REVSTAT – Statistical Journal
Volume 16, Number 1, January 2018, 115–136

HEURISTIC TOOLS FOR THE ESTIMATION OF THE EXTREMAL INDEX: A COMPARISON OF METHODS

Author: Marta Ferreira
– Center of Mathematics of Minho University and
CEMAT, CEAUL — University of Lisbon, Portugal
msferreira@math.uminho.pt

Received: July 2015 Revised: July 2016 Accepted: July 2016

Abstract:

- Clustering of exceedances of a critical level is a phenomenon that concerns risk managers in many areas. The extremal index θ measures the propensity of the large observations in a dataset to cluster. Thus the estimation of θ is an important issue recurrently addressed in literature. Besides a declustering parameter, inference also depends on a threshold. This choice is actually a crucial topic and is transversal to many other extremal parameters. In this paper we analyze a threshold-free heuristic procedure. We also make comparisons with other heuristic procedures already developed within the extremal index estimation. Our study is based on simulation. We illustrate with an application to environmental data.

Key-Words:

- *extreme value theory; extremal index estimation; heuristic methods.*

AMS Subject Classification:

- 62G32, 60G70.

1. INTRODUCTION

In many environmental applications, extreme events are the main aspects of practical concern. Financial time series are increasingly being analyzed to assess the risk from extreme events. A description of extreme events is usually based on observations that exceed a high threshold. Serial dependence leads to large values occurring close in time and thus forming clusters. Clustering of extremes does not take place in an independent and identically distributed (i.i.d.) setting.

Consider $\{X_n\}_{n\geq 1}$ a stationary sequence with common distribution function (df) F and $\{Y_n\}_{n\geq 1}$ an i.i.d. sequence with the same parent df F. We say $\{X_n\}_{n\geq 1}$ has extremal index θ ($0 < \theta \leq 1$) if, for all $\tau > 0$, there is a sequence of levels $u_n \equiv u_n(\tau)$, $n \geq 1$, such that

$$P\big(\max(Y_1,\ldots,Y_n) \leq u_n\big) = F^n(u_n) \underset{n\to\infty}{\longrightarrow} e^{-\tau}$$

$$\text{and} \quad P\big(\max(X_1,\ldots,X_n) \leq u_n\big) \underset{n\to\infty}{\longrightarrow} e^{-\theta\tau} \tag{1.1}$$

(Leadbetter *et al.* [20] 1983). The sequence $u_n \equiv u_n(\tau)$, $n \geq 1$, satisfying $F^n(u_n) \to \exp(-\tau)$ or, equivalently, $n(1-F(u_n)) \to \tau$, as $n \to \infty$, is usually denoted as normalized levels.

There are several characterizations of the extremal index bringing out different estimators. Many of these estimators can be stated as functions of a number k of upper order statistics. Analogous to the semiparametric estimation of various tail measures (e.g., the tail index and tail dependence coefficients in a multivariate framework), there is a proverbial tradeoff between bias and variance. The first increases with k (large bias for a large amount of top order statistics used in estimates) and the second increases as k gets smaller (large variability as fewer top order statistics are considered). A typical path is plotted in Figure 1. After the great variability in the beginning, there is a stable sample path, as function of k, around the true value and then the bias starts to stand out and dominate.

Thus k needs to be chosen from the stability zone that mediates the variance domain and the bias domain. There are several methods developed in literature towards this choice of k concerning the estimation of tail measures. A survey within the tail index estimation can be seen in Beirlant *et al.* ([2] 2004). More recently, a general procedure was introduced in Gomes *et al.* ([14] 2013) for the tail index estimation, which was latter adopted in Neves *et al.* ([23] 2015) to estimate the extremal index. This consists of a pure heuristic procedure to find the "plateau" region of the estimates path from which we may infer the true value of the parameter. The methodology in Frahm *et al.* ([12] 2005), developed within the estimation of the tail dependence coefficient of random pairs, also seeks a stable region but after a smoothing of the sample path; see Frahm *et al.* ([12]

2005) and Ferreira and Silva ([9] 2014). In Ferreira ([5] 2014) and Ferreira ([6] 2015a) it was also adapted to the tail index estimation.

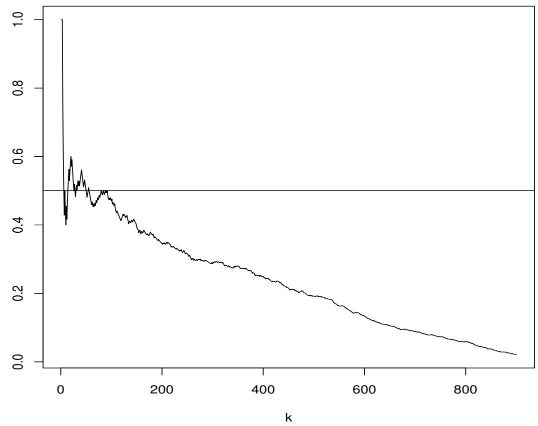

Figure 1: Runs estimates sample path of a moving maximum process, $X_i = \max(a_0 Z_i, a_1 Z_{i-1}, a_2 Z_{i-2})$, $i \geq 1$, where $\{Z_i\}_{i\geq -1}$ is an i.i.d. unit Fréchet sequence, $a_0 = 1/3$, $a_1 = 1/6$ and $a_2 = 1/2$, with run length $r = 2$. The horizontal line corresponds to the true value.

Here we are going to apply the methodology of Frahm *et al.* ([12] 2005) to several estimators of the extremal index. For comparison, we also analyze the performance of the procedure in Neves *et al.* ([23] 2015) applied to those estimators. As these are threshold-free methods, we also compare with the blocks and sliding estimation threshold-free procedure presented in Robert *et al.* ([25] 2009). The description of the methods is addressed in Section 2. The comparison of the procedures is assessed through simulation in Section 3 and an illustration with real data is stated in Section 4. A small discussion is presented in Section 5.

2. ESTIMATION METHODS

The extremal index can be interpreted in different ways, leading to different estimators. In O'Brien ([24] 1974) it is proved that

$$P\big(\max(X_2, ..., X_{r_n} \leq u_n | X_1 > u_n)\big) \underset{n\to\infty}{\longrightarrow} \theta\,, \tag{2.1}$$

where r_n is such that $r_n \to \infty$ and $r_n = o(n)$. Under a mild mixing condition, Hsing *et al.* ([17], 1988) stated that

$$E\big[\textstyle\sum_{i=1}^{r_n} \mathbf{1}_{\{X_j > u_n\}} |\textstyle\sum_{i=1}^{r_n} \mathbf{1}_{\{X_j > u_n\}} \geq 1\big] \to \theta^{-1}, \tag{2.2}$$

with $\mathbf{1}(\cdot)$ denoting the indicator function, i.e., the limiting mean number of exceedances of u_n in an interval of length r_n corresponds to the arithmetic inverse of the extremal index, given that there are exceedances.

Also under a slight mixing condition, Ferro and Segers ([11] 2003) show that

$$P\big(\overline{F}(u_n)\,T(u_n) > t\big) \underset{n\to\infty}{\longrightarrow} \theta e^{-\theta t}\,, \qquad t > 0, \tag{2.3}$$

where $T(u_n) = \min\{n \geq 1\colon X_{n+1} > u_n | X_1 > u_n\}$, i.e., the process of inter-exceedance times normalized by exceedances of u_n follows a mixture of a point mass and an exponential distribution $Exp(\theta^{-1})$.

Relations (2.1)–(2.3) yield the most common approaches to estimate θ, respectively, the runs, the blocks and the intervals method.

The blocks and the runs estimators are based on their own clusters identification procedure and both correspond to the ratio between the number of clusters and the number of exceedances of a high threshold u_n (Hsing [15] 1991; Weissman and Novak [29] 1998; Nandagopalan [22] 1990; Hsing [16] 1993). The intervals estimator is based on an inter-exceedance times method (Ferro and Segers [11] 2003).

More precisely, the runs estimator is expressed as

$$\widehat{\theta}^{\mathrm{R}} = (N_n(u_n))^{-1} \sum_{i=1}^{n-r} \mathbf{1}_{\{X_i > u_n\}}\, \mathbf{1}_{\{X_{i+1} \leq u_n\}} \cdots \mathbf{1}_{\{X_{i+r} \leq u_n\}}, \tag{2.4}$$

where $N_n(u_n)$ is the number of exceedances of u_n. Independent clusters are identified as runs of observations above u_n, separated by r_n consecutive values under u_n.

By considering $b_n = [n/r_n]$ blocks of length r_n ($[\cdot]$ means the integer part), the simple blocks estimator corresponds to

$$\widehat{\theta}^{\mathrm{B}} = \frac{C_n(u_n)}{N_n(u_n)} \tag{2.5}$$

where $C_n(u)$ is the number of clusters, i.e, in this context it corresponds to the number of blocks in which at least one exceedance of u_n occurs. The variant

$$\widehat{\theta}^{\mathrm{BL}} = \frac{\log(1 - C_n(u_n)/k_n)}{r_n \log(1 - N_n(u_n)/n)} \tag{2.6}$$

has been proposed in Smith and Weissman ([27] 1994) as having a better asymptotic behavior of second order.

After some considerations and based on the result in (2.3), the intervals estimator is stated as

$$(2.7)\qquad \widehat{\theta}^{\mathrm{I}} = \begin{cases} 1 \wedge \frac{2\left(\sum_{i=1}^{N-1} T_i\right)^2}{(N-1)\sum_{i=1}^{N-1} T_i^2} & \text{, if } \max\{T_i : 1 \leq i \leq N-1\} \leq 2 \\ 1 \wedge \frac{2\left(\sum_{i=1}^{N-1} (T_i-1)\right)^2}{(N-1)\sum_{i=1}^{N-1} (T_i-1)(T_i-2)} & \text{, if } \max\{T_i : 1 \leq i \leq N-1\} > 2, \end{cases}$$

where T_i denotes the i-th inter-exceedance time, $i = 1, \ldots, N-1$ and $N \equiv N_n(u_n)$.

The analysis of convenient local dependence conditions may eliminate the need for a cluster identification scheme, such as the local dependence condition $\mathrm{D}^{(m)}(u_n)$ of Chernick *et al.* ([4] 1991), with m some positive integer. Consider notation $M_{i,j} = \max\{X_{i+1}, \ldots, X_j\}$, for $i < j$, $M_{i,j} = -\infty$ if $i \geq j$ and $M_{0,j} = M_j$. Under condition $\mathrm{D}(u_n)$ of Leadbetter ([19] 1974) which holds whenever $\alpha_{n,l_n} \to 0$, as $n \to \infty$, for $l_n = o(n)$, where

$$\begin{aligned} \alpha_{n,l} = \sup\Big\{ \Big| & P\big(M_{i_1,i_1+p} \leq u_n, M_{j_1,j_1+q} \leq u_n\big) \\ & - P(M_{i_1,i_1+p} \leq u_n)\, P(M_{j_1,j_1+q} \leq u_n)\Big| : \\ & \qquad 1 \leq i_1 < i_1 + p + l \leq j_1 < j_1 + q \leq n \Big\}, \end{aligned}$$

we say that $\mathrm{D}^{(m)}(u_n)$ is satisfied by $\{X_n\}_{n\geq 1}$ if, for some $\{b_n\}_{n\geq 1}$ such that, as $n \to \infty$,

$$b_n \to \infty\,, \qquad b_n \alpha_{n,l_n} \to 0\,, \qquad b_n l_n/n \to 0,$$

we have

$$nP\big(X_1 > u_n, M_{1,m} \leq u_n < M_{m,r_n}\big) \to 0\,, \qquad n \to \infty,$$

with $\{r_n = [n/b_n]\}_{n\geq 1}$. The stronger conditions

$$n \sum_{j=m+1}^{r_n} P\big(X_1 > u_n, M_{1,m} \leq u_n < X_j\big) \to 0\,, \qquad n \to \infty,$$

also stated in Chernick *et al.* ([4] 1991), lead to $\mathrm{D}'(u_n)$ if $m = 1$ and $\mathrm{D}''(u_n)$ if $m = 2$, considered in Leadbetter *et al.* ([20] 1983) and Leadbetter and Nandagopalan ([21] 1989), respectively. Condition $\mathrm{D}'(u_n)$ inhibits clustering of exceedances and thus resembles an i.i.d. behavior and brings out $\theta = 1$, whilst $\mathrm{D}''(u_n)$ allows clustering but inhibits the occurrence of two or more upcrossings. Moreover, if condition $\mathrm{D}^{(m_0)}(u_n)$ holds then $\mathrm{D}^{(m)}(u_n)$ also holds for all $m \geq m_0$.

Ferreira and Ferreira ([10] 2015) stated a new estimator that works under $\mathrm{D}^{(m)}(u_n)$. More precisely, if $\{X_n\}_{n\geq 1}$ satisfies condition $\mathrm{D}^{(m)}(u_n)$, we can estimate θ by

$$(2.8)\qquad \widehat{\theta}^{\mathtt{FF}} = \frac{U_n^Z(u_n)}{N_n(u_n)}$$

where $U_n^Z(u_n)$ is the number of upcrossings of u_n within $\{Z_1, ..., Z_{[n/(m-1)]}\}$ with $Z_n = M_{(n-1)(m-1),n(m-1)}$, $n \geq 1$. Other estimators developed in the same context were also considered in that work whose overall performance did not surpass $\widehat{\theta}^{\texttt{FF}}$. Estimation approaches working only for series that satisfy condition $\mathrm{D}^{(2)}(u_n)$ can also be seen in Süveges ([28] 2007), Ferreira and Ferreira ([8] 2012) and Ferreira ([7] 2015b).

Chernick *et al.* ([4] 1991) also show that, under $\mathrm{D}^{(m)}(u_n)$, the extremal index exists and can be computed by the limit

$$\theta = \lim_{n\to\infty} P\big(M_{1,m} \leq u_n | X_1 > u_n\big). \tag{2.9}$$

Observe that the runs estimator in (2.4) corresponds to the empirical counterpart of (2.9) by considering $r = m$. Diagnostic tools to analyze condition $\mathrm{D}^{(m)}(u_n)$ may be seen in Süveges ([28] 2007) and Ferreira and Ferreira ([10] 2015).

Observe also that taking $r = 2$ in (2.4) corresponds to the Nandagopalan's runs estimator derived in Nandagopalan ([22] 1990) under $\mathrm{D}''(u_n)$.

The disjoint blocks and the sliding blocks estimators presented in Robert *et al.* ([25] 2009) are derived from the extremal index definition in (1.1).

Consider, for r positive integer,

$$F_r(u) := P(M_r \leq u), \quad \tau_r(u) := r(1 - F(u)) \quad \text{and} \quad \theta_r(u) := -\frac{\log F_r(u)}{\tau_r(u)}.$$

We have $\theta = \lim_{r\to\infty} \theta_r(u_r)$ for normalized levels $u_r = u_r(\tau)$ according to definition in (1.1). The estimation of the block maxima df F_r through $b = [n/r]$ disjoint blocks or $n - r + 1$ sliding blocks, that is

$$\widehat{F}^{\texttt{DJ}}_{n,r}(u) := \frac{1}{b}\sum_{i=1}^{b} \mathbb{1}_{\{M_{(i-1)r,ir} \leq u\}} \quad \text{and} \quad \widehat{F}^{\texttt{SL}}_{n,r}(u) := \frac{1}{n-r+1}\sum_{i=1}^{n-r+1} \mathbb{1}_{\{M_{i-1,i-1+r} \leq u\}}$$

originates the estimators, respectively,

$$\widehat{\theta}^{\texttt{DJ}} = -\frac{\log \widehat{F}^{\texttt{DJ}}_{n,r}(u_n)}{\widehat{\tau}_{n,r}(u_n)} \quad \text{and} \quad \widehat{\theta}^{\texttt{SL}} = -\frac{\log \widehat{F}^{\texttt{SL}}_{n,r}(u_n)}{\widehat{\tau}_{n,r}(u_n)}, \tag{2.10}$$

with

$$\widehat{\tau}_{n,r}(u_n) = \frac{rN_n(u_n)}{n}.$$

In order to achieve consistency in the estimators above, τ must be actually taken as an intermediate sequence τ_n, $n \geq 1$, that is,

$$\tau_n \to \infty \quad \text{and} \quad \tau_n/n \to 0.$$

Gomes *et al.* ([13] 2008) and Neves *et al.* ([23] 2015) considered the levels u_n in the interval between the $k + 1$ and the k-th upper order statistics,

$[X_{n-k:n}, X_{n-k+1:n})$, for the Nandagopalan's runs estimators. The advantage is to move the framework to a similar context of the semiparametric estimation of other important tail measures existing in the literature which allows new estimation methods of the extremal index by adapting the existing ones. In this paper we will consider those levels $u_n \in [X_{n-k:n}, X_{n-k+1:n})$ in the estimators (2.4)–(2.8) and (2.10) and denote, respectively,

$$\widehat{\theta}_k^{\mathtt{R}},\ \widehat{\theta}_k^{\mathtt{B}},\ \widehat{\theta}_k^{\mathtt{BL}},\ \widehat{\theta}_k^{\mathtt{I}},\ \widehat{\theta}_k^{\mathtt{FF}},\ \widehat{\theta}_k^{\mathtt{DJ}} \text{ and } \widehat{\theta}_k^{\mathtt{SL}}. \tag{2.11}$$

Observe that we are replacing τ by k. Indeed, we consider that $k \equiv k_n$, $n \geq 1$, is replacing τ_n and thus it is also an intermediate sequence on behalf of consistency. Estimators in (2.11) are functions of k, the number of order statistics higher than the chosen level, where an increasing/decreasing k increases the bias/variance (see Figure 1). Thus the choice of k is central in the estimation, not only of the extremal index, but also of many other tail measures, making this topic largely addressed in literature (see, e.g., Beirlant *et al.* [2] 2004).

The "plateau-finding" algorithm of Frahm *et al.* ([12] 2005), applied to the estimation of the tail dependence coefficient of random pairs and here adopted to estimate the extremal index, is based on a smoothing of the estimator's sample path by a simple box kernel with integer bandwidth $d > 0$. The resulting trajectory thus corresponds to the moving average of $2d+1$ successive points of the initial one and will be used in the rest of the procedure that consists on the application of a plateau definition and respective finding criterium. In the following we detail the method which we denote Algorithm 1.

Algorithm 1:

For a sample $(X_1, ..., X_n)$, consider bandwidth $d = [wn] \in \mathbb{N}$ and compute the means of $2d+1$ successive points of $\widehat{\theta}_k$, $1 \leq k < n$, with smoothing degree $w = 0.005$ (thus each moving average is about 1% of the data, as suggested in Frahm *et al.* [12] 2005). In the resulting smoothed values, $\overline{\widehat{\theta}}_1, \ldots, \overline{\widehat{\theta}}_{n-2d}$, define the plateaus $p_k = (\overline{\widehat{\theta}}_k, \ldots, \overline{\widehat{\theta}}_{k+m-1})$, $k = 1, \ldots, n-2d-m+1$, with length $m = [\sqrt{n-2d}]$. The algorithm stops at the first plateau satisfying

$$\sum_{i=k+1}^{k+m-1} \left| \overline{\widehat{\theta}}_i - \overline{\widehat{\theta}}_k \right| \leq 2s,$$

where s is the empirical standard deviation of $\overline{\widehat{\theta}}_1, \ldots, \overline{\widehat{\theta}}_{n-2d}$. Estimate θ as the mean of the values of the chosen plane region (consider the estimate zero if no stable region fulfills the stopping condition).

For comparison, we also consider another heuristic procedure introduced in Gomes *et al.* ([14] 2013), also seeking the plane region that presumably includes the "optimal" sample fraction k to be estimated. The algorithm is described below and denoted Algorithm 2.

Algorithms 2 and 3:

For a sample $(X_1, ..., X_n)$, obtain the minimum value j_0, such that the rounded values to j decimal places of $\widehat{\theta}_k$, $1 \leq k < n$, denoted $\widehat{\theta}_k(j)$ are not all equal. Identify the set of values of k associated to equal consecutive values of $\widehat{\theta}_k(j_0)$. Consider the set with largest range $\ell := k_{\texttt{max}} - k_{\texttt{min}}$. Take all the estimates $\widehat{\theta}_k(j_0+2)$ with $k_{\texttt{max}} \leq k \leq k_{\texttt{min}}$, i.e., the estimates with two more decimal points and obtain the mode. Denote $\mathcal{K}$ the set of k-values associated with this mode. Consider $\widehat{\theta}_{\widehat{k}}$, where $\widehat{k}$ is the maximum of $\mathcal{K}$.

We also consider the variant $\widehat{\theta}_{\widetilde{k}}$ by taking $\widetilde{k} = \ell$ as mentioned in Neves *et al.* ([23] 2015). This will be denoted Algorithm 3.

Observe that the described methodologies are all threshold-free. Robert *et al.* ([25] 2009) also presented a threshold-free procedure based on blocks and sliding estimators defined in (2.10). It is described downwards and will be called Algorithm 4:

Algorithm 4:

For a sample $(X_1, ..., X_n)$, choose a block size r, take $b = [n/r]$, $\tau = 1$ and $u = X_{n-[b\tau]+1:n}$. Consider $N_{a,b}(u) := \sum_{a<i\leq b} \mathbf{1}_{\{X_i>u\}}$, $\overline{N}_{n,r}(u) := (1/(n-r+1)) \times \sum_{i=0}^{n-r} N_{i,i+r}(u)$, $\widehat{\sigma}^2_{n,r}(u) := \sum_{i=0}^{n-r}(N_{i,i+r}(u) - \overline{N}_{n,r}(u))^2$ and $\widehat{c}^2_{n,r}(u) := \frac{\widehat{\theta}}{\widehat{\tau}_{n,r}(u)} \times \widehat{\sigma}^2_{n,r}(u) - 1$. Calculate $\widehat{\theta} = \widehat{\theta}^{\Gamma}_{[b\tau]}$, $\Gamma \in \{\texttt{SL}, \texttt{DJ}\}$ and $\widehat{c} = \widehat{c}_{n,r}(u)$. Obtain $\widehat{\mu}$ through

$$\mu = \begin{cases} \mu_{\texttt{SL}} := \theta\alpha^{-2}(e^{\alpha} - 1 - \alpha) + \alpha^{-1}\theta c^2 \\ \mu_{\texttt{DJ}} := \theta(2\alpha)^{-1}(e^{\alpha} - 1) + \alpha^{-1}\theta c^2, \end{cases}$$

replacing θ, c and α by, respectively, $\widehat{\theta}$, $\widehat{c}$ and $\widehat{\theta}\tau$. Obtain the bias-corrected $\widehat{\theta} - \widehat{\mu}/b$ and estimate the variance by evaluating $v = 2(\theta^2/\alpha^3)(e^{\alpha} - 1 - \alpha - \alpha^2/2) + \theta^2 c^2/\alpha$ at $\theta = \widehat{\theta}$, $c = \widehat{c}$ and $\alpha = \widehat{\theta}\tau$. Take $\widehat{\alpha}$ as the value that minimizes v when $\theta = \widehat{\theta}$ and $c = \widehat{c}$. Now repeat the procedure for the founded optimal value $\tau = \widehat{\alpha}/\widehat{\theta}$.

In the sequel, we use the abbreviations A1, A2, A3 and A4, respectively, to refer the algorithms above.

3. SIMULATION STUDY

We are going to analyze through simulation the performance of the estimators in (2.11) within the methodologies A1–A4 described above. This study is based on the following models:

- Max-autoregressive process (MAR), $X_i = \alpha X_{i-1} \vee \epsilon_i$, where $0 < \alpha < 1$ and $\{\epsilon_i\}_{i\geq 1}$ is an i.i.d. sequence of r.v.'s with d.f. $F_\epsilon(x) = \exp(-(1-\alpha)/x)$, $x > 0$. This process has $\theta = 1 - \alpha$. We consider $\alpha = 1/2$ and hence $\theta = 1/2$.
- Moving maxima process (MM), $X_i = \bigvee_{j=0,\dots,m} \alpha_j \epsilon_{i-j}$, with $\sum_{j=0}^{m} \alpha_j = 1$ and $\alpha_j \geq 0$, $\{\epsilon_i\}_{i\geq 1}$ is an i.i.d. sequence of unit Féchet distributed r.v.'s. This process has $\theta = \bigvee_{j=0,\dots,m} \alpha_j$. We consider $m = 3$, $\alpha_0 = 1/3$, $\alpha_1 = 1/6$, $\alpha_2 = 1/2$ leading to $\theta = 1/2$.
- Autoregressive Gaussian process (AR), $X_i = \alpha X_{i-1} + \epsilon_i$, where $\{\epsilon_i\}_{i\geq 1}$ is an i.i.d. sequence of $\mathrm{N}(0, 1-\alpha^2)$ distributed r.v.'s. This process satisfies condition $\mathrm{D}'(u_n)$ and thus $\theta = 1$ (Leadbetter *et al.* [20] 1983).
- A first order autoregressive process, with Cauchy marginals (ARCauchy) of Chernick ([3] 1978), $X_i = sX_{i-1} + \epsilon_i$, with $|s| < 1$. The extremal index is given by $1 - s^2$. We take $s = -3/5$ and thus $\theta = 0.64$.
- A negatively correlated uniform autoregressive process (ARUnif) of Chernick *et al.* ([4], 1991), $X_i = -(1/s)X_{i-1} + \epsilon_i$, where $\{\epsilon_i\}_{i\geq 1}$ is an i.i.d. sequence such that $P(\epsilon_1 = j/s) = 1/s$ for $j = 1, \dots, s$. We have $\theta = 1 - 1/s^2$. Here we consider $s = 2$ and thus $\theta = 3/4$.
- Bivariate extreme value Markov process with standard Gumbel marginals and logistic dependence function, i.e.,

  $$P\big(X_i \leq x, X_{i+1} \leq y\big) \;=\; \exp\big(-(x^{1/\alpha} + y^{1/\alpha})^\alpha\big).$$

  We consider the dependence parameter $\alpha = 0.5$ which gives $\theta = 0.328$ (Smith [26] 1992), and denote the process MCBEV.
- A GARCH(1,1) process, $X_i = \sigma_i \epsilon_i$, with $\sigma_i^2 = \alpha + \lambda X_{i-1}^2 + \beta \sigma_{i-1}^2$, $\alpha, \lambda, \beta > 0$, where $\{\epsilon_i\}_{i\geq 1}$ is an i.i.d. sequence of standard Gaussian r.v.'s. We consider $\alpha = 10^{-6}$, $\lambda = 1/4$ and $\beta = 7/10$ resulting in $\theta = 0.447$ (see details in Laurini and Tawn, [18] 2012).

We consider samples of sizes $n = 100$, 1000, 5000 and generate 100 independent replications of each and for each model. We compare the estimation procedures by computing the absolute mean bias and the root mean square error (rmse).

Remark 3.1. Observe that the methods being compared avoid threshold selection but need a cluster identification parameter, whether be it a block size or a run length. Recall that the dependence condition $\mathrm{D}^{(m)}(u_n)$ of Chernick *et al.* ([4], 1991) is a diagnostic tool for cluster identification within the runs estimator $\widehat{\theta}_k^{\mathtt{R}}$ and estimator $\widehat{\theta}_k^{\mathtt{FF}}$ defined in (2.8). More precisely, we take the run length r equal to m in the first (see discussion concerning (2.9)) and cycles of size $m-1$ in the second as stated in (2.8). The MAR process satisfies condition $\mathrm{D}^{(2)}(u_n)$, whilst the processes MM, ARCauchy and ARUnif satisfy condition $\mathrm{D}^{(3)}(u_n)$. See Ferreira and Ferreira ([10] 2015) and references therein for more details. In this latter reference, we validated conditions $\mathrm{D}^{(4)}(u_n)$ and $\mathrm{D}^{(5)}(u_n)$ for the processes MCBEV and GARCH, respectively. In what concerns the remaining estimators which are based on blocks schemes, the respective cluster parameters were chosen according to an overall good performance found on further simulations.

The results are presented in Tables 1–6 (the bold numbers correspond to the smallest estimates obtained in each model). A high bias is observed in the AR model and also in models ARCauchy, ARUnif and GARCH concerning the runs, the intervals and the FF estimator, under algorithms A2 and A3. The lowest values of rmse rely frequently on blocks $\widehat{\theta}_k^{\mathtt{B}}$ and $\widehat{\theta}_k^{\mathtt{BL}}$ estimators under algorithms A1, A2 and A3, followed by estimators $\widehat{\theta}_k^{\mathtt{R}}$ and $\widehat{\theta}_k^{\mathtt{FF}}$ within algorithm A1. In the AR process, the results differ from the others where the estimators $\widehat{\theta}_k^{\mathtt{DJ}}$ and $\widehat{\theta}_k^{\mathtt{SL}}$ tend to behave better over the four algorithms. Observe that in this case we have the boundary value $\theta = 1$ (as in i.i.d. sequences) where inference is usually problematic (see Ancona-Navarrete and Tawn [1] 2000). The intervals estimator, $\widehat{\theta}_k^{\mathtt{I}}$, is parameter-free under the methods in study and may be considered within MAR, AR, MM and MCBEV models. The worst performances concern mainly estimators $\widehat{\theta}_k^{\mathtt{FF}}$ and $\widehat{\theta}_k^{\mathtt{R}}$ for algorithm A2, where the method is returning a too high k, corresponding to estimates with very large bias. In Gomes *et al.* ([13] 2008) it was presented a reduced-bias version of Nandagopalan's estimator based on the Generalized Jackknife (GJ) methodology, which is given by

$$\widehat{\theta}_k^{\mathtt{NGJ}} = 5\widehat{\theta}_{[k/2]+1}^{\mathtt{R}} - 2\left(\widehat{\theta}_{[k/4]+1}^{\mathtt{R}} + \widehat{\theta}_k^{\mathtt{R}}\right). \tag{3.1}$$

Notice that Nandagopalan's estimator corresponds to the runs estimator whenever we take the run length 2, which in turn requires $\mathrm{D}''(u_n)$. In our examples, only models MAR and AR satisfy this condition. We have also applied the estimator (3.1) to all models within algorithms A2 and A3. Indeed, except in the GARCH case, the rmse of $\widehat{\theta}_k^{\mathtt{NGJ}}$ decreases to about the half of the rmse of the runs estimator, mostly for larger sample sizes ($n \geq 1000$) and with algorithm A2. In the case of algorithm A3, the rmse of $\widehat{\theta}_k^{\mathtt{NGJ}}$ is smaller than the runs estimator only within the largest sample size ($n = 5000$) of models MAR and ARUnif.

Table 1: Root mean squared errors obtained for simulated samples of size $n = 100$. For estimators $\widehat{\theta}_k^{\mathtt{BL}}$ and $\widehat{\theta}_k^{\mathtt{B}}$ we considered blocks of length 3 except in MCBEV and GARCH models where we used blocks of length 4 and 5, respectively. For estimators $\widehat{\theta}_k^{\mathtt{DJ}}$ and $\widehat{\theta}_k^{\mathtt{SL}}$ we considered blocks of length 5. For estimator $\widehat{\theta}_k^{\mathtt{R}}$ ($\widehat{\theta}_k^{\mathtt{FF}}$) we considered runs (cycles) of length 2 in MAR and AR, of length 3 in MM, ARCauchy and ARUnif, of length 4 in MCBEV and length 5 in GARCH. See Remark 3.1.

A1	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.1338	0.3603	0.1215	0.1534	0.1528	0.1249	0.1782
$\widehat{\theta}_k^{\mathtt{I}}$	0.2754	0.2899	0.2271	0.3549	0.2487	0.3071	0.5170
$\widehat{\theta}_k^{\mathtt{FF}}$	0.1469	0.5336	0.1488	0.1831	0.1474	0.1352	0.2018
$\widehat{\theta}_k^{\mathtt{BL}}$	0.1467	0.4548	0.1780	0.2869	0.1820	0.1544	0.1626
$\widehat{\theta}_k^{\mathtt{B}}$	0.1233	0.4826	0.1594	0.1948	0.1459	0.1380	0.1539
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.2971	0.5041	0.2803	0.3296	0.3406	0.2909	0.5316
$\widehat{\theta}_k^{\mathtt{SL}}$	0.2583	**0.2327**	0.2474	0.2907	0.3419	0.3052	0.5185

A2	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.3183	0.7928	0.4133	0.6283	0.7399	0.2663	0.4366
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.3003	0.5448	0.5071	0.3925	0.2500	0.3404	0.5569
$\widehat{\theta}_k^{\mathtt{I}}$	0.2621	0.5387	0.3032	0.2500	0.2500	0.2133	0.5511
$\widehat{\theta}_k^{\mathtt{FF}}$	0.4353	0.9410	0.4627	0.6400	0.7500	0.4447	0.2910
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0996	0.5217	0.0948	0.2176	**0.1079**	0.0734	0.1793
$\widehat{\theta}_k^{\mathtt{B}}$	0.1216	0.6186	0.1162	0.2118	0.3018	**0.0495**	0.2108
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.2646	0.4831	0.2146	0.2801	0.3010	0.2405	0.3230
$\widehat{\theta}_k^{\mathtt{SL}}$	0.3035	0.4431	0.2110	0.3158	0.4289	0.2457	0.4511

A3	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.1199	0.4192	0.1746	0.3026	0.2923	0.1158	0.3151
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.6235	0.5285	0.5323	0.3673	0.2505	0.6538	0.6584
$\widehat{\theta}_k^{\mathtt{I}}$	0.2770	0.2846	0.2527	0.2500	0.2500	0.3095	0.4959
$\widehat{\theta}_k^{\mathtt{FF}}$	0.1801	0.6620	0.2295	0.4697	0.4486	0.3133	0.1766
$\widehat{\theta}_k^{\mathtt{BL}}$	0.1663	0.4521	0.1890	0.2587	0.2051	0.1517	0.1580
$\widehat{\theta}_k^{\mathtt{B}}$	**0.0967**	0.5245	**0.0775**	**0.1105**	0.1179	0.0570	**0.1439**
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.4813	0.2996	0.4017	0.4394	0.3769	0.4782	0.5875
$\widehat{\theta}_k^{\mathtt{SL}}$	0.4802	0.4092	0.4041	0.4609	0.4946	0.5056	0.6435

A4	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.3001	0.2423	0.2748	0.2888	0.3457	0.3496	0.4886
$\widehat{\theta}_k^{\mathtt{SL}}$	0.2482	0.2445	0.2302	0.2644	0.3279	0.3562	0.4811

Table 2: Absolute bias obtained for simulated samples of size $n = 100$. For estimators $\widehat{\theta}_k^{\mathtt{BL}}$ and $\widehat{\theta}_k^{\mathtt{B}}$ we considered blocks of length 3 except in MCBEV and GARCH models where we used blocks of length 4 and 5, respectively. For estimators $\widehat{\theta}_k^{\mathtt{DJ}}$ and $\widehat{\theta}_k^{\mathtt{SL}}$ we considered blocks of length 5. For estimator $\widehat{\theta}_k^{\mathtt{R}}$ ($\widehat{\theta}_k^{\mathtt{FF}}$) we considered runs (cycles) of length 2 in MAR and AR, of length 3 in MM, ARCauchy and ARUnif, of length 4 in MCBEV and length 5 in GARCH. See Remark 3.1.

A1	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.0152	0.3280	0.0833	0.0618	0.0569	0.0076	**0.0303**
$\widehat{\theta}_k^{\mathtt{I}}$	0.1577	0.2040	0.1234	0.3545	0.2487	0.1875	0.5006
$\widehat{\theta}_k^{\mathtt{FF}}$	0.0780	0.5172	0.1243	0.1276	**0.0003**	0.1070	0.1841
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0063	0.4319	0.0709	0.0307	0.0321	0.0439	0.0771
$\widehat{\theta}_k^{\mathtt{B}}$	0.0120	0.4657	0.0432	0.0350	0.0756	0.0265	0.1071
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0454	0.3721	0.1328	0.2646	0.2934	0.2091	0.4143
$\widehat{\theta}_k^{\mathtt{SL}}$	0.2050	0.1729	0.2186	0.2519	0.2993	0.2797	0.4928

A2	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.2971	0.7778	0.4045	0.6283	0.7399	0.2586	0.4366
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.0057	0.4671	0.3341	0.3714	0.2500	0.0889	0.5222
$\widehat{\theta}_k^{\mathtt{I}}$	**0.0031**	0.4575	0.0879	0.2500	0.2500	0.0017	0.5505
$\widehat{\theta}_k^{\mathtt{FF}}$	0.4294	0.9393	0.4599	0.6400	0.7500	0.4445	0.2861
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0371	0.5127	**0.0113**	0.1278	0.0389	0.0097	0.1673
$\widehat{\theta}_k^{\mathtt{B}}$	0.1135	0.6172	0.1051	0.1628	0.2906	0.0311	0.2047
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0107	0.3768	0.0187	0.0198	0.1110	0.0554	0.1470
$\widehat{\theta}_k^{\mathtt{SL}}$	0.1030	0.2913	0.0594	0.0748	0.1952	0.1127	0.2979

A3	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.0488	0.4056	0.1598	0.2875	0.2800	0.0937	0.3110
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.0057	0.4671	0.3341	0.3714	0.2500	0.0889	0.5222
$\widehat{\theta}_k^{\mathtt{I}}$	0.1607	0.2045	0.1672	0.2500	0.2500	0.2044	0.4792
$\widehat{\theta}_k^{\mathtt{FF}}$	0.1593	0.6558	0.2182	0.4652	0.4444	0.3097	0.1711
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0275	0.4163	0.0766	0.0566	0.0702	0.0565	0.0561
$\widehat{\theta}_k^{\mathtt{B}}$	0.0547	0.5184	0.0120	**0.0179**	0.1023	**0.0025**	0.1280
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.3553	**0.0206**	0.2558	0.3441	0.3146	0.3330	0.4882
$\widehat{\theta}_k^{\mathtt{SL}}$	0.2852	0.0424	0.2375	0.2385	0.2952	0.3101	0.4835

A4	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.2433	0.1865	0.2096	0.2462	0.3059	0.3099	0.4571
$\widehat{\theta}_k^{\mathtt{SL}}$	0.2110	0.2031	0.1921	0.2238	0.2821	0.3219	0.4544

Table 3: Root mean squared errors obtained for simulated samples of size $n = 1000$. For estimators $\widehat{\theta}_k^{\mathtt{BL}}$ and $\widehat{\theta}_k^{\mathtt{B}}$ we considered blocks of length 3 except in MCBEV and GARCH models where we used blocks of length 4 and 5, respectively. For estimators $\widehat{\theta}_k^{\mathtt{DJ}}$ and $\widehat{\theta}_k^{\mathtt{SL}}$ we considered blocks of length 20. For estimator $\widehat{\theta}_k^{\mathtt{R}}$ ($\widehat{\theta}_k^{\mathtt{FF}}$) we considered runs (cycles) of length 2 in MAR and AR, of length 3 in MM, ARCauchy and ARUnif, of length 4 in MCBEV and length 5 in GARCH. See Remark 3.1.

A1	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.0666	0.2877	0.0645	0.0874	0.0583	0.0508	0.1059
$\widehat{\theta}_k^{\mathtt{I}}$	0.0795	0.3243	0.0658	0.2312	0.2498	0.0670	0.2217
$\widehat{\theta}_k^{\mathtt{FF}}$	0.0853	0.4148	0.0744	0.1014	0.0566	0.0809	0.1295
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0467	0.4381	0.0767	0.0882	0.0944	0.0836	0.0845
$\widehat{\theta}_k^{\mathtt{B}}$	0.0532	0.4499	0.0680	0.0624	0.1022	0.0675	0.0993
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.2161	0.4323	0.2225	0.1859	0.3059	0.1351	0.2101
$\widehat{\theta}_k^{\mathtt{SL}}$	0.1342	0.3379	0.1101	0.1303	0.1563	0.1057	0.1877

A2	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.1915	0.7503	0.4357	0.6388	0.7489	0.2656	0.4457
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.3003	0.5448	0.5071	0.3925	0.2500	0.3404	0.5569
$\widehat{\theta}_k^{\mathtt{I}}$	0.1227	0.5738	0.0688	0.3600	0.2500	0.0565	0.5202
$\widehat{\theta}_k^{\mathtt{FF}}$	0.4435	0.9582	0.4794	0.6397	0.7500	0.2768	0.4458
$\widehat{\theta}_k^{\mathtt{BL}}$	**0.0283**	0.4765	0.0647	0.2051	0.0631	0.0680	**0.0615**
$\widehat{\theta}_k^{\mathtt{B}}$	0.0926	0.5931	0.0803	0.1206	0.2387	**0.0188**	0.1398
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0932	0.3744	0.3845	0.1216	0.2352	0.1007	0.2247
$\widehat{\theta}_k^{\mathtt{SL}}$	0.1211	0.3330	0.1085	0.1636	0.2534	0.1402	0.2763

A3	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.0950	0.4033	0.1573	0.3373	0.3289	0.1097	0.3671
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.1910	0.2989	0.3553	0.3774	0.2954	0.2495	0.4270
$\widehat{\theta}_k^{\mathtt{I}}$	0.0781	0.3848	0.0640	0.3600	0.4290	0.0573	0.3668
$\widehat{\theta}_k^{\mathtt{FF}}$	0.1726	0.6209	0.1978	0.4741	0.4548	0.1633	0.2951
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0321	0.4584	0.0641	0.0948	**0.0562**	0.0728	0.0657
$\widehat{\theta}_k^{\mathtt{B}}$	0.0839	0.5430	**0.0385**	**0.0320**	0.0868	0.0298	0.1542
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.1130	0.2348	0.1095	0.1169	0.1568	0.0986	0.1913
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0992	0.2507	0.0937	0.1143	0.1497	0.1013	0.1851

A4	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.1203	0.2496	0.1350	0.1409	0.1757	0.1336	0.2330
$\widehat{\theta}_k^{\mathtt{SL}}$	0.1054	**0.2313**	0.0931	0.1300	0.1753	0.1279	0.2279

Table 4: Absolute bias obtained for simulated samples of size $n = 1000$. For estimators $\widehat{\theta}_k^{\mathtt{BL}}$ and $\widehat{\theta}_k^{\mathtt{B}}$ we considered blocks of length 3 except in MCBEV and GARCH models where we used blocks of length 4 and 5, respectively. For estimators $\widehat{\theta}_k^{\mathtt{DJ}}$ and $\widehat{\theta}_k^{\mathtt{SL}}$ we considered blocks of length 20. For estimator $\widehat{\theta}_k^{\mathtt{R}}$ ($\widehat{\theta}_k^{\mathtt{FF}}$) we considered runs (cycles) of length 2 in MAR and AR, of length 3 in MM, ARCauchy and ARUnif, of length 4 in MCBEV and length 5 in GARCH. See Remark 3.1.

A1	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.0301	0.2748	0.0481	0.0607	**0.0128**	0.0134	0.0608
$\widehat{\theta}_k^{\mathtt{I}}$	**0.0007**	0.2901	0.0016	0.1815	0.2498	0.0107	0.1523
$\widehat{\theta}_k^{\mathtt{FF}}$	0.0604	0.4010	0.0593	0.0780	0.0074	0.0707	0.1184
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0032	0.4350	0.0595	0.0400	0.0625	0.0683	0.0703
$\widehat{\theta}_k^{\mathtt{B}}$	0.0146	0.4461	0.0420	0.0174	0.0881	0.0494	0.0886
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0368	0.3275	0.0347	0.0332	0.0237	0.0260	0.0934
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0417	0.2780	0.0459	0.0562	0.1221	0.0800	0.1571

A2	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.1768	0.7390	0.4300	0.6388	0.7489	0.2584	0.4457
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.0437	0.3963	0.2290	0.6301	0.2500	0.1622	0.6187
$\widehat{\theta}_k^{\mathtt{I}}$	0.1075	0.5675	0.0368	0.3600	0.2500	0.0334	0.5136
$\widehat{\theta}_k^{\mathtt{FF}}$	0.4389	0.9578	0.4783	0.6397	0.7500	0.2700	0.4458
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0029	0.4747	0.0533	0.1550	0.0156	0.0629	**0.0570**
$\widehat{\theta}_k^{\mathtt{B}}$	0.0894	0.5922	0.0650	0.0608	0.2316	0.0041	0.1370
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0283	0.3423	0.3701	0.0453	0.1888	0.0847	0.1956
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0872	0.3160	0.0839	0.1281	0.2122	0.1176	0.2455

A3	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.0853	0.4012	0.1554	0.3363	0.3278	0.1070	0.3667
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.0031	0.2511	0.1745	0.3805	0.2954	0.0456	0.3930
$\widehat{\theta}_k^{\mathtt{I}}$	0.0231	0.3732	**0.0015**	0.3373	0.2490	**0.0106**	0.3145
$\widehat{\theta}_k^{\mathtt{FF}}$	0.1703	0.6203	0.1965	0.4735	0.4545	0.1620	0.2945
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0006	0.4565	0.0550	0.0677	0.0254	0.0660	0.0597
$\widehat{\theta}_k^{\mathtt{B}}$	0.0775	0.5422	0.0314	**0.0017**	0.0855	0.0124	0.1522
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0597	0.2013	0.0679	0.0519	0.0959	0.0792	0.1490
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0647	0.2340	0.0673	0.0700	0.1147	0.0848	0.1533

A4	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0658	0.2172	0.0699	0.0860	0.1134	0.1104	0.2088
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0668	**0.1988**	0.0460	0.0833	0.1302	0.1106	0.2057

Table 5: Root mean squared errors obtained for simulated samples of size $n = 5000$. For estimators $\widehat{\theta}_k^{\mathtt{BL}}$ and $\widehat{\theta}_k^{\mathtt{B}}$ we considered blocks of length 3 except in MCBEV and GARCH models where we used blocks of length 4 and 5, respectively. For estimators $\widehat{\theta}_k^{\mathtt{DJ}}$ and $\widehat{\theta}_k^{\mathtt{SL}}$ we considered blocks of length 20. For estimator $\widehat{\theta}_k^{\mathtt{R}}$ ($\widehat{\theta}_k^{\mathtt{FF}}$) we considered runs (cycles) of length 2 in MAR and AR, of length 3 in MM, ARCauchy and ARUnif, of length 4 in MCBEV and length 5 in GARCH. See Remark 3.1.

A1	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.0434	0.2504	0.0395	0.0518	0.0314	0.0409	0.1257
$\widehat{\theta}_k^{\mathtt{I}}$	0.0460	0.2904	0.0327	0.0799	0.2499	0.0387	0.1004
$\widehat{\theta}_k^{\mathtt{FF}}$	0.0541	0.4309	0.0423	0.0558	**0.0303**	0.0608	0.1072
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0231	0.4234	0.0554	0.0585	0.0803	0.0715	0.0683
$\widehat{\theta}_k^{\mathtt{B}}$	0.0290	0.4272	0.0572	0.0457	0.0865	0.0656	0.0823
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0587	0.2439	0.0755	0.0895	0.1030	0.0789	0.1415
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0592	0.2317	0.0557	0.0796	0.0954	0.0781	0.1407

A2	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.1500	0.7498	0.4382	0.6396	0.7498	0.2648	0.4464
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.0700	0.3918	0.2273	0.4596	0.2500	0.1532	0.6150
$\widehat{\theta}_k^{\mathtt{I}}$	0.0940	0.5357	0.0446	0.3600	0.2500	0.0370	0.5351
$\widehat{\theta}_k^{\mathtt{FF}}$	0.4416	0.9505	0.4915	0.6398	0.7500	0.2749	0.4456
$\widehat{\theta}_k^{\mathtt{BL}}$	**0.0113**	0.4786	0.0606	0.2282	0.0510	0.0688	**0.0547**
$\widehat{\theta}_k^{\mathtt{B}}$	0.0838	0.5925	0.0502	0.0505	0.1818	0.0189	0.1308
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0862	0.3085	0.0942	0.1364	0.2416	0.1171	0.2121
$\widehat{\theta}_k^{\mathtt{SL}}$	0.1056	0.2843	0.1056	0.1595	0.2502	0.1350	0.2456

A3	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.0970	0.2576	0.1530	0.3369	0.3316	0.1063	0.3688
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.0700	0.3918	0.2273	0.4596	0.2500	0.1532	0.6150
$\widehat{\theta}_k^{\mathtt{I}}$	0.0611	0.4150	**0.0300**	0.3600	0.2499	0.0378	0.3648
$\widehat{\theta}_k^{\mathtt{FF}}$	0.1663	0.6147	0.1941	0.4709	0.4544	0.1587	0.2940
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0190	0.4818	0.0606	0.1082	0.0253	0.0728	0.0777
$\widehat{\theta}_k^{\mathtt{B}}$	0.0787	0.5461	0.0366	**0.0308**	0.0858	**0.0164**	0.1553
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0716	0.2528	0.0777	0.0992	0.1202	0.1035	0.1692
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0842	0.2695	0.0908	0.1062	0.1312	0.1152	0.1769

A4	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0664	**0.2149**	0.0561	0.0989	0.1147	0.1259	0.2109
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0644	0.2155	0.0523	0.0916	0.1066	0.1224	0.2100

Table 6: Absolute bias obtained for simulated samples of size $n = 5000$. For estimators $\widehat{\theta}_k^{\mathtt{BL}}$ and $\widehat{\theta}_k^{\mathtt{B}}$ we considered blocks of length 3 except in MCBEV and GARCH models where we used blocks of length 4 and 5, respectively. For estimators $\widehat{\theta}_k^{\mathtt{DJ}}$ and $\widehat{\theta}_k^{\mathtt{SL}}$ we considered blocks of length 20. For estimator $\widehat{\theta}_k^{\mathtt{R}}$ ($\widehat{\theta}_k^{\mathtt{FF}}$) we considered runs (cycles) of length 2 in MAR and AR, of length 3 in MM, ARCauchy and ARUnif, of length 4 in MCBEV and length 5 in GARCH. See Remark 3.1.

A1	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.0180	0.2431	0.0293	0.0273	0.0035	**0.0015**	0.1093
$\widehat{\theta}_k^{\mathtt{I}}$	0.0065	0.2784	**0.0024**	0.0522	0.2499	0.0118	**0.0482**
$\widehat{\theta}_k^{\mathtt{FF}}$	0.0360	0.4132	0.0321	0.0343	**0.0015**	0.0500	0.1023
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0044	0.4223	0.0050	0.0363	0.0660	0.0669	0.0638
$\widehat{\theta}_k^{\mathtt{B}}$	0.0015	0.4258	0.0462	0.0238	0.0774	0.0586	0.0779
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0372	0.2241	0.0383	0.0640	0.0752	0.0613	0.1229
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0379	0.2251	0.0469	0.0570	0.0711	0.0648	0.1202

A2	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.1394	0.7498	0.4374	0.6396	0.7498	0.2612	0.4464
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.0475	0.3896	0.2176	0.4050	0.2500	0.1428	0.6102
$\widehat{\theta}_k^{\mathtt{I}}$	0.0852	0.5310	0.0280	0.3600	0.2500	0.0267	0.5330
$\widehat{\theta}_k^{\mathtt{FF}}$	0.4406	0.9501	0.4908	0.6398	0.7500	0.2688	0.4456
$\widehat{\theta}_k^{\mathtt{BL}}$	**0.0009**	0.4783	0.0580	0.2202	0.0349	0.0680	0.0532
$\widehat{\theta}_k^{\mathtt{B}}$	0.0814	0.5919	0.0308	**0.0040**	0.1602	0.0048	0.1284
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0747	0.3002	0.0856	0.1167	0.2156	0.1073	0.1987
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0959	0.2793	0.0976	0.1393	0.2163	0.1240	0.2304

A3	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{R}}$	0.0947	0.2573	0.1526	0.3367	0.3314	0.1055	0.3687
$\widehat{\theta}_k^{\mathtt{NGJ}}$	0.0107	0.2890	0.1773	0.3896	0.2990	0.0859	0.4441
$\widehat{\theta}_k^{\mathtt{I}}$	0.0546	0.4124	0.0164	0.3600	0.2499	0.0210	0.3389
$\widehat{\theta}_k^{\mathtt{FF}}$	0.1657	0.6146	0.1939	0.4708	0.4544	0.1584	0.2939
$\widehat{\theta}_k^{\mathtt{BL}}$	0.0011	0.4813	0.0580	0.1006	0.0152	0.0701	0.0765
$\widehat{\theta}_k^{\mathtt{B}}$	0.0777	0.5460	0.0346	0.0064	0.0856	0.0114	0.1545
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0609	0.2470	0.0692	0.0857	0.1075	0.0990	0.1592
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0758	0.2660	0.0854	0.0988	0.1217	0.1121	0.1688

A4	MAR	AR	MM	ARCauchy	ARUnif	MCBEV	GARCH
$\widehat{\theta}_k^{\mathtt{DJ}}$	0.0562	**0.2079**	0.0451	0.0835	0.0853	0.1212	0.2049
$\widehat{\theta}_k^{\mathtt{SL}}$	0.0570	0.2097	0.0449	0.0812	0.0778	0.1190	0.2065

4. APPLICATION TO REAL DATA

We consider the daily maximum temperatures (in degrees Celsius) at Uccle (Belgium), from 1901 to 1999, on the warmest month of July (thus stationarity is assumed), consisting in $n = 3051$ observations. The data is available at "`http://lstat.kuleuven.be/Wiley/Data/ecad00045TX.txt`" and is plotted in Figure 2. The extremal index of this series was analyzed in Beirlant *et al.* ([2] 2004), where the respective estimates, obtained through parametric modeling, ranged between 0.49 and 0.56.

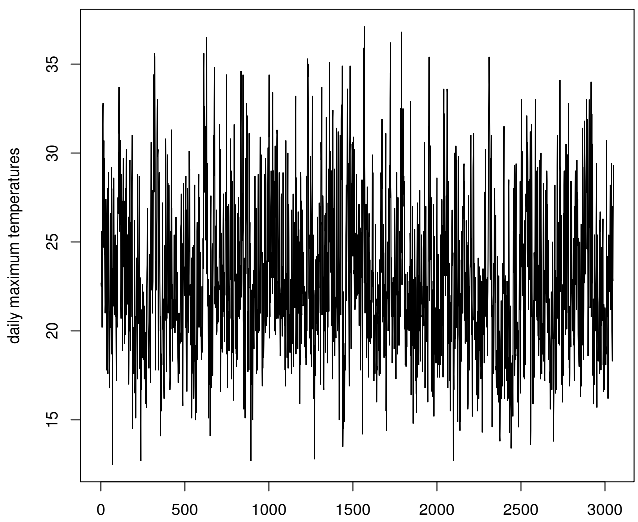

Figure 2: July daily maximum temperatures (in degrees Celsius) at Uccle, over the years 1901–1999.

We start by checking if we can validate some condition $\mathrm{D}^{(k)}(u_n)$. To this end, we use the empirical methodology of Ferreira and Ferreira ([10], 2015) by calculating the proportion of anti-$\mathrm{D}^{(m)}(u_n)$ events among the exceedances for several pairs of normalized levels u_n and block sizes r_n:

$$p(u_n, r_n) = \frac{\sum_{j=1}^{n-r_n+1} \mathbb{1}_{\left\{X_j>u_n, X_{j+1}\le u_n, \ldots, X_{j+m-1}\le u_n, M_{j+m-1, r_n+j-1}>u_n\right\}}}{\sum_{j=1}^{n} \mathbb{1}_{\{X_j>u_n\}}}.$$

More precisely, for each fixed $\tau > 0$, we take u_n as the empirical $(1-\tau/n)$-th quantile for increasing sample sizes n and choose the sequence $\{b_n = [n/r_n]\}_n$ growing at a slower rate than n, e.g., $b_n = [(\log n)^a]$, for some $a > 0$. If $\mathrm{D}^{(m)}(u_n)$ holds with b_n, the points $(n, p(u_n, r_n))$ approach zero as $n \to \infty$. Based on the suggested declustering parameter $r = 4$ in Beirlant *et al.* ([2] 2004), we have analyzed the proportions of anti-D$^{(4)}(u_n)$, plotted in Figure 3 (right panel) for $\tau = 15$ (full line) and $\tau = 20$ (dashed line), with $k_n = [(\log n)^{2.5}]$. Observe that the values are small and almost indistinguishable from the proportions of anti-D$^{(3)}(u_n)$ (left panel). We have also taken $k_n = [(\log n)^3]$ which led to null proportions in both cases. Therefore, we assume the validity of the D$^{(3)}(u_n)$ local condition and consider run length 3 for the runs estimator and cycles of length 2 for the FF estimator in (2.8); see Remark 3.1. We also take block-length 3 in the blocks estimators. The disjoint and slides methods were implemented with block-length 15.

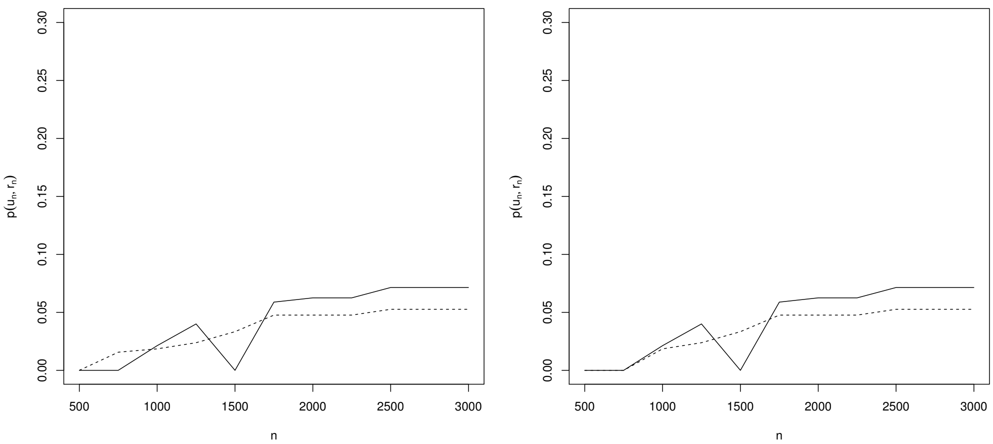

Figure 3: Observed proportions of anti-D$^{(3)}(u_n)$ (left) and anti-D$^{(4)}(u_n)$ conditions for Uccle data, for $\tau = 15$ (full line) and $\tau = 20$ (dashed line), with $k_n = [(\log n)^{2.5}]$.

The sample paths of the considered estimators in (2.11) and (3.1) are in Figure 4. Under algorithm A4, we obtained the estimate 0.51 for both disjoint and slide estimators. We have also applied the bias-reduced GJ Nandagopalan's runs estimator in (3.1) from which the values 0.41 and 0.57 were derived under A2 and A3, respectively. The remaining estimates are summarized in Table 7. The results are mostly in agreement with the simulation study.

Table 7: Extremal index estimates for Uccle data.

	$\widehat{\theta}_k^{\mathtt{R}}$	$\widehat{\theta}_k^{\mathtt{I}}$	$\widehat{\theta}_k^{\mathtt{FF}}$	$\widehat{\theta}_k^{\mathtt{B}}$	$\widehat{\theta}_k^{\mathtt{BL}}$	$\widehat{\theta}_k^{\mathtt{DJ}}$	$\widehat{\theta}_k^{\mathtt{SL}}$
A1	0.49	0.47	0.46	0.50	0.51	0.53	0.57
A2	0.10	0.33	0.05	0.39	0.50	0.52	0.53
A3	0.32	0.30	0.28	0.42	0.50	0.49	0.53

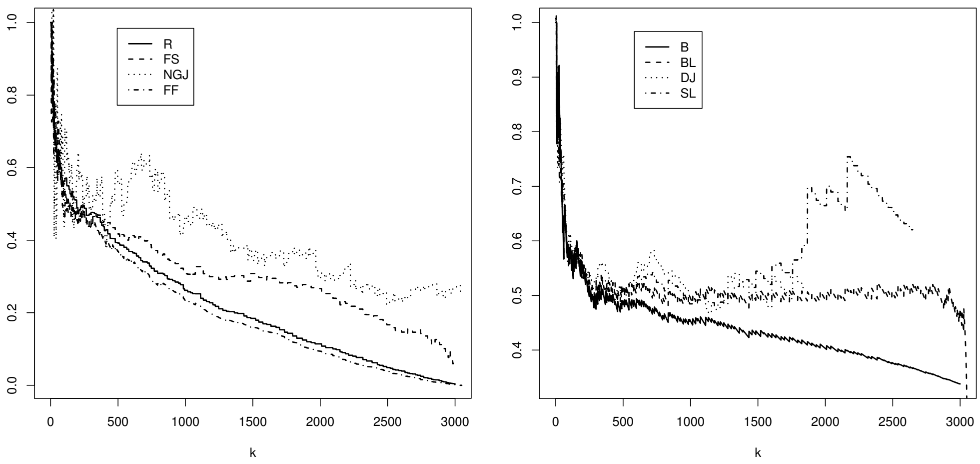

Figure 4: Sample paths of estimators in (2.11) and estimator (3.1) for Uccle data.

5. DISCUSSION

We have analyzed several estimators of the extremal index under different methodologies. The procedure based in Frahm *et al.* ([12] 2005) revealed an overall satisfactory performance. The best results were mostly observed within the blocks estimators, $\widehat{\theta}_k^{\mathsf{B}}$ and $\widehat{\theta}_k^{\mathsf{BL}}$, under the methodology of Neves *et al.* ([23] 2015). The large biases observed in the AR process makes inference within weak dependence, i.e., $\theta = 1$, an open topic to explore in this framework. Other methods to analyze the local dependence D-conditions are also welcome. The bias-reduced GJ Nandagopalan's estimator is sensitive to the restricted condition D'' and a generalization of the method to the broader runs estimator may be more advantageous. These points will be addressed in a future work.

ACKNOWLEDGMENTS

This research was financed by Portuguese Funds through FCT — Fundação para a Ciência e a Tecnologia, through the projects UID/MAT/00013/2013 and UID/MAT/00006/2013 and by the Research Center CEMAT through the Project UID/Multi/04621/2013.
The author also acknowledges the valuable suggestions from the referees.

REFERENCES

[1] Ancona-Navarrete, M.A. and Tawn, J.A. (2000). A comparison of methods for estimating the extremal index, *Extremes*, **3**, 5–38.

[2] Beirlant, J.; Goegebeur, Y.; Segers, J. and Teugels, J. (2004). *Statistics of Extremes: Theory and Application*, John Wiley.

[3] Chernick, M.R. (1978). *Mixing Conditions and Limit Theorems for Maxima of Some Stationary Sequences*, PhD dissertation, Stanford University.

[4] Chernick, M.R.; Hsing, T. and McCormick, W.P. (1991). Calculating the extremal index for a class of stationary sequences, *Advaces in Applied Probability*, **23**, 835–850.

[5] Ferreira, M. (2014). *A heuristic procedure to estimate the tail index.* In "Proceedings of the 14th International Conference in Computational Science and Its Applications — ICCSA 2014" (B. Apduhan, A.M. Rocha, S. Misra, D. Taniar, O. Gervasi and B. Murgante, Eds.), IEEE-Computer Society 4264a241, 241–245.

[6] Ferreira, M. (2015a). Estimating the tail index: Another algorithmic method, *ProbStat Forum*, **8**(5), 45–53.

[7] Ferreira, M. (2015b). Estimating the extremal index through the tail dependence concept, *Discussiones Mathematicae Probability and Statistics*, **35**, 61–74.

[8] Ferreira, M. and Ferreira, H. (2012). On extremal dependence: some contributions, *TEST*, **21**(3), 566–583.

[9] Ferreira, M. and Silva, S. (2014). An Analysis of a Heuristic Procedure to Evaluate Tail (in)dependence, *Journal of Probability and Statistics*, Vol. **2014**, Article ID 913621, 15 pages.

[10] Ferreira, H. and Ferreira, M. (2015). Estimating the extremal index through local dependence. (Submitted.)

[11] Ferro, C.A. and Segers, J. (2003). Inference for clusters of extremes, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **65**, 545–556.

[12] Frahm, G.; Junker, M. and Schmidt, R. (2005). Estimating the tail-dependence coefficient: properties and pitfalls, *Insurance: Mathematics and Economics*, **37**(1), 80–100.

[13] Gomes, M.I.; Hall, A. and Miranda, C. (2008). Subsampling techniques and the jackknife methodology in the estimation of the extremal index, *Journal of Statistical Computation and Simulation*, **52**, 2022–2041.

[14] Gomes, M.I.; Henriques-Rodrigues, L.; Fraga Alves, M.I. and Manjunath, B.G. (2013). Adaptive PORT-MVRB estimation: an empirical comparison of two heuristic algorithms, *Journal of Statistical Computation and Simulation*, **83**(6), 1129–1144.

[15] Hsing, T. (1991). Estimating the parameters of rare events, *Stochastic Processes and their Applications*, **37**, 117–139.

[16] Hsing, T. (1993). Extremal index estimation for a weakly dependent stationary sequence, *Annals of Statistics*, **21**, 2043–2071.

[17] Hsing, T.; Hüsler, J. and Leadbetter, M.R. (1988). On the exceedance point process for a stationary sequence, *Probability Theory and Related Fields*, **78**, 97–112.

[18] Laurini, F. and Tawn, J. (2012). The extremal index for GARCH(1,1) processes, *Extremes*, **15**, 511–529.

[19] Leadbetter, M.R. (2012). On extremes values in stationary sequences, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **15**, 511–529.

[20] Leadbetter, M.R.; Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer, New York.

[21] Leadbetter, M.R. and Nandagopalan, S. (1989). *On exceedance point process for stationary sequences under mild oscillation restrictions.* In "Extreme value theory: proceedings, Oberwolfach 1987" (J. Hüsler and D. Reiss, Eds.), Springer, New York, 69–80.

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

[23] Neves, M.; Gomes, M.I.; Figueiredo, F. and Prata Gomes, D. (2012). Modeling Extreme Events: Sample Fraction Adaptive Choice in Parameter Estimation, *Journal of Statistical Theory and Practice*, **9**(1), 184–199.

[24] O'Brien, G.L. (1974). The maximum term of uniformly mixing stationary processes, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **30**, 57–63.

[25] Robert, C.Y.; Segers, J. and Ferro, C.A. (2009). A sliding blocks estimator for the extremal index, *Electronic Journal of Statistics*, **3**, 993–1020.

[26] Smith, R.L. (1992). The extremal index for a Markov chain, *Journal of Applied Probability*, **29**, 37–45.

[27] Smith, R.L. and Weissman, I. (1994). Estimating the extremal index, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **56**, 515–528.

[28] Süveges, M. (2007). Likelihood estimation of the extremal index, *Extremes*, **10**, 41–55.

[29] Weissman, I. and Novak, S.Y. (1998). On blocks and runs estimators of the extremal index, *Journal of Statistical Planning and Inference*, **66**(2), 281–288.