Adaptive Choice of Thresholds and the Bootstrap Methodology: an Empirical Study*

M. Ivette Gomes, Fernanda Figueiredo and M. Manuela Neves

Abstract

In this paper, we discuss an algorithm for the adaptive estimation of a positive *extreme value index*, γ, the primary parameter in *Statistics of Extremes.* Apart from classical extreme value index estimators, we suggest the consideration of associated second-order corrected-bias estimators, and propose the use of bootstrap computer-intensive methods for the adaptive choice of *thresholds.*

Keywords. Heavy tails; statistics of extremes; extreme value index; adaptive semi-parametric estimation; bias reduction; location/scale invariant estimation.

1 Introduction and outline of the paper

Heavy-tailed models appear often in practice in fields like Telecommunications, Insurance, Finance, Bibliometrics and Biostatistics. We shall deal with the estimation of a positive *extreme value index* (EVI), γ, the primary parameter in *Statistics of Extremes.* Apart from the *classical Hill*, *moment* and *generalized-Hill* semi-parametric estimators of γ, detailed in Section 2, we shall consider the associated classes of *second-order reduced-bias* estimators, based on an adequate estimation of generalized scale and shape second-order parameters, valid for a large class of heavy-tailed underlying parents, and appealing in the sense that we are able to reduce the asymptotic bias of a classical estimator without increasing its asymptotic variance. We shall call these estimators "*classical-variance reduced-bias*" (CVRB) estimators.

*Research partially supported by FCT / POCI 2010 and PTDC / FEDER.

After the introduction, in Section 2, of a few technical details in the area of *Extreme Value Theory* (EVT), related with the EVI-estimators under consideration in this paper, we shall briefly discuss, in Section 3, the kind of second-order parameters' estimation which enables the building of *reduced-bias* estimators with the same asymptotic variance of the associated *classical* estimator. After the discussion, in Section 4, of the asymptotic behaviour of the estimators under consideration, we propose and discuss in Section 5, and in the lines of [13], an algorithm for the adaptive estimation of a positive EVI, through the use of bootstrap computer-intensive methods. The algorithm is described for a classical EVI estimator and associated CVRB estimator, but it works similarly for the estimation of any other parameter of extreme events, like a high quantile, the probability of exceedance or the return period of a high level. Section 6 is entirely dedicated to the application of the algorithm to the analysis of environmental data, the number of hectares burned during all wildfires recorded in Portugal in the period 1999-2003.

2 The EVI-estimators under consideration

In the area of EVT, and for large values, a model F is said to be *heavy-tailed* whenever the *right-tail function*, $\overline{F} := 1 - F$, is a regularly varying function with a negative index of regular variation, denoted $-1/\gamma$, i.e., if for all $x > 0$, there exists $\gamma > 0$, such that $\overline{F}(tx)/\overline{F}(t) \rightarrow x^{-1/\gamma}$, as $t \rightarrow \infty$. If this holds, we use the notation $\overline{F} \in RV_{-1/\gamma}$, and we are working in the whole *domain of attraction* (for maxima) of heavy-tailed models, denoted $\mathcal{D}_{\mathcal{M}}\left(EV_\gamma\right)_{\gamma>0}$. Equivalently, with $U(t) := F^{\leftarrow}\big(1-1/t\big)$=inf $\{x : F(x) \geq 1-1/t\}$, $F \in \mathcal{D}_{\mathcal{M}}\left(EV_\gamma\right)_{\gamma>0} \iff \overline{F} \in RV_{-1/\gamma} \iff U \in RV_\gamma$, the so-called *first-order* condition. For these heavy-tailed parents, given a sample $\underline{\mathbf{X}}_n := (X_1, \ldots, X_n)$ and the associated sample of ascending order statistics (o.s.'s), $(X_{1:n} \leq \cdots \leq X_{n:n})$, the classical EVI estimator is the Hill estimator ([15]),

$$H_k \equiv H_{k,n} := \frac{1}{k} \sum_{i=1}^{k} \{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}, \tag{1}$$

the average of the k log-excesses over a high random threshold $X_{n-k:n}$, an *intermediate* o.s., i.e., with k such that

$$k = k_n \rightarrow \infty \quad \text{and} \quad k/n \rightarrow 0, \text{ as } n \rightarrow \infty. \tag{2}$$

But the Hill-estimator H_k, in (1), reveals usually a high non-null asymptotic bias at optimal levels, i.e., levels k where the mean squaed error (MSE) is minimum. This non-null asymptotic bias, together with a rate of convergence of the order of $1/\sqrt{k}$, leads to sample paths with a high variance for small k, a high bias for large k, and a very sharp MSE pattern, as function of k. Recently, several authors have been dealing with bias reduction in the field of *extremes* (for an overview, see [17], Chapter 6, as well as the more recent paper, [4]). For technical details, we then need to work in a region slightly more restrict than $\mathcal{D}_{\mathcal{M}}\left(EV_\gamma\right)_{\gamma>0}$. In this paper, we shall consider parents such that, with $\gamma > 0$, $\rho < 0$, and $\beta \neq 0$,

$$U(t) = Ct^\gamma\big(1 + A(t)/\rho + O(A^2(t))\big), \text{ as } t \to \infty, \quad A(t) =: \gamma\beta t^\rho. \tag{3}$$

The most simple class of second-order minimum-variance reduced-bias (MVRB) EVI-estimators is the one in [3], used for a semi-parametric estimation of $\ln VaR_p$ in [9]. This class, here denoted $\overline{H} \equiv \overline{H}_k$, is the CVRB-estimator associated with the Hill estimator $H = H_k$, in (1), and depends upon the estimation of the second-order parameters (β, ρ), in (3). Its functional form is

$$\overline{H}_k \equiv \overline{H}_{k,n,\hat\beta,\hat\rho} := H_k\big(1 - \hat\beta(n/k)^{\hat\rho}/(1-\hat\rho)\big), \tag{4}$$

where $(\hat\beta, \hat\rho)$ is an adequate consistent estimator of (β, ρ). Algorithms for the estimation of (β, ρ) are provided, for instance, in [9] , and will be reformulated in Section 3 of this paper.

Apart from the *Hill* estimator, in (1), we suggest the consideration of two other classical estimators, valid for all $\gamma \in \mathbb{R}$, but taken here exclusively for heavy tails, the *moment* ([5]) and the *generalized-Hill* ([1], [2]) estimators. The *moment* estimator (M) has the functional expression

$$M_k \equiv M_{k,n} := M_{k,n}^{(1)} + \tfrac{1}{2}\big\{1 - \big(M_{k,n}^{(2)}/(M_{k,n}^{(1)})^2 - 1\big)^{-1}\big\}, \tag{5}$$

with $M_{k,n}^{(j)} := \frac{1}{k}\sum_{i=1}^k \big(\ln X_{n-i+1:n} - \ln X_{n-k:n}\big)^j$, $j \geq 1$ $\big(M_{k,n}^{(1)} \equiv H_k$, in (1)$\big)$. The *generalized Hill* estimator (GH), based on $H_{k,n}$, in (1), is given by

$$GH_k \equiv GH_{k,n} := H_{k,n} + \frac{1}{k}\sum_{i=1}^k \big\{\ln H_{i,n} - \ln H_{k,n}\big\}. \tag{6}$$

The associated CVRB estimators have similar expressions, due to same dominant component of asymptotic bias of the estimators in (5) and (6), for a positive EVI. Denoting $\overline{W}$ either $\overline{M}$ or $\overline{GH}$, and with the notation W for either M or GH, we get

$$\overline{W}_k \equiv \overline{W}_{k,n,\hat\beta,\hat\rho} := W_k\big(1 - \hat\beta\,(n/k)^{\hat\rho}\,/(1-\hat\rho)\big) - \hat\beta\;\hat\rho\,(n/k)^{\hat\rho}\,/(1-\hat\rho)^2. \tag{7}$$

In the sequel, we generally denote C any of the classical EVI-estimators, in (1), (5) and (6), and $\overline{C}$ the associated CVRB-estimator.

3 Estimation of second-order parameters

The estimation of γ, β and ρ at the same value k leads to a high increase in the asymptotic variance of CVRB-estimators $\overline{C}_{k,\hat{\beta},\hat{\rho}}$, which becomes $\sigma_C^2\,((1-\rho)/\rho)^4$ (see [16], among others). The external estimation of ρ at k_1, but the estimation of γ and β at $k = o(k_1)$, slightly decreases the asymptotic variance to $\sigma_C^2\,((1-\rho)/\rho)^2$, stilll greater than σ_C^2 (see [7], among others). The external estimation of both β and ρ at a level k_1, and the the estimation of γ at a level $k = o(k_1)$, or even $k = O(k_1)$, can lead to a CVRB estimator with an asymptotic variance σ_C^2, provided we choose adequately k_1 (see [3], [10], [11]). Such a choice is theoretically possible (see [12] and [4]), but under conditions difficult to guarantee in practice. As a compromise between theoretical and practical results, we have so far advised any choice $k_1 = n^{1-\epsilon}$, with ϵ small. We shall consider here $\epsilon = 0.001$.

Algorithm (second-order parameters' estimation):

1. Given an observed sample $(x_1,\ldots,x_n)$, plot the observed values of $\hat{\rho}_\tau(k)$, the most simple estimator in [6], for the tuning parameters $\tau = 0$ and $\tau = 1$.

2. Consider $\{\hat{\rho}_\tau(k)\}_{k\in\mathcal{K}}$, with $\mathcal{K} = ([n^{0.995}],\ [n^{0.999}])$, compute their median, denoted η_τ, and compute $I_\tau := \sum_{k\in\mathcal{K}} \left(\hat{\rho}_\tau(k) - \eta_\tau\right)^2$, $\tau = 0, 1$. Next choose the *tuning parameter* $\tau^* = 0$ if $I_0 \le I_1$; otherwise, choose $\tau^* = 1$.

3. Work with $\hat{\rho} \equiv \hat{\rho}_{\tau^*} = \hat{\rho}_{\tau^*}(k_1)$ and $\hat{\beta} \equiv \hat{\beta}_{\tau^*} := \hat{\beta}_{\hat{\rho}_{\tau^*}}(k_1)$, with $k_1 = n^{0.999}$, being $\hat{\beta}_{\hat{\rho}}(k)$, the estimator in [7].

Remark 1. *This algorithm leads usually to the tuning parameter $\tau = 0$ whenever $|\rho| \le 1$ and $\tau = 1$, otherwise. For details on this and similar algorithms, see [9].*

4 Asymptotic distributional behaviour of the estimators

In order to obtain a non-degenerate behaviour for any semi-parametric EVI-estimator, it is convenient to assume a *second-order* condition, measuring the rate of convergence in the first-order condition. Such a condition, valid for all $x > 0$, involves a parameter $\rho \leq 0$, a rate function A, with $|A| \in RV_\rho$ and is given by

$$\lim_{t\to\infty} (U(tx)/U(t) - x^\gamma)/A(t) = x^\gamma\ (x^\rho - 1)/\rho. \tag{8}$$

In this paper, and mainly because of the reduced-bias estimators in (4) and (7), generally denoted $\overline{C}_k \equiv \overline{C}_{k,n,\hat\beta,\hat\rho}$, we shall assume that (3) holds. Then, (8) holds, with $A(t) = \gamma\ \beta t^\rho$. Let C_k be the associated classical estimator of γ. Trivial adaptations of the results in the above-mentioned papers (see also [14]) enable us to state, without proof, the following theorem, again for models with $\gamma > 0$.

Theorem 1. *Assume that condition* (8) *holds, and let* $k = k_n$ *be an intermediate sequence, i.e.* (2) *holds. Then, there exist a sequence* Z_k^C *of asymptotically standard normal random variables,* $\sigma_C > 0$ *and real numbers* $b_{C,1}$ *such that* $C_k \stackrel{d}{=} \gamma + \sigma_C Z_k^C/\sqrt{k} + b_{C,1}\ A(n/k)\ (1+o_p(1))$. *If we further assume that* (3) *holds, and estimate* β *and* ρ *consistently through* $\hat\beta$ *and* $\hat\rho$*, in such a way that* $\hat\rho - \rho = o_p(1/\ln n)$*, we can guarantee that there exists a pair of real numbers* $(\bar b_{C,1}, \bar b_{C,2})$*, with* $b_{\overline{C},1} = 0$*, such that* $\overline{C}_{k,n,\hat\beta,\hat\rho} \stackrel{d}{=} \gamma + \sigma_C Z_k^C/\sqrt{k} + b_{\overline{C},1}\ A(n/k) + b_{\overline{C},2}\ A^2(n/k)\ (1 + o_p(1))$.

As $n \to \infty$, let $k = k_n$ be intermediate such that $\sqrt{k}\ A(n/k) \to \lambda$, finite, the levels k where the MSE of C_k is minimum. Let $\hat\gamma_k$ denote either C_k or $\overline{C}_k$. Then, we have $\sqrt{k}\big(\hat\gamma_k - \gamma\big) \stackrel{d}{\to} Normal(\lambda\ b_{\hat\gamma,1},\ \sigma_C^2)$, even if we work with CVRB EVI-estimators. If $\sqrt{k}\ A(n/k) \to \infty$, with $\sqrt{k}A^2(n/k) \to \lambda_A$, finite, the levels k where the MSE of $\overline{C}_k$ is minimum, $\sqrt{k}\ \ \big(\overline{C}_k - \gamma\big) \stackrel{d}{\to}$ Normal$\big(\lambda_A b_{\overline{C},2},\ \sigma_C^2\big)$. We have $\sigma_H^2 = \gamma^2$, $\sigma_M^2 = \sigma_{GH}^2 = 1 + \gamma^2$, $b_{H,1} = 1/(1-\rho)$, $b_{M,1} = b_{GH,1} = (\gamma - \gamma\rho + \rho)/(\gamma(1-\rho)^2)$, $b_{\overline{H},1} = b_{\overline{M},1} = b_{\overline{GH},1} = 0$. Consequently, since $b_{C,1} \neq 0$ whereas $b_{\overline{C},1} = 0$, the $\overline{C}$-estimators outperform the C-estimators for all k, as can be seen in Figure 1.

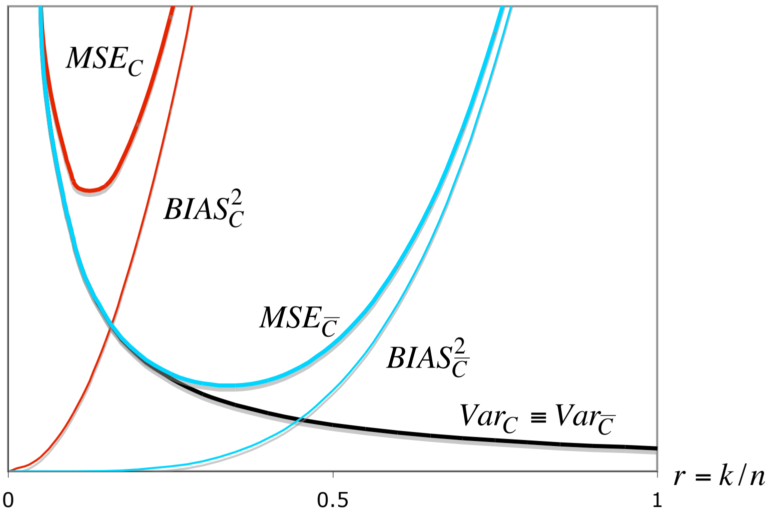

Figure 1: Comparative asymptotic variances (Var), squared bias ($BIAS^2$) and mean squared errors (MSE) of a classical EVI-estimator and associated CVRB estimator.

5 The bootstrap methodology and adaptive EVI-estimation

With AMSE standing for "*asymptotic MSE*", and $k_0^{\widehat{\gamma}}(n) := \arg\min_k MSE(\widehat{\gamma}_k)$,

$$\begin{aligned} k_{0|\widehat{\gamma}}(n) &:= \arg\min_k AMSE\big(\widehat{\gamma}_k\big) \qquad (9)\\ &= \arg\min_k \left\{ \begin{array}{ll} \big(\sigma_C^2/k + b_{C,1}^2\ A^2(n/k)\big) & (\text{if } \widehat{\gamma} = C) \\ \big(\sigma_C^2/k + b_{\overline{C},2}^2\ A^4(n/k)\big) & (\text{if } \widehat{\gamma} = \overline{C}) \end{array} \right. = k_0^{\widehat{\gamma}}(n)(1+o(1)). \end{aligned}$$

The bootstrap methodology can thus enable us to consistently estimate the optimal sample fraction (OSF), $k_0^{\widehat{\gamma}}(n)/n$, on the basis of a consistent estimator of $k_{0|\widehat{\gamma}}(n)$, in (9), in a way similar to the one used for classical EVI-estimation (see, for instance, [8]). We shall here use the most obvious auxiliary statistics, the statistics $T_{k|\hat{\gamma}} \equiv T_{k,n|\hat{\gamma}} := \widehat{\gamma}_{[k/2]} - \widehat{\gamma}_k, \quad k = 2,\ldots,n-1$, which converge in probability to zero, for intermediate k, and have an asymptotic behaviour strongly related with the asymptotic behaviour of $\widehat{\gamma}_k$. Indeed, under the above-mentioned third-order framework in (3), we easily get:

$$T_{k|\hat{\gamma}} \stackrel{d}{=} \frac{\sigma_{\hat{\gamma}}\ P_k^{\widehat{\gamma}}}{\sqrt{k}} + \left\{ \begin{array}{ll} b_{\hat{\gamma},1}(2^{\rho}-1)\ A(n/k)(1+o_p(1)) & (\text{if } \widehat{\gamma} = C) \\ b_{\hat{\gamma},2}(2^{2\rho}-1)\ A^2(n/k)(1+o_p(1)) & (\text{if } \widehat{\gamma} = \overline{C}) \end{array} \right.$$

with $P_k^{\hat\gamma}$ asymptotically standard normal. Consequently, denoting $k_{0|T}(n) := \arg\min_k AMSE(T_k)$, we have

$$k_{0|\widehat{\gamma}}(n) = k_{0|T}(n) \begin{cases} (1-2^{\rho})^{\frac{2}{1-2\rho}}\,(1+o(1)) & (\text{if } \widehat{\gamma}=C) \\ (1-2^{2\rho})^{\frac{2}{1-4\rho}}\,(1+o(1)) & (\text{if } \widehat{\gamma}=\overline{C}). \end{cases}$$

5.1 How does the bootstrap methodology then work?

Given the sample $\underline{\mathbf{X}}_n = (X_1,\ldots,X_n)$ from an unknown model F, and the functional $T_{k,n} =: \phi_k(\underline{\mathbf{X}}_n)$, $1<k<n$, consider for any $n_1 = O(n^{1-\epsilon})$, $0<\epsilon<1$, the bootstrap sample $\underline{\mathbf{X}}^*_{n_1} = (X_1^*,\ldots,X_{n_1}^*)$, from the model $F_n^*(x) = \frac{1}{n}\sum_{i=1}^n I_{[X_i\le x]}$, the empirical d.f. associated to the available sample, $\underline{\mathbf{X}}_n$. Next, associate to the bootstrap sample the corresponding bootstrap auxiliary statistic, $T^*_{k_1,n_1} := \phi_{k_1}(\underline{\mathbf{X}}^*_{n_1})$, $1<k_1<n_1$. Then, with $k^*_{0|T}(n_1) = \arg\min_{k_1} AMSE\big(T^*_{k_1,n_1}\big)$,

$$\frac{k^*_{0|T}(n_1)}{k_{0|T}(n)} = \Big(\frac{n_1}{n}\Big)^{-\frac{c\ \rho}{1-c\ \rho}}(1+o(1)), \qquad c = \begin{cases} 2 & (\text{if } \widehat{\gamma}=C) \\ 4 & (\text{if } \widehat{\gamma}=\overline{C}). \end{cases} \tag{10}$$

Consequently, for another sample size n_2, and for every $\alpha>1$,

$$\frac{\big(k^*_{0|T}(n_1)\big)^{\alpha}}{k^*_{0|T}(n_2)}\ \left(\frac{n_1^{\alpha}}{n^{\alpha}}\frac{n}{n_2}\right)^{-\frac{c\ \rho}{1-c\ \rho}} = \big\{k_{0|T}(n)\big\}^{\alpha-1}(1+o(1)).$$

It is then enough to choose $n_2 = n\,(n_1/n)^{\alpha}$, to have independence of ρ. If we put $n_2 = n_1^2/n$, i.e., $\alpha=2$, we have $\big(k^*_{0|T}(n_1)\big)^2/k^*_{0|T}(n_2) = k_{0|T}(n)(1+o(1))$, as $n\to\infty$. We are now able to estimate $k_0^{\widehat{\gamma}}(n)$, on the basis of any estimate $\hat\rho$ of ρ. With $\hat k^*_{0T}$ denoting the sample counterpart of k^*_{0T}, and $\hat\rho$ the ρ-estimate in Step **3.** of the algorithm, initiated in Section 3, we have the k_0-estimate

$$\hat k_0^{\widehat{\gamma}}(n;n_1) := \min\Big(n-1,\ \big[C_{\hat\rho}\ (\hat k^*_{0|T}(n_1))^2/\hat k^*_{0|T}([n_1^2/n]+1)\big]+1\Big), \tag{11}$$

with $C_{\hat\rho} = \big(1-2^{c\hat\rho/2}\big)^{\frac{2}{1-c\hat\rho}}$, c given in (10).

Again, with $\widehat{\gamma}$ denoting either C or $\overline{C}$, we proceed with the algorithm.

Algorithm (cont.) (bootstrap adaptive estimation of γ):

4. Compute $\widehat{\gamma}_k$, $k=1,2,\ldots,n-1$;

5. Next, consider the sub-sample size $n_1 = n^{0.955}$ and $n_2 = [n_1^2/n] + 1$;

6. For l from 1 till $B = 250$, generate independently, from the empirical d.f. $F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} I_{[X_i \le x]}$, associated with the observed sample, B bootstrap samples $(x_1^*, \ldots, x_{n_2}^*)$ and $(x_1^*, \ldots, x_{n_2}^*, x_{n_2+1}^*, \ldots, x_{n_1}^*)$, of sizes n_2 and n_1, respectively;

7. Denoting $T_{k,n}^*$ the bootstrap counterpart of $T_{k,n}$, obtain $(t_{k,n_1,l}^*,\ t_{k,n_2,l}^*)$, $1 \le l \le B$, the observed values of the statistic T_{k,n_i}^*, $i = 1, 2$, compute $MSE^*(n_i, k) = \frac{1}{B}\sum_{l=1}^{B}\left(t_{k,n_i,l}^*\right)^2$, $k = 1, 2, \ldots, n_i - 1$, and obtain $\hat{k}_{0T}^*(n_i) := \arg\min_{1 \le k \le n_i - 1} MSE^*(n_i, k)$, $i = 1, 2$;

8. Compute $\hat{k}_0^{\widehat{\gamma}}(n; n_1)$ in (11);

9. Compute $\widehat{\gamma}_{n,n_1|T}^* := \widehat{\gamma}_{\hat{k}_0^{\widehat{\gamma}}(n;n_1)}$.

6 An application to burned areas data

Most of the wildfires are extinguished within a short period of time, with almost negligible effects. However, some wildfires go out of control, burning hectares of land and causing significant and negative environmental and economical impacts. The data we analyse here consists of the number of hectares, exceeding 100 ha, burnt during wildfires recorded in Portugal during 14 years (1990-2003). The data (a sample of size $n = 2627$) do not seem to have a significant temporal structure, and we have used it as a whole, although we think also sensible, to try avoiding spatial heterogeneity, considering at least three different regions: the north, the centre and the south of Portugal (a study out of the scope of this note).

The box-plot and a histogram of the available data provide evidence on the heaviness of the right-tail. We shall next consider, for this type of data, the performance of the adaptive CVRB-EVI estimates $\overline{H}$, in (4), which is *minimum variance reduced-bias* (MVRB). These MVRB estimators exhibit stabler sample paths than Hill estimators, as functions of k, and enable us to take a decision upon the estimates to be used, even with the help of any heuristic stability criterion. The algorithm in this paper enables us to adaptively estimate the OSF associated with the MVRB or CVRB estimates. For a sub-sample size $n_1 = [n^{0.955}] = 1843$, and B=250 bootstrap generations, we have got $\hat{k}_{0*}^{\overline{H}} = 1319$ and the bootstrap MVRB-EVI-estimate $\overline{H}^* = 0.658$, the value pictured in the following figure, jointly with the above-mentioned adaptive bootstrap Hill estimate, $H^* = 0.73$. Note the fact that the MVRB EVI-estimators look practically "unbiased" for the data under analysis.

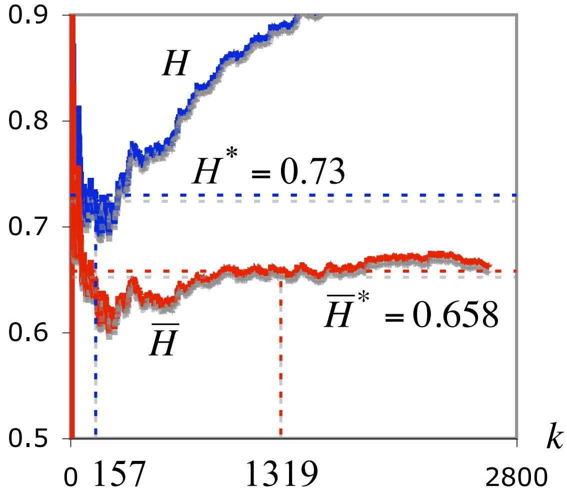

Figure 2: Estimates of the EVI, γ, through the Hill estimator, H, in (1), and the MVRB estimator, $\overline{H}$, in (4), for the burned areas under analysis, together with the bootstrap adaptive estimates H^* and $\overline{H}^*$.

A few practical questions may be raised under the set-up developed: How does the asymptotic method work for moderate sample sizes? Is the method strongly dependent on the choice of n_1? What is the sensitivity of the method with respect to the choice of ρ-estimate? Although aware of the need of $n_1 = o(n)$, what happens if we choose $n_1 = n$? Answers to these questions are expected not to be a long way from the ones given for classical estimation (see [8]), have lightly been addressed in [13], for reduced-bias estimation, but are out of the scope of this paper.

References

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

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

[3] Caeiro, F., Gomes, M.I., Pestana, D.D.: Direct reduction of bias of the classical Hill estimator. Revstat **3**:2, 111-136 (2005)

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

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

[6] Fraga Alves, M.I., Gomes M.I., Haan, L. de: A new class of semi-parametric estimators of the second order parameter. Portugaliae Mathematica **60**:2, 194-213 (2003)

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

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

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

[10] Gomes, M.I., Martins, M.J., Neves, M.M.: Improving second order reduced-bias tail index estimation. Revstat **5**:2, 177-207 (2007)

[11] Gomes, M.I., Haan, L. de, Henriques-Rodrigues, L.: Tail Index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. J. Royal Statistical Society **B70**, Issue 1, 31-52 (2008)

[12] Gomes, M.I., Pestana, D.D., Caeiro, F.: A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator. Statist. and Probab. Letters **79**:3, 295-303 (2009)

[13] Gomes, M.I., Mendonça, S., Pestana, D.D.: Adaptive reduced-bias tail index and value-at-risk estimation. In: Sakalauskas, L., Skiadas, C., Zavadskas, E.K. (eds.) Applied Stochastic Models and Data Analysis – ASMDA 2009, IMI & VGTU Editions, 41-44 (2009)

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

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

[16] Peng, L., Qi, Y.: Estimating the first and second order parameters of a heavy tailed distribution. Australian and New Zealand J. of Statist. **46**, 305-312 (2004)

[17] Reiss, R.-D. and Thomas, M., Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields, 3rd edition, Birkhäuser Verlag, Basel-Boston-Berlin (2007)