WILEY | Hindawi

Research Article

The Use of Generalized Means in the Estimation of the Weibull Tail Coefficient

Frederico Caeiro [ID],[1,2] **Lígia Henriques-Rodrigues** [ID],[3,4] **and M. Ivette Gomes** [ID][5,6]

[1]*NOVA School of Science and Technology (FCT NOVA), NOVA University Lisbon, Caparica, Portugal*
[2]*Centro de Matemática e Aplicações (CMA), NOVA University Lisbon, Caparica, Portugal*
[3]*Department of Mathematics, School of Sciences and Technology, University of Évora, Évora, Portugal*
[4]*Research Center in Mathematics and Applications (CIMA), University of Évora, Évora, Portugal*
[5]*DEIO, Faculty of Sciences, University of Lisbon (FCUL), Lisbon, Portugal*
[6]*Center of Statistics and Applications, University of Lisbon (CEAUL), Lisbon, Portugal*

Correspondence should be addressed to Frederico Caeiro; fac@fct.unl.pt

Due to the specificity of the Weibull tail coefficient, most of the estimators available in the literature are based on the log excesses and are consequently quite similar to the estimators used for the estimation of a positive extreme value index. The interesting performance of estimators based on generalized means leads us to base the estimation of the Weibull tail coefficient on the power mean-of-order-p. Consistency and asymptotic normality of the estimators under study are put forward. Their performance for finite samples is illustrated through a Monte Carlo simulation. It is always possible to find a negative value of p (contrarily to what happens with the mean-of-order-p estimator for the extreme value index), such that, for adequate values of the threshold, there is a reduction in both bias and root mean square error.

1. Introduction and Preliminaries

Statistics of extremes, either univariate or multivariate, have been recently faced with many different challenges, which have enabled to better understand the complexity of extreme events in the most diverse areas of applications, like *biostatistics*, *dynamical systems*, *environment*, *finance*, *insurance*, and *structural engineering*, among other fields. Risky events are commonly in the tails of the underlying distribution, and there are usually only a few observations in those tails. Consequently, and thinking only on the univariate situation, estimates either much above the observed maximum or below the observed minimum are often required. It is thus necessary to consider models for the tails, and those models are most of the times based on asymptotic results.

Let us assume that, possibly after an adequate transformation, the available transformed sample, $\underline{X}_n = (X_1, \cdots, X_n)$, can be regarded as a sample of size n of independent, identically distributed (IID) random variables (RVs) from a cumulative distribution function (CDF) F. More generally, $\underline{X}_n$ can be assumed to be a sample of stationary weakly dependent RVs from F. Let us use the notation $X_{1:n} \leq \cdots \leq X_{n:n}$ for the associated ascending order statistics (OSs). Further assume that there exist sequences of real constants $\{a_n > 0\}$ and $\{b_n \in \mathbb{R}\}$ such that the linearly normalized maximum, $(X_{n:n} - b_n)/a_n$, converges weakly to a nondegenerate RV. Then (see Gnedenko [1]), the limiting CDF is necessarily of the type of the general extreme value (GEV) CDF, given by

$$\mathrm{GEV}_\xi(x) \equiv G_\xi(x) = \begin{cases} e^{-(1+\xi x)^{-1/\xi}}, 1 + \xi x > 0, & \text{if } \xi \neq 0, \\ e^{-e^{-x}}, x \in \mathbb{R}, & \text{if } \xi = 0. \end{cases} \quad (1)$$

The CDF F is then said to belong to the max-domain of attraction of GEV_ξ, and the notation $F \in D_M(\mathrm{GEV}_\xi)$ is

used. The GEV_ξ model, in (1), is perhaps the most relevant univariate asymptotic model in statistical extreme value theory (EVT). For other relevant asymptotic models and different approaches to statistics of univariate extremes, see the reasonably recent overviews by [2–5].

The parameter ξ is the extreme value index (EVI), one of the most relevant parameters of large events. This parameter measures the heaviness of the right-tail function $\bar{F}(x) := 1 - F(x)$, and the heavier the right tail, the larger ξ is. Heavy-tailed models, i.e., Pareto-type underlying CDFs, with a positive EVI, belong to $D_M^+ := D_M(GEV_{\xi>0})$, with $GEV_\xi \equiv G_\xi$ defined in (1). Note that, in a univariate framework and with R_a denoting the class of regularly varying functions at infinity with an index of regular variation a, i.e., positive measurable functions g such that $\lim_{t \to \infty} g(tx)/g(t) = x^a$, for all $x > 0$ (see Bingham et al. [6], for details on regular variation), and with the notation

$$U(t) := F^{\leftarrow}\left(1 - \frac{1}{t}\right), t \geq 1, F^{\leftarrow}(y) := \inf\{x : F(x) \geq y\},$$

$$\tag{2}$$

the following equivalences hold:

$$F \in D_M^+ \Leftrightarrow \bar{F} = 1 - F \in R_{-1/\xi} \Leftrightarrow U \in R_\xi. \tag{3}$$

As an example of a CDF in D_M^+, and among many others, we mention the Fréchet CDF, $F(x) = \exp(-x^{-\alpha}), x \geq 0, \alpha > 0$ ($\xi = 1/\alpha$).

In this paper, our interest lies essentially in the estimation of the Weibull tail coefficient (WTC), another relevant parameter of extreme events. Regularly varying cumulative hazard functions $H(x) := -\ln(1 - F(x))$ will thus be considered. Indeed, the WTC is the parameter θ in a right-tail function of the type:

$$\bar{F}(x) = 1 - F(x) = e^{-H(x)}, \text{with } H \in R_{1/\theta}, \theta \in \mathbb{R}^+. \tag{4}$$

The class of models with a Weibull-type tail is quite broad and includes, among others, the normal, the gamma, the Weibull, and the logistic distributions. This type of models is quite useful in several areas of applications such as hydrology, meteorology, environmental sciences, and nonlife insurance (see de Wet et al. [7]). Further note that condition (4) is equivalent to assume that the inverse cumulative hazard function $H^{\leftarrow}$ is a regularly varying function with index θ. Thus,

$$U(e^t) = H^{\leftarrow}(t) \in RV_\theta \Leftrightarrow U(t) = (\ln t)^\theta L(\ln t), \tag{5}$$

with $L \in R_0$, a slowly varying function.

1.1. Semiparametric Estimators of the WTC. Regarding the estimation of the WTC, one of the first WTC estimators in the literature was based on record values [8]. The use of the k upper order statistics in the sample was considered in Broniatowski [9], Beirlant et al. [10, 11], and Dierckx et al.

[12]. Most WTC estimators are based on the relative excesses:

$$U_{ik} := \frac{X_{n-i+1:n}}{X_{n-k:n}}, \quad 1 \leq i \leq k \leq n-1 \text{ and } X_{n-k:n} > 0, \tag{6}$$

or on the log excesses:

$$V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \leq i \leq k \leq n-1 \text{ and } X_{n-k:n} > 0. \tag{7}$$

Indeed, Beirlant et al. [10] proposed the estimator with functional form:

$$\widehat{\theta}^B(k) := \ln\left(\frac{n}{k}\right)\left(\frac{1}{k}\sum_{i=1}^k U_{ik} - 1\right), \quad k = 1, 2, \cdots, n-1. \tag{8}$$

Beirlant et al. [13] and Girard [14] considered the following estimator of the WTC:

$$\widehat{\theta}^G(k) := \frac{1/k\sum_{i=1}^k V_{ik}}{1/k\sum_{i=1}^k \ln\ln((n+1)/i) - \ln\ln((n+1)/(k+1))},$$

$$k = 1, 2, \cdots, n-1. \tag{9}$$

Weighted versions of the estimator $\widehat{\theta}^G(k)$ can be found in Gardes and Girard [15] and Goegebeur et al. [16]. The following Hill-type WTC estimator was studied in Gardes and Girard [17]:

$$\widehat{\theta}^{GG}(k) := \frac{\ln(n/k)}{k}\sum_{i=1}^k V_{i,k} = \ln\left(\frac{n}{k}\right)H(k), \tag{10}$$

with $H(k)$ being the classical Hill (H) [18] EVI estimators for heavy-tailed models, which can be written as the average of the log excesses, i.e.,

$$H(k) := \frac{1}{k}\sum_{i=1}^k V_{ik}, \quad 1 \leq i \leq k < n, \tag{11}$$

with V_{ik} defined in (7). Consistency of the Hill estimator for ξ holds if $k = k(n)$ is an intermediate sequence, i.e., if

$$k \longrightarrow \infty, \frac{k}{n} \longrightarrow 0, \text{ as } n \longrightarrow \infty. \tag{12}$$

Recent developments in the estimation of the WTC can be found in papers [19–23].

The quite positive performance of most of the EVI estimators based on generalized means (GMs) leads to the consideration of a simple generalization of the H EVI estimators, in (11), studied in Brilhante et al. [24], and almost simultaneously in Paulauskas and Vaiciulis [25] and

in Beran et al. [26] (see also Segers [27]). Such a generalization leads to the so-called power mean-of-order-p (H_p) EVI estimators. Indeed, on the basis of (11), it is possible to write

$$H(k) = \sum_{i=1}^{k} \ln \left(\frac{X_{n-i+1:n}}{X_{n-k:n}} \right)^{1/k} = \ln \left(\prod_{i=1}^{k} \frac{X_{n-i+1:n}}{X_{n-k:n}} \right)^{1/k} =: \ln \left(\prod_{i=1}^{k} U_{ik} \right)^{1/k}.$$

(13)

Since the H EVI estimators are the logarithm of the geometric mean (or mean-of-order-0) of U_{ik}, $1 \le i \le k$, defined in (6), the mean-of-order-p of U_{ik}, for any real p (see Gomes and Caeiro [28] and Caeiro et al. [29], among others), can be more generally considered. This leads to the mean-of-order-p EVI estimators:

$$H_p(k) := \begin{cases} \dfrac{1 - \left((1/k) \sum_{i=1}^{k} U_{ik}^{p} \right)^{-1}}{p}, & \text{if } p < \dfrac{1}{\xi}, p \ne 0, \\ H(k), & \text{if } p = 0. \end{cases}$$

(14)

Just as mentioned above, GMs have recently been used with high success in the estimation of a positive EVI allowing one to obtain reduced bias estimators of ξ. The adequate choice of p, in (14), enables such bias reduction for the mean-of-order-p EVI estimators. Due to the specificity of the WTC, its relevance and its deep link to a positive EVI, the GMs, in (14), will now be used for the estimation of the WTC, with the consideration of

$$\widehat{\theta}_p(k) := \ln \left(\frac{n}{k} \right) H_p(k),$$

(15)

with $H_p(k)$ defined in (14), for any real p. Notice that the estimator $\widehat{\theta}^{GG}(k)$ in (10) is a particular case of $\widehat{\theta}_p(k)$. Indeed, we have $\widehat{\theta}^{GG}(k) \equiv \widehat{\theta}_0(k)$.

In Section 2 of this paper, after a few comments on the role of the WTC and some preliminary results, a few details on the asymptotic behaviour of the WTC estimators in (8), (9), and (15) are provided. Again, a high variance for small k and a high bias for large k can appear, and thus it is necessary to reduce bias and/or properly choose the tuning parameters in play. Section 3 is dedicated to an extensive Monte Carlo simulation of the WTC estimators under study. Regarding the mean-of-order-p estimation, it was always possible to find a value of p (negative, contrarily to what happens with the mean-of-order-p EVI estimation), such that, for adequate values of the threshold, there is a reduction in both bias and root mean square error (RMSE). Finally, in Section 4, a few overall conclusions are drawn. One of the main points of the article is that, as even asymptotically equivalent estimators may exhibit very diversified finite sample properties, it is always sensible to work, in practice, with a few WTC estimators, possibly dependent on tuning parameters, which make them more flexible.

2. Asymptotic Properties

2.1. Preliminary Results. To study the nondegenerate asymptotic behaviour of the estimators, a second-order condition is required to specify the bias term. This condition can be expressed in terms of the slowly varying function $L(\cdot)$ in (5). Let us assume that the rate of convergence of $\ln L(tx) - \ln L(t)$ towards 0 is ruled by a function B. Then, there exists $\beta \le 0$:

$$\frac{\ln L(tx) - \ln L(t)}{B(t)} \xrightarrow[t \to \infty]{} \frac{x^\beta - 1}{\beta},$$

(16)

and $|B| \in RV_\beta$. This second-order parameter β quantifies the rate of convergence of $\ln L(tx)/L(t)$ to 0. The closer β is to 0, the slower is the convergence.

Remark 1. In the context of EVT, the EVI and the second-order parameter ρ are null ($\xi = 0, \rho = 0$). Associated tails are then in the domain of attraction for maxima of Gumbel's law $\Lambda(x) = \exp(-\exp(-x)), x \in \mathbb{R}$, which exhibit a penultimate behaviour, looking more similar either to Weibull or Fréchet tails, according to $\theta < 1$ or $\theta > 1$, respectively. For details on penultimate behaviour, see Gomes [30, 31] and Gomes and de Haan [32], among others. Indeed, notice that $\ln U(t) = \theta \ln (\ln t) + \ln L(\ln t)$, and consequently,

$$\ln U(tx) - \ln U(t) = \theta \ln \left(\frac{\ln t + \ln x}{\ln t} \right) + \ln \frac{L(\ln t + \ln x)}{L(\ln t)}.$$

(17)

Moreover, employing the Taylor expansion to the first term,

$$\begin{aligned} \ln U(tx) - \ln U(t) &= \theta \left\{ \frac{\ln x}{\ln t} + \frac{\ln^2 x}{2 \ln^2 t} (1 + o(1)) \right\} + \ln \frac{L(\ln t + \ln x)}{L(\ln t)} \\ &= \frac{\theta}{\ln t} \left\{ \ln x + \frac{\ln^2 x}{2 \ln t} (1 + o(1)) \right\} \\ &\quad + B(\ln t) \frac{(\ln x / \ln t)^\beta - 1}{\beta} (1 + o(1)). \end{aligned}$$

(18)

Then,

$$\frac{\ln U(tx) - \ln U(t)}{\theta / \ln t} \xrightarrow[t \to \infty]{} \ln x,$$

(19)

i.e., $\xi = 0$. Moreover, there exists a slowly varying function A such that

$$\frac{((\ln U(tx) - \ln U(t))/(\theta / \ln t)) - \ln x}{A(t)} \xrightarrow[t \to \infty]{} \ln^2 x,$$

(20)

i.e., $\rho = 0$.

Next, we provide some information regarding the distributional behavior of V_{ik}, defined in (7). Suppose that $Y_{1:n}$,

$Y_{2:n}, \cdots, Y_{n:n}$ are the order statistics generated by n independent standard Pareto random variables with CDF $F_Y(y) = 1 - 1/y$, $y \geq 1$. Then, $X_{i:n} =^d U(Y_{i:n})$, $1 \leq i \leq n$ and $Y_{n-i+1:n}/Y_{n-k:n} =^d Y_{k-i+1:n}$. If k is intermediate, the following distributional representation:

$$V_{ik} \overset{d}{=} \ln U(Y_{n-i+1:n}) - \ln U(Y_{n-k:n})$$
$$\overset{d}{=} \ln U(Y_{n-k:n} Y_{k-i+1:k}) - \ln U(Y_{n-k:n}), \quad (21)$$

holds. Hence, since $E_i = \ln Y_i$ are independent, identically exponentially distributed with mean one (see [33]), it follows that for $i = 1, 2, \cdots, k$,

$$V_{ik} \overset{d}{=} \theta \ln\left(1 + \frac{E_{k-i+1:k}}{E_{n-k:n}}\right) + B(E_{n-k:n}) \frac{(1 + (E_{k-i+1:k}/E_{n-k:n}))^\beta - 1}{\beta}(1 + o_p(1)). \quad (22)$$

Results for U_{ik} can be easily deduced due to the relation $U_{ik} = e^{V_{ik}}$. Thus, we get

$$U_{ik} \overset{d}{=} 1 + \theta \frac{E_{k-i+1:k}}{E_{n-k:n}} + \frac{\theta(1-\theta)}{2}\left(\frac{E_{k-i+1:k}}{E_{n-k:n}}\right)^2$$
$$\cdot (1 + o_p(1)) + \frac{E_{k-i+1:k}}{E_{n-k:n}} B(E_{n-k:n})(1 + o_p(1)). \quad (23)$$

2.2. Asymptotic Behaviour of the Estimators. The next theorem establishes the limit distribution of $\widehat{\theta}^{GG}(k)$.

Theorem 2. *For intermediate values of k as in (12), the estimator $\widehat{\theta}^{GG}(k)$, in (10), is consistent for the estimation of θ. More than that, the distributional representation*

$$\widehat{\theta}^{GG}(k) \overset{d}{=} \theta\left\{1 + \frac{P_k}{\sqrt{k}} + \left(\frac{B(\ln(n/k))}{\theta} - \frac{1}{\ln(n/k)}\right)(1 + o_p(1))\right\} \quad (24)$$

holds, with $P_k = \sqrt{k}(\sum_{i=1}^k E_i/k - 1)$ an asymptotically standard normal RV.

Proof. Using (22) and the result $E_{n-k:n} \sim \ln(n/k) \longrightarrow \infty$, it is possible to write

$$\widehat{\theta}^{GG}(k) \overset{d}{=} \theta\left(\frac{1}{k}\sum_{i=1}^k E_{k-i+1:k} - \frac{1}{2k\ln(n/k)}\sum_{i=1}^k E_{k-i+1:k}^2\right)(1 + o_p(1))$$
$$+ \frac{1}{k}\sum_{i=1}^k E_{k-i+1:k} B(\ln(n/k))(1 + o_p(1))$$
$$= \theta\left(\frac{1}{k}\sum_{i=1}^k E_i - \frac{1}{2k\ln(n/k)}\sum_{i=1}^k E_i^2\right)(1 + o_p(1))$$
$$+ \frac{1}{k}\sum_{i=1}^k E_i B\left(\ln\left(\frac{n}{k}\right)\right)(1 + o_p(1)). \quad (25)$$

Finally, since $(1/k)\sum_{i=1}^k E_i = 1 + P_k/\sqrt{k}$ and $(1/k)\sum_{i=1}^k E_i^\alpha \longrightarrow^p \alpha$, for $\alpha = 1, 2$, equation (24) and consistency of $\widehat{\theta}^{GG}(k)$ follow straightforwardly. □

The asymptotic behaviour of the new class of mean-of-order-p WTC estimators, in (15), is next stated and proven.

Theorem 3. *Under the validity of the conditions in (4) and (16), with k being a sequence of intermediate values, as in (12), the asymptotic distributional representation*

$$\widehat{\theta}_p(k) \overset{d}{=} \theta\left\{1 + \frac{P_k}{\sqrt{k}} + \left(\frac{B(\ln(n/k))}{\theta} - \frac{1}{\ln(n/k)}\right)(1 + o_p(1))\right\} \quad (26)$$

holds for the mean-of-order-p WTC estimator, $\widehat{\theta}_p(k)$, in (15), with P_k being the standard normal RV in (24).

Proof. It is just needed to prove equation (26) for $p \neq 0$, since the case $p = 0$ was already derived in Theorem 2.

By using (23) and the result $E_{n-k:n} \sim \ln(n/k) \longrightarrow \infty$, we obtain the following distributional representation:

$$U_{ik}^p \overset{d}{=} 1 + \frac{p\theta \ln E_{k-i+1:k}}{\ln(n/k)}\left\{1 + \frac{p\theta(p\theta - 1)\ln E_{k-i+1:k}}{2\ln(n/k)}(1 + o_p(1))\right.$$
$$\left. + \left(\frac{B(\ln(n/k))}{\theta}\right)(1 + o_p(1))\right\}, \quad p \neq 0. \quad (27)$$

Consequently, under the validity of (12) and using (22) and the same results used in the proof of Theorem 2, it is possible to write

$$\frac{1}{k}\sum_{i=1}^k U_{ik}^p \overset{d}{=} 1 + \frac{p\theta}{\ln(n/k)}\left\{1 + \frac{P_k}{\sqrt{k}} + \left(\frac{p\theta - 1}{\ln(n/k)} + \frac{B(\ln(n/k))}{\theta}\right)(1 + o_p(1))\right\}, \quad p \neq 0. \quad (28)$$

Then,

$$\left(\frac{1}{k}\sum_{i=1}^k U_{ik}^p\right)^{-1} \overset{d}{=} 1 - \frac{p\theta}{\ln(n/k)}\left\{1 + \frac{P_k}{\sqrt{k}}\right.$$
$$\left. + \left(\frac{-1}{\ln(n/k)} + \frac{B(\ln(n/k))}{\theta}\right)(1 + o_p(1))\right\}, \quad p \neq 0, \quad (29)$$

and from the definition of $H_p(k)$ in (14), it follows that

$$H_p(k) = \frac{1 - \left((1/k)\sum_{i=1}^k U_{ik}^p\right)^{-1}}{p}$$
$$\overset{d}{=} \frac{\theta}{\ln(n/k)}\left\{1 + \frac{P_k}{\sqrt{k}}\right.$$
$$\left. + \left(\frac{-1}{\ln(n/k)} + \frac{B(\ln(n/k))}{\theta}\right)(1 + o_p(1))\right\}, \quad p \neq 0. \quad (30)$$

FIGURE 1: Simulated mean values (a) and RMSEs (b) of the WTC estimators under study from samples of size $n = 1000$ from a $\exp(1)$ parent ($\theta = 1$).

FIGURE 2: Simulated mean values (a) and RMSEs (b) of the WTC estimators under study from samples of size $n = 1000$ from a $\Gamma(0.75,1)$ parent ($\theta = 1$).

The result in (26) follows straightforward from the definition of $\widehat{\theta}_p(k) = \ln(n/k)H_p(k)$, in (14). □

Remark 4. Under the conditions of Theorem 3, the asymptotic distributional representation of the WTC estimators in (10) and (15) is the same. The independence on the real tuning parameter p, in (26), associated with the mean-of-order-p prevents the determination of the optimal p value, i.e., the value of p that cancels the asymptotic bias, or minimizes the RMSE of the mean-of-order-p WTC esti-

mator. However, dependence on p can appear if higher-order terms are considered in the expansion of the tail quantile function.

Corollary 5. *Under the conditions of Theorem 3 and further assuming that $\sqrt{k}B(\ln(n/k)) \longrightarrow \lambda \in \mathbb{R}$ and $\sqrt{k}/\ln(n/k) \longrightarrow 0$, with $\widehat{\theta}_p(k)$ denoting the estimators in (15), then*

$$\sqrt{k}\left(\widehat{\theta}_p(k) - \theta\right) \xrightarrow{d} \mathcal{N}\left(\lambda, \theta^2\right). \tag{31}$$
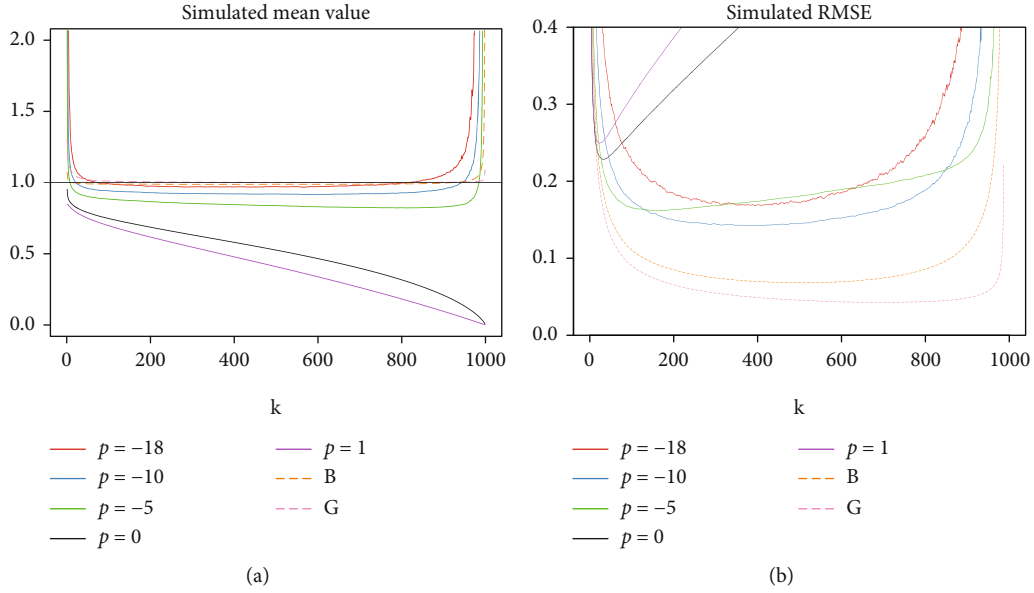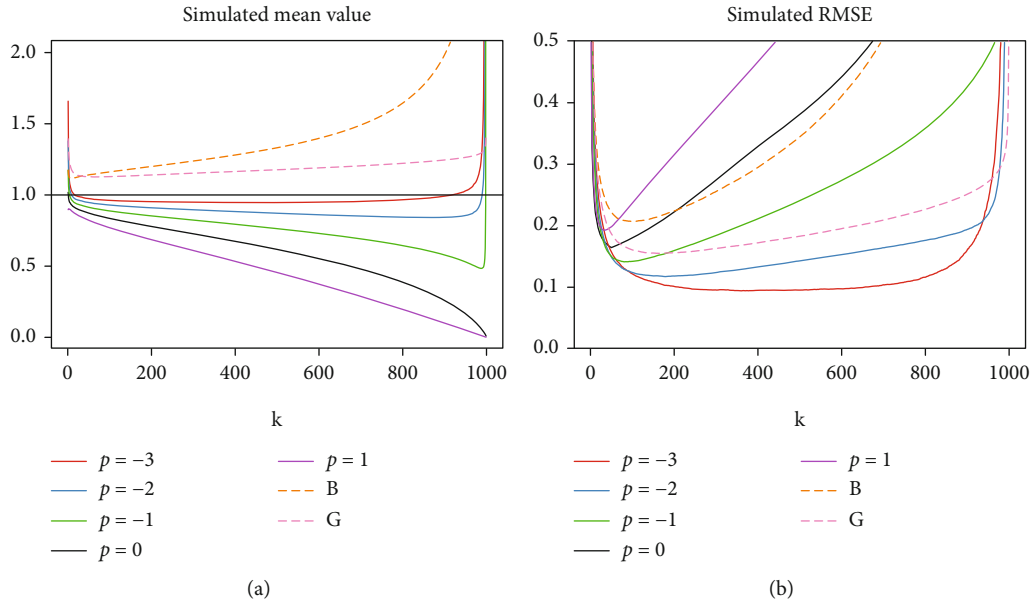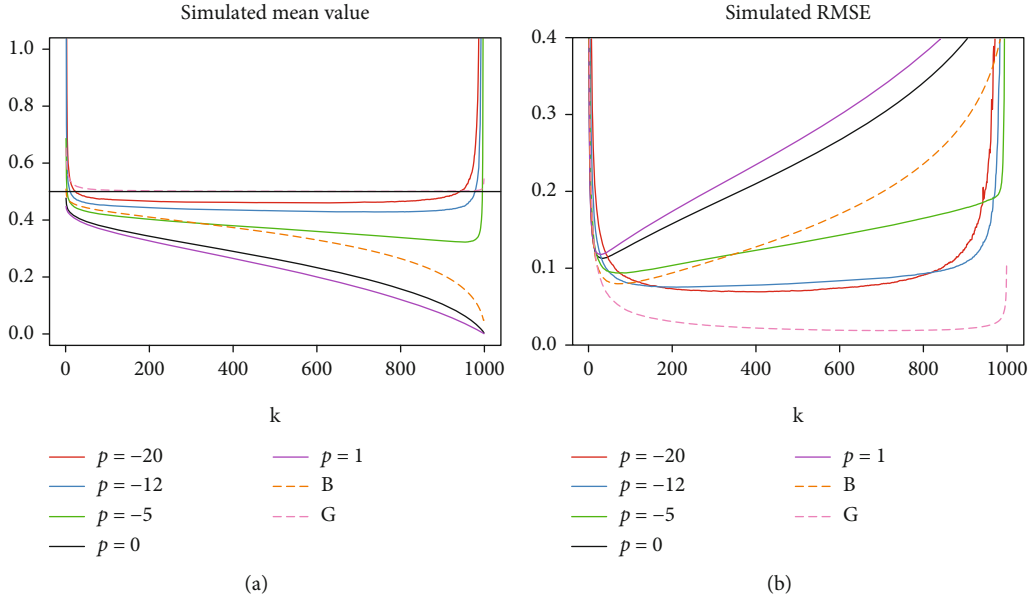
FIGURE 3: Simulated mean values (a) and RMSEs (b) of the WTC estimators under study from samples of size $n = 1000$ from a $\mathscr{W}(2,1)$ parent ($\theta = 0.5$).
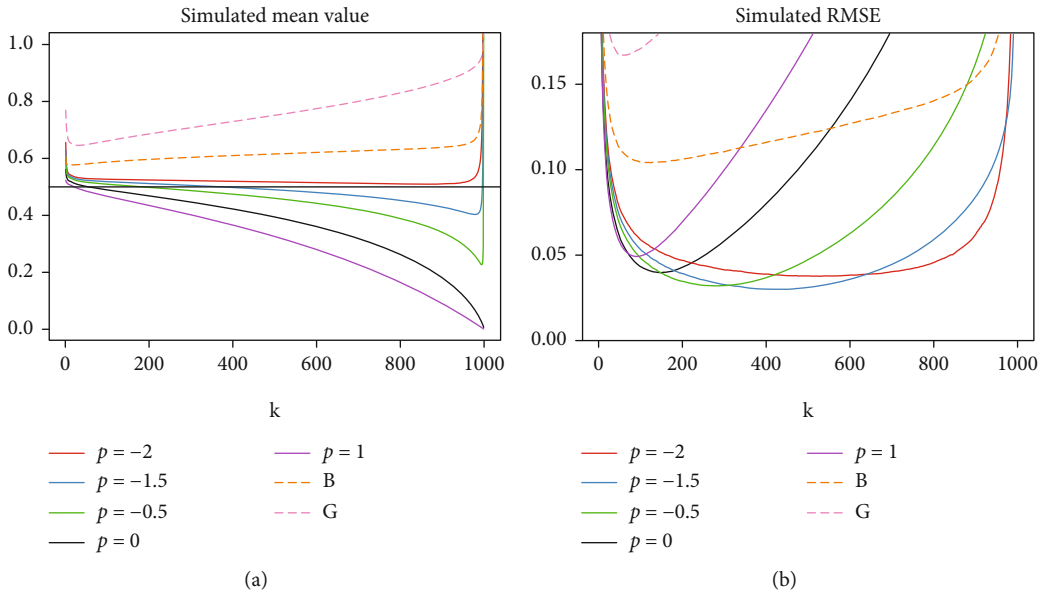
FIGURE 4: Simulated mean values (a) and RMSEs (b) of the WTC estimators under study from samples of size $n = 1000$ from a half-normal parent ($\theta = 0.5$).

Proof. From the distributional representations in (24) and (26), it is possible to write

$$\sqrt{k}\left(\widehat{\theta}_p(k) - \theta\right) \stackrel{d}{=} \theta P_k + \sqrt{k}B\left(\ln\left(\frac{n}{k}\right)\right) - \frac{\theta\sqrt{k}}{\ln(n/k)}\left(1 + o_p(1)\right).$$

(32)

Assuming that $\sqrt{k}B(\ln(n/k)) \longrightarrow \lambda \in \mathbb{R}$ and $\sqrt{k}/\ln(n/k) \longrightarrow 0$ (see [16], Proposition 2.1), with P_k being the standard normal RV in (24), denoted by $\mathcal{N}(0, 1)$, the result follows. □

Next, we state the asymptotic behaviour of the WTC estimators in (8) and (9).

Proposition 6. *Under the same conditions as in Theorem 3, with k being an intermediate sequence satisfying $\sqrt{k}B(\ln(n/k)) \longrightarrow \lambda \in \mathbb{R}$ and $\sqrt{k}/\ln(n/k) \longrightarrow 0$,*

$$\sqrt{k}\left(\widehat{\theta}^{\bullet}(k) - \theta\right) \stackrel{d}{\longrightarrow} \mathcal{N}(\lambda, \theta^2),$$

(33)

Table 1: Simulated optimal sample fraction (OSF), $\widehat{k}_0/n$, mean value, and RMSE (both computed at the optimal level) of $\widehat{\theta}_p(k)$ for $p = -18$, -10, -5, 0, and 1, $\widehat{\theta}^B(k)$, and $\widehat{\theta}^G(k)$ from exp (1) underlying parents $(\theta = 1)$.

	100	200	500	1000	2000	5000
			OSF			
$p = -18$	0.4400	0.3100	0.3180	0.4010	0.3770	0.3640
$p = -10$	0.4500	0.4100	0.3760	0.4190	0.3545	0.1760
$p = -5$	0.4500	0.3800	0.2780	0.1720	0.1060	0.0560
$p = 0$	0.1100	0.0900	0.0460	0.0320	0.0240	0.0128
$p = 1$	0.0900	0.0600	0.0360	0.0240	0.0180	0.0092
B	0.5100	0.5200	0.5380	0.4950	0.4910	0.5282
G	0.6800	0.7300	0.6600	0.6970	0.6885	0.7168
			E			
$p = -18$	1.4161	1.1002	0.9992	0.9725	0.9595	0.9534
$p = -10$	1.0909	0.9881	0.9409	0.9243	0.9185	0.9210
$p = -5$	0.9178	0.8809	0.8668	0.8749	0.8805	0.8883
$p = 0$	0.7569	0.7641	0.7967	0.8111	0.8198	0.8386
$p = 1$	0.7064	0.7369	0.7684	0.7893	0.8011	0.8252
B	1.0293	1.0146	1.0060	1.0030	1.0013	1.0005
G	**1.0134**	**1.0074**	**1.0026**	**1.0011**	**1.0002**	**1.0001**
			RMSE			
$p = -18$	1.7097	0.5313	0.2484	0.1653	0.1178	0.0831
$p = -10$	0.5842	0.3074	0.1815	0.1390	0.1151	0.0981
$p = -5$	0.3113	0.2258	0.1797	0.1587	0.1446	0.1295
$p = 0$	0.3124	0.2817	0.2488	0.2262	0.2076	0.1860
$p = 1$	0.3421	0.3100	0.2725	0.2474	0.2267	0.2023
B	0.2130	0.1464	0.0908	0.0643	0.0454	0.0286
G	**0.1231**	**0.0860**	**0.0541**	**0.0382**	**0.0270**	**0.0170**

with $\widehat{\theta}^\bullet(k)$ generally denoting the estimators $\widehat{\theta}^B(k)$ and $\widehat{\theta}^G(k)$ in (8) and (10), with $\bullet = B$ or $\bullet = G$.

Proof. For the proof of the first limit result, we refer to Theorem 3.2. of Beirlant et al. [10] with some trivial modifications. The second limit result is a particular case of Theorem 1 in Gardes and Girard [15]. □

Remark 7. Although Corollary 5 and Proposition 6 provide similar asymptotic distributions for the WTC estimators considered in this work, the same cannot be guaranteed about their finite sample performance. It is known that asymptotic equivalent estimators of the WTC can provide a different behaviour for small sample sizes (see Goegebeur et al. [16], p. 3697). Indeed, a similar comment applies to estimators of any parameter of rare events.

3. Monte Carlo Simulation Study

In this section, the finite sample performance of the class of estimators $\widehat{\theta}_p(k)$ is evaluated through a Monte Carlo simula-

tion study. For comparative purposes, the WTC estimators $\widehat{\theta}^B(k)$ and $\widehat{\theta}^G(k)$ in (8) and (9), respectively, were also included in the study. The values for the parameter p were selected from a preliminary simulation study. The value $p = 0$ was always used, since it provides the estimator in (10). The value $p = 1$ was also used to illustrate the effect of a positive value parameter. The following Weibull-type models were considered:

(1) Exponential distribution, exp (1), with CDF

$$F(x) = 1 - \exp(-x), \quad x > 0. \tag{34}$$

The WTC is $\theta = 1$.

(2) Gamma distribution, $\Gamma(a, s)$, $a, s > 0$, with density

$$f(x) = \frac{1}{s^a \Gamma(a)} x^{a-1} \exp(-x/s), \quad x > 0, \tag{35}$$

for which $\theta = 1$. Illustration is provided for $(a, s) = (0.75, 1)$.

TABLE 2: Simulated optimal sample fraction (OSF), $\widehat{k}_0/n$, mean value, and RMSE (both computed at the optimal level) of $\widehat{\theta}_p(k)$ for $p = -3$, -2, -1, 0, and 1, $\widehat{\theta}^{\mathrm{B}}(k)$, and $\widehat{\theta}^{\mathrm{G}}(k)$ from $\Gamma(0.75, 1)$ underlying parents ($\theta = 1$).

	100	200	500	1000	2000	5000
			OSF			
$p = -3$	0.4000	0.4350	0.3580	0.3600	0.3425	0.1930
$p = -2$	0.4000	0.3700	0.2340	0.1760	0.1145	0.0668
$p = -1$	0.2800	0.2050	0.1320	0.0810	0.0640	0.0388
$p = 0$	0.1700	0.1150	0.0780	0.0540	0.0390	0.0218
$p = 1$	0.1000	0.0700	0.0480	0.0340	0.0250	0.0140
B	0.2700	0.2250	0.1460	0.1030	0.0660	0.0432
G	0.4700	0.3750	0.2160	0.1620	0.1105	0.0582
			E			
$p = -3$	**1.0072**	**0.9723**	**0.9547**	**0.9478**	**0.9449**	**0.9485**
$p = -2$	0.9245	0.9052	0.9101	0.9142	0.9227	0.9319
$p = -1$	0.8550	0.8648	0.8793	0.8968	0.9013	0.9138
$p = 0$	0.8102	0.8333	0.8530	0.8696	0.8807	0.8972
$p = 1$	0.7691	0.7981	0.8247	0.8438	0.8567	0.8776
B	1.2615	1.2231	1.1827	1.1629	1.1439	1.1307
G	1.1902	1.1715	1.1464	1.1353	1.1250	1.1119
			RMSE			
$p = -3$	0.2808	0.1868	**0.1206**	**0.0942**	**0.0781**	**0.0653**
$p = -2$	0.2311	**0.1768**	0.1369	0.1173	0.1023	0.0867
$p = -1$	**0.2302**	0.1948	0.1619	0.1413	0.1248	0.1068
$p = 0$	0.2547	0.2242	0.1880	0.1651	0.1478	0.1273
$p = 1$	0.2910	0.2573	0.2180	0.1936	0.1738	0.1494
B	0.4048	0.3135	0.2416	0.2069	0.1798	0.1547
G	0.2492	0.2104	0.1748	0.1559	0.1409	0.1254

(3) Weibull distribution $\mathscr{W}(\lambda, k)$, $\lambda, k > 0$, with density

$$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} \exp\left(-(x/\lambda)^k\right), \quad x \geq 0. \tag{36}$$

Illustration is provided for $(\lambda, k) = (2, 1)$. The WTC is $\theta = 1/\lambda$.

(4) Half-normal distribution is the absolute value of a standard normal RV. For this model, the WTC is $\theta = 0.5$.

For each model, 20000 samples of size $n = 100, 200, 500, 1000, 2000,$ and 5000 were simulated. Next, for each model and sample size, n, the simulated values of $\widehat{\theta}_i(k)$, $k = 1, 2, \cdots, n - 1$, $i = 1, 2, \cdots, 20000$, provided by the i-th simulated sample were computed. Next, the Monte Carlo estimates of the mean value (E) and RMSE,

$$E\left[\widehat{\theta}(k)\right] = \sum_{i=1}^{20000} \frac{\widehat{\theta}_i(k)}{20000}, \quad \mathrm{RMSE}\left[\widehat{\theta}(k)\right]$$

$$= \sqrt{\sum_{i=1}^{20000} \frac{\left(\widehat{\theta}_i(k) - \theta\right)^2}{20000}}, \quad 1 \leq k \leq n - 1, \tag{37}$$

were obtained. In addition, the simulated optimum levels,

$$\widehat{k}_0 = \arg\min_k \mathrm{RMSE}\left[\widehat{\theta}(k)\right], E\left[\widehat{\theta}\left(\widehat{k}_0\right)\right], \mathrm{RMSE}\left[\widehat{\theta}\left(\widehat{k}_0\right)\right], \tag{38}$$

were computed.

Figures 1–4 are related to the behaviour of the aforementioned class of WTC estimators $\widehat{\theta}_p(k)$, as a function of k. At the left, the simulated values of the mean value are presented, and, at the right, the corresponding estimates of the

TABLE 3: Simulated optimal sample fraction (OSF), $\widehat{k}_0/n$, mean value, and RMSE (both computed at the optimal level) of $\widehat{\theta}_p(k)$ for $p = -20$, -12, -5, 0, and 1, $\widehat{\theta}^{\mathrm{B}}(k)$, and $\widehat{\theta}^{\mathrm{G}}(k)$ from $\mathcal{W}(2,1)$ underlying parents ($\theta = 0.5$).

	100	200	500	1000	2000	5000
			OSF			
$p = -20$	0.4500	0.4100	0.3760	0.4190	0.3545	0.1760
$p = -12$	0.4500	0.4100	0.3480	0.2110	0.1430	0.0754
$p = -5$	0.3000	0.2000	0.1040	0.0680	0.0455	0.0306
$p = 0$	0.1100	0.0900	0.0460	0.0320	0.0240	0.0128
$p = 1$	0.1000	0.0700	0.0440	0.0300	0.0190	0.0106
B	0.2400	0.1650	0.1040	0.0680	0.0550	0.0312
G	0.6800	0.7300	0.6600	0.6970	0.6885	0.7168
			E			
$p = -20$	0.5454	0.4941	0.4704	0.4622	0.4592	0.4605
$p = -12$	0.4786	0.4534	0.4413	0.4429	0.4446	0.4476
$p = -5$	0.4098	0.4116	0.4207	0.4249	0.4288	0.4322
$p = 0$	0.3785	0.3820	0.3984	0.4055	0.4099	0.4193
$p = 1$	0.3664	0.3774	0.3896	0.3987	0.4072	0.4164
B	0.4103	0.4203	0.4304	0.4379	0.4405	0.4480
G	**0.5067**	**0.5037**	**0.5013**	**0.5006**	**0.5001**	**0.5000**
			RMSE			
$p = -20$	0.2921	0.1537	0.0908	0.0695	0.0576	0.0491
$p = -12$	0.1750	0.1168	0.0867	0.0756	0.0678	0.0608
$p = -5$	0.1370	0.1198	0.1028	0.0935	0.0856	0.0770
$p = 0$	0.1562	0.1409	0.1244	0.1131	0.1038	0.0930
$p = 1$	0.1627	0.1475	0.1300	0.1182	0.1083	0.0970
B	0.1235	0.1071	0.0900	0.0800	0.0715	0.0623
G	**0.0615**	**0.0430**	**0.0271**	**0.0191**	**0.0135**	**0.0085**

RMSE are provided. The horizontal solid line, at the left plot, indicates the true WTC value. A good performance is determined by the flatness of the mean value curve close to the true value of θ and by a small RMSE in such flatness region.

Although the simulation is limited to this selected Weibull tail models, the following comments can be drawn.

(i) The estimator $\widehat{\theta}^{\mathrm{GG}}(k)$ in (10) ($p = 0$ in (15)) has always a negative bias. The bias makes the estimator very sensitive to the choice of the threshold k;

(ii) It appears that, in all the simulated cases, it was always possible to find a negative value of p that drastically reduces the absolute bias and the RMSE. This is the opposite to what typically happens with the mean-of-order-p EVI estimator, where there is a reduction of bias as well as of RMSE, for positive values of p. And for such a value of p, $\widehat{\theta}_p(k)$ strongly beats the estimator $\widehat{\theta}^{\mathrm{GG}}(k)$ considered by Gardes and Girard [17];

(iii) The estimators $\widehat{\theta}^{\mathrm{B}}(k)$ and $\widehat{\theta}^{\mathrm{G}}(k)$, in (8) and (9), beat the class of mean-of-order-p WTC estimators in terms of bias and RMSE for the exponential and Weibull parents under study. For these two parents, the best estimator was the one proposed by Girard [14];

(iv) For the gamma and half-normal parents, here considered, it is always possible to find a value of p such that the estimator $\widehat{\theta}_p(k)$ outperforms in bias and in RMSE both the estimators $\widehat{\theta}^{\mathrm{B}}(k)$ and $\widehat{\theta}^{\mathrm{G}}(k)$;

(v) Algorithmic details on the choice of tuning parameters p and k are still under development but can be easily devised, similarly to what is done in Caeiro and Gomes [34] or Gomes et al. [35].

In Tables 1–4, the simulated values of the optimal sample fraction (OSF, the optimal level divided by the sample size) of the mean value (E) and of the RMSE of the estimators under study are presented. For each model, the mean value closest to the target value θ and the smallest RMSE

Table 4: Simulated optimal sample fraction (OSF), $\widehat{k}_0/n$, mean value, and RMSE (both computed at the optimal level) of $\widehat{\theta}_p(k)$ for $p = -2$, -1.5, -1, 0, and 1, $\widehat{\theta}^{\mathrm{B}}(k)$, and $\widehat{\theta}^{\mathrm{G}}(k)$ from half-normal underlying parents ($\theta = 0.5$).

	100	200	500	1000	2000	5000
			OSF			
$p = -2$	0.4500	0.4500	0.4640	0.5260	0.5645	0.6842
$p = -1.5$	0.5300	0.4650	0.4640	0.4480	0.3920	0.3692
$p = -1$	0.4600	0.3800	0.3400	0.2810	0.2460	0.2200
$p = 0$	0.2900	0.2300	0.1900	0.1440	0.1160	0.0922
$p = 1$	0.1900	0.1550	0.1100	0.0920	0.0750	0.0532
B	0.3400	0.2650	0.1940	0.1250	0.0900	0.0460
G	0.2100	0.1600	0.0960	0.0550	0.0435	0.0210
			E			
$p = -2$	0.5367	0.5262	0.5202	0.5164	0.5144	0.5103
$p = -1.5$	**0.5029**	**0.4991**	**0.4946**	**0.4943**	**0.4975**	**0.4986**
$p = -1$	0.4778	0.4818	0.4840	0.4891	0.4922	0.4945
$p = 0$	0.4571	0.4659	0.4714	0.4803	0.4854	0.4901
$p = 1$	0.4420	0.4513	0.4643	0.4699	0.4752	0.4830
B	0.6212	0.6081	0.5976	0.5897	0.5847	0.5772
G	0.7079	0.6888	0.6657	0.6492	0.6401	0.6260
			RMSE			
$p = -2$	0.1191	0.0814	0.0512	0.0377	0.0280	0.0195
$p = -1.5$	0.0985	0.0678	**0.0419**	**0.0300**	**0.0215**	**0.0137**
$p = -1$	0.0878	**0.0637**	0.0430	0.0320	0.0237	0.0157
$p = 0$	**0.0873**	0.0694	0.0507	0.0398	0.0311	0.0220
$p = 1$	0.0961	0.0792	0.0605	0.0489	0.0396	0.0295
B	0.1732	0.1410	0.1160	0.1041	0.0948	0.0856
G	0.2478	0.2165	0.1852	0.1670	0.1521	0.1368

are written in bold. Observe that to reach the smallest absolute bias or the smallest RMSE, it is necessary to use a larger sample fraction than the one required by $\widehat{\theta}^{\mathrm{GG}}(k)$ ($p = 0$). The smallest absolute bias and RMSE are always achieved by $\widehat{\theta}_p(k)$ with $p < 0$, for the gamma and half-normal models. Also, the optimal p decreases, as the sample size n increases. For large sample sizes, the choices $p = -10$, -3, -20, and -1.5 seem to provide an overall good performance for the exponential, gamma, Weibull, and half-normal models, respectively. For the exponential and Weibull models, the smallest absolute bias and RMSE are always achieved by $\widehat{\theta}^{\mathrm{G}}(k)$.

4. Conclusions

In this paper, the estimation of the WTC, a parameter of high interest when working with Weibull-type models, is the main topic under discussion. Due to the similarity between the WTC estimation and the EVI estimation and the good performance of the EVI estimators based on GMs, a new class of WTC estimators was introduced based on the power mean-of-order-p. The consistency and asymptotic normality of the new class of estimators were obtained under adequate conditions. The finite sample behaviour of the estimators was evaluated through a Monte Carlo simulation study applied to some selected Weibull-type models. The dependence on the tuning parameter p makes the new class highly flexible when compared to the classical WTC estimators $\widehat{\theta}^{\mathrm{B}}(k)$, $\widehat{\theta}^{\mathrm{G}}(k)$, and $\widehat{\theta}^{\mathrm{GG}}(k)$ available in the literature. For the new class of WTC estimators, it is always possible to find a negative value of the tuning parameter p that enables a sharp bias and RMSE reduction for the gamma and half-normal models. For the exponential and Weibull models, the WTC estimators $\widehat{\theta}^{\mathrm{B}}(k)$ and $\widehat{\theta}^{\mathrm{G}}(k)$ outperform the new class of WTC estimators proposed in this paper, with the estimator $\widehat{\theta}^{\mathrm{G}}(k)$ being the one providing the smallest bias and RMSE. A possible improvement to $\widehat{\theta}^{\mathrm{G}}(k)$ could be achieved if we replace $H(k)$ in (10) by $H_p(k)$ in (14). This topic should be addressed in a future work. Anyway and looking at the simulated values, it is possible that a choice of p different from the ones considered in the Monte Carlo

simulations would provide the smallest bias in most situations. Algorithms for the choice of the tuning parameters p and k are under development and out of the scope of this paper.

Data Availability

The simulated data used in this study is available from the authors upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] B. Gnedenko, "Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire," *The Annals of Mathematics*, vol. 44, no. 3, p. 423, 1943.

[2] J. Beirlant, F. Caeiro, and M. I. Gomes, "An overview and open research topics in statistics of univariate extremes," *Revstat–Statistical Journal*, vol. 10, no. 1, pp. 1–31, 2012.

[3] C. Scarrott and A. MacDonald, "A review of extreme value threshold estimation and uncertainty quantification," *Revstat–Statistical Journal*, vol. 10, no. 1, pp. 33–60, 2012.

[4] M. I. Gomes and A. Guillou, "Extreme value theory and statistics of univariate extremes: a review," *International Statistical Review*, vol. 83, no. 2, pp. 263–292, 2015.

[5] A. Davison and R. Huser, "Statistics of extremes," *Annual Review of Statistics and Its Application*, vol. 2, no. 1, pp. 203–235, 2015.

[6] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular Variation. 27*, Cambridge University Press, 1989.

[7] T. de Wet, Y. Goegebeur, A. Guillou, and M. Osmann, "Kernel regression with Weibull-type tails," *Annals of the Institute of Statistical Mathematics*, vol. 68, no. 5, pp. 1135–1162, 2016.

[8] M. Berred, "Record values and the estimation of the Weibull tail-coefficient," *Comptes rendus de l'Académie des sciences. Série 1 Mathématiques*, vol. 312, no. 12, pp. 943–946, 1991.

[9] M. Broniatowski, "On the estimation of the Weibull tail coefficient," *Journal of Statistical Planning and Inference*, vol. 35, no. 3, pp. 349–365, 1993.

[10] J. Beirlant, M. Broniatowski, J. L. Teugels, and P. Vynckier, "The mean residual life function at great age: applications to tail estimation," *Journal of Statistical Planning and Inference*, vol. 45, no. 1-2, pp. 21–48, 1995.

[11] J. Beirlant, C. Bouquiaux, and B. J. Werker, "Semiparametric lower bounds for tail index estimation," *Journal of Statistical Planning and Inference*, vol. 136, no. 3, pp. 705–729, 2006.

[12] G. Dierckx, J. Beirlant, D. de Waal, and A. Guillou, "A new estimation method for Weibull-type tails based on the mean excess function," *Journal of Statistical Planning and Inference*, vol. 139, no. 6, pp. 1905–1920, 2009.

[13] J. Beirlant, J. L. Teugels, and P. Vynckier, *Practical Analysis of Extreme Values*, Leuven University Press, 1996.

[14] S. Girard, "A Hill type estimator of the Weibull tail-coefficient," *Communications in Statistics-Theory and Methods*, vol. 33, no. 2, pp. 205–234, 2004.

[15] L. Gardes and S. Girard, "Estimation of the Weibull tail-coefficient with linear combination of upper order statistics," *Journal of Statistical Planning and Inference*, vol. 138, no. 5, pp. 1416–1427, 2008.

[16] Y. Goegebeur, J. Beirlant, and T. de Wet, "Generalized kernel estimators for the Weibull-tail coefficient," *Communications in Statistics-Theory and Methods*, vol. 39, no. 20, pp. 3695–3716, 2010.

[17] L. Gardes and S. Girard, "Comparison of Weibull tail-coefficient estimators," *Revstat–Statistical Journal*, vol. 4, no. 2, pp. 163–188, 2006.

[18] B. M. Hill, "A simple general approach to inference about the tail of a distribution," *The Annals of Statistics*, vol. 3, no. 5, pp. 1163–1174, 1975.

[19] L. Gardes and S. Girard, "On the estimation of the functional Weibull tail-coefficient," *Journal of Multivariate Analysis*, vol. 146, pp. 29–45, 2016.

[20] T. A. Kpanzou, K. M. Gamado, and H. Hounnon, "A Beran-inspired estimator for theWeibull-type tail coefficient," *Journal of Statistical Theory and Practice*, vol. 13, no. 1, 2018.

[21] C. Gong and C. Ling, "Robust estimations for the tail index of Weibull-type distribution," *Risks*, vol. 6, no. 4, p. 119, 2018.

[22] J. Worms and R. Worms, "Estimation of extremes for Weibull-tail distributions in the presence of random censoring," *Extremes*, vol. 22, no. 4, pp. 667–704, 2019.

[23] F. Caeiro, M. I. Gomes, and L. Henriques-Rodrigues, "Estimation of the Weibull Tail Coefficient through the Power Mean-of-Order-p," in *Recent Developments in Statistics and Data Science*, R. Bispo, L. Henriques-Rodrigues, R. Alpizar-Jara, and M. De Carvalho, Eds., In press, 2022.

[24] M. F. Brilhante, M. I. Gomes, and D. Pestana, "A simple generalisation of the Hill estimator," *Computational Statistics & Data Analysis*, vol. 57, no. 1, pp. 518–535, 2013.

[25] V. Paulauskas and M. Vaiciulis, "On an improvement of Hill and some other estimators," *Lithuanian Mathematical Journal*, vol. 53, no. 3, pp. 336–355, 2013.

[26] J. Beran, D. Schell, and M. Stehlík, "The harmonic moment tail index estimator: asymptotic distribution and robustness," *Annals of the Institute of Statistical Mathematics*, vol. 66, no. 1, pp. 193–220, 2014.

[27] J. Segers, "Residual estimators," *Journal of Statistical Planning and Inference*, vol. 98, no. 1-2, pp. 15–27, 2001.

[28] M. I. Gomes and F. Caeiro, "Efficiency of partially reduced-bias mean-of-order-p versus minimum-variance reduced-bias extreme value index estimation," in *Proceedings of COMPSTAT 2014 International Association for Statistical Computing*, pp. 289–298, Geneve, Switzerland, 2014.

[29] F. Caeiro, M. I. Gomes, J. Beirlant, and T. de Wet, "Mean-of-order p reduced-bias extreme value index estimation under a third-order framework," *Extremes*, vol. 19, no. 4, pp. 561–589, 2016.

[30] M. I. Gomes, *Some Probabilistic and Statistical Problems in Extreme Value Theory, [Ph.D. thesis]*, University of Sheffield, 1978.

[31] M. I. Gomes, "Penultimate behaviour of the extremes," in *Extreme Value Theory and Applications*, J. Galambos, J. Lechner, and E. Simiu, Eds., pp. 403–418, Springer, 1994.

[32] M. I. Gomes and L. de Haan, "Approximation by penultimate extreme value distributions," *Extremes*, vol. 2, no. 1, pp. 71–85, 1999.

[33] R. C. Gupta, "The order statistics of exponential, power function and pareto distributions and some applications," *Series Statistics*, vol. 10, no. 4, pp. 551–554, 1979.

[34] F. Caeiro and M. I. Gomes, "Threshold selection in extreme value analysis," in *Extreme Value Modeling and Risk Analysis: Methods and Applications*, D. Dey and J. Yan, Eds., pp. 71–89, Chapman-Hall/CRC, Boca Raton, FL, USA, 2015.

[35] M. I. Gomes, F. Caeiro, L. Henriques-Rodrigues, and B. Manjunath, "Bootstrap methods in statistics of extremes," in *Extreme Events in Finance: A Handbook of Extreme Value Theory and its Applications*, F. Longin, Ed., vol. 6, pp. 117–138, John Wiley & Sons, Hoboken, NJ, USA, 2016.