(m' = r)$ and $i_j = 1$ for $1 \le j \le k$ originates I;
$m' = n$ $(m = 1)$ and $i_1 = k$ originates II.

IV The *Paretian model* for the excesses, $X_j - u,\ 1 \le j \le k$, of a high deterministic threshold u, suitably chosen [POT approach].

V *Bayesian* approaches.

2. **Semi-parametric approaches**:

VI Under these approaches we work with the k top OSs associated to the whole n observations or with the excesses over a high random threshold, assuming only that 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 $G_\gamma(x)$ provided in (2.2). Conditionally on the intermediate OS $X_{n-k,n}$, the excesses $X_{n-i+1,n} - X_{n-k,n},\ 1 \le i \le k$ are approximately Pareto ["peaks over random thresholds" (PORT) approach].

4 Semi-parametric inference

As mentioned before, under a semi-parametric framework we assume that $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$, i.e., the model F underlying the data is in the domain of attraction for maxima of G_γ in (2.2). Next, we usually base the estimation of γ on the k top OSs in the sample, with k intermediate, i.e., such that (2.6) holds. 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, also under a semi-parametric approach.

4.1 Second (and higher) order conditions

Under a semi-parametric framework, apart from the first order condition in (2.3), we often need a second order condition, specifying the rate of convergence in (2.3). 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

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A^*(t)} = H_{\gamma,\rho}(x) := \frac{1}{\rho^*}\left(\frac{x^{\gamma+\rho^*}-1}{\gamma+\rho^*} - \frac{x^\gamma - 1}{\gamma}\right) \tag{4.1}$$

for all $x > 0$, where $\rho^* \le 0$ is a *second order* parameter controlling the speed of convergence of maximum values, linearly normalized, towards the limit law in (2.2). Then $\lim_{t\to\infty} A^*(tx)/A^*(t) = x^{\rho^*}$, for every $x > 0$, i.e., $|A^*| \in RV_{\rho^*}$ (de Haan and Stadtmüller, 1996).

For heavy tails, it is usually assumed that we know the rate of convergence towards zero of $\ln U(tx) - \ln U(t) - \gamma \ln x$, as $t \to \infty$. The second order condition is then written as

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho}, \tag{4.2}$$

where $\rho \le 0$ 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, Gomes, de Haan and Neves (2007a).

The third order conditions specify, in a parallel way, the rate of convergence either in (4.1) or in (4.2). For further details on the general third order framework, see Fraga Alves, de Haan and Lin (2003b, Appendix; 2006b). Higher order conditions can be similarly postulated, but restrict more and more the distribution functions in $\mathcal{D}_{\mathcal{M}}(G_\gamma)$ considered.

4.2 "Classical" semi-parametric extreme value index estimators

4.2.1 The Hill estimator (H)

For heavy tails, i.e., for $\gamma > 0$, a simple estimator of γ has been proposed by Hill (1975), and from a different perspective by Weissman (1978). Its properties have been thoroughly studied by several authors, among whom we mention Hall (1982b), Mason (1982), Davis and Resnick (1984), Csörgő and Mason (1985), Haeusler and Teugels (1985), Goldie and Smith (1987), Deheuvels, Haeusler and Mason (1988), Beirlant and Teugels (1986; 1989; 1992), Marohn (1997), de Haan and Resnick (1998), de Haan and Peng (1998), among others. The *Hill* estimator has the functional form

$$\widehat{\gamma}^H_{n,k} := \frac{1}{k}\sum_{i=1}^{k} \ln X_{n-i+1,n} - \ln X_{n-k,n} =: M^{(1)}_{n,k}. \tag{4.3}$$

Weak consistency of the estimator in (4.3) is achieved in $\mathcal{D}_{\mathcal{M}}(EV_{\gamma>0})$, whenever (2.3) holds and k is intermediate, i.e., (2.6) holds.

4.2.2 Pickands' estimator (P)

For a general extreme value index, $\gamma \in \mathbb{R}$, and basing here the estimation on the k top OSs, Pickands' estimator (Pickands, 1975) has the functional form

$$\widehat{\gamma}^P_{n,k} := \frac{1}{\ln 2}\, \ln \frac{X_{n-[k/4]+1,n} - X_{n-[k/2]+1,n}}{X_{n-[k/2]+1,n} - X_{n-k+1,n}}, \tag{4.4}$$

where $[x]$ denotes, as usual, the integer part of x. The weak consistency of this EVI estimator, under a first order condition of the type of the one in (2.3) and for k intermediate, has been proved by Pickands (1975) for a general $\gamma \in \mathbb{R}$. A comprehensive study of the asymptotic properties of this estimator, as an estimator based on intermediate and extreme OSs, is provided in Dekkers and de Haan (1989).

4.2.3 The Moment estimator (M)

The *moment* estimator (Dekkers, Einmahl and de Haan, 1989), has the functional expression

$$\widehat{\gamma}_{n,k}^{M} := M_{n,k}^{(1)} + \tfrac{1}{2}\Big\{1 - \big(M_{n,k}^{(2)}/[M_{n,k}^{(1)}]^2 - 1\big)^{-1}\Big\}, \tag{4.5}$$

where

$$M_{n,k}^{(j)} := \frac{1}{k}\sum_{i=1}^{k} \{\ln X_{n-i+1,n} - \ln X_{n-k,n}\}^j, \tag{4.6}$$

is the j-moment of the log-excesses, $j \geq 1$, being $M_{n,k}^{(1)} \equiv \widehat{\gamma}_{n,k}^{H}$ the Hill estimator in (4.3). Again, weak consistency is attained for all $\gamma \in \mathbb{R}$ whenever (2.3) and (2.6) hold.

4.2.4 The POT-ML estimator (ML)

As mentioned in de Haan and Ferreira (2006), the class of distribution function's $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$, for some $\gamma \in \mathbb{R}$, cannot be parameterized with a finite number of parameters, and consequently, there does not exist a ML estimator for γ in such a wide class of models. There exists however an estimator, introduced by Smith (1987) and usually denoted ML estimator. Such an estimator is associated to the approximation provided by the GPD in (2.10), for the excesses over a high observation. The excesses $V_{ik} := X_{n-i+1,n} - X_{n-k,n}$, $1 \leq i \leq k$, are approximately the k top OSs associated to a sample of size k from a distribution function $GP_\gamma(\alpha x/\gamma)$, $\alpha \in \mathbb{R}$, with $GP_\gamma(x)$ given in (2.10). The solution of the ML equations associated to the above mentioned set-up (Davison, 1984) gives rise to the explicit EVI estimator,

$$\widehat{\gamma}_{n,k}^{ML} := \frac{1}{k}\sum_{i=1}^{k} \ln(1 + \hat{\alpha}\ V_{ik}), \tag{4.7}$$

where $\hat{\alpha}$ is the implicit maximum likelihood estimator of the unknown "scale" parameter α. A comprehensive study of the asymptotic properties of the ML estimator in (4.7) has been undertaken in Drees, Ferreira and de Haan (2004). Weak consistency is again attained whenever (2.3) and (2.6) hold and $\gamma > -1/2$.

4.2.5 Asymptotic normal behaviour of the estimators

Under the validity of the second order condition in (4.1), it is possible to guarantee the asymptotic normality of the above mentioned estimators. More precisely, denoting T any of the statistics H, P, M and ML, and with $B(t)$ a bias function converging towards zero as $t \to \infty$ and strongly related with the $A^*(t)$ function in (4.2), it is possible to guarantee for $\gamma \in \mathcal{C}_T$, the existence of real constants (b_T, σ_T), $\sigma_T > 0$, such that:

$$\hat{\gamma}^T_{n,k} \overset{d}{=} \gamma + \sigma_T P^T_k/\sqrt{k} + b_T B(n/k) + o_p(B(n/k)), \tag{4.8}$$

with P^T_k an asymptotically standard normal random variable, $\mathcal{C}_H = \mathbb{R}^+$, $\mathcal{C}_P = \mathcal{C}_M = \mathbb{R}$ and $\mathcal{C}_{ML} = (-\frac{1}{2}, +\infty)$. Consequently, for values k such that $\sqrt{k}\ B(n/k) \to \lambda$, finite, as $n \to \infty$,

$$\sqrt{k}\,\left(\hat{\gamma}^T_{n,k} - \gamma\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \text{Normal}(\lambda b_T, \sigma^2_T). \tag{4.9}$$

The values b_T and σ^2_T are usually called the *asymptotic bias* and *asymptotic variance* of $\hat{\gamma}^T_{n,k}$ respectively.

4.3 Other "classical" EVI semi-parametric estimators

4.3.1 Kernel estimators

A general class of estimators for a positive EVI are the kernel estimators proposed by Csörgő, Deheuvels and Mason (1985), given by

$$\hat{\gamma}^{\mathcal{K}}_{n,k} := \frac{\sum_{i=1}^n \mathcal{K}(i/k)\,\{\ln X_{n-i+1,n} - \ln X_{n-k,n}\}}{\sum_{i=1}^n \mathcal{K}(i/k)}, \tag{4.10}$$

where $\mathcal{K}(\cdot)$ is some non-negative, non-increasing kernel defined on $(0, \infty)$ and integrating to one. As an example, the Hill estimator in (4.3) 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, Lopuha and de Wolf (2003).

4.3.2 QQ-estimators

Several authors have recognized and exploited the potential of probability paper and QQ-plots in estimating $\gamma > 0$. Indeed, the Hill estimator in (4.3) has been obtained from the Pareto quantile plot, through the use of a naïve estimator of the slope in the ultimate right-end of the quantile plot. More flexible regression methods can be applied to the highest k points of the Pareto quantile plot. We refer to Beirlant, Teugels and Vynckier (1996a), Beirlant, Vynckier and Teugels (1996c), Schultze and Steinbach (1996), Kratz and Resnick (1996), Csörgő and Viharos (1998) and Oliveira, Gomes and Fraga Alves (2006). All these estimators can be regarded as kernel estimators.

4.3.3 The generalized Hill estimator

The slope of a generalized quantile plot led Beirlant, Vynckier and Teugels (1996b) to a *generalized Hill* estimator, valid for all $\gamma \in \mathbb{R}$, with the functional form,

$$\widehat{\gamma}_{n,k}^{GH} = \widehat{\gamma}_{n,k}^{H} + \frac{1}{k}\sum_{i=1}^{k}\left\{\ln \widehat{\gamma}_{n,i}^{H} - \ln \widehat{\gamma}_{n,k}^{H}\right\}. \tag{4.11}$$

Further study of this estimator has been performed in Beirlant, Dierckx and Guillou (2005).

4.3.4 Generalized Pickands' and other location-invariant estimators

The large asymptotic variance of Pickands' estimator, together with the three nice features it exhibits — simplicity, validity for all $\gamma \in \mathbb{R}$ and invariance for changes in location —, has motivated different generalizations of Pickands's estimator. We refer Themido Pereira (1993), Fraga Alves (1995) and Yun (2002), who, with slightly different nuances, introduce a *tuning* or *control* parameter θ, and consider generalizations of the type

$$\widehat{\gamma}_{n,k}^{P} := -\frac{1}{\ln\theta}\ \ln\frac{X_{n-[\theta^2 k]+1,n} - X_{n-[\theta k]+1,n}}{X_{n-[\theta k]+1,n} - X_{n-k+1,n}}, \quad 0 < \theta < 1. \tag{4.12}$$

Drees (1995) establishes the asymptotic normality of linear combinations of Pickands' estimators, obtaining optimal weights that can be adaptively estimated from the data. Related work appears in Falk (1994). In Segers (2005), the Pickands estimator for the extreme value index is generalized in a way that includes all of its previously known variants.

Falk (1995a) proposed the location-invariant estimator,

$$\hat{\gamma}_{n,k} := \frac{1}{k}\sum_{i=1}^{k-1}\ln\frac{X_{n,n} - X_{n-i,n}}{X_{n,n} - X_{n-k,n}}, \tag{4.13}$$

as a complement for the ML estimator in the region $\gamma < -1/2$.

The sequence of statistics in (4.12) has also been used in Temido, Gomes and Canto e Castro (2000), to develop methods of estimation in the max-semistable context. In fact, it can be proved that the sequence converges in probability when $\theta = 1/r$ or when θ equals any integer power of $1/r$ and has an oscilatory behaviour in all other cases. This result gives a way of estimating r and r^γ which are the key quantities in representation (2.12). The estimation of the repetitive pattern given by $y(\cdot)$ can also be done using its empirical versions.

The non-invariance for shifts of the Hill estimator in (4.3) led Fraga Alves (2001) to the consideration of the location invariant Hill-type estimator

$$\widehat{\gamma}_{n,k,k_0} := \frac{1}{k_0} \sum_{i=1}^{k} \ln \frac{X_{n-i+1,n} - X_{n-k,n}}{X_{n-k_0+1,n} - X_{n-k,n}}, \tag{4.14}$$

with $k_0 < k$ adequately chosen.

4.3.5 Other estimators

Beirlant, Vynckier and Teugels (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 estimators of γ, 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 the Hill, the Pickands and the kernel estimators, among others. For further references see, e.g., Section 6.4 of Embrechts, Klüppelberg and Mikosch (1997), Beirlant, Teugels and Vynckier (1996a; 1998) and Csörgőand Viharos (1998). See also Chapter 3 of de Haan and Ferreira (2006).

4.4 The Mixed Moment estimator (MM)

Recently, Fraga Alves, Gomes, de Haan and Neves (2007b) considered a combination of Theorems 2.6.1 and 2.6.2 of de Haan (1970): a distribution function F, with right endpoint $x^F \in (0, +\infty]$, is in $\mathcal{D}_{\mathcal{M}}(G_\gamma)$, with G_γ in (2.2), i.e., (2.1) holds with $G = G_\gamma$, if and only if

$$\lim_{t \to x^F} \frac{(1-F(t)) \int_t^{x^F} \int_y^{x^F} (1-F(x)) x^{-2} dx \, dy}{t^2 \left(\int_t^{x^F} (1-F(x)) x^{-2} dx \right)^2} = \varphi(\gamma) := \begin{cases} 1+\gamma & \text{if} \quad \gamma > 0 \\ \frac{1-\gamma}{1-2\gamma} & \text{if} \quad \gamma \le 0. \end{cases} \tag{4.15}$$

Assuming thus that $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$ and $x^F > 0$, we can build a statistic starting from the left hand-side of (4.15), noticing that it can be written as

$$\frac{(1-F(t))^{-1} \left\{ \int_t^\infty \ln \frac{x}{t} dF(x) - \int_t^\infty \left(1 - \frac{t}{x}\right) dF(x) \right\}}{\left\{ (1-F(t))^{-1} \int_t^\infty \left(1 - \frac{t}{x}\right) dF(x) \ \right\}^2} = \frac{E\left(\ln\left(\frac{X}{t}\right) | X > t\right) - E\left(1 - \frac{t}{X} | X > t\right)}{E^2\left(1 - \frac{t}{X} | X > t\right)},$$

and replacing F and t by the empirical distribution function F_n and a high random threshold $X_{n-k,n}$ with $k < n$, respectively. The result is

$$\widehat{\varphi}_{n,k} := \frac{M_{n,k}^{(1)} - L_{n,k}^{(1)}}{\left(L_{n,k}^{(1)}\right)^2}, \qquad \text{with } L_{n,k}^{(1)} := \frac{1}{k} \sum_{i=1}^{k} \left(1 - \frac{X_{n-k,n}}{X_{n-i+1,n}}\right), \tag{4.16}$$

and where $M_{n,k}^{(1)}$ is defined in (4.5).

The statistic $\widehat{\varphi}_{n,k}$ in (4.16) is easily transformed into what has been called the *mixed moment* estimator for the *extreme value index* $\gamma \in \mathbb{R}$:

$$\widehat{\gamma}_{n,k}^{MM} \equiv \widehat{\gamma}_{n,k}^{MM}(X_{n-j+1,n}, 1 \leq j \leq k+1) := \frac{\widehat{\varphi}_n(k) - 1}{1 + 2\min\left(\widehat{\varphi}_n(k) - 1, 0\right)}. \tag{4.17}$$

Under the same conditions as before, we may guarantee that for an adequate $(b_{_{MM}}, \sigma_{_{MM}})$, (4.8) as well as (4.9) hold, with T replaced by MM.

The estimator $\widehat{\gamma}_{n,k}^{MM}$ in (4.17) appears as an interesting alternative to the most popular EVI estimators for a general $\gamma \in \mathbb{R}$. The most attractive features of this new estimator are:

- It is valid for any $\gamma \in \mathbb{R}$ and, contrarily to the ML estimator (valid only for $\gamma > -1/2$), it has a simple explicit functional form, similar to the one of the Moment estimator, also valid for all $\gamma \in \mathbb{R}$.
- Its asymptotic variance is, for $\gamma < 0$, equal to that of the M-estimator and it is equal to that of the ML extreme value index estimator whenever $\gamma \geq 0$.
- It is very close to the ML estimator for a large class of models with $\gamma \geq 0$, and it does not suffer from the disadvantages of this one (lack of convergence of the ML iterative procedure near $\gamma = 0$ and k small).

4.5 Second-order reduced-bias semi-parametric tail index estimation

Let us consider any "classical" semi-parametric estimator $\hat{\gamma}_{n,k}$ of the extreme value index γ. Let us also assume that a distributional representation similar to the one in (4.8), with $(b_{_T}, \sigma_{_T})$ replaced by (b, σ) holds for $\hat{\gamma}_{n,k}$. For intermediate k, i.e., if (2.6) holds, $\hat{\gamma}_{n,k}$ is consistent for the estimation of the extreme value index $\gamma \in \mathbb{R}$, and it is asymptotically normal if we further assume that $\sqrt{k}B(n/k) \to \lambda$, finite. Approximations for the variance and the bias of $\hat{\gamma}_{n,k}$ are then given by σ^2/k and $bB(n/k)$ respectively. Consequently, the pattern of these estimators exhibit the same type of peculiarities:

- high variance for high thresholds $X_{n-k,n}$, i.e., for small values of k;
- high bias for low thresholds, i.e., for large values of k;
- a small region of stability of the sample path (plot of the estimates versus k), as a function of k, making problematic the adaptive choice of the threshold, on the basis of any sample paths' stability criterion;
- a "very peaked" mean-square error, making difficult the choice of the value k_0 where that mean-square error function $MSE\left[\hat{\gamma}_{n,k}\right]$ attains its minimum.

The preceding peculiarities have led researchers to consider the possibility of dealing with the bias term in an appropriate manner, building new estimators $\hat{\gamma}^R_{n,k}$, the so-called reduced-bias extreme value index estimators. Particularly, for heavy tails, i.e., $\gamma > 0$, the reduction of bias is a very important problem involved in the estimation of γ or the Pareto index, $\alpha = 1/\gamma$, in case the slowly varying part of the Pareto type model disappears at a very slow rate. Note however that the use of a logarithmic scale for the k-axis, suggested by Drees, de Haan and Resnick (2000), enlarges the stability regions of "classical" tail index estimators' sample paths. However, such a logarithmic scale masks the regions where second-order reduced-bias estimators have their best performances, as can be seen in Gomes, Hall and Miranda (2006). We consider the following definition.

Definition 4.1. *Under the second order condition* (4.1) *and for k intermediate, i.e., whenever* (2.6) *holds, the statistic* $\hat{\gamma}^R_{n,k}$*, a consistent estimator of the extreme value index* γ*, based on the k top* OSs *in a sample from a distribution function* $F \in \mathcal{D}_{\mathcal{M}}(EV_\gamma)$*, is said to be a reduced-bias semi-parametric estimator of* γ*, 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* $\mathcal{D}_{\mathcal{M}}(EV_\gamma)$*,*

$$\hat{\gamma}^R_{n,k} \stackrel{d}{=} \gamma + \sigma_R P^R_k/\sqrt{k} + o_p(B(n/k)), \tag{4.18}$$

with $B(.)$ *the function in* (4.8).

Notice that for the reduced-bias estimators in (4.18), we no longer have a dominant component of bias of the order of $B(n/k)$, as in (4.8). Therefore,

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

not only when $\sqrt{k}B(n/k) \to 0$ (as for the 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 and a reduction of the mean-square error at the optimal level, in the sense of minimum mean-square error. These optimal levels should be such that $\sqrt{k}B(n/k) \to \infty$ as $n \to \infty$.

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 (4.8), 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. *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 methodology in* Gray and Schucany (1972)*, and* Peng *deals with linear combinations of adequate tail index estimators, in a spirit quite close to the one associated to the Generalized Jackknife* *technique. The latter technique was also used in* Martins, Gomes and Neves (1999)*, where convex mixtures of two Hill's estimators, computed at two different levels, are considered.* Feuerverger and Hall (1999) *discuss the important 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 model. Within the second order framework,* Beirlant, Dierckx, Goegebeur and Matthys (1999) *and* Feuerverger and Hall *investigate the accommodation of bias in the scaled log-spacings and derive approximate "maximum mikelihood" and "least squares" reduced-bias tail index estimators. See also* Pictet, Dacorogna and Müller (1996) *and* Danielsson and de Vries (1997) *for a discussion on this problem, motivated by extreme value applications in finance.* Second order parameters are usually decisive for the bias reduction, and we deal with their estimation in subsection 4.5.1. There is however a bias-corrected estimator for $\gamma > 0$ that, relying on a GP approximation, a model for which $\rho = -\gamma$, does not depend on the estimation of second order parameters, but depends on the choice of an extra control parameter k_0. Such an estimator has been developed in Fraga Alves (2001), and it is based on the statistic $\widehat{\gamma}_{n,k,k_0}$ in (4.14). It has the functional expression

$$\widetilde{\gamma}_{n,k,k_0} := \widehat{\gamma}_{n,k,k_0} - \sqrt{\widehat{\gamma}_{n,k,k_0}/(2k_0)}.$$

For an algorithm providing the adequate adaptive choice of k_0, see Fraga Alves (2001).

4.5.1 Second order parameters estimation for heavy tails

For heavy tails, and under the second order framework in (4.2), we know that $|A| \in RV_\rho$. With $\beta(t)$ any slowly varying function, and assuming $\rho < 0$, we can write $A(t)$ as

$$A(t) = \gamma\ \beta(t)\ t^\rho. \tag{4.19}$$

For simplicity, the function $\beta = \beta(t)$ is often considered as a "scale" parameter $\beta \in \mathbb{R}$. The estimation of second order parameters is thus a practical issue whenever we are interested in the development of statistical inference for extreme values and, in particular, in the estimation of parameters related with rare events.

The estimation of ρ. The first estimator of the parameter ρ in (4.2) 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, Vynckier and Teugels (1996c) and Drees and Kaufmann (1998). Another estimator of ρ appears in Gomes, de Haan and Peng (2002). We elect here particular members of the class of estimators of the second order parameter ρ proposed by Fraga Alves, Gomes and de Haan (2003a). Under adequate general conditions, they are semi-parametric asymptotically normal estimators of ρ, whenever $\rho < 0$, 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 can be parameterised in a tuning real parameter τ, and is defined as,

$$\hat{\rho}_{\tau,k} := -\left|3(T_{n,k}^{(\tau)} - 1)/(T_{n,k}^{(\tau)} - 3)\right|, \quad T_{n,k}^{(\tau)} := \frac{\left(M_{n,k}^{(1)}\right)^{\tau} - \left(M_{n,k}^{(2)}/2\right)^{\tau/2}}{\left(M_{n,k}^{(2)}/2\right)^{\tau/2} - \left(M_{n,k}^{(3)}/6\right)^{\tau/3}}, \quad \tau \in \mathbb{R}, \qquad (4.20)$$

with $M_{n,k}^{(j)}$ given in (4.5) and with the notation $a^{b\tau} = b \ln a$ whenever $\tau = 0$.

The estimation of β. Gomes and Martins (2002) provide an explicit estimator for β, based on the scale log-spacings

$$U_i = i \left\{ \ln X_{n-i+1,n} - \ln X_{n-i,n} \right\}, \quad 1 \le i \le k, \qquad (4.21)$$

and given by

$$\hat{\beta}_{n,k}^{GM(\hat{\rho})} := \left(\frac{k}{n}\right)^{\hat{\rho}} \frac{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\hat{\rho}}\right)\hat{C}_0 - \hat{C}_1}{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\hat{\rho}}\right)\hat{C}_1 - \hat{C}_2}, \quad \hat{C}_j = \frac{1}{k}\sum_{i \le k}\left(\frac{i}{k}\right)^{-j\hat{\rho}} U_i. \qquad (4.22)$$

An additional estimator of β, similar to the estimator of ρ in Fraga Alves, Gomes and de Haan (2003), is provided in Caeiro and Gomes (2006).

4.5.2 A brief review of second-order reduced-bias EVI estimators for heavy tails and the weighted Hill estimator

Generalized jackknife (GJ) estimators of a positive tail index. The pioneering EVI reduced-bias estimators are, in a certain sense, GJ estimators, i.e., affine combinations of well-known estimators of γ. For details on the generalized jackknife methodology, see Gray and Schucany (1972). Whenever we are dealing with semi-parametric estimators of the tail index, or even other parameters of extreme events, we have usually information about the asymptotic bias of those estimators. We may thus choose estimators with similar asymptotic properties, and build the associated GJ random variable or statistic.

Indeed, under the second order condition (4.2), we easily find two statistics $\widehat{\gamma}_{n,k}^{(j)}$, such that, with $P_k^{(j)}$ asymptotically standard normal random variables, we have

$$\widehat{\gamma}_{n,k}^{(j)} \stackrel{d}{=} \gamma + \sigma_j P_k^{(j)}/\sqrt{k} + b_j A(n/k) + o_p(A(n/k)), \quad b_j = b_j(\rho), \quad j = 1, 2. \qquad (4.23)$$

The ratio between the dominant components of bias of $\widehat{\gamma}_{n,k}^{(1)}$ and $\widehat{\gamma}_{n,k}^{(2)}$ is thus $q = b_1/b_2 = q(\rho)$, and we thus get the GJ random variable,

$$\widehat{\gamma}_{n,k}^{GJ(\rho)} := \left(\hat{\gamma}_{n,k}^{(1)} - q(\rho)\ \hat{\gamma}_{n,k}^{(2)}\right)/(1 - q(\rho)), \qquad (4.24)$$

where ρ must be replaced by an estimator $\hat{\rho}$. We can then say that, under the second order condition (4.2), a distributional representation of the type of the one in (4.18) holds for $\widehat{\gamma}_{n,k}^{GJ(\rho)}$, with $\sigma_{GJ} > \gamma$ and $(P_k^R, B(n/k))$ replaced by $(P_k^{GJ}, A(n/k))$. The same result remains true for the GJ estimator $\gamma_{n,k}^{GJ(\hat{\rho})}$, provided that $\hat{\rho} - \rho = o_p(1)$ for all k on which we initially base the tail index estimation, i.e., whenever $\sqrt{k}\ A(n/k) \to \lambda$, finite (Gomes and Martins, 2002).

The maximum likelihood estimation based on the scaled log-spacings. The accommodation of bias in the scaled log-spacings U_i in (4.21) has also been another source of inspiration for the building of second-order reduced-bias tail index estimators. Under the second order condition (4.2), for $\rho < 0$ and with $A(t)$ given in (4.19), Beirlant, Dierckx, Goegebeur and Matthys (1999) motivated the approximation

$$U_i \sim \left(\gamma + A(n/k)\,(i/k)^{-\rho}\right) E_i, \quad 1 \le i \le k, \tag{4.25}$$

where $E_i,\ i \ge 1$, denotes a sequence of i.i.d. standard exponential random variables. In the same context, Feuerverger and Hall (1999) considered the approximation,

$$U_i \sim \gamma \exp\left(A(n/k)\,(i/k)^{-\rho}/\gamma\right) E_i = \gamma \exp\left(A(n/i)/\gamma\right) E_i, \quad 1 \le i \le k. \tag{4.26}$$

The use of the approximation in (4.26) and the joint maximization, in γ, β and ρ, of the approximate log-likelihood of the scaled log-spacings, led Feuerverger and Hall to an explicit expression for $\hat{\gamma}$, as a function of $\hat{\beta}$ and $\hat{\rho}$, given by

$$\hat{\gamma} = \hat{\gamma}_{n,k}^{FH(\hat{\beta},\hat{\rho})} := \frac{1}{k}\sum_{i\le k} e^{-\hat{\beta}(i/n)^{-\hat{\rho}}}\ U_i. \tag{4.27}$$

Then, $\hat{\beta} = \hat{\beta}_{n,k}^{FH(\hat{\rho})}$ and $\hat{\rho} = \hat{\rho}_{n,k}^{FH}$ are numerically obtained, and both computed at the same k used for the γ-estimation. If (2.6) and the second order condition (4.2) hold, the asymptotic variance of $\hat{\beta} = \hat{\beta}_{n,k}^{FH(\hat{\rho})}$ is then ruled by $\sigma_{FH}^2 = \gamma^2\left((1-\rho)/\rho\right)^4$.

A simplified maximum likelihood tail index 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 the explicit estimator for $\hat{\beta}_{n,k}^{GM(\hat{\rho})}$, given by (4.22), and, on the basis of an adequate consistent estimator $\hat{\rho}$ of ρ, they suggest the following approximate ML estimator for the tail index γ,

$$\hat{\gamma}_{n,k}^{GM(\hat{\rho})} := \frac{1}{k}\sum_{i\le k} U_i - \hat{\beta}_{n,k}^{GM(\hat{\rho})}\left(\frac{n}{k}\right)^{\hat{\rho}} \hat{C}_1. \tag{4.28}$$

The estimator in (4.28) is clearly a bias-corrected Hill estimator, i.e., the dominant component of the bias of Hill's estimator, equal to $A(n/k)/(1-\rho) = \gamma\beta(n/k)^\rho/(1-\rho)$ is estimated through $\hat{\beta}_{n,k}^{GM(\hat{\rho})}\,(n/k)^{\hat{\rho}}\,\hat{C}_1$, and directly removed from the Hill estimator in (4.3), which can also be written as $\gamma_{n,k}^{H} = \sum_{i\le k} U_i/k$. Under the same conditions as before, the asymptotic variance of $\hat{\gamma}_{n,k}^{GM(\hat{\rho})}$ is $\sigma^2_{_{GM}} = \gamma^2(1-\rho)^2/\rho^2 < \sigma^2_{_{FH}}$, but still greater than γ^2.

External estimation of β and ρ — the weighted Hill estimator. In a trial to accommodate bias in the excesses over a high random threshold, Gomes, de Haan and Henriques Rodrigues (2004) were led, for heavy tails, i.e., for $\gamma > 0$, to a weighted combination of the log-excesses

$$V_{_{ik}} := \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \le i \le k < n, \tag{4.29}$$

of the type

$$\widehat{\gamma}_{n,k,\hat{\beta},\hat{\rho}}^{WH} := \frac{1}{k}\sum_{i=1}^{k} p_{_{ik}}(\hat{\beta},\hat{\rho})\,\big\{\ln X_{n-i+1:n} - \ln X_{n-k:n}\big\}, \tag{4.30}$$

with WH standing for *weighted Hill*, 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 estimation of the extreme value index. 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, de Haan and Henriques Rodrigues (2004).

Comparatively to second-order reduced-bias estimators available in the literature and published before 2005, this EVI estimator is a *minimum variance reduced bias* (MVRB) estimator, in the sense that, comparatively to the Hill estimator, it keeps the same asymptotic variance $\sigma^2_{_{WH}} = \sigma^2_{_H} = \gamma^2$ and a smaller order asymptotic bias. From a theoretical point of view, we shall expect to have a *mean squared error* (MSE) pattern like the one in Figure 1, i.e., the estimation of γ through $\widehat{\gamma}_{n,k,\hat{\beta},\hat{\rho}}^{WH}$ is, for all k, more reliable than the estimation through the Hill estimator. Related work appears in Caeiro, Gomes and Pestana (2005) and Gomes, Martins and Neves (2007a). For an overview of this subject see Reiss and Thomas (2007), Chapter 6.

How to estimate ρ and β in practice? The theoretical and simulated results in Fraga Alves, Gomes and de Haan (2003a), as well as further results, first in Gomes and Martins (2002), regarding the ρ-estimation and, later on, in Gomes, de Haan and Henriques Rodrigues (2004), regarding the β-estimation, lead to the advice of an external estimation of β and ρ, at a value k_1 of a larger order than the value k used for the γ-estimation. Indeed, for large values k_1, of the type $k_1 = [n^{1-\epsilon}]$ with $\epsilon > 0$ small, and for a large class of heavy-tailed models, we can guarantee that, $\big(\hat{\rho}_{\tau,k_1} - \rho\big)\ln n = o_p(1)$, as $n \to \infty$, a crucial property of the ρ-estimator, if we do not want to increase the asymptotic variance of the random variable,

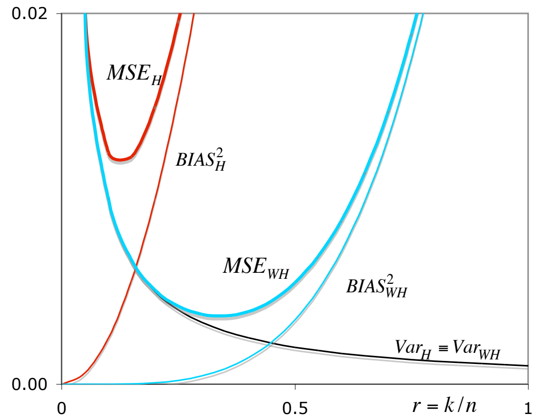

Figure 1: Theoretical pattern of MSE for H and for WH.

function of (β, ρ), underlying the EVI estimator. Such a crucial property can potentially be achieved without any restrictions in the class of models if we compute $\hat{\rho}$ at its optimal level, but the adaptive choice of such a level is still an open research topic. Algorithms for the (β, ρ)-estimation are provided in Gomes, de Haan and Henriques Rodrigues (2004), Gomes and Pestana (2007a; 2007b) and Reiss and Thomas (2007), Chapter 6.

Additional Literature on Bias Reduction. *Other approches to bias reduction, in the estimation of a tail index* $\gamma > 0$, *can be found in* Gomes, Martins and Neves (2000, 2002), Gomes and Martins (2001, 2004), Caeiro and Gomes (2002), Gomes, Caeiro and Figueiredo (2004), Gomes, Figueiredo and Mendonça (2005a), Gomes, Pereira and Miranda (2005b), Canto e Castro and de Haan (2006) *and* Gomes, Miranda and Viseu (2007b).

4.6 Semi-parametric estimation of other parameters of extreme events

High quantiles, or equivalently in financial frameworks the Value at Risk at a level p (VaR_p) are perhaps the most important parameters of extreme events, functions of the EVI, as well as of location and scale parameters. Here again, the work of Laurens de Haan is prominent. 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 (2.3), that provides 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 consider $t = n/k$, $x = k/(np)$ and to estimate χ_{1-p} on the basis of this approximation and adequate estimates of the "shape" γ and the scale $a(n/k)$. 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 estimator of the tail index γ. This estimator is of the type of the one introduced by Weissman (1978). Details on semi-parametric estimation of extremely high quantiles for a general tail index $\gamma \in \mathbb{R}$ may be found in Dekkers and de Haan (1989), de Haan and Rootzén (1993) and more recently in Ferreira, de Haan and Peng (2003). Fraga Alves, Gomes, de Haan and Neves (2007b) also provide, jointly with the MM estimator in (4.17), accompanying shift and scale estimators that make e.g. high quantile estimation almost straightforward. Other approaches to high quantile estimation can be found in the paper by Matthys and Beirlant (2003). Reduced bias quantile estimators have been studied in Matthys, Delafosse, Guillou and Beirlant (2004) and Gomes and Figueiredo (2006), who consider the classical reduced-bias extreme value index estimators. Gomes and Pestana (2007b) and Beirlant, Figueiredo, Gomes and Vandewalle (2006) incorporate the recent minimum variance reduced-bias tail index estimators in Caeiro, Gomes and Pestana (2005) and Gomes, Martins and Neves (2007a) in high quantile semi-parametric estimation.

None of the above mentioned quantile estimators reacts adequately to a shift of the data. Araújo Santos, Fraga Alves and Gomes (2006) provide a class of semi-parametric *high quantile* estimators which enjoy a desirable property in the presence of linear transformations of the data. Such a feature is in accordance with 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 sample of excesses over a random threshold, i.e., on the PORT methodology, providing exact properties for risk measures in finance: translation-equivariance and positively homogeneity.

The dual problem of high quantile estimation, i.e., 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. The estimation of the endpoint of an underlying distribution function has been studied by Hall (1982a), 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).

4.7 Adaptive selection of sample fraction

The semi-parametric estimation of any parameter of extreme events asks for an adaptive selection of the sample fraction to be used in the estimation procedure under consideration, already extensively studied for classical EVI estimators, for which (4.8) holds. In Hall and Welsh (1985), Hall (1990), Beirlant, Vynckier and Teugels (1996c), Drees and Kaufmann (1998) and Danielsson, de Haan, Peng and de Vries (2001), methods for the adaptive choice of k are proposed for the Hill estimator. Gomes and Oliveira (2001) uses also the bootstrap methodology to provide an adaptive choice of the threshold, alternative to the one in Danielsson, de Haan, Peng and de Vries (2001), and easy to generalise to other semi-parametric estimators of parameters of extreme events. Beirlant, Dierckx, Guillou and Starica (2002) consider the exponential regression model (ERM) introduced in Beirlant, Dierckx, Goegebeur and Matthys (1999), and discuss applications of the ERM to the selection of the optimal sample fraction in extreme value 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) also provide a data-driven choice of k for the kernel class of estimators. For a general γ and for the moment estimator in (4.5) and the generalized Pickands' estimator in (4.12), Draisma, de Haan, Peng and Themido Pereira (1999) provide an adaptive choice of the number of OSs involved in an optimal way, balancing variance and bias components. See also Dekkers and de Haan (1993). Apart from the papers by Drees and Kaufmann (1998) and Guillou and Hall (2001), whose choice of the optimal sample fraction is based on bias stability, the other papers make the optimal choice minimizing the estimated mean-square error. For comparisons of different adaptive procedures on the basis of extensive small sample simulations, see Matthys and Beirlant (2000), Gomes and Oliveira (2001) and Beirlant, Dierckx, Guillou and Starica (2002). As far as we know, the optimal sample fraction for second-order reduced-bias estimators is still an open problem. Heuristic selections of these optimal sample fractions are provided in Gomes and Pestana (2007b) and Gomes, Henriques Rodrigues, Vandewalle and Viseu (2006).

4.8 Invariance versus non-invariance

Most of the above mentioned estimators are dependent on the log-excesses, and consequently, are non-invariant with respect to shifts of the data.

The invariance not only to changes in scale but also to changes in location of an EVI estimator is statistically appealing. However, with estimators usually strongly dependent on k, the number of top OSs used, one needs to be open-minded, and it can be sensible to consider, as suggested in Gomes and Oliveira (2003), a deterministic shift τ imposed to the data, as a *tuning parameter* of the statistical procedure. Then, the derived estimators, functions of $X+\tau$ instead of X are not even scale-invariant, but we can play with τ, looking for the τ that, according to any sensible stability criterion, produces the highest stability around a target value, that should be γ, provided that the new "estimator" is consistent for the estimation of γ. This simple trick works very well in many cases, leading to much less disturbing bias, but needs to be used with care.

Indeed, as mentioned by several authors, like for instance Smith (1987), the inadequate use of, for instance, the Hill estimator in conjunction with a shift of the data, can lead to drastic systematic errors. For a discussion on the subject, we recommend the paper by Araújo Santos, Fraga Alves and Gomes (2006), where both the Hill and the moment estimator are transformed into scale/location invariant estimators through the use of the transformed sample

$$X_i^* := X_i - X_{[np]+1,n},\ 0 < p < 1,\ 1 \le i \le n. \tag{4.31}$$

A similar procedure has been used in Fraga Alves, Gomes, de Haan and Neves (2007b), who also propose a class of EVI estimators alternative to the estimator in (4.17), 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 (4.31), $1 \le i \le n$. It is then sensible to define for any *tuning* parameter $p \in (0,1)$,

$$\hat{\gamma}_{n,k}^T(p) := \hat{\gamma}_{n,k}^T(X_{n-j+1,n} - X_{[np]+1,n}, 1 \le j \le k+1). \tag{4.32}$$

The shift invariant versions in (4.32) have properties similar to the ones of the original estimator $\hat{\gamma}_{n,k}^T(X_{n-j+1,n}, 1 \le j \le k+1)$, provided we keep to adequate k-values and choose an adequate *tuning parameter* p.

4.9 Testing the extreme value conditions

Under a semi-parametric framework, it is naturally sensible to test 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. Tests of this nature can already be found in several papers prior to* 2000, *among which we mention* Galambos (1982), Castillo, Galambos and Sarabia (1989), Hasofer and Wang (1992), Fraga Alves and Gomes (1996), Wang, Cooke and Li (1996), Marohn (1998a,b) *and* Fraga Alves (1999). *More recently, further testing procedures of this type can be found in* Segers and Teugels (2000), Neves, Picek and Fraga Alves (2006), Neves and Fraga Alves (2007), *among others. The testing of extreme value conditions can be dated back to* Dietrich, de Haan and Hüsler (2002*), who propose a test static 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* $F \in \mathcal{D}_{\mathcal{M}}(G_{\gamma \ge 0})$. *Further results of this last nature can be found in* Beirlant, de Wet and Goegebeur (2006). Drees, de Haan and Li (2006) *deal with the testing of* $F \in \mathcal{D}_{\mathcal{M}}(G_{\gamma > -1/2})$. *Accurate tables of critical points of this statistic are provided in and* Hüsler and Li (2005).

5 Laurens de Haan: a personal tribute

Once Laurens' mother bought a new, more potent car, to tow her caravan. In the afternoon she went for a ride, and after some time she remarked that there was a car closely following hers. She turned right and the follower turned right, she turned right again and again the other car turned right, and after some more pseudo random driving she had no doubt that she was being followed, and feeling cross and afraid she decided to return home to call the police and complain. When she arrived home, the police was waiting for her, because the other car's owner had complained that someone had stolen the car that he had parked in front of that house — the car that she had been towing around, after "fishing" it with the towing hook when she had started her ride.

With such an upbringing, Laurens couldn't but develop an intimate feeling for rare events! The result, as all the fans of *Extremes* know, has been an extremely successful research career in the field, honoured in this special issue.

Laurens de Haan, in his moving speech when he has been awarded a doctorate *Honoris Causa* by the University of Lisbon, said:

> "*I decided to go into a somewhat marginal subject that would possibly escape the attention of famous people. So I started to read the basic papers in the field of Extreme Values which were written before the Second World War and in French, so that they were not much consulted by the people from the English speaking world.*
>
> *The most prominent mathematician who had published about the subject was* Boris Gnedenko*. He wrote a basic paper in* 1943 *and in the introduction he said: I have solved all the problems but one. By combining with other papers I could solve the problem. This earned me a doctorate, I can now say: my first doctorate and also a head start in the field.*"

One of the "*other papers*" written in French was surely Karamata's pioneering work (Karamata, 1930) on the "*croissance régulire des fonctions*". Doeblin (Doeblin, 1940) and Gnedenko (Gnedenko, 1940), independently characterizing the domains of attraction of additive stable laws, and Gnedenko (Gnedenko, 1943) characterizing the domain of attraction of maxima stable laws, were not aware of Karamata's paper, and they had to work out cumbersome arguments to establish their results. As far as we know, Feller has been the first to observe, in a presentation at a Berkeley Symposium and in the second volume of his very influential monograph (Feller, 1967, 1971), that the presentation of the theory of domains of attraction of additive stable laws can be considerably simplified using regular variation.

Laurens de Haan worked out new results on extended regular variation, and used them to fill the gap in Gnedenko's characterization of the domains of attraction of max-stable laws, renamed Gnedenko-de Haan's characterization by Falk, Hüsler and Reiss (2004), pages

23–24. This was one of the outstanding achievements in Laurens' "*head start in the field*" (de Haan, 1970), establishing the potential of regular variation in the investigation of stable elements and their domains of attraction, a fact plainly confirmed by Bingham's results on stable elements in generalized convolution algebras (Bingham, 1971).

Laurens de Haan's thesis could surely enter the *Guiness Book* for the record number of citations it has from all those working in Extreme Value Theory, and its outstanding inspiring role is widely acknowledged. We quote from influential monographs:

> "*From the theoretical point of view, the* 1970 *doctoral dissertation by* L. de Haan*, On Regular Variation and its Applications to the Weak Convergence of Sample Extremes, seems to be the starting point for the theoretical developments in extreme value theory. For the first time, the probabilistic and stochastic properties of sample extremes were developed into a coherent and attractive theory, comparable to the theory of sums of random variables.*" (Beirlant, Goegebeur, Segers and Teugels, 2004)

> "*The asymptotic behaviour of the largest in a random sample of n from a population, has provided a challenge to many able mathematical statisticians. Noteworthy contributions were made by ....., and finally* Gnedenko (1943) *and* de Haan (1970)*, who give the most complete and rigorous discussion of the problem.*" (David and Nagaraja, 2003)

> "De Haan *applied regular variation analytical tool. His work has been of great importance for the development of modern extreme value theory.*" (Embrechts, Klüppelberg and Mikosch, 1997)

> "*Professor* de Haan *has had enormous influence on the subject, and his* 1970 *monograph remains, despite the huge quantity of research it stimulated, an excellent place to learn about relationship of extreme value theory and regular variation.*" (Resnick, 1987)

> "*Another major stimulus ... was provided by* Laurens de Haan *in his* 1970 *thesis...*" (Bingham, Goldie and Teugels, 1987)

After this auspicious start, L. de Haan had a leading role in broadening the scope of EVT. Quoting again from his Lisbon speech:

> “*Then, over the years our field has become more and more important from a theoretical point of view but also more useful and even fashionable. Thirty years ago it was mainly attractive for the elegant mathematics involved but since about fifteen years the direction of the developments has been dictated by problems that come up in applications. The field is in a healthy state.*”

Modestly he omits that he (co-)worked up most of the elegant mathematics in EVT, and that he has been — is — one of the leading scientists achieving the shift towards applications, combining mathematics and intensive computation methods to obtain useful results for dealing with disastrous events. *On extreme value theory, or how to learn from almost disastrous events* was the title of his Calouste Gulbenkian Professorship talk (de Haan, 2006). This has been an excellent opportunity for organizing an workshop in his honour. In the preface of the associated book of extended abstracts it is said:

> “Laurens de Haan *has one of the most prominent careers in the XX/XXI centuries Statistics. He may indeed be considered as one of the world exponents in the area of Statistics of Extremes and his Ph.D. thesis, written at* 1970*, is still an almost compulsory reference in the field.* Laurens de Haan *has contributed to the development of well-built theories in areas like Extended Regular Variation, Multivariate Extremes, Semi-Parametric Estimation, and Extremes for Dependent Sequences. Recently, he has been paying special attention to the field of Extremes in Infinite-Dimensional Spaces. Beyond the building of a unified and rigorous Extreme Value Theory,* Laurens de Haan *has also had a pioneering work in the solution of important environmental problems, related to the modelling of rivers, sea and dams, and the specification of new standards for the Dutch sea defences. Since* 1997*,* Laurens de Haan *has regularly visited Lisbon, and this has led to the development of joint research work with several members of* CEAUL *(Centro de Estatística e Aplicações da Universidade de Lisboa), as well as inspired the scientific cooperation with other members of the Portuguese statistical community, as well as with international statistical researchers and practitioners.*” (Fraga Alves and Gomes, 2006)

There is no pacific agreement about the *best* statistical methodology for extreme values. In the IID setup, we can say that statistical methodologies associated to a semi-parametric setup for Peaks Over Random Thresholds (PORT) approach have had their main flourishing ideas in the pioneer and present research work of Laurens de Haan. Laurens spread his dutch seeds all over the world in such a way that a flourishing field has been planted, with strong roots on EVT as a rigorous mathematical framework, which has recently culminated in a solid self-contained theoretical basis and applicable framework *Extreme Value Theory: an Introduction* (de Haan and Ferreira, 2006).

We have been very fortunate, indeed, to have Laurens among us, in our research unit in Lisbon University (CEAUL), for most of the time in the last decade. The wealth of ideas he spreads around him are at the root of rewarding joint work with a bunch of members of CEAUL. But his generous share of ideas with junior scientists spreads around the world — Laurens is always taking a plane going to or coming from somewhere, he deserves the title of *Fliegende Holländer* —, and in his Lisbon speech he could claim that he had "*written papers with 43 people*" (the number has been increased since that date). In fact, Laurens has a deep insight of rare stochastic events, and while most of us feel happy to explore one promising idea, he, like a *jongleur*, keeps us fascinated with the way he plays with more than half a dozen lines of research at the same time, with half a dozen co-workers around the world. Quoting again from his *Honoris Causa* Doctorate speech:

> "*I have worked with people who had very different backgrounds and view. But the co-operation invariably not only made the papers much better but also it added much to our personal development in the field and finally it has often resulted in a lasting friendship. I have always tried, when I was confronted with a mathematical problem I could not solve, to seek co-operation with a person whose background allows him or her to solve the problem.*"

We dare to say that one of Laurens great achievements, well known among the "extremists" community but perhaps less evident for "outliers", is his *magisterium*, exerted via the capacity to build up a group of disciples/co-workers, who could learn with him, in the privileged form of direct example, how to create strategies and concepts to solve research problems — with the extra benefit of lasting friendship.

References

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

[2] Achcar, A.J., Bolfarine, H., and Pericchi, L.R. (1987). Transformation of a survival data to an extreme value distribution. *Statistician* **36**, 229-234.

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

[4] 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.

[5] Arnold, B., Balakrishnan, N., and Nagaraja, H.N. (1992). *A First Course in Order Statistics.* Wiley, New York.

[6] Ashour, S.K., and El-Adl, Y.M. (1980). Bayesian estimation of the parameters of the extreme value distribution. *Egypt. Statist. J.* **24**, 140-152.

[7] Balakrishnan, N., and Chan, P.S. (1992). Order statistics from extreme value distribution, II: best linear unbiased estimates and some other uses. *Comm. Statist. – Simul., and Comput.* **21**, 1219-1246.

[8] Balkema, A.A., and Haan, L. de (1974). Residual life time at great age. *Annals of Probability* **2**, 792-804.

[9] Bardsley, W.E. (1977). A test for distinguishing between extreme value distributions. *J. Hydrology* **34**, 377-381.

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

[11] Beirlant, J., Dierckx, G., and Guillou, A. (2005). Estimation of the extreme-value index and generalized quantile plots. *Bernoulli* **11**:6, 949-970.

[12] 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.

[13] Beirlant, J., Figueiredo, F., Gomes, M.I., and Vandewalle, B. (2006). Improved reduced bias tail index and quantile estimators. *J. Statistical Planning and Inference.* DOI:10.1016/j.jspi.2007.07.015, 2007.

[14] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications.* Wiley, England.

[15] Beirlant, J., and Teugels, J. (1986). Asymptotics of Hill's estimator. *Theory Probab. Appl.* **31**, 463-469.

[16] Beirlant, J., and Teugels, J. (1989). Asymptotic normality of Hill's estimator. In Hüsler, J., and Reiss, R.-D. (eds.) *Extreme Value Theory*, Proc. of the 1987 Oberwolfach Conference, 148-155, Springer-Verlag, New, York.

[17] Beirlant, J., and Teugels, J. (1992). Modelling large claims in non-life insurance. *Insurance: Mathematics and Economics* **11**, 17-29.

[18] Beirlant, J., Teugels, J., and Vynckier, P. (1996a). *Practical Analysis of Extreme Values.* Leuven University Press, Leuven, Belgium.

[19] 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.

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

[21] Beirlant, J., Vynckier, P., and Teugels, J. (1996c). Tail index estimation, Pareto quantile plots and regression *J. Amer. Statist. Assoc.* **91**, 1659-1667.

[22] 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.

[23] Bermudez, P. de Zea, Amaral-Turkman, M.A., and Turkman, K. F. (2001). A predictive approach to tail probability estimation. *Extremes* **4**:4, 295-314.

[24] Bermudez, P. de Zea and Amaral-Turkman, M.A. (2003). Bayesian approach to parameter estimation of the generalized pareto distribution. *Test* **12**, 259-277.

[25] Bingham, N. (1971). Factorization theory and domains of attraction for generalized convolution algebras, *Proc. London Math. Soc.* (3), **23**, 16–30.

[26] Bingham, N., Goldie, C.M., and Teugels, J.L. (1987). *Regular Variation.* Cambridge Univ. Press, Cambridge.

[27] Bottolo, L., Consonni, G., Dellaportas, P., and Lijoi, A. (2003). Bayesian analysis of extreme values by mixture modeling. *Extremes* **6** (1), 25-47.

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

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

[30] Caeiro, F., and Gomes, M.I. (2006). A new class of estimators of a "scale" second order parameter. *Extremes* **9**, 193-211.

[31] Caeiro, F., Gomes, M.I., and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat* **3**:2, 113-136.

[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., and Haan. L. de (2006). A class of distribution functions with less bias in extreme value estimation. *Statist., and Probab. Letters* **76**, 1617-1624.

[34] Castillo, E., Galambos, J., and Sarabia, J.M. (1989). The selection of the domin of attraction of an extreme value distribution from a set of data. In Hüsler, J., and Reiss, R.-D. (eds.), *Extreme value Theory, Proc. Oberwolfach* 1987, 181-190, Springer, Berlin-Heidelberg.

[35] Castillo, E., and Hadi, A. (1997). Fitting the generalized Pareto distribution to data. *J. Amer. Statist. Assoc.* **92**, 1609-1620.

[36] 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.

[37] Chaouche, A. and Bacro, J.-N. (2004). Fitting the generalized Pareto distribution to data. *Comput. Statist. and Data Analysis* **45**, 787-803.

[38] Choulakian, V., and Stephens, M.A. (2001). Goodness-of-fit tests for the generalized Pareto distribution. *Technometrics* **43** (4), 478-484.

[39] Christopeit, N. (1994). Estimating parameters of an extreme value distribution by the method of moments. *J. Statist. Plann., and Inference* **41**, 173-186.

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

[41] Coles, S. G., and Dixon, M.J. (1999). Likelihood-based inference for extreme value models. *Extremes* **2** (1), 5-23.

[42] Coles, S., and Powell, E. (1996). Bayesian methods in extreme value modelling: a review and new developments. *Int. Statist. Rev.* **64** (1), 119-136.

[43] Coles, S., and Tawn, J. (1996). A Bayesian analysis of extreme rainfall data. *Appl. Statist.* **45**(4), 463-478.

[44] Crovella, M., and Taqqu, M. (1999). Estimating the heavy tail index from scaling properties. *Methodology and Computing in Appl. Probab.* **1**, 55-79.

[45] Csörgő, S., Deheuvels, P. and Mason, D.M. (1985). Kernel estimates of the tail index of a distribution, *Ann. Statist.* **13**, 1050-1077.

[46] Csörgő, S., and Mason, D.M. (1985). Central limit theorems for sums of extreme events. *Math. Proc. Camb. Phil. Soc.* **98**, 547-558.

[47] 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.

[48] Csörgő, S., and Viharos, L. (1998), Estimating the tail index, In Szyszkowicz, B. ed., *Asymptotic Methods in Probability and Statistics*, 833-881, North-Holland.

[49] Danielsson, J., Haan, L. de, Peng, L., and Vries, C.G. de (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *J. Multivar. Anal.* **76**, 226-248.

[50] Danielsson, J., and de Vries, C.G. (1997). Tail index and quantile estimation with very high frequency data. *J. Empirical Finance* **4**, 241-257.

[51] David, H.A. (1970; 1981). *Order Statistics*, 1st edition; 2nd edition. Wiley, New York.

[52] David, H.A., and Nagaraja, H.N. (2003). *Order Statistics*, 3rd edition. Wiley, Hoboken, New Jersey.

[53] Davis, R., and Resnick, S. (1984). Tail estimates motivated by extreme value theory. *Ann. Statist.* **12**, 1467- 1487.

[54] Davison, A. (1984). Modeling excesses over high threshold with an application. In J. Tiago de Oliveira ed., *Statistical Extremes and Applications*, D. Reidel, 461-482.

[55] Davison, A.C., and Smith, R.L. (1990). Models for exceedances over high thresholds. *J. Royal Statist. Soc.* **B 52**, 393-442.

[56] Deheuvels, P., Haeusler, E., and Mason, D.M. (1988). Almost sure convergence of the Hill estimator. *Mathem. Proc. Cambridge Phil. Soc.* **104**, 371-381.

[57] Dekkers, A.L.M., Einmahl, J.H.J., and Haan, L. de (1989). A moment estimator for the index of an extreme-value distribution. *Ann. Statist.* **17**, 1833-1855.

[58] Dekkers, A.L.M., and Haan, L. de (1989). On the estimation of the extreme-value index and large quantile estimation. *Annals of Statistics* **17**, 1795-1832.

[59] Dekkers, A.L.M., and Haan, L. de (1993). Optimal choice of sample fraction in extreme-value estimation. *J. of Multivariate Analysis* **47**, 173-195.

[60] Diebolt, J., and Guillou, A. (2005). Asymptotic behavior of regular estimators. *Revstat* **3**:1, 19-44.

[61] 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.

[62] Dietrich, D., Haan, L. de, and Hüsler, J. (2002). Testing extreme value conditions. *Extremes* **5**, 71-85.

[63] Dietrich, D., and Hüsler, J. (1996). Minimum distance estimators in extreme value distributions. *Comm. Statist. – Theory and Methods* **25**, 695-703.

[64] 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.

[65] Doeblin, W. (1940). Sur l'ensemble des puissances d'une loi de probabilités, *Studia Math.* **9**, 71–96.

[66] Draisma, G., Haan, L. de, Peng, L., and Themido Pereira, T. (1999). A bootstrap-based method to achieve optimality in estimating the extreme value index. *Extremes* **2**:4, 367-404.

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

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

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

[70] Drees, H., Ferreira, A., and Haan, L. de (2004). On maximum likelihood estimation of the extreme value index. *Ann. Appl. Probab.* **14**, 1179-1201.

[71] 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.

[72] Drees, H., Haan, L. de and Resnick, S. (2000). How to make a Hill plot. *Annals of Statistics* **28**, 254-274.

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

[74] Dupuis, D.J. (1998). Exceedances over high thresholds: a guide to threshold selection. *Extremes* **1** (3), 251-261.

[75] Dupuis, D.J., and Field, C.A. (1998). A comparison of confidence intervals for generalized extreme-value distributions. *J. Statist. Comput. and Simul.* **61**, 341-360.

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

[77] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin, Heidelberg.

[78] Engeland, K., Hisdal, H., and Frigessi, A. (2004). Practical extreme value modelling of hydrological floods and droughts: a case tudy. *Extremes* **7**:1, 5-30.

[79] Engeland, S., and Rackwitz, R. (1992). On predective distribution functions for the three asymptotic extreme value distributions. *Struct. Safety* **11**, 255-258.

[80] Falk, M. (1989). A note on uniform asymptotic normality of intermediate order statistics. *Ann. Inst. Statist. Math.* **41**, 19-29.

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

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

[83] Falk, M. (1995b). On testing the extreme value index via the POT-method. *Ann. Statist.* **23**, 2013-2035.

[84] Falk, M. (1995c). LAN of extreme order statistics. *Ann. Inst. Statist. Math.* **47**, 693-717.

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

[86] Feller, W. (1967). On regular variation and local limit theorems, *Proc. 5th Berkeley Symposium on Mathematical Statistics and Probability*, II, Univ. California Press, Los Angeles, 373–378.

[87] Feller, W. (1971). *An Introduction to Probability Theory and its Applications*, vol. II, Wiley, New York (1st ed. 1966)

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

[89] 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.

[90] Feuerverger, A., and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *Ann. Statist.* **27**, 760-781.

[91] 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.

[92] Fraga Alves, M.I. (1992). The influence of central observations on discrimination among multivariate extremal models. *Theory of Probability and its Applications* **32**, 395-398.

[93] Fraga Alves, M.I (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**, no. 1-2, 143-173.

[94] Fraga Alves, M. I. (1999). Asymptotic distribution of Gumbel statistic in a semi-parametric approach. *Portugaliae Mathematica* **56** (3), 282-298.

[95] Fraga Alves, M.I. (2001). A location invariant Hill-type estimator. *Extremes* **4**:3, 199-217.

[96] Fraga Alves, M.I., and Gomes, M.I. (1996). Statistical choice of extreme value domains of attraction — a comparative analysis. *Communications in Statistics – Theory and Methods* **25** (4), 789-811.

[97] Fraga Alves, M.I., and Gomes, M.I. (2006). *Extremes Day in Honour of Laurens de Haan: Extremes, Risk, Safety and the Environment*, CEAUL Editions, Lisboa.

[98] 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.

[99] Fraga Alves, M.I., Gomes, M.I., Haan, L. de and Neves, C. (2007a). *A Note on Second Order Condition in Extremes: Linking General and Heavy Tails Conditions*. Notas e Comunicações CEAUL 11/2007. Submitted.

[100] Fraga Alves, M.I., Gomes, M.I., Haan, L. de and Neves, C. (2007b). *Mixed Moment Estimator and Location Invariant Alternatives*. Notas e Comunicações CEAUL 14/2007. Submitted.

[101] Fraga Alves, M. I., Haan, L. de and Lin, T. (2003b). Estimation of the parameter controlling the speed of convergence in extreme value theory. *Math. Methods of Statist.* **12**, 155-176.

[102] Fraga Alves, M. I., Haan, L. de and Lin, T. (2006b). Third order extended regular variation. *Publications de lInstitut Mathématique* **80** (94), 109-120.

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

[104] Frigessi, A., Haug, O., and Rue, H. (2002). A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes* **5**:3, 219-235.

[105] Galambos, J. (1978; 1987). *The Asymptotic Theory of Extreme Order Statistics*, 1st edition: Wiley, New York; 2nd edition: Krieger, Malabar, Florida.

[106] Galambos, J. (1982). A statistical test for extreme value distributions. In Gnedenko, B.V. et al. (eds.), *Non-Parametric Statistical Inference*, 221-230, North-Holland, Amsterdam.

[107] Geluk, J., and Haan, L. de (1987). *Regular Variation, Extensions and Tauberian Theorems*. CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam, The Netherlands.

[108] Gnedenko, B.V. (1940). On the theory of domains of attraction of stable laws, *Uchenye Zapiski, Moskov Gos. Univ.*, **30**, 61–72.

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

[110] Goldie, C., and Smith, R. (1987). Slow variation with remainder: theory and applications. *Quart. J. Math. Oxford* **38**, 45-71.

[111] Gomes, M.I. (1978). *Some Probabilistic and Statistical Problems in Extreme Value Theory*. Ph. D. Thesis, Univ. Sheffield.

[112] Gomes, M.I. (1981). An i-dimensional limiting distribution function of largest values and its relevance to the statistical theory of extremes. In C. Taillie et al. (eds.), *Statistical Distributions in Scientific Work*, Vol. **6**, 389-410, D. Reidel.

[113] Gomes, M.I. (1982). A note on statistical choice of extremal models. *Actas IX Jorn. Hispano-Lusas*, 653-655, Salamanca.

[114] Gomes, M.I. (1984). Concomitants in a multidimensional extreme model. In J. Tiago de Oliveira ed., *Statistical Extremes and Applications*, 353-364, D. Reidel, Dordrecht.

[115] Gomes, M.I. (1985a). Concomitants and linear estimators in an i-dimensional extremal model. *Trab. Estad. Inv. Oper.* **36**, 129-140.

[116] Gomes, M.I. (1985b). Statistical theory of extremes — comparison of two approaches. *Statistics and Decision* **2**, 33-37.

[117] Gomes, M.I. (1987). Extreme value theory — statistical choice. In *Colloq. Math. Soc. János Bolyai* **45**, 195-210, Debrecen.

[118] Gomes, M.I. (1989a). Comparison of extremal models through statistical choice in multidimensional backgrounds. In Hüsler, J., and R.-D. Reiss (eds.), *Extreme Value Theory, Proc. Oberwolfach* 1987, Lecture Notes in Statistics 51, 191-203, Springer-Verlag.

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

[120] 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.

[121] 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).

[122] Gomes, M.I., and Alpuim, M.T. (1986). Inference in multivariate generalized extreme value models. *Scandinavian J. Statistics* **13**, 291-300.

[123] Gomes, M. I., Caeiro, F., and Figueiredo, F. (2004). Bias reduction of a extreme value index estimator trough an external estimation of the second order parameter. *Statistics* **38**(6): 497-510.

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

[125] Gomes, M.I., Figueiredo, F., and Mendonça, S. (2005a). Asymptotically best linear unbiased tail estimators under a second order regular variation condition. *J. Statist. Planning and Inference* **134**:2, 409-433.

[126] Gomes, M.I., and Haan, L. de (1999). Approximation by penultimate extreme-value distributions. *Extremes* **2**:1, 71-85.

[127] Gomes, M.I., Haan, L. de, and Henriques Rodrigues, L. (2004). Tail index estimation for heavy-tailed models: accommodation of bias in the weighted log-excesses. *J. Royal Statistical Society* **B**, DOI: 10.1111/j.1467-9869.2007.00620.x, 2007

[128] Gomes, M.I., Haan, L. de, and Peng, L. (2002). Semi-parametric estimation of the second order parameter — asymptotic and finite sample behaviour. *Extremes* **5**:4, 387-414.

[129] Gomes, M.I., Hall, A. and Miranda, C. (2006). Subsampling techniques and the Jackknife methodology in the estimation of the extremal index. *Comput. Statist. and Data Analysis*, DOI: 1016 / j.csda.2007.06.023, 2007.

[130] Gomes, M.I., Henriques Rodrigues, L., Vandewalle, B. and Viseu, C. (2006). A heuristic adaptive choice of the threshold for bias-corrected Hill estimators. *J. Statist. Comput. and Simulation*, DOI: 10.1080 / 10629360600954000, 2007.

[131] Gomes, M.I., and Martins, M.J. (2001). Alternatives to Hill's estimator — asymptotic versus finite sample behaviour. *J. Statist. Planning and Inference* **93**, 161-180.

[132] Gomes, M.I., and Martins, M.J. (2002). "Asymptotically unbiased estimators of the tail index based on external estimation of the second order parameter. *Extremes* **5**:1, 5-31.

[133] 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.

[134] 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**, 207-229.

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

[136] Gomes, M.I., Martins, M.J., and Neves, M.M. (2007a). Improving second order reduced bias extreme value index estimation. *Revstat* **5**:2, 177–207.

[137] Gomes, M.I., Pereira, H., and Miranda, C. (2005b). Revisiting the role of the Jackknife methodology in the estimation of a positive tail index. *Comm. Statistics – Theory and Methods* **34**, 1-20.

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

[139] Gomes, M.I., and van Montfort, M.A.J. (1987). Exponentiality versus Generalized Pareto – quick tests. In *Procc. III Intern. Conf. Statistical Climatology*, 185-195.

[140] Gomes, M.I., and Oliveira, O. (2001). The bootstrap methodology in statistical extremes — the choice of the optimal sample fraction. *Extremes* **4**:4, 331-358.

[141] Gomes, M.I., and Oliveira, O. (2003). How can non-invariant statistics work in our benefit in the semi-parametric estimation of parameters of rare events. *Commun. Statist., Simulation Comput.* **32**(4), 1005-1028.

[142] 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.

[143] Gomes, M.I., and Pestana, D. (2007b). A sturdy reduced-bias extreme quantile (VaR) estimator. *J. American Statistical Association*, Vol. **102**, No. 477, 280-292.

[144] Gray, H.L., and Schucany, W.R. (1972). *The Generalized Jackknife Statistic.* Marcel Dekker.

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

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

[147] Grimshaw, S. (1993). Computing maximum likelihood estimates for the genaralized Pareto distribution. *Technometrics* **35** (2), 185-191.

[148] Groeneboom, P., Lopuha, H.P., and de Wolf, P.P. (2003). Kernel-type estimators for the extreme value index. *Ann. Statist.* **31**, 1956-1995.

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

[150] Gumbel, E.J. (1935). Les valeurs extrêmes des distribution statistiques. *Ann. Inst. Henri Poincaré* **5**, 115-158.

[151] Gumbel, E.J. (1958). *Statistics of Extremes.* Columbia Univ. Press, New York.

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

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

[154] 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.

[155] Haan, L. de (2006). On extreme value theory, or how to learn from almost disastrous events. Gulbenkian lecture (DVD with booklet) and abstract in Fraga Alves, M.I., and Gomes, M.I. (eds.), *Extremes Day in Honour of Laurens de Haan: Extremes, Risk, Safety and the Environment*, 1-2, CEAUL Editions, Lisboa.

[156] Haan, L. de and Ferreira, A. (2006). *Extreme Value Theory: an Introduction.* Springer Science+Business Media, LLC, New York.

[157] Haan, L. de and Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica* **52**, 60-70.

[158] Haan, L. de and Resnick, S. (1998). On asymptotic normality of the Hill estimator. *Stoch. Models* **14**, 849-867.

[159] Haan, L. de and Rootzén, H. (1993). On the estimation of high quantiles. *J. Statist. Planning and Inference* **35**, 1-13.

[160] Haan, L. de and Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Austral. Math. Soc.* **A61**, 381-395.

[161] Haeusler, E., and Teugels, J. (1985). On asymptotic normality of Hill's estimator for the exponent of regular variation. *Ann. Statist.* **13**, 743-756.

[162] Hall, P. (1982a). On estimating the endpoint of a distribution. *Ann. Statist.* **10**, 556-568.

[163] Hall, P. (1982b). On some simple estimates of an exponent of regular variation. *J. Royal Statist. Soc.* **B44**, 37-42.

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

[165] Hall, P., and Welsh, A.W. (1985). Adaptive estimates of parameters of regular variation. *Ann. Statist.* **13**, 331-341.

[166] Hasofer, A.M., and Wang, Z. (1992). A test for extreme value domain of attraction. *J. Amer. Statist. Assoc.* **87**, 171-177.

[167] He, X., and Fung, W.K. (1999). Method of medians for life time data with Weibull models. *Statistics in Medicine* **18**, 1993-2009.

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

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

[170] Hosking, J.R.M. (1985). Algorithm AS 215: Maximum likelihood estimation of the parameters of the generalized extreme value distribution. *Appl. Statist.* **34**, 301-310.

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

[172] 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.

[173] Hüsler, J., and Li, D. (2005). On testing extreme value conditions. *Extremes* **9**, 69-86.

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

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

[176] Juárez, S.F., and Schucany, W.R. (2004). Robust and efficient estimation for the generalized Pareto distribution. *Extremes* **7** (3), 237-251.

[177] Jurečková, J., and Picek, J. (2001). A class of tests on the tail index. *Extremes* **4**:2, 165-183.

[178] Karamata, J. (1930), Sur un mode de croissance régulire des fonctions, *Mathematica (Cluj)*, **4**, 38–53.

[179] Kaufmann, E. (2000). Penultimate approximations in extreme value theory. *Extremes* **3**:1, 39-55.

[180] Kinnison, R. (1989). Correlation coefficient goodness-of-fit test for the extreme value distribution. *Amer. Statist.* **43**, 98-100.

[181] Kratz, M., and Resnick, S. (1996). The qq-estimator of the index of regular variation. *Comm. Statist. — Stochastic Models* **12**, 699-724.

[182] Lamperti, J. (1964). On extreme order statistics. *Ann. Math. Statist.* **35**, 1726-1737.

[183] Lingappaiah, G.S. (1984). Bayesian prediction regions for the extreme order statistics. *Biom. J.* **26**, 49-56.

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

[185] Lye, L.M., Hapuarachchi, K. P., and Ryan, S. (1993). Bayes estimation of the extreme-value reliability function. *IEEE Trans. on Reliab.* 42, 641-644.

[186] Macleod, A.J. (1989). A remark on the algorithm AS 215: Maximum likelihood estimation of the parameters of the generalized extreme value distribution. *Appl. Statist.* **38**, 198-199.

[187] Marohn, F. (1994). On testing the exponential and Gumbel distribution. In Galambos, J. et al. (eds.), *Extreme Value Theory*, 159-174, Kluwer, Dordrecht.

[188] Marohn, F. (1997). Local asymptotic normality in extreme value index estimation. *Ann. Inst. Statist. Math.* **49**, 645-666.

[189] Marohn, F. (1998a). An adaptive efficient test for Gumbel domain of attraction. *Scand. J. Statist.* **25**, 311-324.

[190] Marohn, F. (1998b). Testing the Gumbel hypothesis via the POT-method. *Extremes* **1** (2), 191-213.

[191] Marohn, F. (2000). Testing extreme value models. *Extremes* **3**:4, 363-384.

[192] Martins, M.J., Gomes, M.I., and Neves, M.M. (1999). Some results on the behaviour of Hill's estimator. *J. Statist. Comput., and Simulation* **63**, 283-297.

[193] Martins, E.S., and Stedinger, J.R. (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Res.* **36** (3), 737-744.

[194] Mason, D.M. (1982). Laws of large numbers for sums of extreme values. *Ann. Probab.* **10**, 754-764.

[195] Matthys, G., and Beirlant, J. (2000). Adaptive threshold selection in tail index estimation. In Embrechts, P. (ed.), *Extremes and Integrated Risk Management*, 37-57, UBS, Warburg.

[196] Matthys, G., and Beirlant, J. (2003). Estimating the extreme value index and high quantiles with exponential regression models. *Statistica Sinica* **13**, 853-880.

[197] 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.

[198] McNeil, A.J. (1997). Estimating the tails of loss severity distributions using extreme value theory. *Astin Bulletin* **27**, 117–137.

[199] 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.

[200] Montfort, M.A.J. van (1970). On testing that the distribution of extremes is of type I when type II is the alternative. *J. Hydrology* **11**, 421-427.

[201] Montfort, M.A.J. van (1973). *An asymmetric test on the type of distribution of extremes.* Tech. Rep. 73-18, Medelelugen Landbbouwhoge-School Wageningen.

[202] Montfort, M.A.J. van and Gomes, M.I. (1985). Statistical choice of extremal models for complete and censored data. *J. Hydrology* **77**, 77-87.

[203] Montfort, M.A.J. van and Otten, A. (1978). On testing a shape parameter in the presence of a location and scale parameter. *Math. Operationsforsch. Statist.* **9**, 91-104.

[204] Montfort, M.A.J. van and Witter, J.V. (1985). Testing exponentiality against generalized Pareto distribution. *J. Hydrology* **78**, 305-315.

[205] Neves, C., and Fraga Alves, M.I. (2007). Semi-parametric approach to the Hasofer-Wang and Greenwood statistics in extremes. *Test* **16**, 297-313.

[206] Neves, C., Picek, J., and Fraga Alves, M.I. (2006). The contribution of the maximum to the sum of excesses for testing max-domains of attraction. *J. Statist. Planning and Inference* **136** (4), 1281-1301.

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

[208] Otten, A., and van Montfort, M.A.J. (1978). The power of two tests on the type of distribution of extremes. *J. Hydrology* **37**, 195-199.

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

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

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

[212] Peng, L., and Welsh, A.W. (2001). Robust estimation of the generalized Pareto distribution. *Extremes* **4**(1), 53-65.

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

[214] Pickands III, J. (1994). Bayes quantile estimation and threshold selection for the generalized Pareto family. In Galambos, J., Leigh, S., and Simiu, E. (eds.), *Extreme Value Theory and Applications*, 123-138, Kluwer, Amsterdam.

[215] Pictet, O.V., Dacorogna, M.M., and Müller, U.A. (1998). Hill, bootstrap and Jackknife estimators for heavy tails. In Adler, R., Feldman, R., and Taqqu, M. S. (eds.), *A Practical Guide to Heavy Tails*, 55-79, Birkhäuser Publishers.

[216] Prescott, P., and Walden, A.T. (1980). Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika* **67**, 723-724.

[217] Prescott, P., and Walden, A.T. (1983). Maximum likelihood estimation of the parameters of the three-parameter generalized extreme-value distribution from censored samples. *J. Statist. Comput., and Simulation* **16**, 241-250.

[218] 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.

[219] Reiss, R.-D., and Thomas, M. (1999). A new class of Bayesian estimators in paretian excess-of-loss reinsurance. *Astin J.* **29** (2), 339-349.

[220] 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.

[221] Resnick, S.I. (1987). *Extreme Values, Regular Variation and Point Processes.* Springer.Verlag, New York.

[222] Resnick, S.I. (1997). Heavy tail modelling and teletraffic data. *Ann. Statist.* **25** (5), 1805-1869.

[223] Rootzén, H., and Tajvidi, N. (1997). Extreme value statistics and wind storm losses: a case study. *Scand. Actuarial J.* **1**, 70-94.

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

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

[226] Segers, J., and Teugels, J. (2000). Testing the Gumbel hypothesis by Galton's ratio. *Extremes* **3** (3), 291-303.

[227] Singh, V.P., and Guo, H. (1997). Parameter estimation for 2-parameter generalized Pareto distribution. *Stochastic Hydrology and Hydraulics* **11**:2, 11-227.

[228] Smirnov, N. (1949; 1952). Limit distributions for terms of a variational series (in russian). *Trudy Mat. Inst. Steklov* **25**; translation: *Amer. Math. Soc. Trans.* **67**, 82-143.

[229] Smirnov, N. (1967). Some remarks on limit laws for order statistics. *Theory Probability Appl.* **12**, 337-339.

[230] Smith, R.L. (1984). Threshold methods for sample extremes. In Tiago de Oliveira, J. (ed.), *Statistical Extremes and Applications*, 621-638, D. Reidel, Dordrecht.

[231] Smith, R.L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika* **72**, 67-90.

[232] Smith, R.L. (1986). Extreme value theory based on their largest annual events. *J. Hydrology* **86**, 27-43.

[233] Smith, R.L. (1987). Estimating tails of probability distributions. *Ann. Statist.* **15**, 1174-1207.

[234] Smith, R.L. (1999). Bayesian and frequencist approaches to parametric predictive inference. In Bernardo, J. O., Berger, J. O., Dawid, A. P., and Smith, A. F. (eds.), 589-612, Oxford Press, U.K..

[235] Smith, R.L., and Goodman, D.J. (2000). Bayesian risk analysis. In Embrechts, P. (ed.), *Extremes and Integrated Risk Management*, 235.251, Risk Books, London.

[236] Smith, R.L., and Naylor, J.C. (1987). A comparison of maximum likelihood and Bayesian estimators for the three parameter Weibull distribution. *Appl. Statist.* **36**, 358-369.

[237] Stephens, M.A. (1976). Asymptotic results for goodness-of-fit statistics with unknown parameters. *Annals of Statistics* **4**, 357-369.

[238] Stephens, M.A. (1977). Goodness-of-fit for the extreme value distribution. *Biometrika* **64**, 583-588.

[239] Stephens, M.A. (1986). Tests for the exponential distribution. In D'Agostinho, R, B., and Stephens, M. A. (eds.), *Goodness-of-Fit Techniques*, 421-459, Marcel Dekker, New York.

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

[241] Tawn, J. (1988). An extreme value theory model for dependent observations. *J. of Hydrology* **101**, 227-250.

[242] Temido, M.G., Gomes, M.I., and Canto e Castro, L. (2000). Inferência estatística em modelos max-semiestáveis. In P. Oliveira and E. Athayde eds., *Um Olhar sobre a Estatística*, 291-305, S.P.E. Editions (in Portuguese).

[243] 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.).

[244] Tiago de Oliveira, J. (1981). Statistical choice of univariate extreme models. In C. Taillie et al. (eds.), *Statistical Distributions in Scientific Work*, Vol. **6**, 367-387.

[245] Tiago de Oliveira, J. (1984). Univariate extremes: statistical choice. In Tiago de Oliveira, J. (ed.), *Statistical Extremes and Application*, 91-107, D. Reidel, Dordrecht.

[246] Tiago de Oliveira, J., and Gomes, M.I. (1984). Two test statistics for choice of univariate extreme models. In Tiago de Oliveira, J. (ed.), *Statistical Extremes and Application*, 651-668, D. Reidel, Dordrecht.

[247] Walshaw, D. (2000). Modelling extreme wind speeds in regions prone to hurricanes. *Appl. Statist.* **49**, 51-62.

[248] Wang, J.Z. (1995). Selection of the k largest order statistics for the domain of attraction of the Gumbel distribution. *J. Amer. Statist. Assoc.* **90**, 1055-1061.

[249] Wang, J.Z., Cooke, P., and Li, S. (1996). Determination of domains of attraction based on a sequence of maxima. *Austral. J. Statist.* **38**, 173-181.

[250] Weissman, I. (1975). Multivariate extremal processes generated by independent non-identically distributed random variables. *J. Appl. Probab.* **12**, 477-487.

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

[252] Weissman, I. (1984). Statistical estimation in extreme value theory. In Tiago de Oliveira, J. (ed.), *Statistical Extremes and Applications*, 109-115, D. Reidel.

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

Index