Modeling Extreme Events: Sample Fraction Adaptive Choice in Parameter Estimation*

M. Manuela Neves

Universidade Técnica de Lisboa, ISA, and CEAUL; email:manela@isa.utl.pt

M. Ivette Gomes

Universidade de Lisboa, FCUL, DEIO and CEAUL; email:migomes@fc.ul.pt

Fernanda Figueiredo

Universidade do Porto, FEP, and CEAUL; email: otilia@fep.up.pt

Dora Prata Gomes

Universidade Nova de Lisboa, FCT and CMA; email:dsrp@fct.unl.pt

Abstract

When modeling extreme events there are a few primordial parameters, among which we refer the extreme value index and the extremal index. The extreme value index measures the right tail-weight of the underlying distribution and the extremal index characterizes the degree of local dependence in the extremes of a stationary sequence.

*Research partially supported by National Funds through **FCT**—Fundação para a Ciência e a Tecnologia, projects PEst-OE/MAT/UI0006/2011 (CEAUL), PEst-OE/MAT/UI0297/2011 (CMA/UNL) and EXTREMA, PTDC/MAT/101736/2008.

Most of the semi-parametric estimators of these parameters shows the same type of behavior: nice asymptotic properties, but a high variance for small values of k, the number of upper order statistics used in the estimation, and a high bias for large values of k. This brings a real need for the choice of k. Choosing some well-known estimators of those two parameters we revisit the application of a heuristic algorithm for the adaptive choice of k. A simulation study illustrates the performance of the proposed algorithm.

AMS Subject Classification: Primary 62G32; Secondary 65C05
Key words: adaptive choice, extremal index, extreme value index, sample fraction, semi-parametric estimation

1 Introduction and preliminaries

In *Extreme Value Theory* (EVT) we deal essentially with the estimation of parameters of extreme or rare events. A large number of applications in areas such as hydrology, telecommunications, finance, biology and environmental studies, reveals the need for an adequate estimation of those parameters.

The classical assumption in EVT is that we have a set of independent and identically distributed (i.i.d.) random variables (r.v.'s), $X_1, \ldots, X_n$, from an unknown distribution function (d.f.) F and we are concerned with the limit behavior of either $M_n \equiv X_{n:n} = \max(X_1, \ldots, X_n)$ or $m_n \equiv X_{1:n} = \min(X_1, \ldots X_n)$ as $n \to \infty$. Thinking on maximum values, whenever it is possible to linearly normalize M_n so that we get a non-degenerate limit, as $n \to \infty$, such a limit is of the type of the extreme value (EV) d.f.,

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

We then say that F is in the domain of attraction for maxima of EV_γ, denoting this by $F \in D_{\mathcal{M}}(EV_\gamma)$.

There is a large variety of parameters of extreme events, but the estimation of the *extreme value index*, the parameter denoted γ, in (1), is of primordial importance by itself and because it is the basis for the estimation of all other parameters of extreme events. Among the most relevant parameters of extreme events, and assuming that we are interested in large values, i.e. in the right tail of the underlying model F, we can mention:

- the *probability of exceedance* of a high level $x \equiv x_{_H}$, $p_x := \mathbb{P}(X > x) = 1 - F(x)$;
- the *return period* of a high level x, which is given by $r_x := 1/(1-F(x))$ in any i.i.d. scheme;
- the *right endpoint* of an underlying model F, $x^* \equiv x^F := \sup\{x : F(x) < 1\}$;
- a *high quantile* of probability $1-p$, p small, situated in the border or even beyond the range of the available data, defined as $\chi_{1-p} := \inf\{x : F(x) \geq 1-p\} =: F^{\leftarrow}(1-p) = U(1/p)$, $p < 1/n$, where $U(t) = F^{\leftarrow}(1-1/t)$, $t \geq 1$ is the associated reciprocal quantile function. In financial frameworks χ_{1-p} is known as *the Value at Risk* at a level p (VaR_p).

The EV_γ d.f., in (1), incorporates the three Fisher-Tippett types:

- Gumbel: $\Lambda(x) = EV_0(x) = \exp(-\exp(-x))$, $x \in \mathbb{R}$, $(\gamma = 0)$, the limit for exponential tailed distributions;
- Fréchet: $\Phi_\alpha(x) = EV_{1/\alpha}(\alpha(x-1)) = \exp(-x^{-\alpha})$, $\alpha > 0$, $x > 0$, $(\gamma = 1/\alpha > 0)$, the limit for heavy tailed distributions;
- Weibull: $\Psi_\alpha(x) = EV_{-1/\alpha}(\alpha(x+1)) = \exp(-(-x)^{\alpha})$, $\alpha > 0$, $x < 0$, $(\gamma = -1/\alpha < 0)$, the limit for short tailed distributions.

We can consider location and scale parameters, $\mu \in \mathbb{R}$ and $\sigma > 0$, respectively, in the EV_γ d.f., denoting it by $EV_\gamma(x; \mu, \sigma) = EV_\gamma((x - \mu)/\sigma)$. To say that $F \in D_{\mathcal{M}}(EV_\gamma)$ means that for large values of n we can consider the approximation $P[X_{n:n} \leq x] = F^n(x) \approx EV_\gamma((x - b_n)/a_n)$, for adequate $a_n > 0$ and $b_n \in \mathbb{R}$. The shape parameter γ is directly related to the weight of the right tail, $\overline{F} := 1 - F$, of the underlying model F. As γ increases the right tail becomes heavier and heavier. Figure 1 shows the behavior of the right-tails for the three different types of EV models, and the Gauss model for comparison.

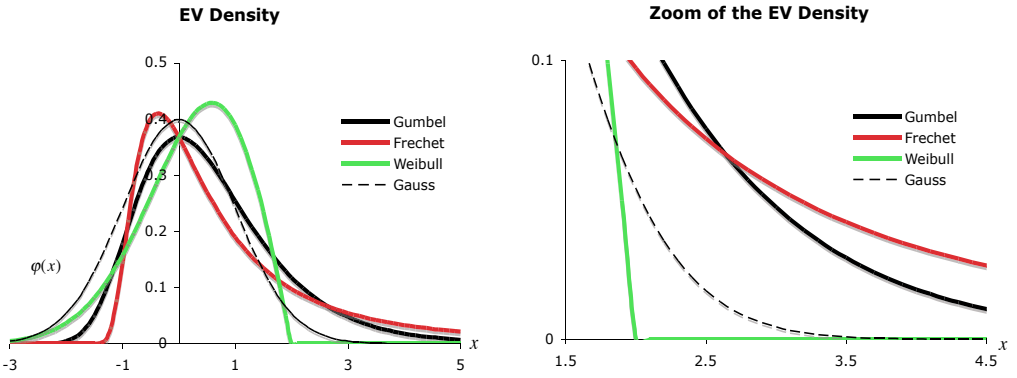

Figure 1: Gumbel, Fréchet, Weibull and Gauss p.d.f.'s (left) and a zoom of the right tail of those p.d.f.'s (right).

- If $\gamma = 0$, the right tail is of an exponential type. The *right endpoint* can then be either finite or infinite;
- If $\gamma > 0$, the right tail is heavy, of a negative polynomial type, and F has an infinite *right endpoint*;
- If $\gamma < 0$, the right tail is light, and F has a finite *right endpoint*, $x^* < +\infty$.

However in many practical applications extreme conditions often persist over several consecutive observations, i.e. the random variables are no longer

independent. If $\{Y_n\}_{n\geq 1}$ is a stationary dependent sequence coming from an underlying d.f. F such that some general local and asymptotic dependence conditions hold, the limiting process of exceedances of high levels, after a suitable normalization is a compound Poisson process, i.e. there is a clustering of exceedances of high levels. The underlying Poisson values represent cluster positions and the multiplicities are the clusters sizes. In such a situation it appears a new parameter that needs to be taken into account in any inferential procedure. Such a parameter, called the *extremal index* and usually denoted by θ, has an important role in the obtention of the sizes of the clusters of exceedances.

Let us consider the following example:

Example 1.1. *Let $\{X_n\}_{n\geq 1}$ be a sequence of i.i.d. variables from the model $F(x) = (1-\exp(-x))^2$, $x \geq 0$ and $\{Y_n\}_{n\geq 1}$ the two-dependent sequence defined as $Y_n = \max(Z_{n+1}, Z_n)$, $n \geq 1$, where $\{Z_i\}_{i\geq 1}$ are exponential i.i.d. random variables with d.f. $H(z) = 1-\exp(-z)$, $z \geq 0$. The underlying model for Y_n is then also given by $F(y) = (1-\exp(-y))^2$, $y \geq 0$.* Figure 2 *shows a size equal to 2 for the clusters of exceedances of high levels by the $\{Y_n\}$ sequence. We shall see that this implies an extremal index $\theta = 1/2 = 0.5$. It can also be seen a shrinkage of the largest observations for the 2-dependent sequence, despite of the fact that we have the same model underlying both sequences.*

For the *extremal index* θ, $0 \leq \theta \leq 1$, directly related to the clustering of exceedances, we have $\theta = 1$ for i.i.d. sequences and $\theta \to 0$ whenever dependence increases. The case $\theta = 0$ appears in pathological situations. For 'almost all cases' of interest, we have $\theta > 0$.

If the stationary sequence $\{Y_n\}_{n\geq 1}$ comes from an underlying d.f. F, and $\{X_n\}_{n\geq 1}$ is the associated i.i.d. sequence, i.e. an i.i.d. sequence from the same model F, then, under general asymptotic and local dependence restrictions as, e.g., both the long-range dependence condition D, in Lead-

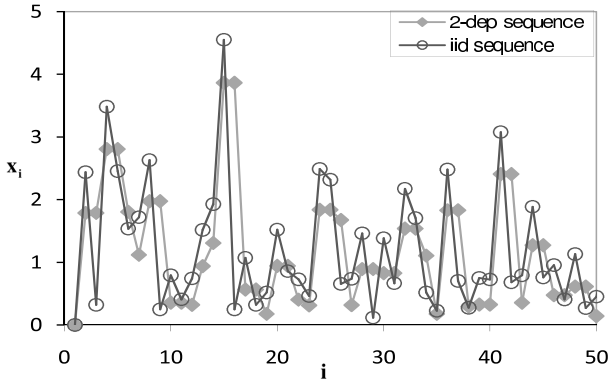

Figure 2: Sample paths from an i.i.d. process and a 2-dependent process.

better *et al.* (1983), and the local dependence condition D', in Leadbetter and Nandagopalan (1989), the limiting d.f. of the maximum $Y_{n:n} = \max(Y_1, \ldots, Y_n)$ may be directly related to the limiting d.f. of the maximum, $X_{n:n}$, of the i.i.d. associated sequence, through the *extremal index*.

Definition 1.1 (Leadbetter *et. al.*, 1983)**.** *The stationary sequence* $\{Y_n\}_{n\geq 1}$ *is said to have an extremal index* $\theta \in (0, 1]$ *if, for all* $\tau > 0$*, there exists a sequence of levels* $(u_n(\tau))_{n\in N}$*, such that*

$$\lim_{n\to\infty} P[X_{n:n} \leq u_n(\tau)] = \lim_{n\to\infty} F^n(u_n(\tau)) = e^{-\tau}$$

$$\textit{and} \quad \lim_{n\to\infty} P[Y_{n:n} \leq u_n(\tau)] = e^{-\theta\tau}.$$

From this definition we see that the *extremal index*, θ, can be informally defined by the approximation

$$P[Y_{n:n} \leq x] \approx F^{n\theta}(x) \approx EV_\gamma^\theta \left(\frac{x - a_n}{b_n}\right) = EV_\gamma \left(\frac{x - a'_n}{b'_n}\right),$$

$$\text{where} \quad \begin{cases} b'_n &= b_n \theta^\gamma \\ a'_n &= a_n + b_n \left(\frac{\theta^\gamma - 1}{\gamma}\right) \end{cases}.$$

So, as a consequence of the stability for maxima of the EV d.f., the limiting d.f. of $Y_{n:n}$, linearly normalized, is still an EV d.f. It is what is established in the following theorem:

Theorem 1.1 (Leadbetter *et. al.*, 1983)**.** *Let* $\{Y_n\}_{n\geq 1}$ *be a stationary sequence with marginal distribution* F*,* $\{X_n\}_{n\geq 1}$ *an i.i.d. sequence of r.v.'s with the same distribution* F*,* $M_n = \max(Y_1, \cdots, Y_n)$ *and* $\widetilde{M}_n = \max(X_1, \cdots, X_n)$*. Under the* $D(u_n)$ *condition, with* $u_n = a_n x + b_n$*,*

$$Pr\left\{(\widetilde{M}_n - b_n)/a_n \leq x\right\} \underset{n\to\infty}{\longrightarrow} G_1(x)$$

as $n \longrightarrow \infty$*, for normalizing sequences* $\{a_n > 0\}$ *and* $\{b_n\}$*, if and only if,*

$$Pr\left\{(M_n - b_n)/a_n \leq x\right\} \underset{n\to\infty}{\longrightarrow} G_2(x)$$

where $G_2(x) = G_1^{\theta}(x)$*, for* $0 < \theta \leq 1$*.*

So, being $G_1(\cdot) \equiv EV_\gamma(\cdot)$ an extreme value d.f., the limit law $G_2(\cdot) \equiv EV_\gamma^\theta(\cdot)$ is a extreme value d.f. with location, scale and shape parameters $(\mu_\theta, \sigma_\theta, \gamma_\theta)$ given by

$$\mu_\theta = \mu - \sigma\frac{1-\theta^\gamma}{\gamma}, \quad \sigma_\theta = \sigma\theta^\gamma \quad \text{and} \quad \gamma_\theta = \gamma.$$

The *extremal index* estimation is then important not only by itself but also because of its influence in the estimation of other important parameters of rare events.

2 Extreme value index and extremal index estimation

Here, we consider models with a heavy right tail, i.e. models for which $\gamma > 0$. We write $F \in \mathcal{D}_{\mathcal{M}}(EV_{\gamma>0}) =: {\mathcal{D}_{\mathcal{M}}}^+$.

Given a sample $(X_1, \ldots, X_n)$, let $X_{1:n} \leq \cdots \leq X_{n:n}$ be the associated ascending order statistics. The semi-parametric estimation of γ is usually based on the k top order statistics in the sample. Several estimators have been extensively studied by many authors. For heavy right tail models, we

consider here a classical estimator, the Hill estimator (Hill, 1975), H, and a more recent very well behaved estimator, the minimum-variance reduced-bias estimator (Caeiro *et al.*, 2005), $\overline{H}$, defined, respectively, as

$$H \equiv H(k) := \frac{1}{k} \sum_{i=1}^{k} \ln X_{n-i+1:n} - \ln X_{n-k:n} \tag{2}$$

and

$$\overline{H} \equiv \overline{H}(k) := H(k) \left(1 - \hat{\beta}(n/k)^{\hat{\rho}}/(1-\hat{\rho})\right), \tag{3}$$

where $\hat{\beta}$ and $\hat{\rho}$ are adequate consistent estimators of $\beta \neq 0$ and $\rho < 0$, 'scale' and 'shape' second-order parameters, respectively. For the estimation of ρ we refer to Fraga Alves *et al.* (2003) and for the estimation of β see Gomes and Martins (2002), Gomes *et al.* (2008a) and Caeiro *et al.* (2009), among others.

For consistent semi-parametric γ-estimation in the whole domain of attraction, ${\mathcal{D}_{\mathcal{M}}}^{+}$, we need to consider a first order condition, $1 - F \in RV_{-1/\gamma}$, where the notation RV_α stands for the class of regularly-varying functions with an index of regular variation equal to $\alpha \in R$, i.e. positive measurable functions $g(\cdot)$ such that for any $x > 0$, $g(tx)/g(t) \to x^\alpha$, as $t \to \infty$ (Bingham *et al.*, 1987). We further need to work with intermediate k, i.e. k verifying $k \equiv k_n \to \infty$ and $k/n \to 0$, as $n \to \infty$.

Classical estimators of θ have been developed based on characterizations of θ given by Leaddbetter (1983) and O'Brien (1987). We consider the most common interpretation of θ, as being *the reciprocal of the 'mean time of duration of extreme events'* what is directly related to the exceedances of high levels (Hsing *et al.*, 1988, Leadbetter and Nandagopalan,1989),

$$\theta = \frac{1}{\text{limiting mean size of clusters}}.$$

Identifying clusters by the occurrence of downcrossings or upcrossings, we can write

$$\theta = \lim_{n\to\infty} Pr[X_2 \leq u_n | X_1 > u_n] = \lim_{n\to\infty} Pr[X_1 \leq u_n | X_2 > u_n]. \tag{4}$$

The classical up-crossing UC-estimator, $\widehat{\Theta}^{UC}$ (Nandagopalan, 1990; Gomes, 1990, 1992, 1993), is a naive estimator that comes directly as an empirical counterpart of (4),

$$\widehat{\Theta}^{UC} \equiv \widehat{\Theta}^{UC}(u_n) := \frac{\sum_{i=1}^{n-1} I\left(X_i \leq u_n < X_{i+1}\right)}{\sum_{i=1}^{n} I(X_i > u_n)}, \tag{5}$$

for a suitable threshold u_n, where $I(A)$ denotes, as usual, the indicator function of A.

Consistency of this estimator is obtained provided that the high level u_n is a normalized level, i.e. if with $\tau \equiv \tau_n$ fixed, the underlying d.f. F verifies $F(u_n) = 1 - \tau/n + o(1/n), \quad n \to \infty$ and $\tau/n \to 0$.

A deterministic level $u \in [X_{n-k:n}, X_{n-k+1:n}[$ is considered. The $UC-$estimator can now be written as a function of k, the number of top order statistics above the chosen threshold,

$$\widehat{\Theta}^{UC}(k) := \frac{1}{k} \sum_{i=1}^{n-1} I\left(X_i \leq X_{n-k:n} < X_{i+1}\right). \tag{6}$$

For many dependent structures, the bias of $\widehat{\Theta}^{UC}(k)$ has two dominant components of orders k/n and $1/k$ (see Gomes *et al.*, 2008b),

$$Bias[\widehat{\Theta}^{UC}(k)] = \varphi_1(\theta)\left(\frac{k}{n}\right) + \varphi_2(\theta)\left(\frac{1}{k}\right) + o\left(\frac{k}{n}\right) + o\left(\frac{1}{k}\right) \tag{7}$$

whenever $n \to \infty$ and $k \equiv k(n) \to \infty$, $k = o(n)$.

The Generalized Jackknife methodology has the properties of estimating the bias and the variance of any estimator, helping the building of estimators with bias and mean squared error often smaller than those of an initial set of estimators.

The Generalized Jackknife methodology states that if the bias has two main terms we would like to reduce, we need to have access to three estimators, with the same type of bias.

Definition 2.1 (Gray and Schucany, 1972)**.** *Given three biased estimators of* θ, $T_n^{(1)}$, $T_n^{(2)}$ *and* $T_n^{(3)}$ *such that*

$$E[T_n^{(i)} - \theta] = b_1(\theta)\varphi_1^{(i)}(n) + b_2(\theta)\varphi_2^{(i)}(n) \qquad i = 1, 2, 3,$$

the Generalized Jackknife Statistic (of order 2) is given by

$$T_n^{GJ} := \frac{\begin{Vmatrix} T_n^{(1)} & T_n^{(2)} & T_n^{(3)} \\ \varphi_1^{(1)}(n) & \varphi_1^{(2)}(n) & \varphi_1^{(3)}(n) \\ \varphi_2^{(1)}(n) & \varphi_2^{(2)}(n) & \varphi_2^{(3)}(n) \end{Vmatrix}}{\begin{Vmatrix} 1 & 2 & 3 \\ \varphi_1^{(1)}(n) & \varphi_1^{(2)}(n) & \varphi_1^{(3)}(n) \\ \varphi_2^{(1)}(n) & \varphi_2^{(2)}(n) & \varphi_2^{(3)}(n) \end{Vmatrix}}.$$

Using the information obtained from (7) and based on the estimator $\widehat{\Theta}^{UC}$ computed at the three levels, k, $[k/2]+1$ and $[k/4]+1$, where [x] denotes, as usual, the integer part of x, Gomes *et al.* (2008b) derived a reduced-bias estimator for θ, the Generalized Jackknife estimator of order 2, $\widehat{\Theta}^{GJ}$, defined as

$$\widehat{\Theta}^{GJ} \equiv \widehat{\Theta}^{GJ}(k) := 5\widehat{\Theta}^{UC}([k/2]+1) - 2\big(\widehat{\Theta}^{UC}([k/4]+1) + \widehat{\Theta}^{UC}(k)\big). \quad (8)$$

This is an asymptotically unbiased estimator of θ, in the sense that it can remove the two dominant components of bias referred to in (7).

3 Monte Carlo simulations

Asymptotically, the reduced-bias estimators in (3) and (8) present very nice properties outperforming the associated classical estimators. However for finite samples, those estimators still present the aforementioned difficulties, high variance for small values of k, the number of upper order statistics used in the estimation, and a high bias for very large values of k. Procedures for the choice of k are again required. Recently the use of adequate bootstrap procedures revealed an improvement in the finite sample properties of the estimators. Let us refer to Prata Gomes and Neves (2011) and Gomes *et al.* (2012a). However the choice of the level k still remains an interesting research topic.

Monte Carlo simulation studies give us the possibility of analyzing the behavior of our procedures, given that the true values of the parameters are known. In this study, the Fréchet model for an i.i.d. setup and some stationary models, described below, will be considered. For those models, and on the basis of 1000 replicates, patterns of Mean Values (E) and Root Mean Squared Errors (RMSE) of estimators in (2), (3), (6) and (8) are shown, for some values of the unknown parameters, and for samples of size 1000. In this article the following models were considered:

Model I. (IID sequence) Let $\{X_i\}_{i\geq 1}$ be a sequence of independent random variables, with d.f. $F(x) = \exp(-x^{-1/\gamma})$, $x > 0$, with $\gamma = 1$ and $\gamma = 0.25$. Then $\theta = 1$.

Model II. (Moving Maximum Process I). The stationary two-dependent model referred to in Example 1.1, defined by $Y_i = \max\{Z_{i+1}, Z_i\}$, for $i \geq 1$, where $\{Z_i\}_{i\geq 1}$ is an i.i.d. sequence of standard exponential random variables. Then, $\{Y_i\}_{i\geq 1}$ has a d.f. $F(y) = (1 - \exp(-y))^2, y > 0$. Choosing $u_n(\tau)$ such that $nP(Y_1 > u_n) \to \tau$ as $n \to \infty$, then $nP(Z_1 > u_n) \to \tau/2$ as $n \to \infty$ and $P\{Y_{n:n} \leq u_n(\tau)\} = P\{\max(Z_1, \cdots, Z_n) \leq u_n(\tau)\} P\{Z_{n+1} \leq u_n(\tau)\} \to \exp(-\tau/2)$, as $n \to \infty$, i.e. $\{Y_n\}$ has then *extremal index* $\theta = 1/2$ (see Leadbetter and Rootzén, 1988).

Model III. (Max-Autoregressive Process I). Let $\{Z_i\}_{i\geq 1}$ be a sequence of independent, unit-Fréchet distributed random variables. For $0 < \theta \leq 1$, let $Y_1 = Z_1$ and $Y_i = \max\{(1 - \theta)Y_{i-1}, \theta Z_i\}$, for $i \geq 2$. The marginal distribution of the process $\{Y_i\}_{i\geq 1}$ is unit-Fréchet and for $u_n = ny$, $0 < y < \infty$, $P\{M_n \leq u_n\} \to \exp(-\theta/y)$, as $n \to \infty$. The *extremal index* of the sequence is equal to θ (see Beirlant *et al.*, 2004).

Model IV. (Max-Autoregressive Process II). Let Z_0 be a random variable with d.f. $H_0(z) = \exp(-\theta z^{-\alpha}(1 - \theta)^{-1})$, $\alpha > 0$, and $\{Z_i\}_{i\geq 1}$ be a sequence of independent random variables, independent of Z_0, with d.f. $F(z) = \exp(-z^{-\alpha})$. For $0 < \theta < 1$, let $Y_0 = Z_0$ and $Y_i =$

$(1-\theta)^{1/\alpha}\max\{Y_{i-1}, Z_i\}$, for $i \geq 1$. The *extremal index* of this process is θ (see Alpuim, 1989, and Canto e Castro, 1992).

Model V. (Moving Maximum Process II). Let $\{Z_i\}_{i\geq 0}$ be a sequence of independent, unit-Fréchet distributed random variables. For $a \geq 0$, let $Y_0 = Z_0$ and $Y_i = (a+1)^{-1}\max\left\{aZ_{i-1}, Z_i\right\}$. If $a \leq 1$ we have $P\{M_n \leq nx\} = P\{\widetilde{M}_n \leq nx\}^\theta \to \{\exp\left(-x^{-1}\right)\}^{1/(a+1)}$ as $n \to \infty$, otherwise $P\{M_n \leq nx\} \to \{\exp\left(-x^{-1}\right)\}^{a/(a+1)}$ as $n \to \infty$. Then $\theta = \max\{1, a\}/(a+1)$ and $1/2 \leq \theta \leq 1$ (see Davidson, 2011).

3.1 An illustration of the finite sample behavior of the extreme value index estimators

Extensive comparative simulation studies of the *extreme value index* estimators H and $\overline{H}$, in (2) and (3), respectively, can be seen in Caeiro *et al.* (2005), among others. The pattern for dependent processes does not differ much from the pattern associated to independent structures, as can be seen in Figure 3, where the simulated mean values and root mean squared errors of the aforementioned estimators are pictured, as a function of k, for **Model III** above and $\theta = 0.1$.

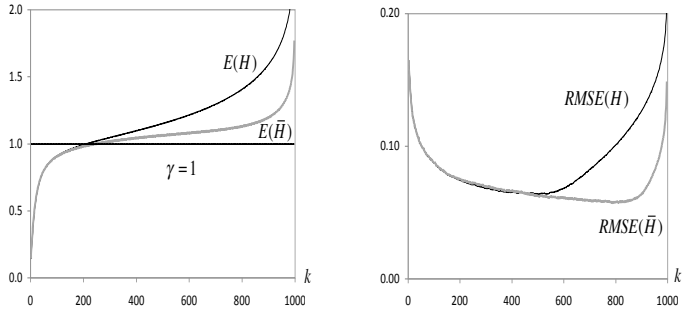

Figure 3: Simulated mean values (left) amd RMSE's (right) of the *extreme value index* estimators under study, for **Model III** ($\theta = 0.1$) and a sample $n = 1000$.

The main advantage of the corrected Hill estimators in (3) lies on the fact that we can adequately estimate the second-order parameters β and ρ so that the MSE of $\overline{H}$ is smaller than the MSE of Hill's estimator for all k. And this happens together with a higher stability of the sample paths around the target value not only for i.i.d. processes, but also for all dependent processes considered.

3.2 Simulation of the finite sample behavior of the extremal index estimators

Figures 4, 5, 6, 7, 8, 9 and 10 show an illustration of the results obtained for the simulated mean values and root mean squared errors of the *extremal index* estimators under study, for the above mentioned models, here provided only for a sample size $n = 1000$ and for a few values of the parameters. Further values for the sample size as well as for the parameters are available from the authors.

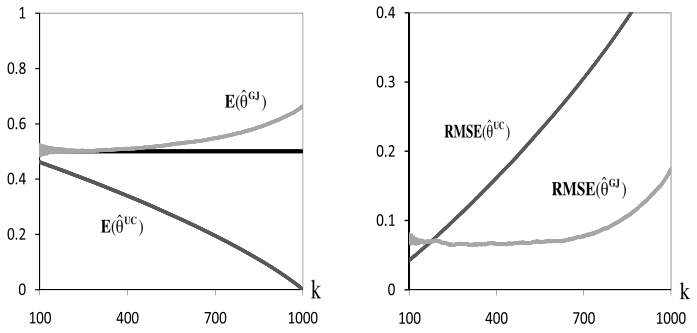

Figure 4: Simulated mean values (left) amd RMSE's (right) of the estimators under study, for **Model II** ($\theta = 0.5$) and a sample $n = 1000$.

As already noticed in Gomes *et al.* (2008b), the generalized jackknife extremal index estimator in (8), has very stable sample paths, around the target value θ, but at expenses of a very high variance, which does not enable

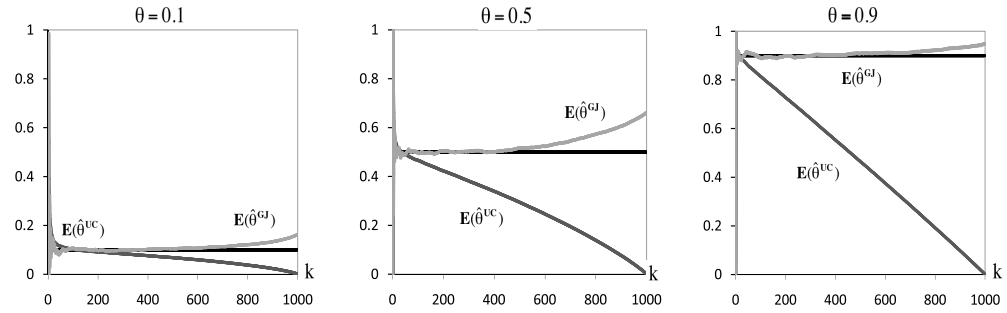

Figure 5: Simulated mean values of the estimators under study, for **Model III** with $\theta = 0.1, 0.5, 0.9$ and a sample $n = 1000$.

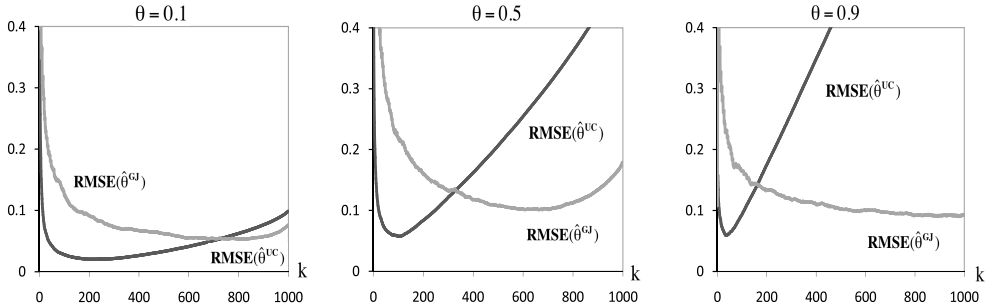

Figure 6: Simulated RMSE's of the estimators under study, for **Model III** with $\theta = 0.1, 0.5, 0.9$ and a sample $n = 1000$.

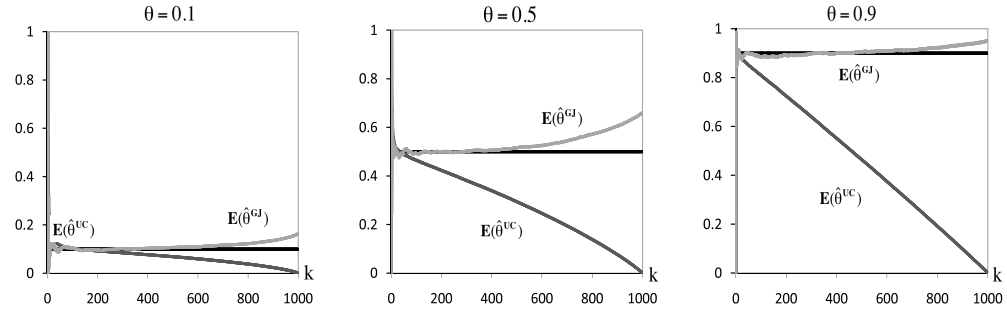

Figure 7: Simulated mean values of the estimators under study, for **Model IV** with $\theta = 0.1, 0.5, 0.9$ and a sample $n = 1000$.

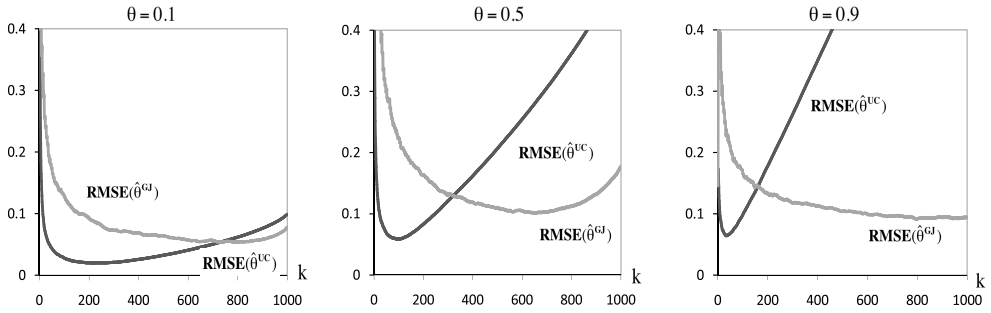

Figure 8: Simulated RMSE's of the estimators under study, for **Model IV** with with $\theta = 0.1, 0.5, 0.9$ and a sample $n = 1000$.

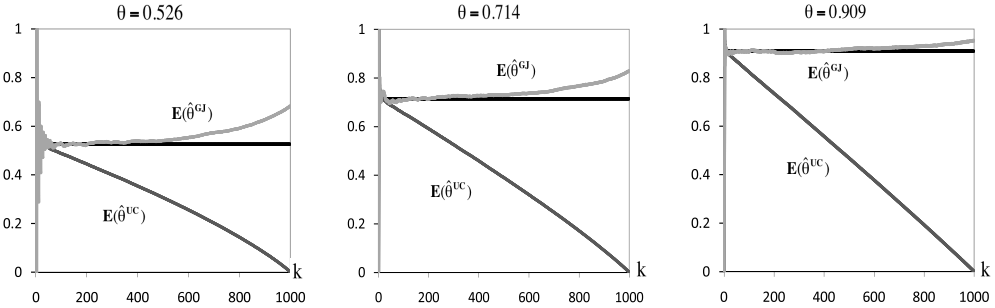

Figure 9: Simulated mean values of the estimators under study, for **Model V** with $a = 0.9, 0.4, 0.1 (\theta = 0.526, 0.714, 0.909)$ and a sample $n = 1000$.

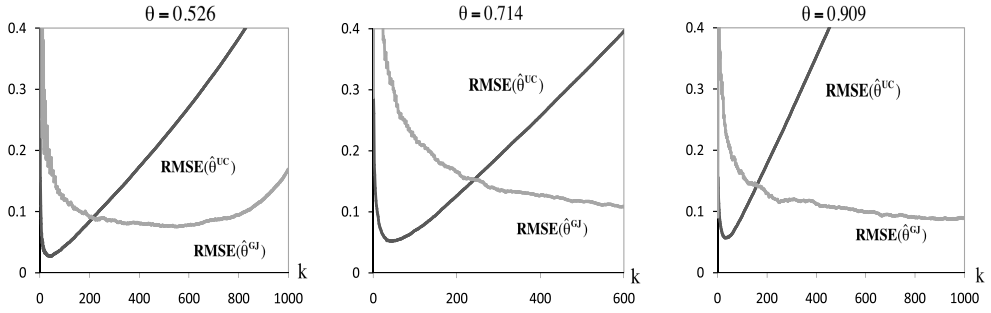

Figure 10: Simulated RMSE's of the estimators under study, for **Model V** with $a = 0.9, 0.4, 0.1 (\theta = 0.526, 0.714, 0.909)$ and a sample $n = 1000$.

it to overpass the original estimator, regarding MSE at optimal levels. It is there suggested the consideration of the levels k, $[\delta k]+1$ and $[\delta^2 k]+1$, dependent of a *tuning parameter* δ, $0<\delta<1$, or the use of subsampling techniques similar to the ones in Robinson and Tawn (2000), Scotto *et al.* (2003) and Martins and Ferreira (2004), a topic deserving further research, but out of the scope of this paper (see Gomes *et al.*, 2008b, for further details on the subject).

3.3 Finite sample behaviour of adaptive extreme value index and extremal index estimators

Our interest lies now on the choice of the level k to be used in the estimation of either γ or θ. Let us generally denote by $T(k)$ any of the above mentioned estimators. The heuristic algorithm in Gomes *et al.* (2012b) is now applied to the adaptive choice of the sample fraction needed for estimating each one of the parameters. This algorithm is the following:

Heuristic choice of the optimal sample fraction

1. Given an observed sample $(x_1,\ldots,x_n)$, compute, for $k=1,\ldots,n-1$, the observed values of $T(k)$.
2. Obtain j_0, the minimum value of j, a non-negative integer, such that the rounded values, to j decimal places, of the estimates in Step **1** are distinct. Define $a_k^{(T)}(j)=\text{round}(T(k),j)$, $k=1,2,\ldots,n-1$, the rounded values of $T(k)$ to j decimal places.
3. Consider the sets of k values associated to equal consecutive values of $a_k^{(T)}(j_0)$, obtained in Step **2.** Set $k_{min}^{(T)}$ and $k_{max}^{(T)}$ the minimum and maximum values, respectively, of the set with the largest range. The largest run size is then $l_{_T}:=k_{max}^{(T)}-k_{min}^{(T)}$.
4. Consider all those estimates, $T(k)$, $k_{min}^{(T)}\le k\le k_{max}^{(T)}$, now with two extra decimal places, i.e. compute $T(k)=a_k^{(T)}(j_0+2)$. Obtain the

mode of $T(k)$ and denote $\mathcal{K}_{_T}$ the set of k-values associated with this mode.

5. Take $\hat{k}_T$ as the maximum value of $\mathcal{K}_{_T}$, and consider the adaptive estimate $T(\hat{k}_{_T})$.

6. The best estimate is the value of T that corresponds to the maximum run size $l_{_T}$ computed in Step **3**.

From the models described above, samples were generated for some values of the parameters. Table 1 shows, for each model, the values of the parameters considered in the simulation and the results obtained from the application of the adaptive algorithm: the largest run size, $l_{(\bullet)}$, the estimate of k, $\hat{k}_{(\bullet)}$, where $(\bullet)$ denotes the estimators under study, and finally the parameters' estimates, either of γ for the i.i.d. structures or of θ for the dependent processes.

	Parameter	l_H	$l_{\overline{H}}$	$\hat{k}_H$	$\hat{k}_{\overline{H}}$	$\hat{\gamma}^H$	$\hat{\gamma}^{\overline{H}}$
Model I	$\gamma = 1$	278	302	583	790	1.220	0.960
	$\gamma = 0.25$	540	379	593	792	0.305	0.242
	Parameter	l_{UC}	l_{GJ}	$\hat{k}_{UC}$	$\hat{k}_{GJ}$	$\hat{\theta}^{UC}$	$\hat{\theta}^{GJ}$
Model II	$\theta = 0.5$	284	242	191	646	0.435	0.594
(another sample)	$\theta = 0.5$	261	96	196	585	0.439	0.582
Model III	$\theta = 0.1$	387	190	545	621	0.059	0.122
	$\theta = 0.5$	246	119	206	886	0.495	0.525
	$\theta = 0.9$	148	184	142	623	0.826	0.886
Model V	$\theta = 0.526$	228	107	473	550	0.331	0.537
	$\theta = 0.714$	146	154	265	990	0.580	0.726
	$\theta = 0.909$	125	186	227	945	0.708	0.975

Table 1: Results from the application of the adaptive algorithm to the models described. Random samples were generated considering the true values of parameters pointed out.

In Figure 11, we present the sample paths of the estimators of γ for two independent samples from Fréchet models with $\gamma = 1$ and $\gamma = 0.25$, as a function of k, together with the adaptive choices provided by the algorithm. Finally, Figures 12, 13 and 14 are similar to Figure 11, but for the estimation of the *extremal index* θ and three different dependent sample paths, respectively associated to **Model II**, **III** and **V**.

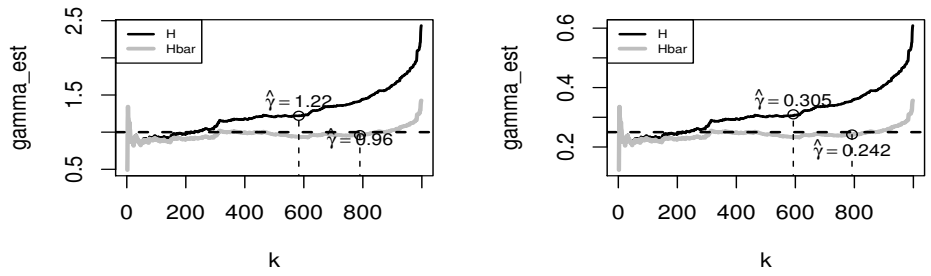

Figure 11: Adaptive choice of the level k for estimating γ for two samples from a **Fréchet Model,** with $\gamma = 1$ (left) and $\gamma = 0.25$ (right).

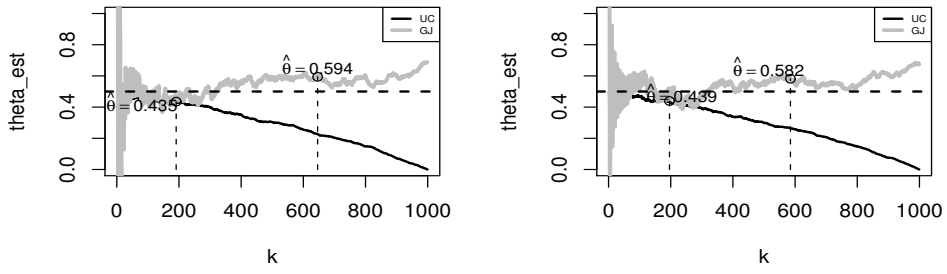

Figure 12: Adaptive choice of the level k for estimating θ for two random samples from a **2-dependent Model.**

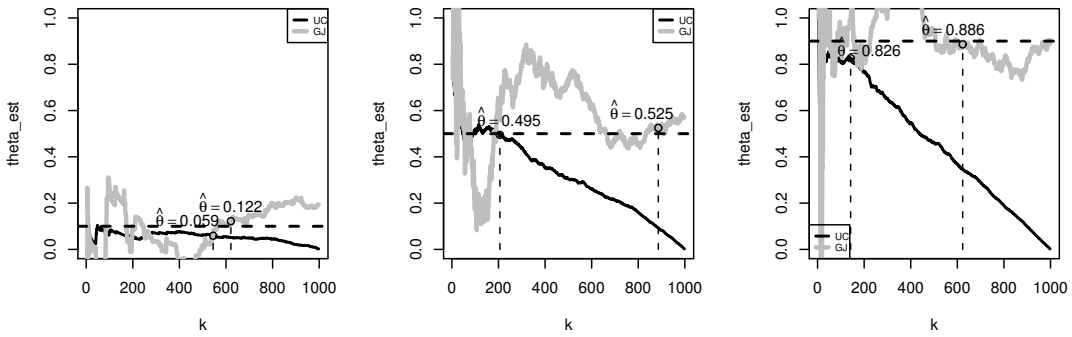

Figure 13: Adaptive choice of the level k for estimating θ for random samples generated with $\theta = 0.1, 0.5, 0.9$ (left to right) from **Model III.**

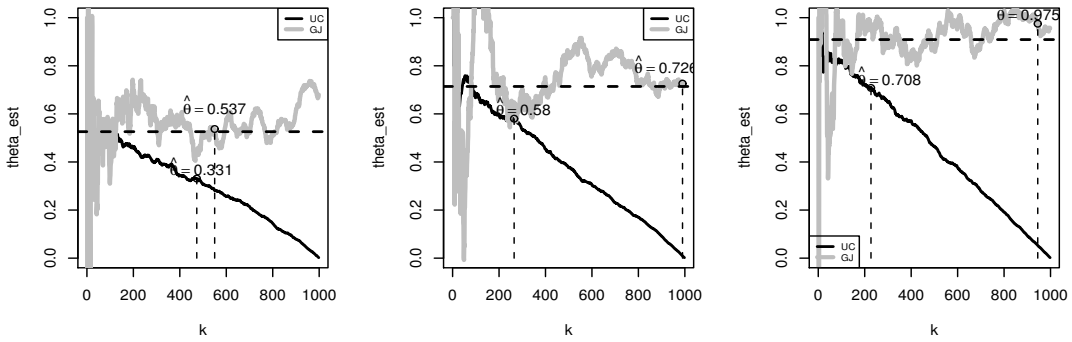

Figure 14: Adaptive choice of the level k for estimating θ for random samples generated with $\theta = 0.526, 0.714, 0.909$ (left to right) from **Model V**.

4 Concluding remarks

The heuristic algorithm described above seems to perform very well in the choice of the level k to be used in the estimation of both γ and θ. Further simulation studies are now in progress. For the estimation of γ, and as expected, the algorithm leads to large k-estimates and consequently to more

reliable estimates of γ. A similar comment applies to the adaptive estimation of θ.

However we think that future research is still welcome for improving the estimators of θ, so that more stable patterns can be obtained and possibly other alternative adaptive estimators too.

References

[1] Alpuim, M.T. (1989). An extremal markovian sequence. *J. Appl. Prob.* **26**, 219–222.

[2] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., Waal, D. and Ferro, C. (2004). *Statistics of Extremes: Theory and Applications*. John Wiley & Sons.

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

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

[5] Caeiro, F., Gomes, M.I. and Henriques-Rodrigues, L. (2009). Reduced-bias tail index estimators under a third order framework. *Commun. in Statist.—Theory and Methods* **38**:7, 1019–1040.

[6] Canto e Castro, L. (1992). *Sobre a Teoria Assintótica de Extremos*. Tese de Doutoramento, DEIO, Faculdade de Ciências da Universidade de Lisboa.

[7] Davidson, A. (2011). *Statistics of Extremes*. Courses 2011-2012. École Polytechnique Fédérale de Lausanne EPFL.

[8] 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**:2, 194–213.

[9] Gomes, M.I. (1990). Statistical inference in an extremal markovian model. *COMPSTAT*, 257–262.

[10] Gomes, M.I. (1992). Modelos extremais em esquemas de dependência. *I Congresso Ibero-Americano de Esdadistica e Investigacion Operativa*, 209–220.

[11] Gomes, M.I. (1993). On the estimation of parameters of rare events in environmental time series. *Statistics for the Environment* (Barnett *et. al.* eds.), 226–241.

[12] 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.

[13] Gomes, M.I., de Haan, L. and Henriques-Rodrigues, L. (2008a). Tail Index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. *J. Royal Statist. Soc.* **B70**:1, 31–52.

[14] Gomes, M.I., Hall, A. and Miranda, C. (2008b). Subsampling techniques and the Jackknife methodology in the estimation of the extremal index. *J. Comput. Statist. and Data Analysis* **52**:4, 2022–2041.

[15] Gomes, M.I., Figueiredo, F. and Neves, M.M. (2012a). Adaptive estimation of heavy right tails: resampling-based methods in action. *Extremes* **15**, 463–489.

[16] Gomes, M.I., Martins, M.J. and Neves, M.M. (2012b). Generalised Jackknife-Based Estimators for Univariate Extreme-Value

Modelling. *Commun. Statist.—Theory and Methods* **42**, DOI: 10.1080/03610926.2012.725263

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

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

[19] Hsing, T., Husler, J. and Leadbetter, M.R. (1988). On exceedance point process for a stationary sequence. *Probab. Theory Related Fields* 78, 97–112.

[20] Leadbetter, M.R. (1983). Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **65**:2, 291–306.

[21] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and related properties of random sequences and series*. Springer-Verlag, New York.

[22] Leadbetter, M.R. and Nandagopalan, L. (1989). On exceedance point process for stationary sequences under mild oscillation restrictions. In *Extreme Value Theory*: Proceedings, Oberwolfach 1987, J. Hüsler and R.D. Reiss (eds.), Lecture Notes in Statistics **52**, 69-80. Springer-Verlag, Berlim.

[23] Martins, A.P. and Ferreira, H. (2004). The extremal index of sub-sampled processes. *J. Statist. Planning and Inference* **124**, 145–152.

[24] Nandagopalan, S. (1990). *Multivariate Extremes and Estimation of the Extremal Index*. PhD Thesis, University of North Carolina, Chapel Hill.

[25] Nandagopalan, S. and Rootzén, H. (1988). Extremal Theory for Stochastic processes. *Ann. Probab.* **16**:2, 431-478.

[26] O'Brien, G. (1987). Extreme values for stationary and Markov sequences. *Ann. Probab.* **15**:1, 281–289.

[27] Prata Gomes, D. and Neves, M.M. (2011). Resampling Methodologies and the Estimation of Parameters of Rare Events, *In Numerical Analysis and Applied Mathematics (ICNAAM 2011)*, AIP Conf. Proc. 1389, 1475–1478.

[28] Robinson, M.E. and Tawn, J.A. (2000). Extremal analysis of processes sampled at different frequencies. *J. Royal Statist. Soc. B* **62**, 117–135.

[29] Scotto, M., Turkman, K.F. and Anderson, C.W. (2003). Extremes of some sub-sampled time series. *J. Time Series Analysis* **24**:5, 505–512.