WILEY

Reduced-bias kernel estimators of a positive extreme value index

Frederico Caeiro[1,2] 🆔 | **Lígia Henriques-Rodrigues**[3,4] 🆔

[1]Faculdade de Ciências e Tecnologia (FCT), Universidade Nova de Lisboa, Caparica, Portugal

[2]Centro de Matemática e Aplicações (CMA), Universidade Nova de Lisboa, Caparica, Portugal

[3]IME, University of São Paulo, Rua do Matão 1010, São Paulo, Brazil

[4]Centro de Estatística e Aplicações, Universidade de Lisboa (CEAUL), Lisbon, Portugal

Correspondence
Frederico Caeiro, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.
Email: fac@fct.unl.pt

Communicated by: J Vigo Aguiar

In this paper, we deal with the semi-parametric estimation of the extreme value index, an important parameter in extreme value analysis. It is well known that many classic estimators, such as the Hill estimator, reveal a strong bias. This problem motivated the study of two classes of kernel estimators. Those classes generalize the classical Hill estimator and have a tuning parameter that enables us to modify the asymptotic mean squared error and eventually to improve their efficiency. Since the improvement in efficiency is not very expressive, we also study new reduced bias estimators based on the two classes of kernel statistics. Under suitable conditions, we prove their asymptotic normality. Moreover, an asymptotic comparison, at optimal levels, shows that the new classes of reduced bias estimators are more efficient than other reduced bias estimator from the literature. An illustration of the finite sample behaviour of the kernel reduced-bias estimators is also provided through the analysis of a data set in the field of insurance.

KEYWORDS
asymptotic behaviour, bias estimation, heavy tails, optimal levels, semi-parametric estimation, statistics of extremes

MSC CLASSIFICATION
62G32; 62G20; 62G05; 62P05

1 | INTRODUCTION

Suppose we have a sample $X_1, X_2, \ldots, X_n$, of independent and identically distributed (i.i.d.) random variables (r.v.'s) with a common distribution function (d.f.) F with F a heavy tailed model. Then $1 - F$ is of regular variation with index of regular variation equal to $-1/\xi > 0$ (see Bingham et al[1] for details); ie, for every $x > 0$,

$$\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\xi},$$

and we write $1 - F \in RV_{-1/\xi}$. In the case $\xi > 0$, it is known that the quantile function $U(t) := F^{\leftarrow}(1 - 1/t)$, $t \geq 1$, with $F^{\leftarrow}(y) := \inf \{x : F(x) \geq y\}$ the generalized inverse of F, is of regular variation with index $\xi > 0$, ie,

$$\lim_{t \to \infty} \frac{U(tx)}{U(t)} = x^{\xi}. \tag{1}$$

Consequentially, F is in the max-domain of attraction of the extreme value d.f., given by

$$\mathrm{EV}_{\xi}(x) = \begin{cases} \exp(-(1 + \xi x)^{-1/\xi}), & 1 + \xi x > 0, \text{ if } \xi \neq 0, \\ \exp(-\exp(-x)), & x \in \mathbb{R}, \quad \text{ if } \xi = 0, \end{cases} \tag{2}$$

and the notation $F \in \mathcal{D}_M \left(\mathrm{EV}_\xi \right)$ is used. The parameter ξ, in (2), is usually called the *extreme value index* (EVI), and it is the primary parameter of extreme and rare events. This index measures the weight of the right tail. Let $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ be the associated ascending order statistics (o.s.) of the sample of size n. For $\xi > 0$, the most popular semi-parametric EVI-estimator is the Hill estimator,[2] the average of the k log-excesses, $V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}$, $1 \leq i \leq k < n$, above the high level $X_{n-k:n}$ or the average of the k-scaled log-spacings of consecutive order statistics $U_i := i(\ln X_{n-i+1:n} - \ln X_{n-i:n})$, $1 \leq i \leq k < n$,

$$\hat{\xi}_n^H(k) \equiv H(k) := \frac{1}{k} \sum_{i=1}^{k} U_i, \quad 1 \leq k < n. \tag{3}$$

Consistency is achieved for an intermediate threshold k, ie, for sequences of integers $k = k_n$, $1 \leq k < n$, such that

$$k \to \infty \quad \text{and} \quad k/n \to 0, \quad \text{as} \quad n \to \infty. \tag{4}$$

If we assume a second-order condition that measures the rate of convergence in (1), the Hill estimator in (3) is asymptotically normally distributed with asymptotic variance equal to ξ^2/k. Moreover, if F is not the strict Pareto model, with d.f. $F(x) = 1 - (x/c)^{-1/\xi}$, $x > c > 0$, then $H(k)$ is also asymptotically biased. The bias of the Hill estimator can be very small, moderate, or large and increases as k increases. Thus, we always have to deal with the bias-variance trade-off problem when we choose the threshold k. This problem has motivated several researchers to improve the Hill estimator by dealing with the correction of the dominant component of asymptotic bias or through the introduction of new EVI estimators (see Beirlant et al,[3,4] Reiss and Thomas,[5] and Gomes and Guillou[6] and references therein).

In this paper, we shall consider the following generalizations of the classical Hill estimator, defined as

$$\hat{\xi}_n^{\mathcal{P}(\omega)}(k) := \frac{\omega}{k} \sum_{i=1}^{k} \left(\frac{i}{k} \right)^{\omega-1} U_i, \quad 1 \leq k < n \tag{5}$$

and

$$\hat{\xi}_n^{\mathcal{L}(\omega)}(k) := \frac{1}{k\Gamma(\omega)} \sum_{i=1}^{k} \left(-\ln \left(\frac{i}{k} \right) \right)^{\omega-1} U_i, \quad 1 \leq k < n \tag{6}$$

where $\omega > 0$ is a tuning parameter. Different classes of generalized Hill estimators can be found in previous studies.[7-12] The classes of EVI-estimators in (5) and (6) belong to the class of kernel EVI-estimators studied in Csörgő et al[13] and Goegebeur et al,[14]

$$\hat{\xi}_n^{\mathcal{K}}(k) = \frac{1}{k} \sum_{i=1}^{n} \mathcal{K} \left(\frac{i}{k} \right) U_i, \quad 1 \leq k < n, \tag{7}$$

where $\mathcal{K}$ is a kernel function integrating to one. Indeed, the statistics in (5) and (6) are kernel EVI estimators, of the type (7), built with the power ($\mathcal{P}$) and log ($\mathcal{L}$) kernel functions,

$$\mathcal{P}(w) \equiv \mathcal{P}(\omega, u) := \omega u^{\omega-1}, \quad \text{and} \quad \mathcal{L}(w) \equiv \mathcal{L}(\omega, u) := \frac{1}{\Gamma(\omega)} (-\ln u)^{\omega-1}, \quad 0 < u < 1, \tag{8}$$

respectively, where ω is a tuning parameter and Γ denotes the complete Gamma function, defined by $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$, $t > 0$. Also, the class of estimators in (5) was studied in Gomes and Martins[15] and Gomes et al[16] for any real $\omega \geq 1$ and in Diop and Lo,[17] Lo et al,[18] and Caeiro et al[19] for any real parameter $\omega > 0$. Ratios of differences of the statistics in (5) with $\omega \geq 1$ were considered in Goegebeur et al[14] and Caeiro and Gomes[20] to estimate high-order tail parameters.

Remark 1. The class of kernel EVI-estimators in (5) and (6) generalize the Hill estimator in (3). Indeed, $\mathcal{P}(1, u)$ and $\mathcal{L}(1, u)$ are the uniform kernel function, with $\mathcal{P}(\omega, u)$ and $\mathcal{L}(\omega, u)$ given in (8).

This paper is organized as follows: After the introduction, in Section 2, of a few technical details in the field of EVT, we study the asymptotic behaviour of the classes of kernel EVI estimators in (5) and (6) and compare them asymptotically. Since the improvement in the efficiency of the aforementioned classes of kernel EVI-estimators is not expressive, in Section 3, we propose new classes of reduced bias EVI estimators based on the classes of kernel estimators. Those classes

are compared asymptotically with a bias-corrected Hill estimator from the literature. In Section 4, we illustrate the finite sample behaviour of the kernel reduced-bias estimators through the analysis of a data set in the field of insurance. Finally, some conclusions are presented in Section 5.

2 | RESULTS FOR THE CLASSES OF KERNEL EVI ESTIMATORS

2.1 | Second- and third-order conditions for a heavy right tail model

To study the asymptotic distributional behaviour of the EVI estimators under study, we need to assume a second-order condition for the quantile function U that measures the rate of convergence of $\ln U(tx) - \ln U(t)$ towards $\xi \ln x$, in the relation

$$\lim_{t \to \infty} \frac{\ln U(tx) - \ln U(t) - \xi \ln x}{A(t)} = \frac{x^\rho - 1}{\rho} \quad \forall x > 0, \tag{9}$$

where $\rho < 0$ is a second-order shape parameter, which measures the rate of convergence and $|A(t)|$ must then be a regular variation function with index ρ. To obtain additional information on the asymptotic bias of the EVI estimators, we shall further assume a third-order condition, ruling now the rate of convergence in (9), and which guarantees that

$$\lim_{t \to \infty} \frac{\frac{\ln U(tx) - \ln U(t) - \xi \ln x}{A(t)} - \frac{x^\rho - 1}{\rho}}{B(t)} = \frac{x^{\rho + \rho'} - 1}{\rho + \rho'}, \quad \forall x > 0, \tag{10}$$

where $|B|$ must then be of regular variation with index ρ'. For more details on the second- and third-order conditions, see Beirlant et al[3] and De Haan and Ferreira.[21]

For simplicity, we shall further assume a slightly more restrictive third-order condition. We assume that the third-order condition in (10) is satisfied with $\rho' = \rho < 0$,

$$A(t) = \xi \beta t^\rho, \qquad B(t) = \beta' t^{\rho'} = \frac{\delta A(t)}{\xi}, \quad \delta = \frac{\beta'}{\beta}, \tag{11}$$

with $\beta \in \mathbb{R} \neq 0$ and $\beta' \in \mathbb{R}$ scale second-order parameters. This is equivalent to assume that the tail quantile function of the underlying model can be written as

$$U(t) = ct^\xi \left(1 + \xi \beta t^\rho / \rho + O(t^{2\rho}) \right), \quad \text{as} \quad t \to \infty.$$

Among the models, used in applications, under slightly restrictive third-order condition, we mention the Fréchet, the Student t, the Burr, or the generalized Pareto. Algorithms for the estimation of the second-order parameters (ρ, β) can be found in previous studies,[22-24] among others.

2.2 | Asymptotic behaviour of the estimators under consideration

We will now derive the asymptotic behaviour of the above generalizations of the Hill estimator in (5) and (6). Theorem 1 in Goegebeur et al[14] enables us to state the following theorem and corollary. The proof of Theorem 1 follows closely the proof of theorem 1 in Goegebeur et al.[14]
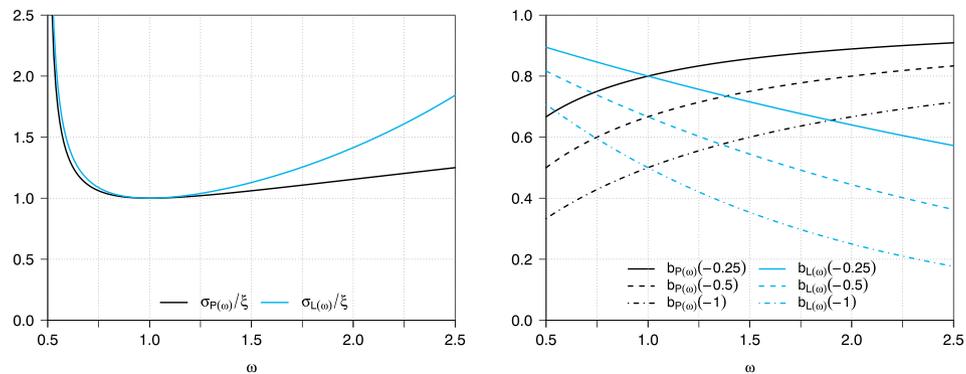
Theorem 1. *Suppose the second-order framework in (9) holds and let $k = k_n$ be an intermediate sequence. Then, with $\hat{\xi}_n^\bullet(k)$ generally denoting the classes of estimators in (5) and (6), we have the validity of the following asymptotic distributional representation for any real number $\omega > 0.5$*

$$\hat{\xi}_n^\bullet(k) \stackrel{d}{=} \xi + \frac{\sigma_\bullet}{\sqrt{k}} Z_k^\bullet + b_\bullet A(n/k)(1 + o_p(1)) \tag{12}$$

with

$$\sigma_{\mathcal{P}(\omega)}^2 = \frac{\xi^2 \omega^2}{2\omega - 1}, \quad \sigma_{\mathcal{L}(\omega)}^2 = \frac{\xi^2 \Gamma(2\omega - 1)}{\Gamma^2(\omega)}, \tag{13}$$

FIGURE 1 Graph of the normalized standard deviation, $\sigma_\bullet/\xi$, as a function of ω (left) and of the bias component, $b_\bullet$, for $\rho = -1, -0.5$ and -0.25, as a function of ω (right) [Colour figure can be viewed at wileyonlinelibrary.com]

$$b_{P(\omega)} = \frac{\omega}{\omega - \rho}, \quad b_{\mathcal{L}(\omega)} = \frac{1}{(1 - \rho)^\omega}, \tag{14}$$

and

$$Z_k^{P(\omega)} = \sqrt{k}\frac{\frac{1}{k}\sum_{i=1}^{k}\omega\left(\frac{i}{k}\right)^{\omega-1}(E_i - 1)}{\sigma_{P(\omega)}}, \quad Z_k^{\mathcal{L}(\omega)} = \sqrt{k}\frac{\frac{1}{k}\sum_{i=1}^{k}\frac{1}{\Gamma(\omega)}\left(-\ln\left(\frac{i}{k}\right)\right)^{\omega-1}(E_i - 1)}{\sigma_{\mathcal{L}(\omega)}},$$

are asymptotically standard normal r.v.'s, with E_i, $i = 1, 2, \dots$ a sequence of i.i.d. exponential random variables.

Notice that the asymptotic variance and bias of the EVI-estimator $\hat{\xi}_n^\bullet(k)$, in (12), are given by $Var(\hat{\xi}_n^\bullet(k)) = \sigma_\bullet^2/k$ and $Bias(\hat{\xi}_n^\bullet(k)) = b_\bullet A(n/k)$, respectively. Then the asymptotic mean squared error (AMSE) of $\hat{\xi}_n^\bullet(k)$ is

$$AMSE(\hat{\xi}_n^\bullet(k)) = \frac{\sigma_\bullet^2}{k} + b_\bullet^2 A^2(n/k). \tag{15}$$

Moreover, if we assume $A(t) = \xi\beta t^\rho$, $\rho < 0$ and ω a fixed value, the threshold k_0 that minimizes the AMSE, the so-called optimal level for the estimation of ξ (Brilhante et al,[8] Hall,[25] and Caeiro and Gomes,[26] among others), is given by

$$k_0^\bullet = k_0^\bullet(n) = \arg\min_k AMSE(\hat{\xi}_n^\bullet(k)) = \left(\frac{\sigma_\bullet^2}{(-2\rho)b_\bullet^2\xi^2\beta^2}\right)^{\frac{1}{1-2\rho}} n^{\frac{-2\rho}{1-2\rho}}(1 + o(1)).$$

For the estimators under consideration, it should be noted that $\sigma_\bullet^2/\xi^2$, and consequently, $k_0^\bullet$ are independent of ξ.

Corollary 1. *Under the assumptions of Theorem 1 and assuming that the intermediate sequence $k = k_n$, in (4), satisfies $\sqrt{k}A(n/k) \to \lambda$, finite, as $n \to \infty$, we have*

$$\sqrt{k}(\hat{\xi}_n^\bullet(k) - \xi)\xrightarrow[n\to\infty]{d} \mathcal{N}\left(\lambda b_\bullet, \sigma_\bullet^2\right),$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes a Normal r.v. with mean value μ and variance σ^2.

Remark 2. By replacing ω by 1 in the expressions of $b_\bullet$ and $\sigma_\bullet^2$ of both classes of EVI estimators considered in the previous corollary, we obtain the well-known asymptotically normal limit behaviour of the Hill estimator in (3),[27] ie,

$$\sqrt{k}(\hat{\xi}_n^H(k) - \xi) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda}{1 - \rho}, \xi^2\right), \quad \text{as} \quad n \to \infty.$$

Further note that the normalized standard deviations, $\sigma_{P(\omega)}/\xi$ and $\sigma_{\mathcal{L}(\omega)}/\xi$, with $\sigma_{P(\omega)}^2$ and $\sigma_{\mathcal{L}(\omega)}^2$ given in (13), are independent of ξ and have a global minimum for $\omega = 1$. Additionally, the classes of generalized Hill estimators in (5) and (6) are asymptotically biased. Indeed, both $b_{P(\omega)}$ and $b_{\mathcal{L}(\omega)}$ bias components in (14) are also independent of ξ and take only positive values for $\omega > 0$. Also, $b_{P(\omega)}$ ($b_{\mathcal{L}(\omega)}$) increases (decreases) as the ω-value increases and $b_{P(\omega)} \to 1$, and $b_{\mathcal{L}(\omega)} \to 0$, as $\omega \to \infty$. In Figure 1, we illustrate the behaviour of the normalized standard deviations and bias components.

We expect that the class of EVI estimators $\hat{\xi}_n^{\mathcal{P}(\omega)}(k)$, defined in (5), to be less efficient than the Hill estimator for $\omega > 1$. The same happens with the class of EVI estimators $\hat{\xi}_n^{\mathcal{L}(\omega)}(k)$ in (6) for $\omega < 1$.

2.3 | Asymptotic comparison at optimal levels

Following De Haan and Peng[28] (see also Gomes and Martins,[7] Gomes et al,[16,29] and Gomes and Henriques-Rodrigues,[30] among others) for the classical EVI estimators and recently in Caeiro and Gomes[31] for reduced-bias (RB) EVI estimators, we now proceed to an asymptotic comparison of the estimators under study at their optimal levels, $k_0^{\bullet}$. As a first step, we minimize the AMSE with respect to the threshold k, assuming that the tuning parameter ω is fixed. Next, we improve the efficiency of the classes of EVI-estimators by dealing with the choice of the parameter ω.

Let us assume that $\hat{\xi}_n^{\bullet}(k)$ denotes any arbitrary semi-parametric estimator of the EVI with an AMSE given in (15). The use of regular variation theory[1] enables us to guarantee that, whenever $b_{\bullet} \neq 0$, there exists a function $\varphi(n, \xi, \rho)$, dependent only on the underlying model, and not on the estimator,[28] such that

$$\lim_{n \to \infty} \varphi(n, \xi, \rho) \, \text{AMSE}\left(\hat{\xi}_0^{\bullet}\right) = \left(\sigma_{\bullet}^2\right)^{-\frac{2\rho}{1-2\rho}} \left(b_{\bullet}^2\right)^{\frac{1}{1-2\rho}} =: \, \text{LMSE}\left(\hat{\xi}_0^{\bullet}\right),$$

with $\hat{\xi}_0^{\bullet} = \hat{\xi}_n^{\bullet}(k_0^{\bullet})$ the EVI-estimator computed at its optimal level and with LMSE standing for limiting mean square error. By considering the ratio of the LMSE for two different EVI estimators, we get a measure independent of the $\varphi(.)$ function. We can then consider the following measure of efficiency:

Definition 1. Given two biased estimators $\hat{\xi}_n^{(1)}(k)$ and $\hat{\xi}_n^{(2)}(k)$, for which a distributional representation of the type of the one in (12) holds, with constants (σ_1, b_1) and (σ_2, b_2), $b_1, b_2 \neq 0$, respectively, both computed at their optimal levels, the *asymptotic root efficiency* (AREFF) of $\hat{\xi}_0^{(1)}$ relatively to $\hat{\xi}_0^{(2)}$ is

$$\text{AREFF}_{1|2} \equiv \text{AREFF}_{\hat{\xi}_0^{(1)}|\hat{\xi}_0^{(2)}} := \sqrt{\frac{\text{LMSE}[\hat{\xi}_0^{(2)}]}{\text{LMSE}[\hat{\xi}_0^{(1)}]}} = \left(\left(\frac{\sigma_2}{\sigma_1}\right)^{-2\rho} \left|\frac{b_2}{b_1}\right|\right)^{\frac{1}{1-2\rho}}. \tag{16}$$

The measure, in (16), was devised so that if AREFF is equal to 1, the estimators have the same efficiency at their optimal levels. The higher the AREFF measure is, the better is the estimator identified with the superscript (1). If the asymptotic comparison of the aforementioned classes of EVI estimators is made with the Hill estimator, we have

$$\text{AREFF}_{\mathcal{P}(\omega)|H} \equiv \text{AREFF}_{\hat{\xi}_0^{\mathcal{P}(\omega)}|\hat{\xi}_0^H} := \left(\frac{2\omega-1}{\omega^2}\right)^{\frac{-\rho}{1-2\rho}} \left(\frac{\omega-\rho}{\omega(1-\rho)}\right)^{\frac{1}{1-2\rho}}, \tag{17}$$

and

$$\text{AREFF}_{\mathcal{L}(\omega)|H} \equiv \text{AREFF}_{\hat{\xi}_0^{\mathcal{P}(\omega)}|\hat{\xi}_0^H} := \left(\frac{\Gamma^2(\omega)}{\Gamma(2\omega-1)}\right)^{\frac{-\rho}{1-2\rho}} \left((1-\rho)^{\omega-1}\right)^{\frac{1}{1-2\rho}}, \tag{18}$$

for $\omega > 0.5$. It should be noted that both AREFF indicators in (17) and (18) are independent of ξ and β and depend only on the values of ρ and ω. Figure 2 shows a contour plot of $\text{AREFF}_{\mathcal{P}(\omega)|H}$ and $\text{AREFF}_{\mathcal{L}(\omega)|H}$ in (17) and (18), respectively, in the (ω, ρ)-plane. As may be seen in the figure, there is a region where both $\hat{\xi}_0^{\mathcal{P}(\omega)}$ and $\hat{\xi}_0^{\mathcal{L}(\omega)}$ outperforms asymptotically the Hill estimator. For an arbitrary values of the second-order parameter $\rho < 0$, the asymptotic efficiency of $\hat{\xi}_0^{\mathcal{P}(\omega)}$ and $\hat{\xi}_0^{\mathcal{L}(\omega)}$ relatively to $\hat{\xi}_0^H$ is larger than 1 in a region of ω-values, dependent on ρ. The region includes more values of w as ρ approaches 0. However, the improvement is not impressive. The maximal value of the measures $\text{AREFF}_{\mathcal{P}(\omega)|H}$ and $\text{AREFF}_{\mathcal{L}(\omega)|H}$ is 1.02505 and 1.04211, respectively.

Remark 3. By considering the possibility of working with a tuning parameter ω in the interval (0.5, 1), we were able to improve the conclusions presented in Gomes et al,[16] for the class of EVI estimators in (17), in which $\omega = 1$ was considered optimal. Indeed, Figure 2 shows that we can have an $\text{AREFF}_{\mathcal{P}(\omega)|H}$ indicator greater than 1 for $\omega < 1$.
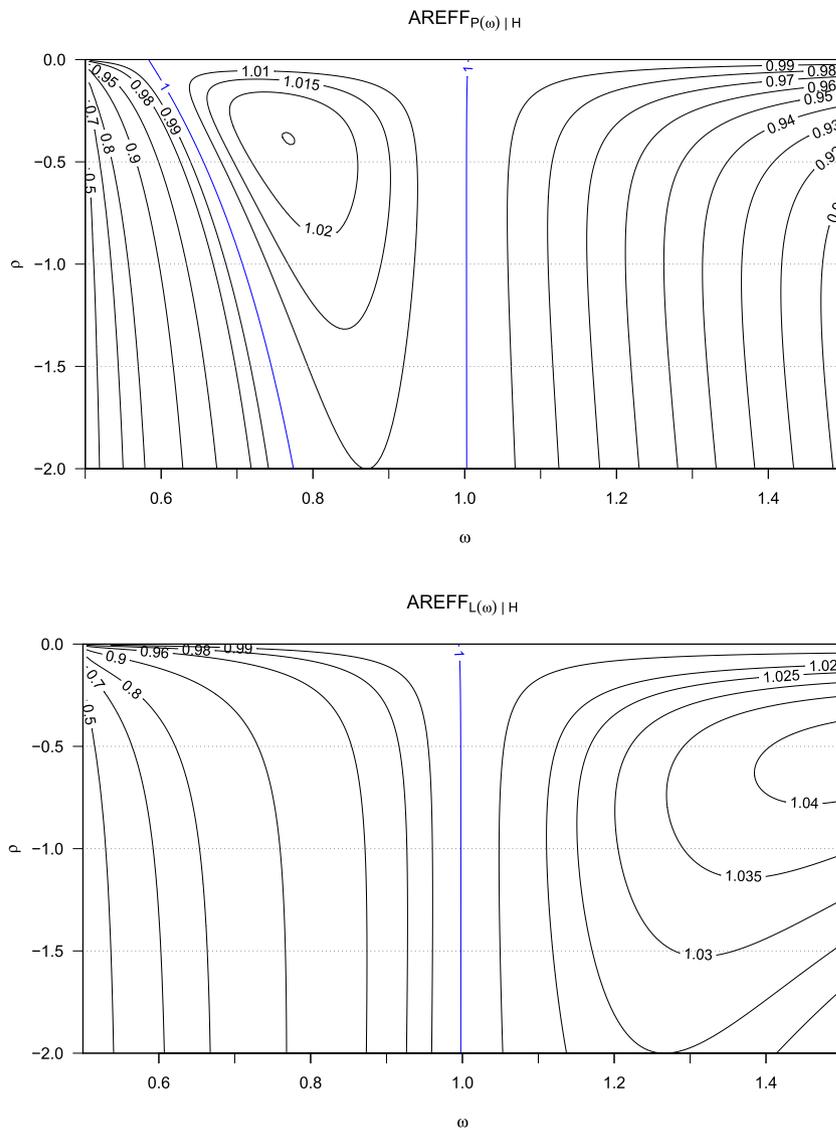
$\text{AREFF}_{\mathcal{P}(\omega)\,|\,H}$

$\text{AREFF}_{L(\omega)\,|\,H}$

FIGURE 2 Contour plot of $\text{AREFF}_{\mathcal{P}(\omega)|H}$ (top) and $\text{AREFF}_{\mathcal{L}(\omega)|H}$ (bottom) as a function of (ω, ρ) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Values of $\omega_{0|\mathcal{P}}$ and $\omega_{0|\mathcal{L}}$ for some selected values of ρ

ρ	$-\infty$	-3	-2.5	-2	-1.5	-1	-0.75	-0.5	-0.25	-0.1	-0.01	0
$\omega_{0|\mathcal{P}}$	1	0.898	0.886	0.871	0.851	0.823	0.804	0.781	0.750	0.726	0.709	$\frac{1}{\sqrt{2}}$
$\omega_{0|\mathcal{L}}$	1	1.204	1.231	1.266	1.318	1.400	1.464	1.557	1.708	1.855	1.983	2

Regarding the optimal choice of ω, it is easy to verify that the maximization of $\text{AREFF}_{\mathcal{P}(\omega)|H}$ or $\text{AREFF}_{\mathcal{L}(\omega)|H}$ is equivalent to the minimization of the LMSE of the classes of EVI estimators $\hat{\xi}_0^{\mathcal{P}(\omega)}$ and $\hat{\xi}_0^{\mathcal{L}(\omega)}$, respectively. Therefore, with respect to $\hat{\xi}_0^{\mathcal{P}(\omega)}$, it follows that the optimal value of ω, denoted by $\omega_{0|\mathcal{P}}$, is the solution of the equation $\omega^2 - \rho\omega + \rho - 0.5 = 0$; ie, it is explicitly given by

$$\omega_{0|\mathcal{P}} = \arg\min_{\omega} \text{LMSE}\left(\hat{\xi}_0^{\mathcal{P}(\omega)}\right) = \arg\max_{\omega} \text{AREFF}_{\mathcal{P}(\omega)|H} = \frac{\rho}{2} + \sqrt{\left(1 - \frac{\rho}{2}\right)^2 - \frac{1}{2}}.$$

As for the class of EVI estimators $\hat{\xi}_0^{\mathcal{L}(\omega)}$, the optimal value of ω, now denoted $\omega_{0|\mathcal{L}}$, is a function of ρ and is the solution of the equation $\psi(2\omega) - \psi(\omega) = -\ln(1-\rho)/(2\rho)$, where $\psi(t) = \Gamma'(t)/\Gamma(t), t > 0$ denotes the Digamma function. This equation can be easily solved numerically. In Table 1, we present the optimal values $\omega_{0|\mathcal{P}}$ and $\omega_{0|\mathcal{L}}$ for some selected values of ρ. The limits of the optimal levels, when $\rho \to -\infty$ or $\rho \to 0$, are also presented in the aforementioned table.

Let us now denote $\mathcal{P}_0 := \mathcal{P}(\omega_{0|\mathcal{P}})$ and $\mathcal{L}_0 := \mathcal{L}(\omega_{0|\mathcal{L}})$. Then we have $\text{AREFF}_{\mathcal{P}_0|H} > 1$ and $\text{AREFF}_{\mathcal{L}_0|H} > 1$ for every $\rho < 0$. Consequently, it is thus sensible to compare directly both estimators $\hat{\xi}_0^{\mathcal{P}_0}$ and $\hat{\xi}_0^{\mathcal{L}_0}$, ie, the classes of EVI estimators
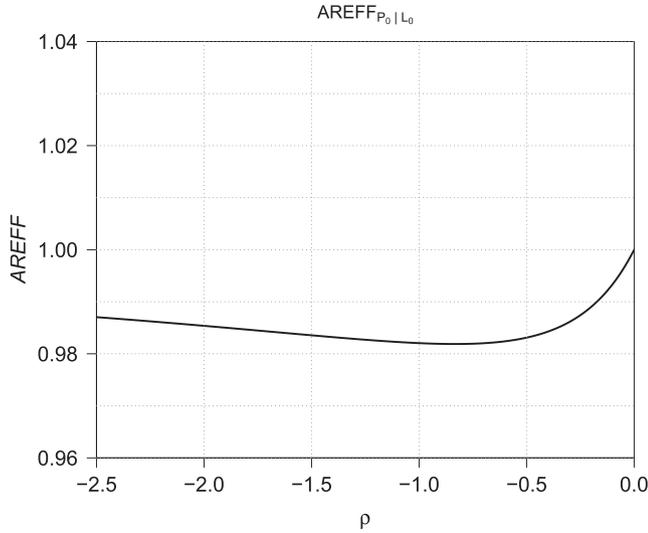
FIGURE 3   Graph of $\mathrm{AREFF}_{\mathcal{P}_0|\mathcal{L}_0}$, as a function of ρ

(5) and (6) at their optimal values of the parameters k and ω. In Figure 3, we provide a plot of $\mathrm{AREFF}_{\mathcal{P}_0|\mathcal{L}_0}$. The plot shows that $\mathrm{AREFF}_{\mathcal{P}_0|\mathcal{L}_0} < 1$ for $-2.5 < \rho < 0$; ie, at the corresponding optimal levels k and ω, $\hat{\xi}_0^{\mathcal{L}_0}$ is asymptotically more efficient than $\hat{\xi}_0^{\mathcal{P}_0}$.

The asymptotic results here presented provide us important information about the behaviour of the estimators for very large samples. Empirical simulation results such as the ones in Gomes and Martins[7] and Brilhante et al[8] show that the empirical efficiency of many estimators, comparatively with the Hill estimator, is usually much larger than the asymptotic efficiency.

Remark 4. On the basis of the theoretical results here presented, we can suggest the following adaptive EVI-estimators:

$$\hat{\xi}_n^{\hat{\mathcal{P}}_0}(k) = \hat{\xi}_n^{\mathcal{P}(\hat{\omega}_{0|\mathcal{P}})}(k) \qquad \text{and} \qquad \hat{\xi}_n^{\hat{\mathcal{L}}_0}(k) = \hat{\xi}_n^{\mathcal{L}(\hat{\omega}_{0|\mathcal{L}})}(k), \quad 1 \le k < n, \tag{19}$$

where $\hat{\rho}$ is an adequate estimator of the second-order parameter ρ in (9),

$$\hat{\omega}_{0|\mathcal{P}} = \frac{\hat{\rho}}{2} + \sqrt{\left(1 - \frac{\hat{\rho}}{2}\right)^2 - \frac{1}{2}},$$

and $\hat{\omega}_{0|\mathcal{L}}$ is the solution of the equation $\psi(2\hat{\omega}_{0|\mathcal{L}}) - \psi(\hat{\omega}_{0|\mathcal{L}}) = -\ln(1 - \hat{\rho})/(2\hat{\rho})$. The study of the estimators in (19) is beyond the scope of this paper and will be be addressed in future work.

3 | NEW CLASSES OF REDUCED BIAS KERNEL ESTIMATORS

The problem of the bias of classic EVI estimators, such as the estimators in (3), (5), and (6), has stimulate researchers to consider the correction of the dominant term of asymptotic bias. In the first papers, the classical bias-variance trade-off always appears. Such a trade-off was only removed with an appropriate estimation of the second-order parameters (ρ, β), as done in previous studies.[23,32-35] Those second-order reduced-bias estimators are usually called minimum variance reduced-bias (MVRB) EVI estimators because they have an asymptotic variance equal to that of the Hill estimator but an asymptotic bias of smaller order and thus outperform the Hill estimator for all k. The simplest MVRB estimator is the one introduced in Caeiro et al,[32] with the functional form

$$\hat{\xi}_n^{CH}(k) = \hat{\xi}_n^{CH}(k|\hat{\beta}, \hat{\rho}) \equiv CH(k) = \hat{\xi}_n^H(k)\left(1 - \frac{\hat{\beta}}{1 - \hat{\rho}}\left(\frac{n}{k}\right)^{\hat{\rho}}\right), \quad 1 \le k < n, \tag{20}$$

with $(\hat{\beta}, \hat{\rho})$ adequate estimators of the vector of second-order parameters, $(\beta, \rho) \in (\mathbb{R}, \mathbb{R}^-)$, computed at the high level $k_1 = \lfloor n^{1-\epsilon} \rfloor$, with $\epsilon \to 0$, a threshold used in Caeiro et al[32] and Gomes and Pestana,[22] among others. As usual, for the estimation of ρ, it has been advised the use of the simplest class of estimators in Fraga Alves et al,[36] denoted now

$\hat{\rho} = \hat{\rho}(k_1)$. The estimation of β has been performed through the use of a statistics dependent on a consistent estimator of ρ, denoted $\hat{\beta}(k; \hat{\rho})$, introduced in Gomes and Martins,[37] and also computed at the same k_1, ie, $\hat{\beta} = \hat{\beta}(k_1; \hat{\rho})$. Algorithms for the estimation of the second-order parameters (ρ, β) can be found in Gomes and Pestana,[22] Gomes et al,[23] and Caeiro and Gomes,[26] among others.

Since there is not any RB or MVRB estimator that can always dominate all the alternative estimators, the study of new EVI-estimators is still an important and relevant research topic. Using the method of direct reduction of the dominant component of asymptotic bias used in Caeiro et al,[32] in this paper, we further consider the RB estimators

$$\bar{\xi}_n^{\mathcal{P}(\omega)}(k) = \bar{\xi}_n^{\mathcal{P}(\omega)}(k|\hat{\beta}, \hat{\rho}) = \hat{\xi}_n^{\mathcal{P}(\omega)}(k)\left(1 - \frac{\omega\hat{\beta}}{\omega - \hat{\rho}}\left(\frac{n}{k}\right)^{\hat{\rho}}\right), \quad 1 \le k < n \tag{21}$$

and

$$\bar{\xi}_n^{\mathcal{L}(\omega)}(k) = \bar{\xi}_n^{\mathcal{L}(\omega)}(k|\hat{\beta}, \hat{\rho}) = \hat{\xi}_n^{\mathcal{L}(\omega)}(k)\left(1 - \frac{\hat{\beta}}{(1 - \hat{\rho})^{\omega}}\left(\frac{n}{k}\right)^{\hat{\rho}}\right), \quad 1 \le k < n \tag{22}$$

where $\omega > 0.5$ is a tuning parameter and $(\hat{\beta}, \hat{\rho})$ adequate estimators of the vector of second order parameters (β, ρ). Further note that the class of RB EVI-estimators (21) and (22) extends the MVRB estimators in (20).

3.1 | Asymptotic behaviour of the reduced-bias estimators

Adaptations of the theorem 2 in Goegebeur et al[14] and proposition 4.1.1. in Caeiro and Gomes[35] enable us to state the following proposition.

Proposition 1. *Suppose the third-order framework in (10) with $A(t)$ and $B(t)$ satisfying (11) and let $k = k_n$ be an intermediate sequence. Then, with $\hat{\xi}_n^{\bullet}(k)$ generally denoting any of the classes of estimators in (5) and (6), we have the validity of the following asymptotic distributional representation for any real number $\omega > 0.5$*

$$\hat{\xi}_n^{\bullet}(k) \overset{d}{=} \xi + \frac{\sigma_{\bullet}}{\sqrt{k}}Z_k^{\bullet} + b_{\bullet}^{(1)}A(n/k) + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right) + b_{\bullet}^{(2)}A^2(n/k)(1 + o_p(1))$$

with $\sigma_{\bullet}$, $Z_k^{\bullet}$ and $b_{\bullet}^{(1)} = b_{\bullet}$ defined in Theorem 1 and

$$b_{\mathcal{P}(\omega)}^{(2)} = \frac{\delta\omega}{\xi(\omega - 2\rho)}, \quad b_{\mathcal{L}(\omega)}^{(2)} = \frac{\delta}{\xi(1 - 2\rho)^{\omega}}.$$

Next, we formulate the main results about the asymptotic behaviour of the EVI estimators in Equations (21) and (22).

Theorem 2. *Under the conditions of Proposition 1, and if we consider consistent estimators $(\hat{\beta}, \hat{\rho})$ of (β, ρ) both computed at a level k_1 such that $k = o(k_1)$, and assuming $(\hat{\rho} - \rho)\ln n = o_p(1)$ and $(\hat{\beta} - \beta)/\beta \overset{p}{\sim} -(\hat{\rho} - \rho)\ln(n/k_1)$, a condition that holds for several estimators of β, then for intermediate k such that $\sqrt{k}A(n/k) \to \lambda_1$, $\bar{\xi}_n^{\bullet}(k)$ have a null dominant component of asymptotic bias. More specifically,*

$$\sqrt{k}\left(\bar{\xi}_n^{\bullet}(k) - \xi\right) \overset{d}{\to} \mathcal{N}\left(0, \sigma_{\bullet}^2\right), \quad as \quad n \to \infty. \tag{23}$$

Moreover, for levels k such that $\sqrt{k}A(n/k) \to \infty$ and $\sqrt{k}A^2(n/k) \to \lambda_2$, finite, and further assuming that k_1 is optimal for the estimation of ρ, ie, $\hat{\rho} - \rho = O_p(1/\sqrt{k_1}A(n/k_1))$,

$$\sqrt{k}\left(\bar{\xi}_n^{\bullet}(k) - \xi\right) \overset{d}{\to} N(\lambda_2\tilde{b}_{\bullet}, \sigma_{\bullet}^2), \quad as \quad n \to \infty, \tag{24}$$

with

$$\tilde{b}_{\mathcal{P}(\omega)} = \frac{1}{\xi}\left(\frac{\delta\omega}{\omega - 2\rho} - \frac{\omega^2}{(\omega - \rho)^2}\right), \quad \tilde{b}_{\mathcal{L}(\omega)} = \frac{1}{\xi}\left(\frac{\delta}{(1 - 2\rho)^{\omega}} - \frac{1}{(1 - \rho)^{2\omega}}\right).$$

Proof. Trivial computations allow us to get

$$\tilde{\xi}_n^\bullet(k|\beta,\rho) \stackrel{d}{=} \xi + \frac{\sigma_\bullet}{\sqrt{k}} Z_k^\bullet + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right) + \tilde{b}_\bullet A^2(n/k)(1 + o_p(1)).$$

Consequently, $\sqrt{k}(\tilde{\xi}_n^\bullet(k|\beta,\rho) - \xi) \stackrel{d}{\to} N\left(0, \sigma_\bullet^2\right)$ if $\sqrt{k}A(n/k) \to \lambda_1$ and $\sqrt{k}(\tilde{\xi}_n^\bullet(k|\beta,\rho) - \xi) \stackrel{d}{\to} N\left(\lambda_2 \tilde{b}_\bullet, \sigma_\bullet^2\right)$ if $\sqrt{k}A(n/k) \to \lambda_2$.

Cramer delta method allows us to check that it holds

$$b_\bullet^{(1)}(\hat\rho)\hat\beta\left(\frac{n}{k}\right)^{\hat\rho} = b_\bullet^{(1)}(\hat\rho)\frac{A(n/k)}{\xi}\left[1 + (a_\bullet + \ln(n/k))(\hat\rho - \rho)(1 + o_p(1)) + \frac{\hat\beta - \beta}{\beta}(1 + o_p(1))\right]$$

for some suitable $a_\bullet = a_\bullet(\rho, \omega)$. Next, we obtain

$$\tilde{\xi}_n^\bullet(k|\beta,\rho) - \tilde{\xi}_n^\bullet(k|\beta,\rho) \stackrel{d}{=} b_\bullet^{(1)}(\hat\rho)\frac{A(n/k)}{\xi}\left(\ln(n/k)(\hat\rho - \rho)(1 + o_p(1)) + \frac{\hat\beta - \beta}{\beta}(1 + o_p(1))\right)$$

and the first part of the theorem, in (23), follows straightforwardly.

To prove (24), we use $(\hat\beta - \beta)/\beta \stackrel{p}{\sim} -(\hat\rho - \rho)\ln(n/k_1)$ and obtain

$$D_{k,k_1,n} := \tilde{\xi}_n^\bullet(k|\hat\beta,\hat\rho) - \tilde{\xi}_n^\bullet(k|\beta,\rho) \stackrel{d}{=} -b_\bullet^{(1)}(\hat\rho)\frac{A(n/k)}{\xi}\ln(k_1/k)(\hat\rho - \rho)(1 + o_p(1)).$$

By using $\hat\rho - \rho = O_p(1/\sqrt{k_1}A(n/k_1))$, we deduce $\sqrt{k}D_{k,k_1,n} = o_p(1)$, and the second part of the theorem follows. $\square$

Remark 5. The asymptotic behaviour in Theorem 2 is also true for the CH estimator in (20). Indeed, we have $\sigma_{CH}^2 = \sigma_{P(1)}^2 = \sigma_{\mathcal{L}(1)}^2 = \xi^2$ and

$$\tilde{b}_{CH} = \tilde{b}_{P(1)} = \tilde{b}_{\mathcal{L}(1)} = \frac{1}{\xi}\left(\frac{\delta}{1 - 2\rho} - \frac{1}{(1 - \rho)^2}\right).$$

It is worth to mention that all RB EVI estimators considered in this section can be third-order reduced bias estimators, ie, can have $\tilde{b}_\bullet = 0$ in (24) for certain values of the parameters (δ, ρ). For the estimator $\tilde{\xi}_n^{CH}(k)$, in (20), the asymptotic bias is null if and only if the parameters (δ, ρ) satisfy the relation

$$\rho = 1 - \frac{1}{1 - \sqrt{1 - \delta}}, \qquad 0 < \delta < 1.$$

With respect to $\tilde{\xi}_n^{P(\omega)}(k)$ and $\tilde{\xi}_n^{\mathcal{L}(\omega)}(k)$, in (21) and (22), respectively, the adequate choice of the tuning parameter ω allow us to have a larger region of the plane (δ, ρ) where $\tilde{b}_\bullet = 0$. For the RB EVI estimators in (21) and (22), the asymptotic bias is null whenever we choose $\omega = \tilde{\omega}_{P_0}$ and $\omega = \tilde{\omega}_{\mathcal{L}_0}$, respectively, with $\tilde{\omega}_{P_0}$ and $\tilde{\omega}_{\mathcal{L}_0}$ given by

$$\tilde{\omega}_{P_0} = \rho\left(1 - \frac{1}{\sqrt{1 - \delta}}\right), \qquad \text{with} \qquad \rho < \frac{-\sqrt{1 - \delta}}{2(\sqrt{1 - \delta} - 1)}, \qquad 0 < \delta < 1, \qquad (25)$$

and

$$\tilde{\omega}_{\mathcal{L}_0} = -\frac{\ln(\delta)}{\ln(1 - \rho)^2 - \ln(1 - 2\rho)}, \qquad \text{with} \qquad \rho > \frac{\delta^2 - 1 - \sqrt{1 - \delta^2}}{\delta^2}, \qquad 0 < \delta < 1. \qquad (26)$$

The direct estimation of the tuning parameters in (25) and (26) is outside the scope of this work. This is due to the fact that the estimation of the third-order parameter β' in (11) is still an open topic in statistics of extremes.

In Figure 4, we present the region in the plane (δ, ρ), where the aforementioned RB EVI estimators can be third-order reduced bias estimators. Notice that such region is a curve for the CH estimator in (20). For the two other RB EVI estimators in (21) and (22), there is a large region on the plane (δ, ρ) where we can have can have $\tilde{b}_\bullet = 0$. Moreover, if $\tilde{b}_{P(\omega)} = 0$
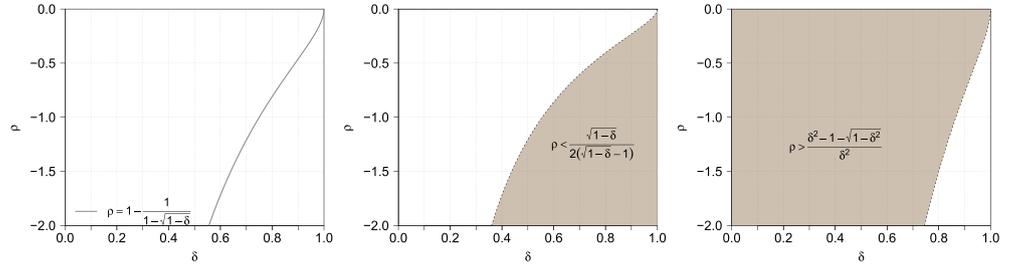
or $\tilde{b}_{\mathcal{L}(\omega)} = 0$ and $\tilde{b}_{CH} \neq 0$, then the new RB EVI estimators in (21) and (22) outperform the state of the art in RB EVI estimation by having lower mean squared error. Such behaviour holds, for example, for Student t distribution with ν degrees of freedom.

3.2 | Asymptotic comparison at optimal levels

We now proceed to an asymptotic comparison of the RB estimators under study at their optimal levels. This is again done in a way similar to the one used in Section 2.3. Let us now assume that $\tilde{\xi}_n^{\bullet}(k)$ denotes any arbitrary RB semi-parametric estimator of the EVI. Then the AMSE of $\tilde{\xi}_n^{\bullet}(k)$ is given by $\sigma_{\bullet}^2/k + \tilde{b}_{\bullet}^2 \quad A^4(n/k)$. Let $\tilde{k}_0^{\bullet}(n):=\arg\min_k \text{AMSE}(\tilde{\xi}_n^{\bullet}(k))$ be the optimal level for the estimation of ξ through $\tilde{\xi}_n^{\bullet}(k)$ and let us denote $\tilde{\xi}_0^{\bullet} := \tilde{\xi}_n^{\bullet}(\tilde{k}_0)$, the estimator computed at its optimal level. Then, whenever $\tilde{b}_{\bullet} \neq 0$, there exists a function $\varphi(n, \xi, \rho)$ such that

$$\lim_{n\to\infty} \varphi(n, \xi, \rho) \text{AMSE}\left(\tilde{\xi}_0^{\bullet}\right) = \left(\sigma_{\bullet}^2\right)^{-\frac{4\rho}{1-4\rho}} \left(\tilde{b}_{\bullet}^2\right)^{\frac{1}{1-4\rho}} =: \text{LMSE}\left(\tilde{\xi}_0^{\bullet}\right).$$

Let us now consider two estimators $\tilde{\xi}_0^{(1)}$ and $\tilde{\xi}_0^{(2)}$, for which the distributional representation in (24) holds, with constants $(\sigma_1, \tilde{b}_1)$ and $(\sigma_2, \tilde{b}_2)$, $\tilde{b}_1, \tilde{b}_2 \neq 0$, respectively, both computed at their optimal levels. Then the AREFF measure of efficiency, presented in Definition 1, of $\tilde{\xi}_0^{(1)}$ relatively to $\tilde{\xi}_0^{(2)}$, is

$$\text{AMSE}_{\tilde{1}|\tilde{2}} \equiv \text{AMSE}_{\tilde{\xi}_0^{(1)}|\tilde{\xi}_0^{(2)}} := \sqrt{\frac{\text{LMSE}[\tilde{\xi}_0^{(2)}]}{\text{LMSE}[\tilde{\xi}_0^{(1)}]}} = \left(\left(\frac{\sigma_2}{\sigma_1}\right)^{-4\rho} \left|\frac{\tilde{b}_2}{\tilde{b}_1}\right|\right)^{\frac{1}{1-4\rho}}.$$

Now, we are able to compare the RB EVI estimators $\tilde{\xi}_0^{P(\omega)}$ and $\tilde{\xi}_0^{\mathcal{L}(\omega)}$ with the $\hat{\xi}_0^{CH}$ RB EVI estimator. It is not difficult to obtain

$$\text{AMSE}_{\tilde{P}(\omega)|CH} := \left(\left(\frac{2\omega-1}{\omega^2}\right)^{-2\rho} \left|\frac{\delta(1-2\rho)^{-1} - (1-\rho)^{-2}}{\delta\omega(\omega-2\rho)^{-1} - \omega^2(\omega-\rho)^{-2}}\right|\right)^{\frac{1}{1-4\rho}} \tag{27}$$

and

$$\text{AMSE}_{\tilde{\mathcal{L}}(\omega)|CH} := \left(\left(\frac{\Gamma^2(\omega)}{\Gamma(2\omega-1)}\right)^{-2\rho} \left|\frac{(1-2\rho)^{-1} - (1-\rho)^{-2}}{\delta(1-2\rho)^{-w} - (1-\rho)^{-2w}}\right|\right)^{\frac{1}{1-4\rho}}. \tag{28}$$

We observe that the AREFF measures in (27) and (28) depends on the the values of the parameters ω, δ, and ρ, and it is convenient to consider their maximization with respect to ω. For a large region of the parameter space of (δ, ρ), such maximization must be done by using numerical algorithms. Moreover, the adequate asymptotic comparison of the RB EVI estimators in (20), (21), and (22) is not trivial when $0 < \delta < 1$, due to the fact that the asymptotic bias for these estimators can be null at some values of the parameters. The complete study of the estimators for this type of δ-values would require a fourth-order framework, like the one considered in Goegebeur and De Wet,[38] and such study is beyond the scope of this work.

For $\delta \geq 1$, we can maximize the AREFF measure with respect to ω and perform a comparison of the estimators. For the sake of simplicity, we shall only perform the comparison for $\delta = 1$, the true value of this parameter for important models such as the Burr or the generalized Pareto (GP) models with d.f. $F(x) = 1 - (1 + x^{-\rho/\xi})^{1/\rho}$, $x > 0$, and $F(x) = 1 + \ln EV_{\xi}(x)$, $x > 0$ ($\rho = -\xi$), respectively. In Figure 5, we present a contour plot of $\text{AREFF}_{\tilde{P}(\omega)|CH}$ and $\text{AREFF}_{\tilde{\mathcal{L}}(\omega)|CH}$ defined in (27) and (28), respectively, in the (ω, ρ)-plane.
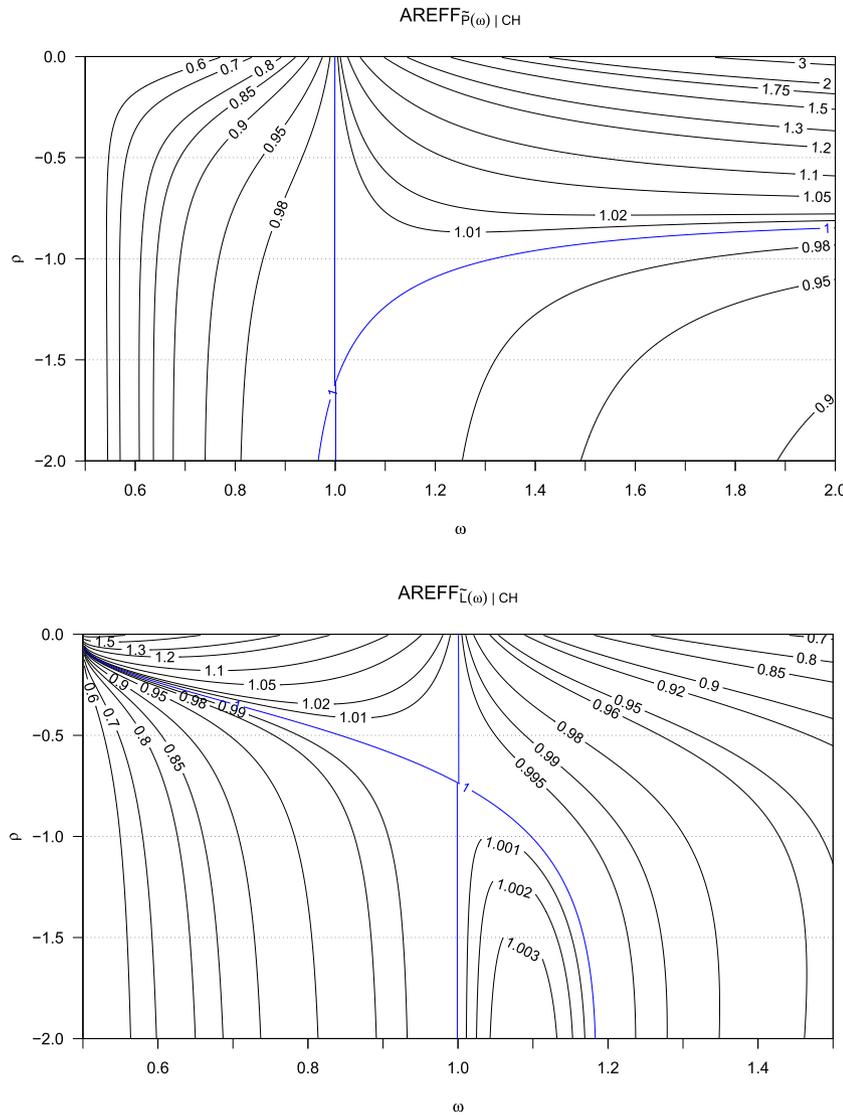
AREFF$_{\tilde{\mathcal{P}}(\omega)\,|\,CH}$

AREFF$_{\tilde{\mathcal{L}}(\omega)\,|\,CH}$

FIGURE 5 Contour plot of AREFF$_{\tilde{\mathcal{P}}(\omega)|CH}$ (top) and AREFF$_{\tilde{\mathcal{L}}(\omega)|CH}$ (bottom) as a function of (ω, ρ) [Colour figure can be viewed at wileyonlinelibrary.com]

As may be seen in this figure, there is a region of ω-values where both $\tilde{\xi}_0^{\mathcal{P}(\omega)}$ and $\tilde{\xi}_0^{\mathcal{L}(\omega)}$ outperforms asymptotically the estimator $\hat{\xi}_0^{CH}$. With an adequate choice of ω, the improvement of the new classes of reduced bias estimators is more expressive for values of the shape second order parameter ρ close to 0.

4 | A CASE-STUDY IN THE FIELD OF INSURANCE

We next consider the application of the proposed reduced bias kernel estimators in (21) and (22) to the analysis of the $n = 371$ automobile claim amounts exceeding 1,200,000 euros from several insurance companies operating with the re-insurer Secura Belgian Re. This data set is censored to the left, and the underlying model has a heavy right tail (see Beirlant et al[3,39] and Brilhante et al[8]).

Since it is not yet possible to estimate all third-order parameters, we cannot estimate directly the optimal parameters (ω, k) and extreme value index the associated reduced bias estimator. Therefore, we shall consider a two-step data-driven heuristic procedure based on the stability of the sample path of the estimates, similar to the ones we can find in Gomes et al[40] and Neves et al.[41] Bootstrap procedures like the ones described in previous studies[42-45] could also be used. The estimates of the second-order parameters (ρ, β) were computed at the level $k_1 = [n^{0.995}] = 360$ and are given by $(\hat{\rho}, \hat{\beta}) = (-0.65, 0.78)$.[39]

In Figure 6, we present, at the left, the sample paths of the estimates provided by (21) generally denoted by $\tilde{\mathcal{P}}(\omega)$, for the selected values of ω, $\omega = 0.75, 1, 1.5, 2$, and at the right, the sample paths of the estimates provided by (22) generally denoted by $\tilde{\mathcal{L}}(\omega)$ also for the selected values of ω, $\omega = 1, 1.5, 2, 2.5$. From Remark 1, we know that the MVRB estimator in

FIGURE 6 Estimates of the extreme value index provided by the estimators $\tilde{\xi}_n^{\mathcal{P}(\omega)}$ (left) and $\tilde{\xi}_n^{\mathcal{L}(\omega)}$ (right) for several values of ω [Colour figure can be viewed at wileyonlinelibrary.com]
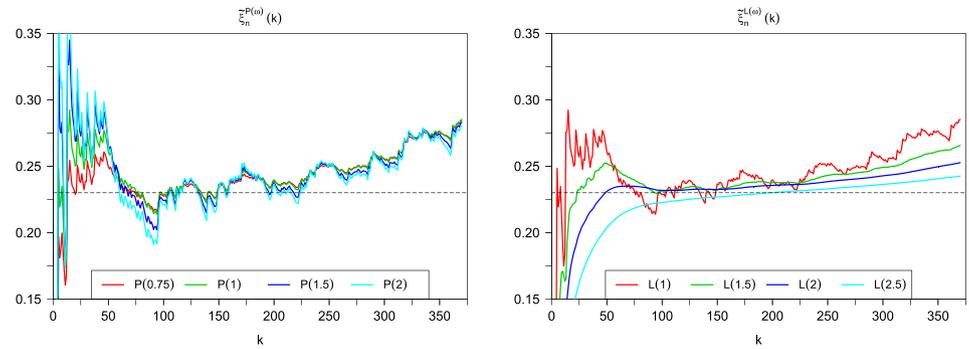
(20) is a particular case of both kernel reduced bias estimators when $\omega = 1$; ie, $\tilde{\xi}_n^{\mathcal{P}(1)}(k) \equiv \tilde{\xi}_n^{\mathcal{L}(1)}(k) \equiv CH(k)$. For this data set, the sample paths associated with the estimator $\tilde{\xi}_n^{\mathcal{P}(\omega)}$, in (21), almost overlap for a wide region of k values, independent of tuning parameter ω chosen, while the ones associated with $\tilde{\xi}_n^{\mathcal{L}(\omega)}$, in (22), exhibit the same pattern for all values of $\omega \neq 1$ and for a wide region of k values. Also, the sample paths of the new class of estimators, $\tilde{\mathcal{L}}(\omega)$, $\omega > 0.5$, and $\omega \neq 1$, are less volatile and more stable than the one associated with the estimator $CH \equiv \tilde{\mathcal{L}}(1)$ (see Figure 6, right). Using a heuristic procedure based on the stability of the sample paths, we get for the $\tilde{\mathcal{P}}(\omega)$ estimator $(\hat{\omega}, \hat{k}) = (0.75, 186)$, leading to an adaptive estimate $\tilde{\xi}^{\mathcal{P}} = 0.24$, pictured in Figure 6 (left), and for the $\tilde{\mathcal{L}}(\omega)$ estimator $(\hat{\omega}, \hat{k}) = (2.0, 169)$, leading to an adaptive estimate $\tilde{\xi}^{\mathcal{L}} = 0.23$, pictured in Figure 6 (right). The EVI estimates obtained for these new classes of kernel RB estimators are in line with the EVI estimate obtained in Beirlant et al.[39]

5 | CONCLUSIONS

In this paper, we considered two classes of kernel estimators that generalize the Hill EVI estimator. We proved their asymptotic normality and demonstrate that those classes of EVI estimators can be more efficient than the Hill estimator. In addition, we studied new classes of reduced bias versions of the aforementioned classes of kernel estimators. The asymptotic results show that with an adequate choice of the parameters ω and k, the new reduced bias estimators have a good performance. The choice of the nuisance parameters ω and k can be done on the basis of data-driven heuristic procedures related to the sample path stability or based on the double-bootstrap procedure. In real case studies, we advise the practitioners to plot the sample paths of the new estimators for several values of ω to identify patterns behaviour and to hint the best adaptive procedure to estimate ω and k. Although the results here presented for the new reduced bias EVI-estimators are appealing, future work on the behaviour of the estimators for finite sample sizes and applications of the estimators in other real case studies are still needed.

CONFLICT OF INTEREST

There are no conflicts of interest to this work.

ORCID

Frederico Caeiro 🆔 https://orcid.org/0000-0001-8628-7281
Lígia Henriques-Rodrigues 🆔 https://orcid.org/0000-0003-4881-4188

REFERENCES

1. Bingham NH, Goldie CM, Teugels JL. *Regular Variation*. Cambridge university press; 1989.
2. Hill BM. A simple general approach to inference about the tail of a distribution. *Ann Statist*. 1975;3(5):1163-1174.
3. Beirlant J, Goegebeur Y, Segers J, Teugels JL. *Statistics of Extremes: Theory and Applications*: John Wiley & Sons; 2004.

4. Beirlant J, Caeiro F, Gomes MI. An overview and open research topics in statistics of univariate extremes. *Revstat*. 2012;10(1):1-31.

5. Reiss R-D, Thomas M. *Statistical Analysis of Extreme Values With Applications to Insurance, Finance, Hydrology and Other Fields*: Birkhäuser Verlag; 2007.

6. Gomes MI, Guillou A. Extreme value theory and statistics of univariate extremes: a review. *Int Stat Rev*. 2015;83(2):263-292.

7. Gomes MI, Martins MJ. Generalizations of the Hill estimator—asymptotic versus finite sample behaviour. *J Statist Plan Inference*. 2001;93(1-2):161-180.

8. Brilhante MF, Gomes MI, Pestana D. A simple generalization of the Hill estimator. *Comput Stat Data Anal*. 2013;57(1):518-535.

9. Paulauskas V, Vaičiulis M. On an improvement of Hill and some other estimators. *Lith Math J*. 2013;53(3):336-355.

10. Paulauskas V, Vaičiulis M. A class of new tail index estimators. *Ann Inst Statist Math*. 2017;69(2):461-487.

11. Beran J, Schell D, Stehlík M. The harmonic moment tail index estimator: asymptotic distribution and robustness. *Ann Inst Stat Math*. 2014;66(1):193-220.

12. Penalva H, Gomes MI, Caeiro F, Neves MM. A couple of non reduced bias generalized means in extreme value theory: an asymptotic comparison. *Revstat*. in press.

13. Csörgő S, Deheuvels P, Mason D. Kernel estimates of the tail index of a distribution. *Ann Statist*. 1985;13(3):1050-1077.

14. Goegebeur Y, Beirlant J, De Wet T. Kernel estimators for the second order parameter in extreme value statistics. *J Statist Plan Inference*. 2010;140(9):2632-2652.

15. Gomes MI, Martins MJ. Bias reduction and explicit semi-parametric estimation of the tail index. *J Statist Plan Inference*. 2004;124(2):361-378.

16. Gomes MI, Miranda C, Viseu C. Reduced-bias tail index estimation and the Jackknife methodology. *Statistica Neerlandica*. 2007;61(2):243-270.

17. Diop A, Lo GS. Generalized Hill's estimator. *Far East J Theor Stat*. 2006;20(2):129-149.

18. Lo GS, Deme EH, Diop A. On the generalized Hill process for small parameters and applications. *J Stat Theory Appl*. 2012;11(4):397-418.

19. Caeiro F, Henriques-Rodrigues L, Prata Gomes D. A simple class of reduced bias kernel estimators of extreme value parameters. *Comput Math Methods*. 2019;1(3):e1025.

20. Caeiro F, Gomes MI. Bias reduction in the estimation of a shape second order parameter of a heavy tailed model. *J Stat Comput Simul*. 2015;85(17):3405-3419.

21. De Haan L, Ferreira A. *Extreme Value Theory: An Introduction*: Springer; 2006.

22. Gomes MI, Pestana D. A sturdy reduced-bias extreme quantile (VaR) estimator. *J Am Stat Assoc*. 2007;102(477):280-292.

23. Gomes MI, De Haan L, Rodrigues LH. Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. *J R Stat Soc Ser B (Stat Methodol.)*. 2008;70(1):31-52.

24. Caeiro F, Prata Gomes D. Adaptive estimation of a tail shape second order parameter: a computational comparative study. *AIP Conf Proc*. 2015;1702(1):030005.

25. Hall P. On some simple estimates of an exponent of regular variation. *J R Stat Soc Ser B Methodol*. 1982;44:37-42.

26. Caeiro F, Gomes MI. Threshold selection in extreme value analysis. In: Dey D. K., Yan J., eds. *Extreme Value Modeling and Risk Analysis*: Chapman and Hall/CRC; 2016:69-86.

27. Peng L. Asymptotically unbiased estimators for the extreme-value index. *Statist Probab Lett*. 1998;38(2):107-115.

28. De Haan L, Peng L. Comparison of extreme value index estimators. *Statistica Neerlandica*. 1998;52(1):60-70.

29. Gomes MI, Pereira H, Miranda MC. Revisiting the role of the Jackknife methodology in the estimation of a positive tail index. *Commun Stat - Theory Methods*. 2005;34(2):319-335.

30. Gomes MI, Henriques-Rodrigues L. Comparison at optimal levels of classical tail index estimators: a challenge for reduced-bias estimation?. *Discussiones Mathematica: Probab Stat*. 2010;30(1):35-51.

31. Caeiro F, Gomes MI. Asymptotic comparison at optimal levels of reduced-bias extreme value index estimators. *Statistica Neerlandica*. 2011;65(4):462-488.

32. Caeiro F, Gomes MI, Pestana D. Direct reduction of bias of the classical Hill estimator. *Revstat*. 2005;3(2):113-136.

33. Caeiro F, Gomes MI, Henriques-Rodrigues L. Reduced-bias tail index estimators under a third order framework. *Commun Stat - Theory Methods*. 2009;38(7):1019-1040.

34. Gomes MI, Martins MJ, Neves M. Improving second order reduced bias extreme value index estimation. *Revstat*. 2007;5(2):177-207.

35. Caeiro F, Gomes MI. Revisiting the maximum likelihood estimation of a positive extreme value index. *J Stat Theory Appl*. 2015;9(1):200-218.

36. Fraga Alves MI, Gomes MI, De Haan L. A new class of semi-parametric estimators of the second order parameter. *Port Math*. 2003;60(2):193-214.

37. Gomes MI, Martins MJ. "asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes*. 2002;5(1):5-31.

38. Goegebeur Y, De Wet T. Estimation of the third-order parameter in extreme value statistics. *Test*. 2012;21(2):330-354.

39. Beirlant J, Figueiredo F, Gomes MI, Vandewalle B. Improved reduced-bias tail index and quantile estimators. *Journal of Statistical Planning and Inference*. 2008;138(6):1851-1870.

40. Gomes MI, Henriques-Rodrigues L, Fraga Alves MI, Manjunath B. Adaptive PORT–MVRB estimation: an empirical comparison of two heuristic algorithms. *J Stat Comput Simul*. 2013;83(6):1129-1144.

41. Neves MM, Gomes MI, Figueiredo F, Prata Gomes D. Modeling extreme events: sample fraction adaptive choice in parameter estimation. *J Stat Theory Practice*. 2015;9(1):184-199.

42. Gomes MI, Mendonça S, Pestana D. Adaptive reduced-bias tail index and VaR estimation via the bootstrap methodology. *Commun Stat - Theory Methods*. 2011;40(16):2946-2968.

43. Gomes MI, Figueiredo F, Neves MM. Adaptive estimation of heavy right tails: resampling-based methods in action. *Extremes*. 2012;15(4):463-489.

44. Gomes MI, Figueiredo F, Martins MJ, Neves MM. Resampling methodologies and reliable tail estimation. *South African Statistical Journal*. 2015;49:1-20.

45. Gomes MI, Caeiro F, Henriques-Rodrigues L, Manjunath B. Bootstrap methods in statistics of extremes. In: Longin F, ed. *Extreme Events in Finance*: John Wiley & Sons, Ltd; 2016:117-138.