Discussiones Mathematicae
Probability and Statistics xx (xxxx) 1–14

ESTIMATING THE EXTREMAL INDEX THROUGH THE TAIL DEPENDENCE CONCEPT

Marta Ferreira

Center of Mathematics
University of Minho, Portugal

e-mail: msferreira@math.uminho.pt

Abstract

The extremal index θ is an important parameter in extreme value analysis when extending results from independent and identically distributed sequences to stationary ones. A connection between the extremal index and the tail dependence coefficient allows the introduction of new estimators. The proposed ones are easy to compute and we analyze their performance through a simulation study. Comparisons with other existing methods are also presented. Case studies within environment are considered in the end.

Keywords: Extreme value theory, extremal index, tail dependence coefficient.

2010 Mathematics Subject Classification: 62G32.

1. Introduction

The central result in classical Extreme Value Theory states that, for an i.i.d. sequence, $\{X_n\}_{n\geq 1}$, having common distribution function (d.f.) F, if there are constants $a_n > 0$ and $b_n \in \mathrm{R}$ such that,

$$P(\max(X_1,\ldots,X_n) \leq a_n x + b_n) \underset{n\to\infty}{\longrightarrow} G(x)\,, \tag{1}$$

for some non degenerate function G, then it must be the Generalized Extreme Value function (GEV),

$$G(x) = \exp(-(1+\gamma x)^{-1/\gamma}),\ 1+\gamma x > 0,\ \gamma \in \mathbb{R},$$

($G(x) = \exp(-e^{-x})$ for $\gamma = 0$) and we say that F belongs to the max-domain of attraction of G, in short, $F \in \mathcal{D}(G)$. The parameter γ, known as the tail index, is a shape parameter determining the tail behavior of F: if $\gamma > 0$ we are in the domain of attraction Fréchet corresponding to a heavy tail, $\gamma < 0$ indicates the Weibull domain of attraction of light tails and $\gamma = 0$ means a Gumbel domain of attraction and an exponential tail.

In a multivariate context, it is possible to extend the convergence given in (1), but the class of models in the limit is much wider than model GEV. For simplicity, we consider the bivariate case, but everything can be rewritten for the more general d-variate case, $d \geq 2$. More precisely, let $\{(X_1^{(n)}, X_2^{(n)})\}_{n\geq 1}$ be a sequence of i.i.d. copies of the random pair (X_1, X_2), with common d.f. $\mathbf{F}$, and let $M_j^{(n)} = \max_{1\leq i\leq n} X_j^{(i)}$, $j = 1, 2$, be the maximum of each marginal. If there exist sequences of real constants $a_j^{(n)} > 0$ and $b_j^{(n)}$, for $j = 1, 2$ and $n \geq 1$, and a d.f. $\mathbf{G}$ with non-degenerate margins, such that,

$$\begin{aligned} & P(M_1^{(n)} \leq a_1^{(n)} x_1 + b_1^{(n)}, M_2^{(n)} \leq a_2^{(n)} x_2 + b_2^{(n)}) \\ = \ & \mathbf{F}^n(a_1^{(n)} x_1 + b_1^{(n)}, a_2^{(n)} x_2 + b_2^{(n)}) \\ \underset{n\to\infty}{\longrightarrow} \ & \mathbf{G}(x_1, x_2)\,, \end{aligned}$$

for every continuity points of $\mathbf{G}$, then this latter is said to be a bivariate extreme value distribution (BEV) and is defined by expression

$$\mathbf{G}(x_1, x_2) = \exp[-l\{-\log G_1(x_1), -\log G_2(x_2)\}], \tag{2}$$

for some bivariate function l, where G_j, $j = 1, 2$, is the marginal d.f. of $\mathbf{G}$. In this case, we have that $\mathbf{F}$ belongs to the max-domain of attraction of $\mathbf{G}$, in short $\mathbf{F} \in \mathcal{D}(\mathbf{G})$. The function l in (2), usually called *stable tail dependence function* is convex and homogeneous of order 1, and we have $\max(x_1, x_2) \leq l(x_1, x_2) \leq x_1 + x_2$, for all $(x_1, x_2) \in [0, \infty)^2$, where the upper limit corresponds to independence and the lower one means complete dependence (see, e.g. Beirlant *et al.* [2], Section 8.2.2).

The result in (1) may also be extended to study the maximum of a wide class of dependent processes, a more realistic assumption for several data. Here we concentrate on stationary sequences where the dependence is restricted by distributional mixing conditions.

The condition $\mathrm{D}(u_n)$ of Leadbetter ([14], 1983), providing a short range dependence for which at long lags the extremes are independent, is sufficient to extend the result in (1) to stationary sequences. More precisely, for a stationary sequence $\{X_n\}_{n\geq 1}$ satisfying $\mathrm{D}(u_n)$ with $u_n = a_n x + b_n$, we have that

$$P(\max(X_1, \ldots, X_n) \leq u_n) \underset{n\to\infty}{\longrightarrow} G^{\theta}(x)\,, \tag{3}$$

where $0 \leq \theta \leq 1$ is the extremal index. The extremal index is the primary measure of extremal dependence in such processes, with $\theta = 1$ indicating independence at asymptotically high levels.

There are different interpretations of the extremal index. This concept, originated in papers by Loynes ([15], 1965), O'Brien ([17], 1974) and developed in detail by Leadbetter ([14], 1983), reflects the effect of clustering of extreme observations on the limiting distribution of the maximum. O'Brien (1987) proved that the presence of clustering affects the limiting distribution of block maxima:

$$P(\max(X_2, \ldots, X_{r_n}) \leq u_n | X_1 > u_n) \underset{n\to\infty}{\longrightarrow} \theta\,, \tag{4}$$

with r_n such that $r_n \to \infty$ and $r_n = o(n)$. Under a mixing condition slightly restrictive than $\mathrm{D}(u_n)$, Hsing *et al.* ([13], 1988) showed that the limiting mean number of exceedances of u_n in an interval of length r_n is the inverse of the extremal index:

$$E\big[\textstyle\sum_{i=1}^{r_n} \mathbb{1}_{\{X_j > u_n\}} | \sum_{i=1}^{r_n} \mathbb{1}_{\{X_j > u_n\}} \geq 1\big] \to \theta^{-1}, \tag{5}$$

with $\mathbb{1}(\cdot)$ the indicator function. By stationarity this property is satisfied for any block of r_n consecutive elements defined in the sequence. By rewriting (3) as

$$P(\max(X_1, \ldots, X_n) \leq u_n) \underset{n\to\infty}{\longrightarrow} e^{-\theta\tau(x)}\,, 0 < \tau(x) < \infty,$$

Ferro and Segers ([9], 2003) found that the process of inter-exceedance times normalized by exceedances of u_n follows a mixture of a point mass and an exponential distribution $Exp(\theta^{-1})$, i.e.,

$$P(\overline{F}(u_n)T(u_n) > t) \underset{n\to\infty}{\longrightarrow} \theta e^{-\theta t}\,, t > 0, \tag{6}$$

with $T(u_n) = \min\{n \geq 1 : X_{n+1} > u_n | X_1 > u_n\}$, also under a slightly stricter mixing condition than $\mathrm{D}(u_n)$.

Inference about θ has been extensively studied, with the most popular estimators being the runs method obtained from equation (4), the blocks method derived from (5) and the intervals method developed from (6). More precisely, the runs estimator is given by

$$\widehat{\theta}^{(R)} = (N)^{-1} \sum_{i=1}^{n-1} \mathbb{1}_{\{X_i > u\}} \mathbb{1}_{\{X_{i+1} \leq u\}} \cdots \mathbb{1}_{\{X_{i+r} \leq u\}},$$

where N is the total number of exceedances of a high threshold u. The blocks estimator for a sample divided into b blocks of length r (so $n \approx br$), can be stated as

$$\widehat{\theta}^{(B)} = \frac{\log(1 - C_n(u)/b)}{r \log(1 - N/n)}$$

where $C_n(u)$ is the number of blocks in which at least one exceedance of u occurs. After some considerations, the result in (6) yields the intervals estimator

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

with T_i denoting the ith inter-exceedance time, $i = 1, \ldots, N-1$. For a survey, see for instance, Ancona-Navarrete and Tawn ([1], 2000) and Beirlant *et al.* ([2], 2004).

Imposing some convenient local dependence condition may eliminate the need for a cluster identification scheme as in the case of the blocks or the runs estimators. An example of such condition is the local dependence condition $\mathrm{D}^{(2)}(u_n)$ of Chernick *et al.* (1991), which holds whenever

$$nP(X_j > u_n, X_{j+1} \leq u_n, M_{j+2,r_n} > u_n) \to 0,\ n \to \infty,$$

with $M_{i,j} = \max\{X_i, \ldots, X_j\}$, for $i \leq j$ ($M_{i,j} = -\infty$ if $i > j$), the block sizes sequence $\{r_n\}$ is such that $n/r_n \to \infty$ and condition $\mathrm{D}(u_n)$ is simultaneously satisfied. Condition $\mathrm{D}^{(2)}(u_n)$ restricts the occurrence of an observation again exceeding the high threshold u_n after dropping below it within a cluster.

Under $\mathrm{D}^{(2)}(u_n)$, and considering a log-likelihood based on the limiting d.f. obtained in (6), Süveges ([22], 2007) presents the maximum likelihood estimator

$$\widehat{\theta}^{(ML)} = \frac{\sum_{i=1}^{N-1} qS_i + N - 1 + N_C - \left[\left(\sum_{i=1}^{N-1} qS_i N - 1 + N_C\right)^2 - 8N_C \sum_{i=1}^{N-1} qS_i\right]^{1/2}}{2\sum_{i=1}^{N-1} qS_i},$$

where q is the estimate of $\overline{F}(u)$, $S_i = T_i - 1$ and $N_C = \sum_{i=1}^{N-1} \mathbb{1}_{\{S_i \neq 0\}}$.

Considering a lightly stronger condition $\mathrm{D}''(u_n)$ that restricts the occurrence of two or more upcrossings by imposing that $n\sum_{j=2}^{r_n-1} P(X_1 > u_n, X_j \leq u_n < X_{j+1}) \to 0$, as $n \to \infty$, Nandagopalan ([16], 1990) derives the estimator

$$\widehat{\theta}^{(N)} = \frac{\sum_{j=1}^{n-1} \mathbb{1}_{\{X_j \leq u < X_{j+1}\}}}{\sum_{j=1}^{n} \mathbb{1}_{\{X_j > u\}}},$$

for a suitable high threshold u. This is a special case of the runs estimator when $r = 1$.

A recent result in Ferreira and Ferreira ([7], 2012a), allow us to state $\theta = 1-\lambda$ under condition $\mathrm{D}^{(2)}(u_n)$, where λ is the tail dependence coefficient introduced by Sibuya ([21], 1960). Here we shall analyze the estimation of θ based on some λ estimation methodologies of the literature. This will be done through a simulation study. The performance of our approach will be also assessed by comparing with the simulation results obtained for the above exposed existing estimators of the extremal index. At the end, we illustrate with applications to real environmental data.

2. Tail dependence

The *tail-dependence coefficient* (TDC), usually denoted λ and first introduced in Sibuya ([21], 1960), measures the probability of occurring extreme values for one random variable (r.v.) given that another assumes an extreme value too, i.e.,

$$\lambda = \lim_{t\to\infty} P(F_1(X_1) > 1 - 1/t | F_2(X_2) > 1 - 1/t), \tag{7}$$

where F_1 and F_2 are the distribution functions (d.f.'s) of r.v.'s X_1 and X_2, respectively. It characterizes the dependence in the tail of a random pair (X_1, X_2), in the sense that, $\lambda > 0$ corresponds to tail dependence whereas $\lambda = 0$ means tail independence.

The relation $\theta = 1 - \lambda$ stated in Proposition 4 of Ferreira and Ferreira ([7], 2012a) under the local dependence condition $\mathrm{D}^{(2)}$, lead to new estimators for θ through the TDC. A wide study concerning TDC estimation is presented in Frahm *et al.* (2005). Parametric estimators are more accurate but may have disastrous performances under wrong model assumptions. Here we will focus on nonparametric approach.

Schmidt and Stadtmüller ([19], 2006) considered the estimator based on (7) by plugging-in the respective empirical counterparts,

$$\widehat{\lambda}^{(SS)} \equiv \widehat{\lambda}^{(SS)}(k_n) = \frac{1}{k_n}\sum_{i=1}^{n} \mathbb{1}_{\{\widehat{F}_1(X_1)>1-\frac{k_n}{n},\widehat{F}_2(X_2)>1-\frac{k_n}{n}\}}, \tag{8}$$

where $\widehat{F}_j$ is the empirical d.f. of F_j, $j = 1, 2$, and $\{k_n\}$ is an intermediate sequence, i.e., $k_n \to \infty$ and $k_n/n \to 0$, as $n \to \infty$. Concerning estimation accuracy, some modifications of this latter may be used, like replacing the denominator n by $n+1$, i.e., considering

$$\widehat{F}_j(u) = \frac{1}{n+1}\sum_{k=1}^{n} \mathbb{1}_{\{X_j^{(k)} \le u\}}$$

(for a discussion on this topic see, for instance, Beirlant et al. 2004). The choice of the value k in the sequence $\{k_n\}$ that allows the better trade-off between bias and variance is of major difficulty, since small values of k come along with a large variance whenever an increasing k results in a strong bias. The true value is usually located at a stable region of the plot $(k, \widehat{\lambda}^{(SS)}(k))$, for $1 \leq k < n$.

In order to avoid the variance-bias problem, we will use an heuristic procedure presented in Frahm *et al.* ([10], 2005), consisting on a "plateau finding algorithm" applied to a smoothed version of $(k, \widehat{\lambda}^{(SS)}(k))$, $1 \leq k < n$.

Based on the approach considered in Capéraà *et al.* ([3], 1997), which assumes that the underlying distribution approximates a BEV model given in (2), Frahm *et al.* ([10], 2005) have proposed the following estimator:

$$\widehat{\lambda}^{(CFG)} = 2 - 2\exp\left\{ \tfrac{1}{n} \textstyle\sum_{i=1}^{n} \log\left(\frac{\sqrt{\log \widehat{F}_1(X_1) \log \widehat{F}_2(X_2)}}{\log(\widehat{F}_1(X_1) \vee \widehat{F}_2(X_2))^{-2}} \right) \right\}, \tag{9}$$

where $x \vee y = \max(x, y)$. Another estimator developed in Ferreira and Ferreira ([8], 2012b) under the same assumption but with a simpler form, is given by

$$\widehat{\lambda}^{(FF)} = 3 - (1 - \overline{\widehat{F}_1(X_1) \vee \widehat{F}_2(X_2)})^{-1},$$

where $\overline{\widehat{F}_1(X_1) \vee \widehat{F}_2(X_2)}$ is the sample mean of $\widehat{F}_1(X_1) \vee \widehat{F}_2(X_2)$, i.e.,

$$\overline{\widehat{F}_1(X_1) \vee \widehat{F}_2(X_2)} = \frac{1}{n} \sum_{i=1}^{n} \left[\widehat{F}_1(X_1^{(i)}) \vee \widehat{F}_2(X_2^{(i)}) \right].$$

For a discussion about the asymptotic properties of these estimators see, respectively, Genest and Segers ([11], 2009) and Ferreira ([6], 2013).

From now on, we will use notation $\widehat{\theta}^{(SS)}$, $\widehat{\theta}^{(CFG)}$ and $\widehat{\theta}^{(FF)}$, whenever we refer to estimators $\widehat{\lambda}^{(SS)}$, $\widehat{\lambda}^{(CFG)}$ and $\widehat{\lambda}^{(FF)}$, that is,

$$\widehat{\theta}^{(SS)} = 1 - \widehat{\lambda}^{(SS)}, \quad \widehat{\theta}^{(CFG)} = 1 - \widehat{\lambda}^{(CFG)} \quad \text{and} \quad \widehat{\theta}^{(FF)} = 1 - \widehat{\lambda}^{(FF)}.$$

3. Simulation study

We are going to analyze the performance of the estimators described above, through a simulation study based on the following models:

- Independent sequence which have $\theta = 1$ (with unit Fréchet margins).
- Markov Gaussian dependence process, $Z_j = \alpha Z_{j-1} + \epsilon_j$, where the ϵ_j are i.i.d. $\mathrm{N}(0, 1-\alpha^2)$ r.v.'s, for $j \geq 2$ and Z_1 is $\mathrm{N}(0,1)$ distributed. This process has $\theta = 1$ and shall be denoted AR.

- Bivariate extreme value Markov process with logistic dependence function, i.e.,
$$P(X_j \leq x, X_{j+1} \leq y) = \exp(-(x^{1/\alpha} + y^{1/\alpha})^{\alpha}).$$
As in Ancona-Navarrete and Tawn ([1], 2000), we consider the dependence parameter $\alpha = 0.5$ which gives $\theta = 0.328$, and denote the process *BEV*.

- Autoregressive maximum process, $X_i = \alpha X_{i-1} \vee \epsilon_i$, where $0 < \alpha < 1$ and $\{\epsilon_i\}$ are i.i.d. r.v.'s with d.f. $F_\epsilon(x) = \exp(-(1-\alpha)/x)$, $x > 0$. This process has $\theta = 1 - \alpha$. We consider $\alpha = 0.5$ and hence $\theta = 0.5$, and denote the process *MAR*.

- Moving maxima process, $X_i = \bigvee_{j=0,\ldots,m} \alpha_j \epsilon_{i-j}$, with $\sum_{j=0}^{m} \alpha_j = 1$ and $\alpha_j \geq 0$, $\{\epsilon_i\}$ are i.i.d. unit Féchet r.v.'s. This process has $\theta = \vee_{j=0,\ldots,m} \alpha_j$. We consider $m = 3$, $\alpha_1 = \alpha_2 = 0.2$, $\alpha_0 = \alpha_3 = 0.3$ and so $\theta = 0.3$, and denote the process *MM*.

We consider samples of size $n = 10000$ and compare the estimators using the absolute mean bias and the root mean square error (rmse) criteria, obtained using 200 independent replications of the estimation procedures. The results of the proposed estimators, $\widehat{\theta}^{(FF)}$, $\widehat{\theta}^{(CFG)}$ and $\widehat{\theta}^{(SS)}$, are presented in Table 1. For comparison, we also include the simulation results obtained from estimators $\widehat{\theta}^{(ML)}$ and $\widehat{\theta}^{(N)}$ derived under similar local dependence conditions, i.e., $\mathrm{D}^{(2)}$ and D'', respectively (see Table 2). The estimates derived from the runs, the blocks and the intervals methods were also computed and can be found in Table 3. We remark that the values considered for the number of blocks/runs were derived through additional simulation studies conducted in Ancona-Navarrete and Tawn, ([1], 2000).

Observe that the worst performance of the estimators coincides with the AR process. In this case, estimator $\widehat{\theta}^{(SS)}$ followed by $\widehat{\theta}^{(ML)}$, $\widehat{\theta}^{(N)}$, $\widehat{\theta}^{(B)}$ and $\widehat{\theta}^{(I)}$ for $u = q_{0.99}$ exceed the remaining. In particular, the bad performance of the proposed estimators $\widehat{\theta}^{(FF)}$ and $\widehat{\theta}^{(CFG)}$ is due to the bad behavior of the respective tail dependence coefficient estimators $\widehat{\lambda}^{(FF)}$ in (8) and $\widehat{\lambda}^{(CFG)}$ in (9) under tail independent non-BEV models, i.e., models for which $\lambda = 0$ and whose dependence structure for consecutive pairs can not be formulated as in (2), such as the case of AR (see Ferreira, [6] 2013). Indeed, estimators $\widehat{\theta}^{(FF)}$ and $\widehat{\theta}^{(CFG)}$ are not robust. They present the worst performances also within the BEV and MM processes, missing the $\mathrm{D}^{(2)}$ condition. Therefore, concerning robustness, the best of the three here proposed estimators is $\widehat{\theta}^{(SS)}$, which only demands the $\mathrm{D}^{(2)}$ condition and behaves better whenever this latter is violated (see the results for BEV and AR in Table 1). All the estimators behave quite well in the MAR process, with the best performances occurring for our proposals $\widehat{\theta}^{(FF)}$ and $\widehat{\theta}^{(CFG)}$, as well as, for $\widehat{\theta}^{(ML)}$ and $\widehat{\theta}^{(N)}$ with $u = q_{0.99}$. We remark that this process satisfies condition

$D^{(2)}$ as well as the BEV dependence assumption (see, e.g., Ferreira and Ferreira [7] 2012a and Ancona-Navarrete and Tawn [1] 2000). Regarding the MM case, the best performance lies with the runs, blocks and intervals estimators, which is not surprising since it is easy to identify independent clusters in this process.

Table 1. Sample absolute mean bias and rmse (in brackets) of estimators $\widehat{\theta}^{(FF)}$, $\widehat{\theta}^{(CFG)}$ and $\widehat{\theta}^{(SS)}$.

	$\widehat{\theta}^{(FF)}$	$\widehat{\theta}^{(CFG)}$	$\widehat{\theta}^{(SS)}$
Indep.	0.00 (0.010)	0.00 (0.010)	0.05 (0.050)
AR	0.40 (0.403)	0.36 (0.364)	0.12 (0.131)
BEV	0.09 (0.088)	0.09 (0.089)	0.06 (0.063)
MAR	0.00 (0.010)	0.00 (0.010)	0.03 (0.041)
MM	0.10 (0.100)	0.10 (0.101)	0.07 (0.073)

Table 2. Sample absolute mean bias and rmse (in brackets) of estimators $\widehat{\theta}^{(ML)} \equiv \widehat{\theta}_u^{(ML)}$ and $\widehat{\theta}^{(N)} \equiv \widehat{\theta}_u^{(N)}$, by considering thresholds $u = q_{0.95}, q_{0.99}$, respectively, the empirical quantiles 0.95 and 0.99.

	$\widehat{\theta}_{q_{0.95}}^{(ML)}$	$\widehat{\theta}_{q_{0.99}}^{(ML)}$	$\widehat{\theta}_{q_{0.95}}^{(N)}$	$\widehat{\theta}_{q_{0.99}}^{(N)}$
Indep.	0.05 (0.045)	0.01 (0.000)	0.05 (0.055)	0.01 (0.000)
AR	0.24 (0.237)	0.13 (0.130)	0.24 (0.245)	0.13 (0.134)
BEV	0.08 (0.089)	0.10 (0.114)	0.08 (0.077)	0.09 (0.114)
MAR	0.01 (0.032)	0.00 (0.045)	0.02 (0.032)	0.00 (0.045)
MM	0.10 (0.095)	0.11 (0.118)	0.09 (0.089)	0.11 (0.114)

3.1. Case studies

3.1.1. Wooster temperatures

We consider the daily minimum temperatures (in degrees Fahrenheit) at Wooster (Ohio), from 1983 to 1988, more precisely, the period of November-February winter months in order to achieve some stationarity (see Figure 1). This series was analyzed in Coles ([5], 2001) and blocks estimates were computed

Table 3. Sample absolute mean bias and rmse (in brackets) of runs estimator $\widehat{\theta}^{(R)} \equiv \widehat{\theta}_u^{(R)}$, blocks estimator $\widehat{\theta}^{(B)} \equiv \widehat{\theta}_u^{(B)}$ and intervals estimator $\widehat{\theta}^{(I)} \equiv \widehat{\theta}_u^{(I)}$ by considering thresholds $u = q_{0.95}, q_{0.99}$, respectively, the empirical quantiles 0.95 and 0.99. In the blocks and runs estimators it was used the suggested number of runs/blocks in Ancona-Navarrete and Tawn ([1], 2000).

	$\widehat{\theta}^{(R)}_{q_{0.95}}$	$\widehat{\theta}^{(R)}_{q_{0.99}}$	$\widehat{\theta}^{(B)}_{q_{0.95}}$	$\widehat{\theta}^{(B)}_{q_{0.99}}$	$\widehat{\theta}^{(I)}_{q_{0.95}}$	$\widehat{\theta}^{(I)}_{q_{0.99}}$
Indep.	0.05 (0.055)	0.01 (0.000)	0.00 (0.008)	0.01 (0.014)	0.01 (0.000)	0.03 (0.055)
AR	0.37 (0.370)	0.19 (0.183)	0.24 (0.241)	0.13 (0.135)	0.22 (0.224)	0.13 (0.155)
BEV	0.03 (0.028)	0.04 (0.063)	0.07 (0.064)	0.03 (0.090)	0.04 (0.055)	0.03 (0.084)
MAR	0.02 (0.032)	0.00 (0.045)	0.03 (0.044)	0.02 (0.034)	0.03 (0.045)	0.03 (0.084)
MM	0.03 (0.027)	0.00 (0.031)	0.02 (0.030)	0.03 (0.041)	0.03 (0.045)	0.02 (0.055)

for the extremal index. In particular, it was considered the threshold $u = -10$ with number of blocks $b = 20, 31$ leading to, respectively, $\widehat{\theta}^{(B)} = 0.27, 0.42$.

Since we have a sample of minimum values we assume that an approximation to a BEV model dependence structure between consecutive pairs is plausible. In order to check condition $\mathrm{D}^{(2)}$, we use the empirical methodology of Süveges ([22], 2007) by calculating the proportion of anti-$\mathrm{D}^{(2)}$ events among the exceedances for a range of block sizes and thresholds:

$$p(u,r) = \frac{\sum_{j=1}^{n} \mathbb{1}_{\{X_j>u, X_{j+1}\le u, M_{j+2,r}>u\}}}{\sum_{j=1}^{n} \mathbb{1}_{\{X_j>u\}}}.$$

Observe in Figure 2 that $p(u,r) \approx 0$ as u and r increase, which leads to an informal validation of $\mathrm{D}^{(2)}$. Thus we assume the validity of estimators $\widehat{\theta}^{(ML)}$ and $\widehat{\theta}^{(N)}$, as well as the here presented $\widehat{\theta}^{(FF)}$, $\widehat{\theta}^{(CFG)}$ and $\widehat{\theta}^{(SS)}$.

In Figure 3 are plotted, for several thresholds, the obtained estimates from $\widehat{\theta}^{(B)}$ (for $b = 20, 31$), $\widehat{\theta}^{(R)}$ (for $r = 2, 4$) and $\widehat{\theta}^{(I)}$ (left), and from $\widehat{\theta}^{(ML)}$ and $\widehat{\theta}^{(N)}$ (right). Considering again $u = -10$, we have $\widehat{\theta}^{(R)} = 0.35, 0.23$, for $r = 2, 4$, respectively, $\widehat{\theta}^{(I)} = 0.26$, $\widehat{\theta}^{(ML)} = 0.43$ and $\widehat{\theta}^{(N)} = 0.4$. By applying our estimators, we have $\widehat{\theta}^{(FF)} = 0.36$, $\widehat{\theta}^{(CFG)} = 0.38$ and $\widehat{\theta}^{(SS)} = 0.38$, more closer to the ones obtained for $\widehat{\theta}^{(ML)}$, $\widehat{\theta}^{(N)}$, $\widehat{\theta}^{(B)}$ with $b = 31$ and $\widehat{\theta}^{(R)}$ with $r = 2$.

3.1.2. Ozone pollution

We now consider $n = 120$ weekly maxima of hourly averages of ozone concentrations measured in parts per million, in the San Francisco bay area, San Jose, available in the package Xtremes (Reiss and Thomas, [18] 2007). These data have

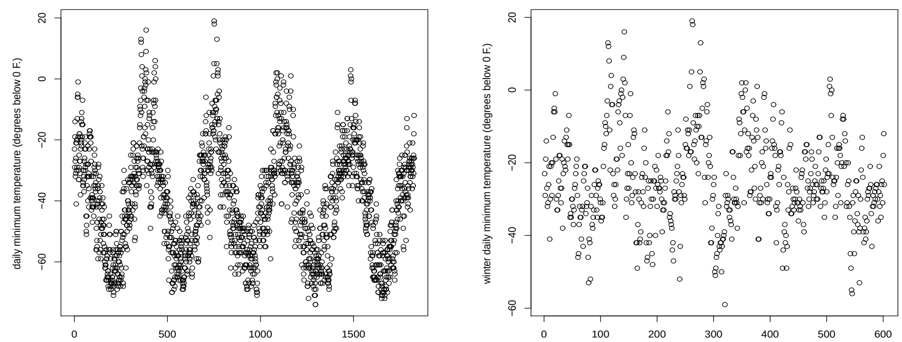

Figure 1. Negated Wooster daily minimum temperatures (in degrees Fahrenheit) on the left, and considering winters only on the right.

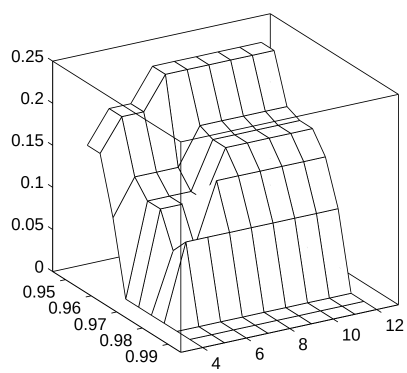

Figure 2. The observed proportion of anti $\mathrm{D}^{(2)}(u_n)$ condition for winters negated Wooster daily minimum temperatures (in degrees Fahrenheit).

been analyzed in Gomes *et al.* ([12], 2008) and Sebastião *et al.* ([20], 2013). We assume stationarity as in the latter reference (see also Figure 4). Gomes *et al.* ([12], 2008) argued the plausibility of condition $\mathrm{D}^{(2)}$ to hold, based on the fact that these type of meteorological data is usually modeled by processes that satisfy this latter. See also Figure 5 and the conclusions in Sebastião *et al.* ([20], 2013) which corroborates this assumption. A sample of maxima makes us comfortable with the hypothesis of an underlying model approximately BEV for consecutive pairs of observations. The extremal index was evaluated in 0.7 in Gomes *et al.* ([12], 2008). In what concerns estimators $\widehat{\theta}^{(FF)}$, $\widehat{\theta}^{(CFG)}$ and $\widehat{\theta}^{(SS)}$, we have obtained,

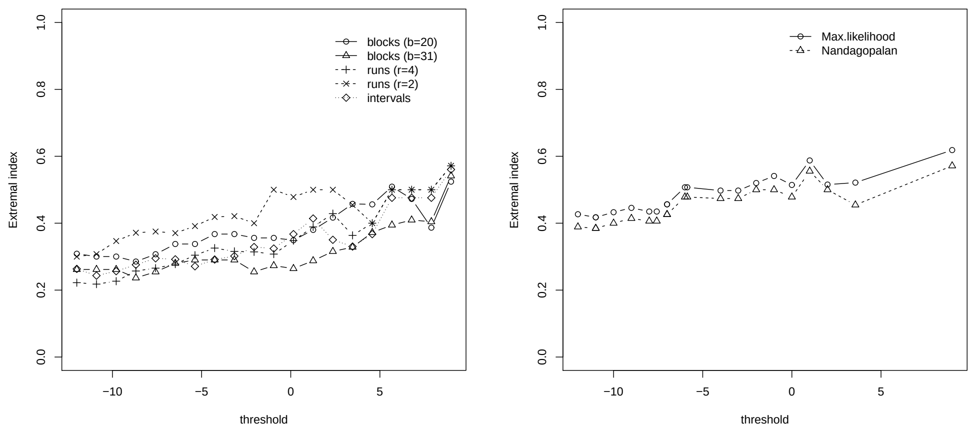

Figure 3. The blocks, runs and intervals estimators (left) and the maximum likelihood and Nandagopalan estimators (right), against threshold, for winters negated Wooster daily minimum temperatures (in degrees Fahrenheit).

respectively, 0.74, 0.74 and 0.75. In analyzing Figure 6, the value 0.7 is a possible estimate, except in the case of the blocks estimator.

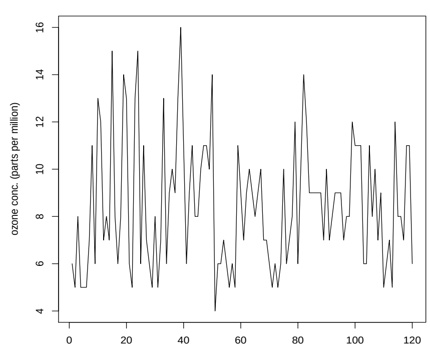

Figure 4. Weekly maxima of hourly averages of ozone concentrations (in parts per million), in the San Francisco bay area, San Jose.

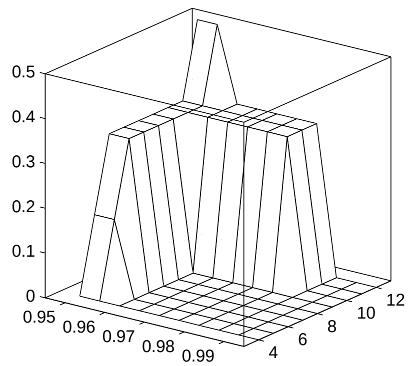

Figure 5. The observed proportion of anti $\mathrm{D}^{(2)}(u_n)$ condition for weekly maxima of hourly averages of ozone concentrations (in parts per million), in the San Francisco bay area, San Jose.

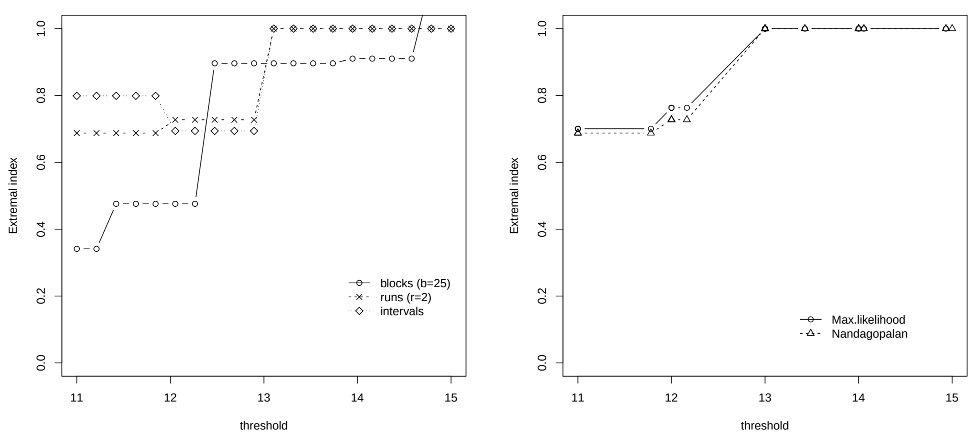

Figure 6. The blocks, runs and intervals estimators against threshold for weekly maxima of hourly averages of ozone concentrations (in parts per million), in the San Francisco bay area, San Jose.

4. Concluding remarks

Here we have considered new estimators for the extremal index based on the tail dependence coefficient estimation, under the validity of condition $\mathrm{D}^{(2)}(u_n)$ of Chernick *et al.* ([4], 1991). Estimators $\widehat{\theta}^{(FF)}$ and $\widehat{\theta}^{(CFG)}$ also require that the underlying distribution of consecutive random pairs can be approximated by a BEV model dependence structure. These latter are not robust whenever one of the two assumptions is breached. On the other hand, estimator $\widehat{\theta}^{(SS)}$ presents comparable biases and rmse's to estimators $\widehat{\theta}^{(ML)}$ and $\widehat{\theta}^{(N)}$ which were also derived under condition $\mathrm{D}^{(2)}(u_n)$, in some cases, even outperforming these two latter. Estimator $\widehat{\theta}^{(SS)}$ has also comparable performances to the ones of the runs and the blocks estimators in some models. Observe that it depends on only one parameter (the number k of observations to consider in the estimation), while the runs and blocks estimators depend on a high threshold u and the number of runs r or blocks b, respectively. Since $\mathrm{D}^{(2)}(u_n)$ is a crucial requisite in the new approach, it is important to develop a more reliable diagnostic statistical tool for this condition. This will be the aim of a future work.

Acknowledgements

The author was financed by FEDER Funds through "Programa Operacional Factores de Competitividade - COMPETE" and by Portuguese Funds through FCT - “Fundação para a Ciência e a Tecnologia", within the project PEst-OE/MAT/UI0013/2014.

REFERENCES

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

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

[3] P. Capéraà, A.L. Fougères and C. Genest, *A nonparametric estimation procedure for bivariate extreme value copulas*, Biometrika **84** (1997) 567–577.

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

[5] S.G. Coles, An Introduction to Statistical Modelling of Extreme Values (London, Springer, 2001).

[6] M. Ferreira, *Nonparametric estimation of the tail dependence coefficient*, REVSTAT **11(1)** (2013) 1–16.

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

[8] H. Ferreira and M. Ferreira, *On extremal dependence of block vectors*, Kybernetika **48(5)** (2012b) 988–1006.

[9] C.A. Ferro and J. Segers, *Inference for clusters of extremes*, J. R. Stat. Soc. Ser. B Stat. Methodol. **65** (2003) 545–556.

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

[11] C. Genest and J. Segers J., *Rank-based inference for bivariate extreme-value copulas*, Ann. Statist. **37** (2009) 2990–3022.

[12] M.I. Gomes, A. Hall and C. Miranda, *Subsampling techniques and the jackknife methodology in the estimation of the extremal index*, J. Stat. Comput. Simul. **52** (2008) 2022–2041.

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

[14] M.R. Leadbetter, *Extremes and local dependence in stationary sequences*, Z. Wahrsch. Ver. Geb. **65** (1983) 291–306.

[15] R.M. Loynes, *Extreme Values in Uniformly Mixing Stationary Stochastic Processes*, Annals of Mathematical Statistics **36** (1965) 993–999.

[16] S. Nandagopalan, Multivariate extremes and estimation of the extremal index (Ph.D. Thesis, University of North Carolina at Chapel Hill, 1990).

[17] G.L. O'Brien, *The maximum term of uniformly mixing stationary sequences*, Z. Wahrsch. Ver. Geb. **30** (1974) 57–63.

[18] R.D. Reiss, M. Thomas, Statistical analysis of extreme values with applications to insurance, finance, hydrology and other fields (Birkhäuser, Basel, 2007).

[19] R. Schmidt and U. Stadtmüller, *Nonparametric estimation of tail dependence*, Scandinavian J. Statist. **33** (2006) 307–335.

[20] J.R. Sebastião, A.P. Martins, H. Ferreira and L. Pereira, *Estimating the upcrossings index*, TEST **22(4)** (2013) 549–579.

[21] M. Sibuya, *Bivariate extreme statistics*, Ann. Inst. Stat. Math. **11** (1960) 195–210.

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