REVSTAT – Statistical Journal
Volume 14, Number 2, April 2016, 193–215

EXTREME VALUE ANALYSIS – A BRIEF OVERVIEW WITH AN APPLICATION TO FLOW DISCHARGE RATE DATA IN A HYDROMETRIC STATION IN THE NORTH OF PORTUGAL

Authors: Helena Penalva
– Department of Economics and Management,
School of Business Administration, and CEAUL,
Polytechnic Institute of Setúbal, Portugal
helena.penalva@esce.ips.pt

Sandra Nunes
– Department of Economics and Management,
School of Business Administration, CEAUL and CMA/FCT/UNL,
Polytechnic Institute of Setúbal, Portugal
sandra.nunes@esce.ips.pt

M. Manuela Neves
– Department of Sciences and Engineering of Biosystems,
Institute of Agronomy, and CEAUL, University of Lisbon, Portugal
manela@isa.ulisboa.pt

Received: October 2015 Revised: February 2016 Accepted: February 2016

Abstract:

- Extreme value theory is dedicated to characterise the behaviour of the extreme observations. The interest is then focused in the tails of the underlying distribution. It is important to test for the adequate shape of the tail, because it influences the estimation of parameters of extreme or even rare events. The aim of this work is to present a brief overview on several tests and parameter estimation procedures available in the literature. They will be applied to daily mean flow discharge rate values in the hydrometric station of Fragas da Torre in the river Paiva, collected from 1946/47 to 2005/06.

Key-Words:

- *extreme value theory; parametric and semi-parametric estimation; statistical choice of the tail; statistical tests.*

AMS Subject Classification:

- 62G32, 62G10, 62E20, 62P12.

1. INTRODUCTION

Extreme value theory (EVT) is concerned with the stochastic behaviour of extremes values. In EVT we need to deal with events that are more extreme than any that have already been observed. The question is how to make inference beyond the sample data. Obviously, statistical inference can be deduced only from those observations which are extreme in some sense.

There are a few parameters whose estimation is of major importance. The extreme value index (EVI), which is directly related with the heaviness of the right tail of the underlying distribution of the data, is a crucial parameter. It influences the estimation of other parameters of extreme values, such as, high quantiles of probability $1-p$, with p "small", i.e., the *high levels* usually designed by the *return levels* associated with the *return periods*, $1/p$, i.e. the expected waiting time between independent exceedances of a specific high level.

In all areas of application it is of major importance to use adequate and accurate statistical methods. The R software (R Development Core Team, [55]) is an open source environment that incorporates in its base a huge amount of statistical packages built and made freely available by the scientific community.

Penalva *et al.* ([52], [53]) have illustrated the application of some procedures of modelling and estimating in EVT, under a parametric framework. Some R packages were explained and some data sets were considered. The *Block Maxima* (BM), the *Peaks Over Threshold* (POT) and the *k Largest Observations* (k-LO) methods were described and applied. Different methodologies for parameter estimation were also considered. In Neves *et al.* ([50]) R procedures for the semi-parametric estimation in EVT have been presented and discussed. A real data set of daily mean flow discharge rate values from the hydrometric station of Fragas da Torre in the river Paiva during the years from 1946/47 to 1996/97 was considered.

In this paper parametric and semi-parametric frameworks are briefly reviewed. In both cases EVT theory relies on certain assumptions that should be validated when dealing with an application. Regardless the framework followed statistical inference will be improved if one makes the choice of the most adequate tail previously. A brief overview of some testing procedures for the so-called *extreme value condition* and for the *statistical choice of the tail* will be given. An application to a larger data set than the one mentioned above will be performed, now considering the years from 1946/47 to 2005/06.

Section 2 provides a brief review on the basic notions in EVT. In Section 3 parametric and semi-parametric statistical approaches in EVT are summarized and the main statistical methods for the estimation of parameters are described.

Section 4 and 5 are dedicated to a brief reference to testing issues and finally Section 6 presents a case study and the application of some of the methods described in the previous sections, still giving some attention to the main packages available in R software for the extreme value analysis.

2. PRELIMINARIES IN EVT LIMITING LAWS

Classic theory of extremes is concerned with the limiting behaviour of the maximum $M_n := \max\big(X_1, ..., X_n\big)$ or the minimum $m_n := \min\big(X_1, ..., X_n\big)$, as $n \to \infty$, of a sample $(X_1, ..., X_n)$ of independent and identically distributed (i.i.d.) or possibly stationary, weakly dependent, random variables with unknown distribution function (d.f.) F. It is well known that in those conditions the distribution of the maximum M_n is $F^n(\cdot)$, and also for the minimum m_n, i.e., $1-[1-F(\cdot)]^n$. However the d.f. F^n is of little help in practice since F is itself unknown and should F be misspecified, this can lead to large errors in the distribution of the maximum.

First results in EVT date back to Fréchet ([27]), Fisher and Tippet ([22]), Gumbel ([39]) and von Mises ([60]), but Gnedenko ([30]) and de Haan ([41]) have solved the problems related with the asymptotic behaviour of statistical extremes, giving conditions for the existence of sequences $\{a_n\} \in \mathbb{R}^+$ and $\{b_n\} \in \mathbb{R}$ such that

$$\lim_{n\to\infty} P\left(\frac{M_n - b_n}{a_n} \le x\right) = \lim_{n\to\infty} F^n(a_n x + b_n) \;=\; EV_\xi(x) \qquad \forall x \in \mathbb{R}\,, \tag{2.1}$$

where EV_ξ is a nondegenerate distribution function.

This function, known as the Extreme Value d.f., is usually denoted by EV_ξ and is given by

$$EV_\xi(x) \;=\; \begin{cases} \exp\big\{-[1+\xi x]^{-1/\xi}\big\}\,, & 1+\xi x > 0\,, \quad \text{if} \;\; \xi \neq 0\,, \\ \exp\big\{-\exp[-x]\big\}\,, & x \in \mathbb{R}\,, \quad \text{if} \;\; \xi = 0\,, \end{cases} \tag{2.2}$$

where $\xi \in \mathbb{R}$ is the shape parameter.

Definition 2.1. We say that F is in the domain of attraction (for maxima) of EV_ξ and write $F \in \mathcal{D}_{\mathcal{M}}(EV_\xi)$, whenever (2.1) holds.

As a consequence of the existence of that limit, when $n \to \infty$ we may consider the approximation, $P[M_n \le x] = F^n(x) \approx EV_\xi\big((x-b_n)/a_n\big)$.

The EV_ξ incorporates the three (Fisher–Tippett) families: the *Gumbel family*, that is the limit for exponential tailed distributions, $\Lambda(x) = EV_0(x) =$

$\exp\bigl(-\exp(-x)\bigr)$, $x \in \mathbb{R}$, $\xi = 0$; the *Fréchet family*, that is the limit for heavy tailed distributions, $\Phi_\alpha(x) = EV_{1/\alpha}\bigl(\alpha(x-1)\bigr) = \exp(-x^{-\alpha})$, $x > 0$, $\xi = 1/\alpha > 0$ and the *Weibull family*, that is the limit for short tailed distributions, $\Psi_\alpha(x) = EV_{-1/\alpha}\bigl(\alpha(x+1)\bigr) = \exp\bigl(-(-x)^\alpha\bigr)$, $x < 0$, $\xi = -1/\alpha < 0$.

The shape parameter, ξ, is the so-called *extreme value index* (EVI), it is the primary parameter in EVT and it measures the heaviness of the right-tail, $\overline{F} := 1 - F$. If $\xi = 0$, the right tail is of an exponential type; if $\xi > 0$, the right tail is heavy, it is of a negative polynomial type and if $\xi < 0$, the right tail is short and F has a finite right endpoint.

The limit distribution family, EV_ξ in (2.2), seems to present some difficulties due to the normalizing constants, $\{a_n\}$ and $\{b_n\}$ be unknown. However that limit can be interpreted, for sufficiently large n, as

$$P\left(\frac{M_n - b_n}{a_n} \le x\right) \approx EV_\xi(x) \iff P\bigl(M_n \le x\bigr) \approx EV_\xi\left(\frac{x - b_n}{a_n}\right). \tag{2.3}$$

We can further consider location and scale parameters, $\lambda \in \mathbb{R}$ and $\delta \in \mathbb{R}^+$, respectively, in the EV_ξ d.f., denoting it by $EV_\xi(x;\lambda,\delta) \equiv EV_\xi\bigl((x-\lambda)/\delta\bigr)$, so the constants in (2.3) can incorporate this location/scale version.

Instead of just considering the maximum value of a sample as an extreme value, we may consider all the observations, X_i, above a high level or threshold, u, established previously, as extremes. The differences $X_i - u$, are called exceedances over that threshold. Balkema and de Haan ([1]) and Pickands ([54]) proved that if $F \in \mathcal{D}_{\mathcal{M}}(EV_\xi)$, see Definition (2.1), then for large enough u, $Y = \bigl((X-u)\,|\,X > u\bigr)$ is approximately the generalized Pareto (GP) d.f.,

$$H_\xi(y) = 1 - \bigl(1 + \xi y/\widetilde{\delta}\bigr)^{-1/\xi}, \qquad \text{for } y > 0 \text{ and } \bigl(1 + \xi y/\widetilde{\delta}\bigr) > 0\,, \tag{2.4}$$

where ξ is the shape parameter, equal to that of the corresponding EV distribution, and the scale parameter $\widetilde{\delta} = \delta + \xi(u - \lambda)$, where λ is the location parameter in the EV d.f.. The reciprocal of the stated above is also true.

We can also consider the joint distribution of the k top order statistics. More specifically, if X is a random variable with d.f. F belonging to the domain of attraction of an EV d.f. then, for fixed k, the limiting distribution, as $n \to \infty$, of the k-dimensional random vector, suitably normalized by constants $\{a_n\} \in \mathbb{R}^+$ and $\{b_n\} \in \mathbb{R}$, $\left(\frac{M_n^{(1)} - b_n}{a_n}, ..., \frac{M_n^{(k)} - b_n}{a_n}\right)$, where $M_n^{(k)} \equiv X_{n-k+1:n} := k$ largest of $\{X_1, ..., X_n\}$ and the joint probability density function is given by

$$g(w_1, ..., w_k) = EV_\xi(w_k) \prod_{i=1}^{k} \frac{ev_\xi(w_i)}{EV_\xi(w_i)}\,, \qquad w_1 > \cdots > w_k\,, \tag{2.5}$$

with $EV_\xi(w)$ defined in (2.2) and where $ev_\xi(w) = \frac{\partial EV_\xi(w)}{\partial w}$ is the probability density function of the EV model. This model is known as the Multivariate-EV_ξ model, also known as the extremal process, Dwass ([20]).

3. MODELLING AND ESTIMATING IN EVT

Statistical inference in EVT is based on extreme observations, however there are different ways of defining such observations leading to the application of different models. Classical parametric approaches for modelling and estimation were the first to appear, based on limiting distributions defined in the previous section. In the late seventies, estimation procedures in EVT began to be performed on a semi-parametric approach based on probabilistic asymptotic results in the tail of the unknown distribution.

3.1. Parametric statistical approaches and estimation

The first approach for modelling extremes is the so-called *Block Maxima* (BM), *Annual Maxima* or *Gumbel's* approach, Gumbel ([40]). In this approach the n-sized sample is splitted into m sub-samples (usually m corresponds to the number of the observed years) of size l ($n = m \times l$ for a sufficiently large l). EV_ξ or one of the models, Gumbel, Fréchet or Weibull, with unknown $\xi \in \mathbb{R}$, $\lambda \in \mathbb{R}$ or $\delta \in \mathbb{R}^+$ are then fitted to the m maxima values of the m sub-samples.

However, in many applications there is no natural way of defining blocks of observations. Besides it may occur that the maximum within a block has a lower value than some values in another block. Thus, some extreme values contained in a block may not be included in data for the analysis. So, BM methodology may not be the best method for studying the behaviour of extreme values.

Another methodology consists of setting a high level or threshold, u, and defining as extremes all the observations above that value. The idea is then to fit the model referred to in (2.4) to the *excesses* over such a high level, u. This method, known as *Peaks Over Thresholds* (POT) method, uses relevant information that can be lost by the BM method. Details of this procedure can be seen in Davison ([15]), Davison and Smith ([16]) and Smith ([57]).

Another approach, in some sense parallel to the previous one, consists of considering the k top order statistics of the sample. In this methodology, usually denoted as the *k-Largest Observations* (k-LO), inference can be done when the size n of the sample is large and k fixed, based on the multivariate structure of the k top order statistics, referred to in (2.5). This model was developed and studied by Weissman ([61]) and Gomes ([31]).

Note that the use of POT method needs the choice of a suitable threshold, u, what is equivalent to the choice of the number, k, of upper order statistics to be taken on the k-LO approach.

We can also think of combining the BM and the k-LO approaches. In each of the m sub-samples, we can collect a few top order statistics and, in this case, inference is based on the m random k-dimensional vectors. These m random k-dimensional vectors, after being suitably normalized by constants $\{a_n\} \in \mathbb{R}^+$ and $\{b_n\} \in \mathbb{R}$, are well modelled by the Multivariate-EV_ξ defined in (2.5). This methodology is known as *Multidimensional-*EV_ξ approach.

For estimating extreme value parameters several procedures have been proposed: (i) graphical methods; (ii) moment-based methods and (iii) likelihood methods. All these procedures have been extensively studied and applied in classical parametric modelling. In this work we will review parameter estimation using the maximum likelihood (ML) method , the profile likelihood (PL) method and the probability weighted moments (PWM) method.

Difficulties that arose with the "regularity conditions" for the maximum likelihood estimation were solved by Smith ([58]), who showed that the usual property of asymptotic normality holds provided the extreme value parameter ξ is larger than -0.5. Recently, Zhou ([62], [63]) showed that the ML estimators verify the property of asymptotic normality for $\xi > -1$. This condition, that is not verified for very light tailed distributions, is satisfied for most environmental applications.

The asymptotic normality, that would allow to obtain confidence intervals, is not very accurate because the normal approximation to the true sampling distribution of the estimator is rather poor. An alternative, and usually more accurate method of estimation is based on the profile likelihood function. Given a parameter vector $\boldsymbol{\theta}$ the *profile log-likelihood* function of the component θ_i is defined as $\log L_p(\theta_i) := \max_{\boldsymbol{\theta}_{-i}} \log L(\theta_i, \boldsymbol{\theta}_{-i})$ where $\boldsymbol{\theta}_{-i}$ denotes a vector with all components of vector $\boldsymbol{\theta}$ excluding θ_i. For each value of θ_i, the profile log-likelihood is defined as the maximized log-likelihood with respect to the other components of the parameter vector $\boldsymbol{\theta}$.

So, for example, for the estimation of ξ in the EV model,

$$\log L_p(\xi) := \max_{\lambda,\delta\,|\,\xi} \log L(\lambda, \delta, \xi) .$$

Under suitable regularity conditions, see Beirlant *et al.* ([3]), for large n, the *deviance function* is:

$$D_p(\xi) := 2\left\{\log L(\widehat{\lambda}, \widehat{\delta}, \widehat{\xi}) - \log L_p(\xi)\right\} \;\dot{\sim}\; \chi^2_{(1)} ,$$

where $\widehat{\lambda}, \widehat{\delta}$ and $\widehat{\xi}$ are the maximum likelihood estimators of λ, δ and ξ, respectively. This property is used to obtain the $(1-\alpha)\times 100\%$ confidence interval for the parameters of the underlying distribution. Particularly, for a singular component, for example ξ, the $(1-\alpha)\times 100\%$ confidence interval is $\left\{\xi\colon D_p(\xi) \le q_{1-\alpha}\right\} =$

$\{\xi\colon \log L_p(\xi) \geq \log L(\widehat{\lambda}, \widehat{\delta}, \widehat{\xi}) - \frac{q_{1-\alpha}}{2}\}$, where $q_{1-\alpha}$ is the $(1-\alpha)$ quantile of $\chi^2_{(1)}$. Therefore the profile log-likelihood ratio statistic

$$-2\log\Lambda \;=\; -2\log\left\{\frac{L_p(\xi_0)}{L_p(\widehat{\xi})}\right\} \;=\; 2\left\{\log L_p(\widehat{\xi}) - \log L_p(\xi_0)\right\},$$

to test $H_0\colon \xi = \xi_0$ versus $H_1\colon \xi \neq \xi_0$ has, under the hypothesis H_0, asymptotic distribution $\chi^2_{(1)}$, when $n \to \infty$. H_0 is rejected at a level of significance α if $-2\log\Lambda > q_{1-\alpha}$, see Coles ([13]) and Beirlant *et al.* ([3]) for more details.

The probability-weighted moments (PWM) (Greenwood *et al.*, [38]) of a random variable X, with d.f. F are defined as

$$M_{p,r,s} \;:=\; E\left\{X^p\left[F(X)\right]^r\left[1-F(X)\right]^s\right\}, \qquad p,r,s \in \mathbb{R}\,.$$

For the EV d.f., these moments were extensively studied by Hosking *et al.* ([44]). Considering a random sample $(X_1,...,X_m)$ from a EV population, the PWM estimator, $(\widehat{\lambda},\widehat{\delta})$, when $\xi = 0$, is the solution of the system of equations:

$$\begin{cases} \widehat{M}_{1,0,0} = \lambda + \delta\,\Gamma'(1) \\ 2\widehat{M}_{1,1,0} - \widehat{M}_{1,0,0} = \log 2\delta \end{cases} \qquad \text{where} \quad \widehat{M}_{1,r,0} = \frac{1}{m}\sum_{i=1}^{m}\left(\prod_{l=1}^{r}\frac{i-l}{m-l}\right)X_{i:m}\,,$$

with $X_{1:m} \leq X_{2:m} \leq \cdots \leq X_{m:m}$ the ascending order statistics associated with the random sample $(X_1, X_2, ..., X_m)$.

For $0<\xi<1$, we can obtain the PWM estimator, $(\widehat{\lambda},\widehat{\delta},\widehat{\xi})$ solving the equations system,

$$\begin{cases} \widehat{M}_{1,0,0} = \lambda - \dfrac{\delta}{\xi}\big(1-\xi(1-\xi)\big) \\ 2\widehat{M}_{1,1,0} - \widehat{M}_{1,0,0} = \dfrac{\delta}{\xi}\,\xi(1-\xi)\,(2^\xi-1) \\ \dfrac{3\widehat{M}_{1,2,0} - \widehat{M}_{1,0,0}}{2\widehat{M}_{1,1,0} - \widehat{M}_{1,0,0}} = \dfrac{3^\xi-1}{2^\xi-1} \end{cases}.$$

Also in this method the asymptotic normality for the PWM estimator (λ,δ,ξ) holds provided that $\xi < 0.5$ and $m \to \infty$ (see Beirlant *et al.*, [3]).

3.2. Semi-parametric statistical framework and EVI estimation

In the late seventies estimation in EVT began to be performed in a semi-parametric approach. Here it is not necessary to fit a specific parametric model, dependent upon a location, scale and shape parameters, but only assume that the underlying distribution function F belongs to $\mathcal{D}_{\mathcal{M}}(EV_\xi)$, for an appropriate value of ξ in specific sub-domain of $\mathcal{D}_{\mathcal{M}}(EV_\xi)$, being ξ the primordial parameter to be estimated. Estimates are based on the k top order statistics in the sample, or on the excesses over a high random threshold, u. For the consistence of the estimators we need to work with an intermediate sequence $k \equiv k_n$, i.e., $k \equiv k_n \to \infty$ and $k/n \to 0$ as $n \to \infty$.

In this framework several EVI estimators have been proposed. We will refer to the classical ones, such as, the Hill estimator ([43]), the Moment estimator, Dekkers *et al.* ([17]), the Generalized Hill estimator, introduced in Beirlant *et al.* ([4]) and studied later in Beirlant *et al.* ([2]), the Mixed Moment estimator, Fraga Alves *et al.* ([26]) and also a recent estimator of reduced bias and minimum variance, (MVRB), Caeiro *et al.* ([10]). A family of estimators based on the logarithm of the *mean of order p* (MOP) of $X_{n-i-1:n}/X_{n-k:n}$, $1 \le i \le k < n$, has been very recently proposed by Brilhante *et al.* ([7]). See also other related estimators such as the harmonic mean estimator introduced in Beran *et al.* ([5]) and a family of estimators introduced in Paulauskas and Vaiciulis ([51]).

Let $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ the ascending order statistics associated with the random sample $(X_1, X_2, ..., X_n)$, and for $r \ge 1$ let us define

$$L^{(r)}_{k,n} := \frac{1}{k}\sum_{i=1}^{k}\left[1-\frac{X_{n-k:n}}{X_{n-i+1:n}}\right]^r \quad\text{and}\quad M^{(r)}_{k,n} := \frac{1}{k}\sum_{i=1}^{k}\left[\ln\frac{X_{n-i+1:n}}{X_{n-k:n}}\right]^r . \tag{3.1}$$

Among the aforementioned estimators we will consider:

The Hill estimator ($\xi > 0$)

$$\widehat{\xi}^{H}_{k,n} := M^{(1)}_{k,n}\,, \qquad k = 1, 2, ..., n-1\ ; \tag{3.2}$$

The Moments estimator ($\xi \in \mathbb{R}$)

$$\widehat{\xi}^{M}_{k,n} := M^{(1)}_{k,n} + 1 - \frac{1}{2}\left(1 - \frac{\left(M^{(1)}_{k,n}\right)^2}{M^{(2)}_{k,n}}\right)^{-1}, \qquad k = 1, 2, ..., n-1\ ; \tag{3.3}$$

The Generalized Hill estimator ($\xi \in \mathbb{R}$)

$$\widehat{\xi}^{GH}_{k,n} := M^{(1)}_{k,n} + \frac{1}{k}\sum_{i=1}^{k}\left[\ln\frac{M^{(1)}_{i,n}}{M^{(1)}_{k,n}}\right], \qquad k = 1, 2, ..., n-1\ ; \tag{3.4}$$

The Mixed Moment estimator ($\xi \in \mathbb{R}$)

$$\widehat{\xi}^{MM}_{k,n} := \frac{\widehat{\varphi}_{k,n} - 1}{1 + 2\min(\widehat{\varphi}_{k,n} - 1, 0)}\,, \qquad k = 1, 2, ..., n-1\,, \tag{3.5}$$

$$\widehat{\varphi}_{k,n} := \frac{M^{(1)}_{k,n} - L^{(1)}_{k,n}}{\left(L^{(1)}_{k,n}\right)^2}\,.$$

The **MVRB** estimators, Caeiro *et al.* ([10]) have revealed a better performance than the classical estimators in the context of heavy tails ($\xi > 0$). This class of estimators has the functional form

$$\overline{\widehat{\xi}^H_{k,n}}(\widehat{\beta},\widehat{\rho}) := \widehat{\xi}^H_{k,n}\Big(1 - \widehat{\beta}(n/k)^{\widehat{\rho}}/(1-\widehat{\rho})\Big)\,, \tag{3.6}$$

with $\widehat{\xi}^H_{k,n}$ the Hill estimator and $(\widehat{\beta},\widehat{\rho})$ consistent estimators of second order parameters $(\beta,\rho) \in (\mathbb{R},\mathbb{R}^-)$. About reduced bias estimation, we may also refer to Gomes *et al.* ([36]), Gomes *et al.* ([33]) and Caeiro *et al.* ([9]), among others.

For the estimation of ρ we consider a particular member of a class of estimators introduced in Fraga Alves *et al.* ([24]). This class, parametrized in a control parameter $\tau \in \mathbb{R}$, which here we will take as $\tau = 0$, see Gomes *et al.* ([37]), is defined as: $\widehat{\rho}(k) \equiv \widehat{\rho}_0(k) := \min\Big(0, \frac{3\,(T_n^{(0)}(k)-1)}{T_n^{(0)}(k)-3}\Big)$, being $T_n^{(0)}(k)$ defined as

$$T_n^{(0)}(k) := \left[\ln\Big(M^{(1)}_{k,n}\Big) - \frac{1}{2}\ln\Big(M^{(2)}_{k,n}/2\Big)\right] \Big/ \left[\frac{1}{2}\ln\Big(M^{(2)}_{k,n}/2\Big) - \frac{1}{3}\ln\Big(M^{(3)}_{k,n}/6\Big)\right],$$

with $M^{(j)}_{k,n}(k),\ j = 1,2,3$, defined above.

For the estimation of the second order scale parameter, β, we will consider

$$\widehat{\beta}_{\widehat{\rho}}(k) := \left(\frac{k}{n}\right)^{\widehat{\rho}} \Big[d_{\widehat{\rho}}(k)\,D_0(k) - D_{\widehat{\rho}}(k)\Big] \Big/ \Big[d_{\widehat{\rho}}(k)\,D_{\widehat{\rho}}(k) - D_{2\widehat{\rho}}(k)\Big]\,,$$

with $\widehat{\rho} = \widehat{\rho}_0(k)$, $d_\alpha(k) := \frac{1}{k}\sum_{i=1}^k \big(\frac{i}{k}\big)^{-\alpha}$ and $D_\alpha(k) := \frac{1}{k}\sum_{i=1}^k \big(\frac{i}{k}\big)^{-\alpha} U_i$, for $\alpha \le 0$, with $U_i := i\big[\ln(X_{n-i+1:n}/X_{n-i:n})\big],\ 1 \le i \le k$.

In order not to have an increase in the variance of the estimator $\overline{\widehat{\xi}^H_{k,n}}$, estimators $\widehat{\rho}_0(k)$ and $\widehat{\beta}_{\widehat{\rho}}(k)$ must be calculated at $k = k_1$, with $k_1 = \lfloor n^{1-\epsilon}\rfloor$, $\epsilon = 0.001$, see Gomes and Martins ([35]), Gomes *et al.* ([33]) and Caeiro *et al.* ([9]), for more details. Alternative estimators for β can be seen in Caeiro and Gomes ([8]) and Gomes *et al.* ([34]).

4. TESTING EXTREME VALUE CONDITIONS

In any of the aforementioned procedures it is assumed that the underlying d.f. F belongs to $\mathcal{D}_{\mathcal{M}}(EV_\xi)$, for an appropriate value of ξ, or it is in a specific sub-domain of $\mathcal{D}_{\mathcal{M}}(EV_\xi)$. This condition is called the *extreme value condition* and is not always fulfilled. So, before performing an application, it is important to check whether the extreme value condition is reasonable for a data set or not. So, we must test the hypothesis:

$$H_0:\quad F \in \mathcal{D}_{\mathcal{M}}(EV_\xi) \quad \text{for some } \xi \in \mathbb{R}\,.$$

Dietrich *et al.* ([18]) proposed the E, PE tests (if we assume $\xi \geq 0$) and Drees *et al.* ([19]) proposed the T test (assuming $\xi > -1/2$).

Let $X_1, X_2, ..., X_n$ be independent random variables with d.f. F and suppose that some additional second order conditions hold, then for $\eta > 0$ the corresponding test statistics are:

$$E_n := k \int_0^1 \left(\frac{\log X_{n-\lfloor kt \rfloor,n} - \log X_{n-k,n}}{\widehat{\xi}_+} - \frac{t^{-\widehat{\xi}_-} - 1}{\widehat{\xi}_-} \left(1 - \widehat{\xi}_-\right) \right)^2 t^\eta \, dt \,, \tag{4.1}$$

$$PE_n := k \int_0^1 \left(\frac{\log X_{n-\lfloor kt \rfloor,n} - \log X_{n-k,n}}{\widehat{\xi}_+} + \log t \right)^2 t^\eta \, dt \,, \tag{4.2}$$

$$T_n := k \int_0^1 \left(\frac{n}{k} \overline{F}_n \left(\widehat{a}_{n/k} \frac{x^{-\widehat{\xi}} - 1}{\widehat{\xi}} + \widehat{b}_{n/k} \right) - x \right)^2 x^{\eta-2} \, dx \,, \tag{4.3}$$

where the estimates for ξ_+ and ξ_- are obtained through the moment estimators in Dekkers *et al.* ([17]), and k is again an intermediate sequence, $k = k_n \to \infty$, $k/n \to 0$ and $k^{1/2} A(n/k) \to 0$ as $n \to \infty$. A is related to the second order condition. Hüsler and Li ([45]) present an algorithm for testing H_0 using the test statistic E_n in (4.1). They have carried out an extensive simulation study with guidelines for obtaining the value of η and have provided tables of critical values. See also Neves and Fraga Alves ([48]) for a description of those tests.

5. STATISTICAL CHOICE OF EXTREME DOMAINS OF ATTRACTION — SEMI-PARAMETRIC APPROACH

In a semi-parametric framework, ξ is the primordial parameter since determines the shape of the tail of the underlying distribution function F. A negative value for ξ is associated to the Weibull domain of attraction in which all the d.f.'s are short tailed with finite right endpoint. If $\xi > 0$ we have the Fréchet domain of attraction to which the heavy tailed d.f.'s with polynomially decaying tail belong. The case of $\xi = 0$ is particularly important, due to the simplicity of inference, within the Gumbel domain which contains a great variety of d.f.'s with an exponential tail having finite right end point or not. Whenever we intend to perform a statistical inference in extreme values we should look for the most adequate procedures according to the domain of attraction selected. Therefore, it is of great benefit to test the Gumbel domain against the Fréchet or Weibull domains. The hypothesis to test is:

$$H_0\colon\ F \in \mathcal{D}_{\mathcal{M}}(EV_0) \quad vs. \quad H_1\colon\ F \in \mathcal{D}_{\mathcal{M}}(EV_\xi)_{\xi \neq 0} \;; \tag{5.1}$$

or *versus* the one-sided alternatives $F \in \mathcal{D}_{\mathcal{M}}(EV_\xi)_{\xi<0}$ or $F \in \mathcal{D}_{\mathcal{M}}(EV_\xi)_{\xi>0}$.

Several tests have been proposed in the literature, among which we can mention Galambos ([28]), Castilho *et al.* ([12]), Hasofer and Wang ([42]), Falk ([21]), Fraga Alves and Gomes ([23]) and Correia and Neves ([14]), that have proposed a slight modification of the Hasofer and Wang statistic, Marohn ([46, 47]), Fraga Alves ([25]) and Segers and Teugels ([56]). More recently Brilhante ([6]) derived a resistant and robust test for the exponential *versus* the generalized Pareto, Neves and Fraga Alves ([48]) introduced three tests statistics based on the reformulation of the Hasofer and Wang statistic. Those tests were later studied in Neves and Fraga Alves ([49]). Castillo *et al.* ([11]) provided a test based on the properties of the *coefficient of variation.*

In this work the tests introduced in Neves and Fraga Alves ([48]) will be considered. The statistics for testing (5.1) are based on the k excesses over the $(n-k)$-th ascending intermediate order statistic $X_{n-k:n}$, Thus, under the null hypothesis of Gumbel domain of attraction and further assuming: (i) second order conditions on the upper tail of F and (ii) the intermediate sequence $k \equiv k_n$, such that $k^{1/2}A(n/k) \to 0$ as $n \to \infty$ where A is related to the second order condition, Neves and Fraga Alves ([48]) have defined the following tests:

The ratio-test

$$R_n^* := \frac{X_{n:n} - X_{n-k:n}}{\frac{1}{k}\sum_{i=1}^{k}\big(X_{n-i+1:n} - X_{n-k:n}\big)} - \log k \;\xrightarrow[n\to\infty]{d}\; \Lambda \;; \tag{5.2}$$

The GT-test

$$\begin{aligned} G_n(k) &:= \frac{\frac{1}{k}\sum_{i=1}\big(X_{n-i+1:n} - X_{n-k:n}\big)^2}{\left(\frac{1}{k}\sum_{i=1}^{k} X_{n-i+1:n} - X_{n-k:n}\right)^2} \,, \\ G_n^*(k) &= \sqrt{k/4}\,\big(G_n(k) - 2\big) \;\xrightarrow[n\to\infty]{d}\; \mathcal{N}(0,1) \;; \end{aligned} \tag{5.3}$$

The HW-test

$$\begin{aligned} W_n(k) &:= \frac{1}{k}\left[1 - \frac{G_n(k) - 2}{1 + \big(G_n(k) - 2\big)}\right] , \\ W_n^*(k) &= \sqrt{k/4}\,\big(k\,W_n(k) - 1\big) \;\xrightarrow[n\to\infty]{d}\; \mathcal{N}(0,1) \,, \end{aligned} \tag{5.4}$$

where Λ is a Gumbel random variable.

The null hypothesis in (5.1) is rejected if $T_n^* < \chi_{\alpha/2}$ or $T_n^* > \chi_{1-\alpha/2}$, where T^* has to be replaced by R^*, G^* or W^* and χ_p is the p probability quantile of the corresponding distribution.

If we are interested in the one-sided tests, and being χ_p the p probability quantile of the corresponding distribution, the critical regions for:

Gumbel vs Weibull domain of attraction are:

$$R_n^*(k) < \chi_\alpha\ , \qquad G_n^*(k) < \chi_\alpha\ , \qquad W_n^*(k) > \chi_{1-\alpha}\ ; \tag{5.5}$$

Gumbel vs Fréchet domain of attraction are:

$$R_n^*(k) > \chi_{1-\alpha}\ , \qquad G_n^*(k) > \chi_{1-\alpha}\ , \qquad W_n^*(k) < \chi_\alpha\ . \tag{5.6}$$

As an illustration of the methodologies reviewed in the previous sections and also for showing some functions available in the R software, a real data set will be studied in the next section.

6. A CASE STUDY: DAILY MEAN FLOW DISCHARGE RATE

Here we will focus our attention on the estimation of the EVI. Packages and/or functions available in the R environment will be used and mentioned. R software contains already a large number of packages with several functions for modelling extreme data, such as `evd`, `ismev`, `evir`, `POT`, `fExtremes`, `evdbayes`, `copula`, `SpatialExtremes`, among others. Gilleland *et al.* ([29]) give an excellent software review for extreme value analysis. They describe and compare packages available in R with other software.

6.1. A preliminary data analysis

Our data set consists of daily mean flow discharge rate in the hydrometric station of Fragas da Torre in the river Paiva. The source of this river is in the Serra de Leomil, in the north of Portugal, it is an effluent of the river Douro, with a watershed area of approximately 700 Km. More precisely the data set studied is the daily mean flow discharge rate values (m^3/s) from 1 October 1946 to 30 September 2006, collected from the "SNIRH: Sistema Nacional de Informação dos Recursos Hídricos" and the interest is to analyse the extreme values.

After some previous graphical analyses on the empirical tail behaviour of the different months showing the occurrence of the maximum values, advices of hydrologists and taking into account a previous work that considered a few initial years of these data, Gomes ([32]), only the months from November until April were used in each year. We had then a total of 10 860 daily mean flow discharge rate values. The results of a preliminary graphical and descriptive analysis are shown in Figure 1 and Table 1.

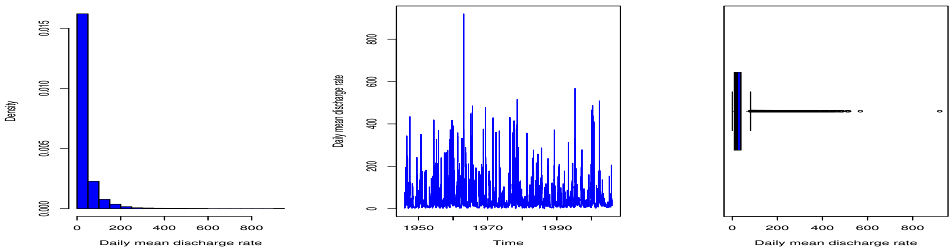

Figure 1: Histogram (left); chronogram (center) and boxplot (right).

Table 1: Basic descriptive statistics for the data.

n	Min	1st Quart.	Median	Mean	3rd Quart.	Max	St Dev.	Skew.	Kurt.
10860	0	9.20	17.30	34.83	38.00	920.00	50.92	4.15	27.31

The stationarity was also studied by the Augmented Dickey–Fuller Test through the function `adf.test()`, available in the *package* `tseries`. The boxplot, the histogram and the descriptives statistics, in particular the skewness = 4.15 and the kurtosis = 27.31 indicate a tail heavier than the normal one.

6.2. Testing extreme value conditions

Following the brief introduction given in section 4, we will use here the test E, Dietrich *et al.* ([18]) and Hüsler and Li ([45]). The function `MTestEVC1d()` in the *package* `TestEVC1d` gave the results shown in Figure 2. We observe that the values

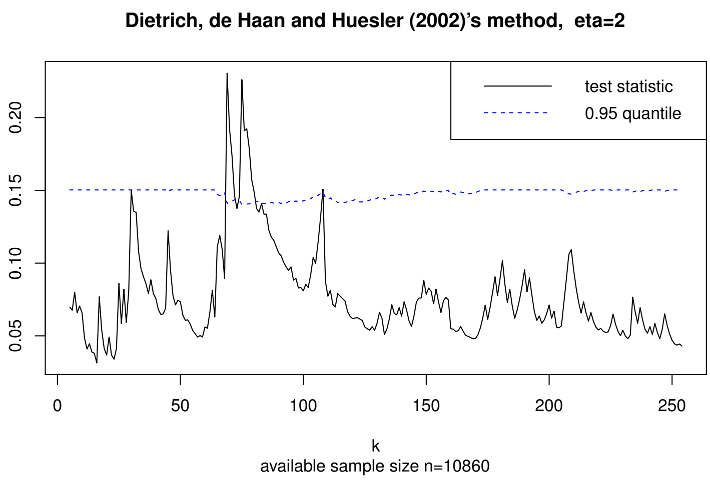

Figure 2: Sample path of E statistic.

of the test statistic E are smaller than the corresponding asymptotic 0.95-quantile for a large range of k-values. So, since the sample path of the test statistic is almost always outside the rejection region, except for a small range of k, we find no evidence to reject the null hypothesis.

6.3. Parametric framework

The BM methodology

In this framework, we have considered the years as blocks of observations and have picked the maximum values up in each block. So, we will use the maximum values of each of 60 years — these are all the years available in SNIRH: "Sistema Nacional de Informação dos Recursos Hídricos" for the hydrometric station of Fragas da Torre in river Paiva.

We have now obtained the skewness = 0.998 and the kurtosis = 2.265. Graphical analyses are shown in Figure 3.

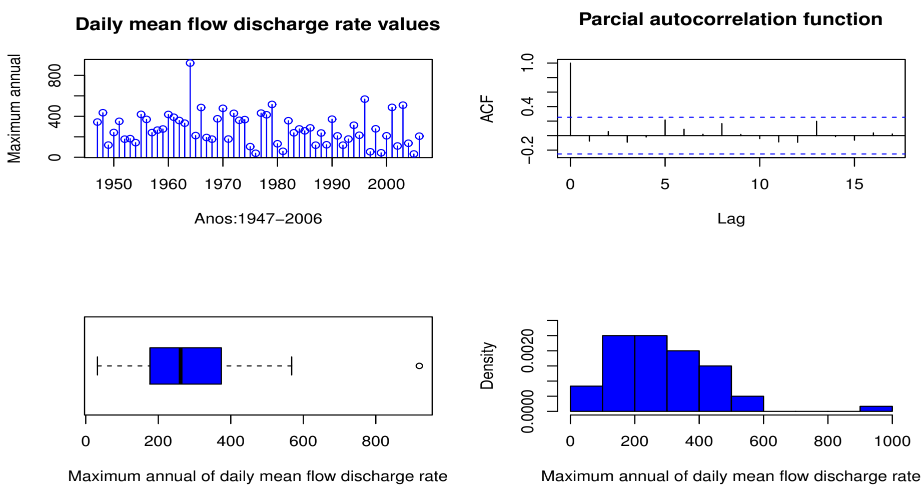

Figure 3: Chronogram (top left); ACF (top right); boxplot (bottom left) and histogram (bottom right).

The histogram, the boxplot and the skewness indicate a moderate positive asymmetry. From the autocorrelation partial function (ACF), it seems reasonable to assume that these data are not correlated. So an EV_ξ was fitted to the maxima in each year.

The ML fitting for the EV distribution for all the parameters can be obtained through the *package* `evd` and the function `fgev()`, see Table 2. The parameter estimates by the PWM method can be obtained using the *package* `fExtreme` and running `gevFit( ,type="pwm")`[1]. See results in Table 2.

Table 2: Parameters estimates (standard errors in parenthesis) and the profile Log-Likelihood (pLog-L) 95% confidence intervals.

	$\widehat{\lambda}$	$\widehat{\delta}$	$\widehat{\xi}$
ML	210.08 (18.77)	129.81 (13.62)	-0.03 (0.09)
PWM	213.65	137.37	-0.09

	λ	δ	ξ
pLog-L	(174.15; 248.23)	(106.41; 160.79)	(-0.16; 0.19)

Using the same *package*, Wald confidence intervals of level $1-\alpha$, can be obtained through `confint(fgvev(), level=`$1-\alpha$`)`. Greater accuracy for the confidence intervals is usual attained by the profile log-likelihood. Plots for the profile log-likelihood for all parameters can be obtained by `plot(profile(fgev()),ci = c(0.95, 0.99))`. The confidence interval limits can be obtained through `confint(profile(fgvev()),level=`$1-\alpha$`)` and are given in Table 2.

Notice that the confidence intervals for ξ include zero, so lead to not reject the null hypothesis, $\xi = 0$.

The POT methodology

The POT method is based on fitting the statistical model in (2.4) to the *excesses* over a given threshold u. A challenge here is the choice of u. Choosing a value too high can lead to a very small number of observations in the tail resulting in estimators with high variance, but a small threshold may lead to the violation on the Pickands Theorem.

The most traditional methods for the choice of u are graphical procedures. A graph widely used is the *mean residual life* (mrl) plot, based on the mean value of the GP distribution, which is a linear function of u. If the GP model is valid for the excesses above u_0 then will also be valid for all $u > u_0$. So, this graph should show a linear behaviour above a suitable choice of the threshold u. Another graphical method is based on the *threshold choice* (tc) plot, which represents the estimated values of the GP model over a set of thresholds. The threshold u will be a "good'" choice if the parameter estimates appear approximately constant

[1]`gevFit`() function can also determine the maximum likelihood estimates, setting type="mle".

above u. The function `mrlplot()` in the *package* `evd` plots the mean excess plot, and the function `tcplot()` plots two graphs for both parameters, see Figure 4.

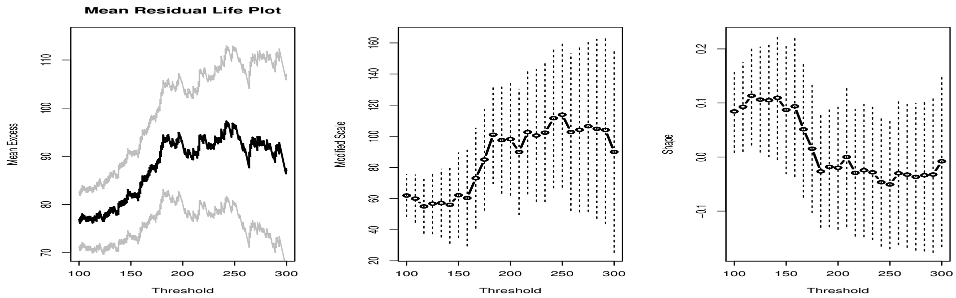

Figure 4: The mean residual life plot (left) and the tc plots (centre and right).

A threshold around 200 is suggested. We have chosen $u = 180$, corresponding to a number of 254 exceedances. Figure 5 shows those exceedances, no correlation of the exceedances and the asymmetry of the data. Using the function `gpd()` in *package* `evir`, we got similar results to those by the BM method, see Table 3.

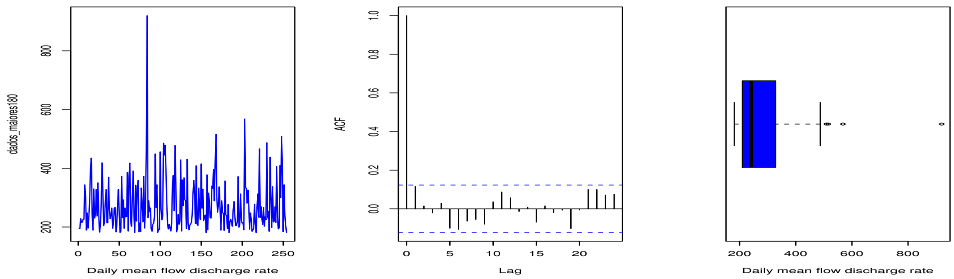

Figure 5: Chronogram (left) with u=180, partial autocorrelation function (center) and boxplot (right).

Table 3: Parameters Estimates (standards errors in parenthesis) and the profile Log-Likelihood (pLog-L) 95% Confidence Intervals.

	$\hat{\hat{\delta}}$	$\hat{\xi}$
ML	94.69 (7.83)	-0.02 (0.05)
PWM	95.47(9.47)	-0.03 (0.07)

	δ	ξ
pLog-L	(80.01;110.80)	(-0.10;0.11)

Note again that the results obtained indicate a value for ξ close to zero.

6.4. Semi-parametric framework

In this approach ξ is the primordial parameter to be estimated. As we referred to in section 3.2, estimates are based on the k top order statistics in the sample, with k an intermediate sequence, assuming that the underlying distribution function F belongs to $\mathcal{D}_{\mathcal{M}}(EV_\xi)$, for an appropriate value of ξ. Since, there are specific estimation procedures according to the signal of ξ, we should start this framework by testing the Gumbel max-domain against Fréchet or Weibull max-domains.

The choice of the tail

To test $H_0: F \in \mathcal{D}_{\mathcal{M}}(EV_0)$ *vs.* $H_1: F \in \mathcal{D}_{\mathcal{M}}(EV_\xi)_{\xi \neq 0}$ or against the one-sided alternatives $F \in \mathcal{D}_{\mathcal{M}}(EV_\xi)_{\xi<0}$ or $F \in \mathcal{D}_{\mathcal{M}}(EV_\xi)_{\xi>0}$ we will consider the *Ratio-test*, the *Gt-test* and the *HW-test*, mentioned in section 5. Figure 6 presents the sample paths of G^*, R^* and W^* for several values of k. As we can see in Figure 6, for a large range of k-values, the three tests statistics present values that belong to the corresponding region of no rejection. So we find no evidence to reject the null hypothesis, $F \in D_M(EV_0)$.

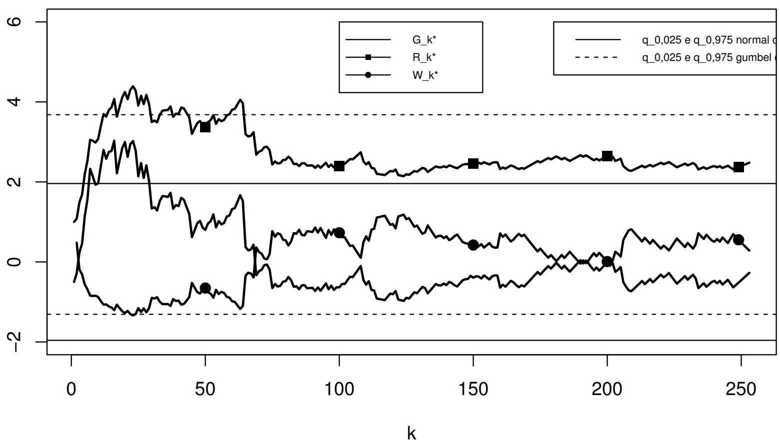

Figure 6: Sample paths of G^*, R^* and W^* statistics.

Some semi-parametric estimates

As specified in Section 3.2, we will consider here the Hill estimator, the Moment estimator, the Generalized Hill estimator, the Mixed Moment estimator and the MVRB estimator. Although having been led above to the non rejection of the Gumbel domain of attraction we present here the results of application of all those estimators.

Figure 7 shows the sample paths of the estimates obtained for each k. It is worthwhile to mention that the Hill estimator and the MVRB estimator, specifically built for $\xi > 0$ show results that are far from those previously obtained (notice that the MVRB estimates show a very stable path, but around positive values of $\widehat{\xi}$). The other estimators present sample paths near $\xi = 0$.

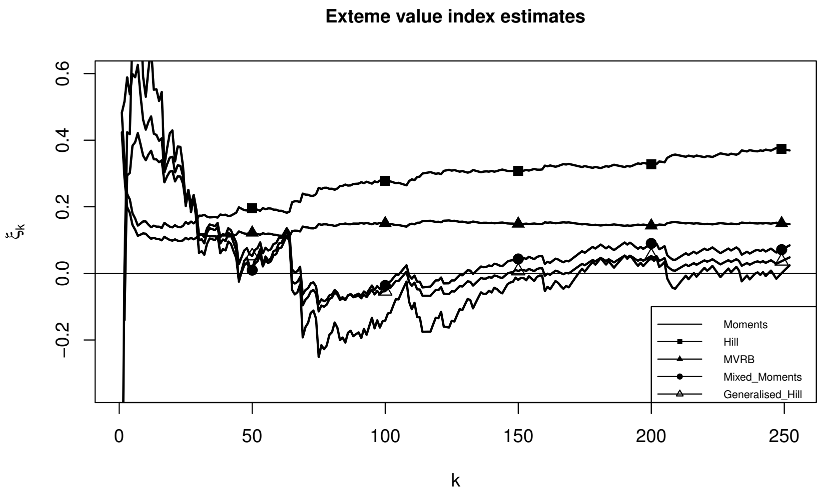

Figure 7: Sample paths of ξ estimates.

7. A FEW OVERALL COMMENTS

Testing whether $F \in \mathcal{D}_{\mathcal{M}}(EV_\xi)$, for a certain ξ, is a crucial topic when an application of extreme values procedures is needed to be considered. This subject has been dealt in several articles mentioned along this paper. However, several times a real problem in the area of EVT is studied without that previous analysis.

With the study of this application we intended to motivate the discussion regarding the need of a previous analysis on the choice of the tail before applying the well theoretically studied estimators. The influence of the estimate of the tail index parameter in the estimation of high quantiles, parameters of major interest for preventing catastrophes that can occur in this domain of application, is also another important issue, however out of the scope of this study.

ACKNOWLEDGMENTS

Research partially supported by National Funds through **FCT** — Fundação para a Ciência e a Tecnologia, projects PEst-OE/MAT/UI0006/2014 (CEAUL).

REFERENCES

[1] Balkema, A.A. and de Haan, L. (1974). Residual life time at great age, *Ann. Probab.*, **2**, 792–804.

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

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

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

[5] Beran, J.; Schell, D. and Stehlik, M. (2014). The harmonic moment tail index estimator: asymptotic distribution and robustness, *Annals of the Institute of Statistical Mathematics*, **66**(1), 193–220.

[6] Brilhante, M.F. (2004). Exponentiality versus generalized Pareto — a resistant and robust test, *RevStat*, **2**(1), 1–13.

[7] Brilhante, M.F.; Gomes, M.I. and Pestana, D. (2013). A simple generalization of the Hill estimator, *Comput. Stat. Data Anal.*, **57**(1), 518–535.

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

[9] Caeiro, F.; Gomes, M.I. and Henriques-Rodrigues, L. (2009). Reduced-bias tail index estimators under a third order framework, *Comm. Statist. Th. Meth.*, **38**(7), 1019–1040.

[10] Caeiro, F.; Gomes, M.I. and Pestana, D.D. (2005). Direct reduction of bias of the classical Hill estimator, *Revstat*, **3**, 111–136.

[11] Castillo, J. del; Daoudi, J. and Lockhart, R. (2014). Methods to distinguish between polynomial and exponential tails, *Scandinavian Journal of Statistics*, **41**, 382–393.

[12] Castillo, E.; Galambos, J. and Sarabia, J.M. (1989). *The selection of the domain of attraction of an extreme value distribution from a set of data.* In "Extreme value theory (Oberwolfach, 1987) — Lecture Notes in Statistics" (J. Häusler and R.-D. Reiss, Eds.), Springer, Berlin-Heidelberg, **51**, 181–190.

[13] Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, London.

[14] Correia, A.L. and Neves, M. (1996). *Escolha estatística em modelos extremais — testes de ajustamento.* In "Bom Senso e Sensibilidade" (J. Branco, P. Gomes e J. Prata, Eds.), Actas do III Congresso Anual da Sociedade Portuguesa de Estatística, Edições Salamandra, 223–236.

[15] Davison, A.C. (1984). *Modelling excesses over high threshold, with an application.* In "Statistical Extremes and Applications" (J. Tiago de Oliveira, Ed.), Reidel, Dordrecht, 389–410.

[16] Davison, A.C. and Smith, R.L. (1990). Models for exceedances over high thresholds (with discussions), *J. Royal Statistical Society*, **52**(3), 393–442.

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

[18] Dietrich, D.; De Haan, L. and Hüsler, J. (2002). Testing extreme value conditions, *Extremes*, **5**(1), 71–85.

[19] Drees, H.; De Haan, L. and Li, D. (2006). Approximation to the tail empirical distribution function with application to testing extreme value conditions, *J. Stat. Plan. Inference*, **136**, 3498–3538.

[20] Dwass, M. (1964). Extremal processes, *Ann. Math. Stat.*, **35**, 1718–1725.

[21] Falk, M. (1995). On testing the extreme value index via the POT-method, *Annals of Statistics*, **23**, 2013–2035.

[22] Fisher, R.A. and Tippett, L.H.C. (1928). On the estimation of the frequency distributions of the largest or smallest member of a sample, *Proceedings of the Cambridge Philosophical Society*, **24**, 180–190.

[23] Fraga Alves, M.I. and Gomes, M.I. (1996). Statistical choice of extreme value domains of attraction — a comparative analysis, *Comm. in Stat. Th. and Meth.*, **25**(4), 789–811.

[24] Fraga Alves, M.I.; Gomes M.I. and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter, *Portugaliae Mathematica*, **60**, 194–213.

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

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

[27] Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum, *Ann. Soc. Polon. Math. (Cracovie)*, **6**, 93–116.

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

[29] Gilleland, E.; Ribatet, M. and Stephenson, A.G. (2012). A software review for extreme value analysis, *Springer (Springerlink.com). Published online: 20 July 2012)*, DOI 10.1007/s10687-012-0155-0.

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

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

[32] Gomes, M.I. (1993). On the estimation parameters of rare events in environmental times series, *Statistics for the Environment*, 226–241.

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

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

[35] 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**, 5–31.

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

[37] Gomes, M.I.; Martins, M.J. and Neves, M.M. (2013). Generalized Jackknife-Based Estimators for Univariate Extreme-Value Modeling, *Comm. Statist. Th. Meth.*, **42**(7), 1227–1245.

[38] Greenwood, J.A.; Landwehr, J.M.; Matalas, N.C. and Wallis, J.R. (1979). Probability-weighted moments: definition and relation to parameters of several distributions expressable in inverse form, *Water Resources Research*, **15**, 1049–1054.

[39] Gumbel, E.J. (1935). Les valeurs extrêmes des distributions statistiques, *Ann. Inst. Henri Poincaré*, **5**(2), 115–158.

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

[41] de Haan, L. (1970). *On Regular Variation and its Applications to the Weak Convergence of Sample Extremes*, Mathematical Centre Tract 32, Amsterdam, Dordrecht: D. Reidel.

[42] Hasofer, A. and Wang, J.Z. (1992). A test for extreme value domain of attraction, *Journal of the American Statistical Association*, **87**, 171–177.

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

[44] 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.

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

[46] Marohn, F. (1998a). An adaptive efficient test for Gumbel domain of attraction, *Scandinavian Journal of Statistics*, **25**, 311–324.

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

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

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

[50] Neves, M.M.; Penalva, H. and Nunes, S. (2015). *Extreme value analysis of river levels in a hydrometric station in the North of Portugal.* In "Current Topics on Risk Analysis: ICRA6 and RISK 2015 Conference proceedings" (M. Guillén, A. Juan, H. Ramalhinho, I. Serra and C. Serrat, Eds.), 533–538.

[51] Paulauskas, V. and Vaiciulis, M. (2013). On the improvement of Hill and some others estimators, *Lith. Math. J.*, **53**, 336–355.

[52] Penalva, H.; Neves, M. and Nunes, S. (2013). Topics in data analysis using R in extreme value theory, *Advances in Methodology & Statistics / Metodoloski zvezki*, **10**(1), 17–29.

[53] Penalva, H.; Neves, M. and Nunes, S. (2014). *Statistical Modeling and Inference in Extremes: Applications with R.* In "Statistical and Biometrical Challenges: Theory and Applications, Biometrie und Medizinische Informatik — Greifswalder Seminarberichte" (T. Oliveira, K.-E. Biebler, A. Oliveira and B. Jager, Eds.), Shaker-Verlag, Aachen, 281–309.

[54] Pickands, J. (1975). Statistical inference using extreme order statistics, *Annals of Statistics*, **3**, 119–131.

[55] R Core Team (2013). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. `http://www.R-project.org/`

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

[57] Smith, R.L. (1984). *Threshold methods for sample extremes.* In "Statistical Extremes and Applications" (J. Tiago de Oliveira, Ed.), Reidel, Dordrecht, 621–638.

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

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

[60] von Mises, R. (1936). La distribution de la plus grande de n valeurs, *American Mathematical Society*, Reprinted in Selected Papers Volumen II, Providence, R.I. (1954), 271–294.

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

[62] Zhou, C. (2009). Existence and consistency of the maximum likelihood estimator for the extreme value index, *J. Multivariate Anal.*, **100**(4), 794–815.

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