



HAL
open science

Estimation of the conditional tail moment for Weibull-type distributions

Yuri Goegebeur, Armelle Guillou, Jing Qin

► **To cite this version:**

Yuri Goegebeur, Armelle Guillou, Jing Qin. Estimation of the conditional tail moment for Weibull-type distributions. 2024. hal-04050303v4

HAL Id: hal-04050303

<https://hal.science/hal-04050303v4>

Preprint submitted on 21 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimation of the conditional tail moment for Weibull-type distributions

Yuri Goegebeur⁽¹⁾, Armelle Guillou⁽²⁾, Jing Qin⁽¹⁾

⁽¹⁾ University of Southern Denmark

⁽²⁾ Université de Strasbourg et CNRS

Abstract

We consider the estimation of the conditional tail moment at extreme levels for the class of Weibull-type distributions. A two-step procedure is introduced where in the first stage one estimates the conditional tail moment at an intermediate level, followed by an extrapolation in the second stage. The asymptotic properties of the estimators introduced in the two stages are derived under suitable assumptions. The finite sample properties of the proposed estimator are examined with a simulation experiment. We conclude with two applications on real life data: wind speed measurements collected at an offshore wind farm and PM_{2.5} air pollution data.

Keywords: Empirical process, extreme-value statistics, hazard rate function, Weibull-type distribution.

1 Introduction

Quantifying and estimating the risk of extreme events is of interest in many scientific disciplines. Floods due to sea storms can entail massive losses, both in terms of human lives and economical losses, so the modelling of the distribution of wave heights and other sea state parameters is of crucial importance in the design of coastal protection. In environmental science, extreme levels of pollutants can have adverse effect on human health. Over the years several risk measures have been introduced, the most popular of these being the Value-at-Risk (also called return

level) and the conditional tail expectation. We refer to, e.g., Jorion (2007), Singh et al. (2007), Coles (2013), Cooley (2013) and Salas and Obeyesekera (2014). In this paper we consider the conditional tail moment (*CTM*), being a natural generalisation of the conditional tail expectation, and study its estimation for the class of Weibull-type distributions. Note that moment estimation is of general interest in statistics. Although we are often primarily interested in the first and second moments for practical reasons, in some applications one may also be interested in the conditional tail skewness or kurtosis, see, e.g., Hong and Elshahat (2010) or Eini and Khaloozadeh (2021) for such examples, and more generally El Methni et al. (2014).

Let $\beta > 0$ and $p \in (0, 1)$. We want to estimate

$$CTM_{\beta,p} := \mathbb{E} \left[X^\beta \mid X > U_X(1/p) \right],$$

when the distribution of X is of Weibull-type, i.e., the distribution function satisfies

$$F_X(x) = 1 - e^{-x^{1/\theta} \ell_X(x)}, x > 0, \tag{1}$$

with θ a positive parameter and ℓ_X a slowly varying function at infinity, i.e., a positive measurable function such that

$$\lim_{t \rightarrow \infty} \frac{\ell_X(tx)}{\ell_X(t)} = 1 \quad \text{for all } x > 0,$$

and U_X the tail quantile function of X defined as $U_X(x) = \inf\{z : F_X(z) \geq 1 - 1/x\}, x > 1$. This risk measure $CTM_{\beta,p}$ uses the information in the upper tail, i.e., above the high quantile $U_X(1/p)$. Specific interest is here in the extreme case, that is, the situation where p is very small, typically less than $1/n$, where n is the size of the random sample that is available for estimation. In this case, estimation is challenging as the conditional expectation in the definition of $CTM_{\beta,p}$ cannot simply be estimated by an empirical average. Indeed, for $p < 1/n$ the natural estimator for $U_X(1/p)$ is the largest observation in the sample, but then there are no further data to base the estimation of the conditional expectation upon. The role of the power parameter β is, e.g., to allow the computation of both the mean and the second moment and thus the variance of X given that X is extreme (above its quantile). Note that when $\beta = 1$, we recover the usual

conditional tail expectation.

The considered class of Weibull-type distributions forms a rich subclass of the Gumbel max-domain of attraction. The tail heaviness of Weibull-type distributions is clearly governed by the parameter θ , where larger values of θ correspond with a slower decay of the upper tail. Popular members of this family include the Weibull (corresponding to a constant function ℓ_X), normal, gamma and logistic distributions, to mention only a few. Loosely speaking, the upper tail decays exponentially fast, in contrast to the Pareto-type models which have a tail decay corresponding to a decreasing power function, and thus are heavier tailed. The Weibull-type distributions find important applications in areas like environmental science, climate science, hydrology, engineering and medicine. As the Weibull-type model depends only explicitly on the parameter θ , it is not unexpected that the estimation of this parameter received a lot of attention in the extreme-value literature. Girard (2004) introduced a Hill-type estimator for θ ; see also Gardes and Girard (2005) for the related problem of extreme quantile estimation. Generalisations to estimators based on weighted sums (of spacings) of upper order statistics were considered in Broniatowski (1993), Gardes and Girard (2008) and Goegebeur et al. (2010). Beirlant et al. (1995), Dierckx et al. (2009) and Goegebeur and Guillou (2011) developed alternative estimation methods for θ based on the mean excess function, where the latter two papers highlighted the improved performance with respect to bias compared to the estimators based on properties of upper order statistics; see also Diebolt et al. (2008) where a dedicated bias-reducing estimation procedure was developed. Further extensions where θ is estimated in presence of random covariates can be found in Goegebeur et al. (2015), Gardes and Girard (2016) and de Wet et al. (2016).

The estimation of the conditional tail moment is already studied for the context of Pareto-type distributions in Goegebeur et al. (2022), who consider the classical setup of independent identically distributed (i.i.d.) random variables. Goegebeur et al. (2024) studied this in an i.i.d. setting with random right censoring. El Methni et al. (2014) studied the estimation of the conditional tail moment in a regression setup where the conditional distribution of the response variable given the covariate is of Pareto-type. Estimation of the conditional tail moment for the

class of Weibull-type distributions is to the best of our knowledge unexplored.

We organise this paper as follows. In Section 2 we provide as a first result an approximation of $CTM_{\beta,p}$ for $p \downarrow 0$, which will be used to derive the estimator for this risk measure under extrapolation. The proposed estimator depends on an estimator for the conditional tail moment in the intermediate case, i.e., for $p = k/n$, where $k \rightarrow \infty$ as the sample size $n \rightarrow \infty$ but in such a way that $k/n \rightarrow 0$. This intermediate case is studied in Section 3 where we introduce an estimator and derive its limiting distribution under suitable assumptions. Section 4 is devoted to the study of the asymptotic properties of the estimator for $CTM_{\beta,p}$ in the extreme case. The finite sample properties are examined by a simulation experiment in Section 5 while the practical applicability of the proposed method is illustrated in Section 6 on wind speed and air pollution data. Concluding remarks are given in Section 7. Section 8 contains the proofs of the results. Some additional simulation results are provided in an online Supporting Information.

2 Construction of the estimator

Let RV_ψ denote the class of regularly varying functions at infinity with index $\psi \in \mathbb{R}$, i.e., positive measurable functions f satisfying $f(tx)/f(t) \rightarrow x^\psi$, as $t \rightarrow \infty$, for all $x > 0$. Denote $H(x) = -\log \bar{F}_X(x)$, and hence $\bar{F}_X(x) = e^{-H(x)}$. In the case where F_X is of Weibull-type we have then $H \in RV_{1/\theta}$. Note that if H is differentiable with derivative H' then $H' = h$ where h is the hazard rate function.

We start to expand $CTM_{\beta,p}$ in case $p \downarrow 0$ in order to have an idea about how to estimate this risk measure under extrapolation, i.e., for $p < 1/n$, where n is the size of the sample on which the estimation is based.

Lemma 2.1 *Assume F_X satisfies (1), H is differentiable with derivative $h \in RV_{1/\theta-1}$. Then, as $p \downarrow 0$, we have*

$$\frac{CTM_{\beta,p}}{\left[U_X\left(\frac{1}{p}\right)\right]^\beta} \rightarrow 1.$$

Then, we propose an estimator for $CTM_{\beta,p}$ based on a two-step procedure.

Let k be an intermediate sequence, i.e., a sequence such that $k \rightarrow \infty$ and $k/n \rightarrow 0$ as $n \rightarrow \infty$.

Lemma 2.1 yields the approximation

$$\frac{CTM_{\beta,p}/\left[U_X\left(\frac{1}{p}\right)\right]^\beta}{CTM_{\beta,k/n}/\left[U_X\left(\frac{n}{k}\right)\right]^\beta} \sim 1,$$

from which we deduce that

$$CTM_{\beta,p} \sim \left(\frac{U_X\left(\frac{1}{p}\right)}{U_X\left(\frac{n}{k}\right)}\right)^\beta CTM_{\beta,k/n}.$$

Note that (1) is equivalent to assume that U_X satisfies

$$U_X(x) = (\log x)^\theta \ell_U(\log x),$$

where ℓ_U is also a slowly varying function at infinity. Thus, we deduce the following estimator for $CTM_{\beta,p}$

$$\widehat{CTM}_{\beta,p} := \left(\frac{\log(1/p)}{\log(n/k)}\right)^{\beta \hat{\theta}_k} \overline{CTM}_{\beta,k/n}, \quad (2)$$

where $\hat{\theta}_k$ is an estimator of θ and $\overline{CTM}_{\beta,k/n}$ an intermediate estimator of $CTM_{\beta,k/n}$.

3 Asymptotic properties of $\overline{CTM}_{\beta,k/n}$

Assume we have at our disposal a sample X_1, \dots, X_n of independent copies of the random variable X and denote by $X_{1,n} \leq \dots \leq X_{n,n}$ the order statistics. We start with considering the estimation of the conditional tail moment in the intermediate case. In this situation, the natural estimator for $U_X(n/k)$ is then $X_{n-k,n}$, which is within the data range, and hence $CTM_{\beta,k/n}$ can be estimated empirically by the sample mean of the X_i^β for which $X_i > X_{n-k,n}$. We define

$$\overline{CTM}_{\beta,k/n} = \frac{1}{k} \sum_{i=1}^n X_i^\beta \mathbb{1}_{\{X_i > X_{n-k,n}\}},$$

and introduce

$$\overline{CTM}_{\beta,k/n}(x) := \frac{1}{k} \sum_{i=1}^n X_i^\beta \mathbb{1}_{\{X_i > U_X\left(\frac{n}{k}\right) + a_X\left(\frac{n}{k}\right)x\}},$$

where

$$a_X(t) := \frac{1}{h(U_X(t))}.$$

Then, $\overline{CTM}_{\beta,k/n}$ can be rewritten as

$$\overline{CTM}_{\beta,k/n} = \overline{CTM}_{\beta,k/n}(\hat{u}_{n,k}),$$

where $\hat{u}_{n,k} := (X_{n-k,n} - U_X(n/k))/a_X(n/k)$.

The weak convergence of $\overline{CTM}_{\beta,k/n}$, after normalisation, is given in the following theorem.

Theorem 3.1 *Assume F_X satisfies (1), H is differentiable with derivative h which is ultimately monotone, and h is differentiable with derivative h' satisfying $xh'(x)/h(x) \rightarrow 1/\theta - 1$ as $x \rightarrow \infty$.*

Then, if $k \rightarrow \infty$ as $n \rightarrow \infty$ such that $k/n \rightarrow 0$ and $\sqrt{k}a_X(n/k)/U_X(n/k) \rightarrow 0$, we have

$$\sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} \left(\frac{\overline{CTM}_{\beta,k/n}}{CTM_{\beta,k/n}} - 1 \right) \xrightarrow{d} \Theta := \beta \int_0^\infty W(e^{-v}) dv,$$

where $W(z)$ is a zero centered Gaussian process with covariance function

$$\mathbb{E}\{W(z)W(\bar{z})\} = \min(z, \bar{z}).$$

Note that this theorem implies the convergence of the rescaled estimator of $CTM_{\beta,k/n}$ towards the $\mathcal{N}(0, 2\beta^2)$ distribution. As is clear, the limiting distribution in Theorem 3.1 does not depend on θ . At first sight, this can appear unexpected, but this can be explained by the fact that the first order expansion of the conditional tail moment (Lemma 2.1), as well as the limiting distribution of the intermediate quantile (see Lemma 8.2 in Section 8) are independent on θ in case of Weibull-type distributions. Note that this was also observed in Goegebeur et al. (2023) for the intermediate estimator of the marginal excess moments for Weibull-type distributions. Note also that under the assumption

$$\lim_{x \rightarrow \infty} x \frac{\ell'_X(x)}{\ell_X(x)} = 0, \tag{3}$$

our condition $\sqrt{k}a_X(n/k)/U_X(n/k) \rightarrow 0$ is equivalent to $\sqrt{k}/\log(n/k) \rightarrow 0$, a condition independent on the value of θ . The convergence (3) is not a restrictive condition since it is automatically

satisfied for a normalized slowly varying function ℓ_X , i.e., a function of the form

$$\ell_X(x) = c \exp \left\{ \int_a^x \frac{\varepsilon(u)}{u} du \right\}, x \geq a,$$

with c a positive constant and $\varepsilon(x) \rightarrow 0$, see, e.g., Bingham et al. (1987, page 15).

4 Asymptotic properties of $\widehat{CTM}_{\beta,p}$

In this section we study the asymptotic properties of the estimator $\widehat{CTM}_{\beta,p}$ given in (2). As estimator for the Weibull-tail coefficient θ we use the estimator proposed by Girard (2004), given by

$$\widehat{\theta}_k := \frac{\frac{1}{k} \sum_{i=1}^k (\log X_{n-i+1,n} - \log X_{n-k,n})}{\frac{1}{k} \sum_{i=1}^k (\log \log \frac{n}{i} - \log \log \frac{n}{k})}.$$

Below, we recall the limiting distribution of this estimator, properly normalised, in terms of a stochastic process, as established in Theorem 4.1 in Goegebeur et al. (2023).

Lemma 4.1 *Assume F_X satisfies (1), H is differentiable with derivative h which is ultimately monotone, and h is differentiable with derivative h' satisfying $xh'(x)/h(x) \rightarrow 1/\theta - 1$ as $x \rightarrow \infty$. Then, if $k \rightarrow \infty$ as $n \rightarrow \infty$ such that $\sqrt{k}a_X(n/k)/U_X(n/k) \rightarrow 0$ and $\sqrt{k} \frac{U_X(n/k)\ell'_X(U_X(n/k))}{\ell_X(U_X(n/k))} \rightarrow 0$, we have*

$$\sqrt{k} \left(\widehat{\theta}_k - \theta \right) \xrightarrow{d} \Gamma := \theta \left\{ \int_0^\infty W(e^{-v}) dv - W(1) \right\}.$$

Note that the limiting distribution in Lemma 4.1 is $\mathcal{N}(0, \theta^2)$. As is clear from Theorem 3.1 and Lemma 4.1, the limiting behavior of $\widehat{CTM}_{\beta,k/n}$ and $\widehat{\theta}_k$ involves the same stochastic process W . For more details we refer to Section 8 of the present paper and the proof of Theorem 4.1 in Goegebeur et al. (2023).

We can now state the main result of the paper, namely the weak convergence of $\widehat{CTM}_{\beta,p}$, properly normalised.

Theorem 4.1 *Assume F_X satisfies (1), H is differentiable with derivative h which is ultimately monotone, and h is differentiable with derivative h' satisfying $xh'(x)/h(x) \rightarrow 1/\theta - 1$ as $x \rightarrow \infty$. Then, if $k \rightarrow \infty$ as $n \rightarrow \infty$ such that $\sqrt{k} \frac{a_X(n/k)}{U_X(n/k)} \rightarrow 0$ and $\sqrt{k} \sup_{x \geq U_X(n/k)} \frac{x|\ell'_X(x)|}{\ell_X(x)} \rightarrow 0$, we have for p satisfying $p \leq \frac{k}{n}$ such that $\frac{\log 1/p}{\log n/k} \rightarrow \lambda \in (1, \infty)$ that*

$$\sqrt{k} \left(\frac{\widehat{CTM}_{\beta,p}}{CTM_{\beta,p}} - 1 \right) \xrightarrow{d} \beta \log(\lambda) \Gamma.$$

From the result of Theorem 4.1, it is clear that the estimator $\widehat{CTM}_{\beta,p}$, after normalisation, inherits its limiting distribution, up to the factor $\beta \log(\lambda)$, from the estimator $\widehat{\theta}_k$ for the Weibull-tail coefficient. Indeed, as the proof reveals

$$\sqrt{k} \left(\frac{\widehat{CTM}_{\beta,p}}{CTM_{\beta,p}} - 1 \right) = \beta \log(\lambda) \sqrt{k} (\widehat{\theta}_k - \theta) + o_{\mathbb{P}}(1),$$

for $n \rightarrow \infty$. In particular, we obtain the $\mathcal{N}(0, (\theta\beta \log(\lambda))^2)$ distribution as limiting distribution for the rescaled estimator of $CTM_{\beta,p}$. As expected, more severe extrapolations, corresponding with larger values for λ , lead to an increase in the variance of the limiting distribution.

Note that an alternative estimator for $CTM_{\beta,p}$ could have been

$$\widetilde{CTM}_{\beta,p} := \left(\frac{\log(1/p)}{\log(n/k)} \right)^{\beta \widehat{\theta}_k} X_{n-k,n}^{\beta},$$

i.e., an estimator of the form (2) with the conditional tail moment at an intermediate level estimated by $X_{n-k,n}^{\beta}$ (according to our Lemma 2.1). However, this method has two main drawbacks: first, the estimator $X_{n-k,n}^{\beta}$ clearly underestimates $CTM_{\beta,k/n}$ by definition, and also, since

$$\widetilde{CTM}_{\beta,p} = \left[\widetilde{CTM}_{1,p} \right]^{\beta},$$

it does not allow to estimate the conditional tail variance $V_p := CTM_{2,p} - CTM_{1,p}^2$ (or other higher order central moments), since it leads to a trivial zero estimate.

Below we give some examples of distributions satisfying the assumptions of Theorem 4.1, in

particular the conditions

$$\lim_{x \rightarrow \infty} x \frac{h'(x)}{h(x)} = \frac{1}{\theta} - 1, \quad (4)$$

$$\lim_{n \rightarrow \infty} \sqrt{k} \frac{a_X(n/k)}{U_X(n/k)} = 0, \quad (5)$$

$$\lim_{n \rightarrow \infty} \sqrt{k} \sup_{x \geq U_X(n/k)} x \frac{|\ell'_X(x)|}{\ell_X(x)} = 0. \quad (6)$$

The distributions are:

- Weibull distribution (*Weibull*(θ), $\theta > 0$): We have $\bar{F}_X(x) = e^{-x^{1/\theta}}$ and thus $H(x) = x^{1/\theta}$. Clearly H is differentiable with derivative $h(x) = \frac{1}{\theta} x^{\frac{1}{\theta}-1}$ which is monotone and which satisfies (4). Also since $\ell_X(x) = 1$, (6) is trivially satisfied. Since $U_X(x) = (\log x)^\theta$, (5) is equivalent to the condition $\frac{\sqrt{k}}{\log n/k} \rightarrow 0$.

- Gamma distribution ($\Gamma(\alpha, \lambda)$, $\alpha, \lambda > 0$). Denote by $\Gamma_I(a, z) := \int_z^\infty t^{a-1} e^{-t} dt$ the upper incomplete Gamma function and recall its expansion for $z \rightarrow \infty$

$$\Gamma_I(a, z) \sim z^{a-1} e^{-z} \left[1 + \frac{a-1}{z} + O\left(\frac{1}{z^2}\right) \right], \quad (7)$$

see Abramowitz and Stegun (1983, p. 263). Then we have

$$\bar{F}_X(x) = \frac{\Gamma_I(\alpha, \lambda x)}{\Gamma(\alpha)},$$

and for x large enough

$$x \frac{h'(x)}{h(x)} = x \frac{f'_X(x)}{f_X(x)} + x \frac{f_X(x)}{\bar{F}_X(x)} = O\left(\frac{1}{x}\right),$$

from which (4) follows with $\theta = 1$. In that case $h(x) = \ell_X(x) \left[1 + x \frac{\ell'_X(x)}{\ell_X(x)} \right]$, from which we deduce that

$$\begin{aligned} x \frac{\ell'_X(x)}{\ell_X(x)} &= \frac{h(x)}{\ell_X(x)} - 1 = \frac{f_X(x)}{\bar{F}_X(x) \ell_X(x)} - 1 \\ &= \frac{1 + (\alpha - 1) \frac{\log H(x)}{H(x)} + O\left(\frac{1}{H(x)}\right)}{1 + O\left(\frac{1}{x}\right)} - 1 \\ &= (\alpha - 1) \frac{\log H(x)}{H(x)} + O\left(\frac{1}{H(x)}\right) + O\left(\frac{1}{x}\right), \end{aligned}$$

using (7) combining with the fact that, according to Girard (2004):

$$\ell_X(x) = \left\{ \frac{1}{\lambda} + \frac{\alpha - 1}{\lambda} \frac{\log H(x)}{H(x)} + O\left(\frac{1}{H(x)}\right) \right\}^{-1}.$$

This implies that

$$U_X\left(\frac{n}{k}\right) \frac{\ell'_X(U_X(\frac{n}{k}))}{\ell_X(U_X(\frac{n}{k}))} = (\alpha - 1) \frac{\log \log \frac{n}{k}}{\log \frac{n}{k}} + O\left(\frac{1}{\log \frac{n}{k}}\right).$$

Since the right-hand side of the above display tends to 0, (5) is equivalent to the condition $\frac{\sqrt{k}}{\log n/k} \rightarrow 0$. Further, we have, for x large enough, that

$$x \frac{|\ell'_X(x)|}{\ell_X(x)} \leq C \frac{\log x}{x},$$

where C is some positive constant, and thus, for n large,

$$\sqrt{k} \sup_{x \geq U_X(n/k)} x \frac{|\ell'_X(x)|}{\ell_X(x)} \leq C \sqrt{k} \frac{\log \log \frac{n}{k}}{\log \frac{n}{k}},$$

and hence (6) is satisfied if $\sqrt{k} \frac{\log \log \frac{n}{k}}{\log \frac{n}{k}} \rightarrow 0$. This condition is slightly stronger than the one due to (5) and similar to the condition $\sqrt{k} b(\log(\frac{n}{k})) \rightarrow 0$ imposed by Girard (2004) for the asymptotic normality of the Weibull tail coefficient, since for this distribution $b(x) = (1 - \alpha) \frac{\log x}{x}$ (see Girard, 2004).

- Extended Weibull distribution (*EW*eibull(θ, ξ), $\theta > 1$, $\xi \in \mathbb{R}$) with survival function

$$1 - F_X(x) = r(x)e^{-x^{1/\theta}}, \quad x > 0,$$

where $r \in RV_\xi$ (Kluppelberg and Villaseñor, 1993). We consider the simplified case where $r(x) = Dx^\xi$ instead of being RV_ξ , then direct computations yield:

$$\begin{aligned} \ell_X(x) &= 1 - \xi x^{-\frac{1}{\theta}} \log x - \log D x^{-\frac{1}{\theta}}, \\ \ell'_X(x) &= \frac{\xi}{\theta} x^{-\frac{1}{\theta}-1} \log x \left\{ 1 - \frac{\theta - (\log D)/\xi}{\log x} \right\}, \\ h(x) &= \frac{1}{\theta} x^{\frac{1}{\theta}-1} \left\{ 1 - \theta \xi x^{-\frac{1}{\theta}} \right\}, \\ h'(x) &= \frac{1}{\theta} \left(\frac{1}{\theta} - 1 \right) x^{\frac{1}{\theta}-2} \left\{ 1 + \frac{\theta^2 \xi}{1 - \theta} x^{-\frac{1}{\theta}} \right\}, \end{aligned}$$

from which we deduce that (4) is satisfied, (5) is equivalent to the condition $\frac{\sqrt{k}}{\log n/k} \rightarrow 0$. Concerning (6) we have, for x large enough,

$$x \frac{|\ell'_X(x)|}{\ell_X(x)} \leq C \frac{\log x}{x^{1/\theta}},$$

where C is a positive constant and thus, for n large enough,

$$\sqrt{k} \sup_{x \geq U_X(n/k)} x \frac{|\ell'_X(x)|}{\ell_X(x)} \leq C \sqrt{k} \frac{\log \log \frac{n}{k}}{\log \frac{n}{k}}.$$

Hence, (6) is satisfied if $\sqrt{k} \frac{\log \log \frac{n}{k}}{\log \frac{n}{k}} \rightarrow 0$. This is similar to the condition $\sqrt{k} b(\log(\frac{n}{k})) \rightarrow 0$ imposed by Girard (2004) since $b(x) = -\theta^2 \xi \frac{\log x}{x}$ for this distribution (see Girard, 2004).

- Absolute normal distribution ($|\mathcal{N}(\mu, \sigma^2)|$, $\mu \in \mathbb{R}$, $\sigma > 0$), i.e., the distribution of $|X|$ where $X \sim \mathcal{N}(\mu, \sigma^2)$. We verify for brevity the case $|\mathcal{N}(0, 1)|$. Let X be a random variable from a $\mathcal{N}(0, 1)$ distribution, with density f_X and survival function \bar{F}_X . We have $\bar{F}_{|X|}(x) = 2\bar{F}_X(x)$, from which we deduce that

$$\begin{aligned} h(x) &= \frac{f_X(x)}{\bar{F}_X(x)}, \\ x \frac{h'(x)}{h(x)} &= x \frac{f'_X(x)}{f_X(x)} + x \frac{f_X(x)}{\bar{F}_X(x)} = 1 + O\left(\frac{1}{x^2}\right), \end{aligned} \quad (8)$$

since according to Abramowitz and Stegun (1983, p. 932), we have the approximation

$$\bar{F}_X(x) = \frac{f_X(x)}{x} \left\{ 1 - \frac{1}{x^2} + O\left(\frac{1}{x^4}\right) \right\}.$$

The right-hand side of (8) tends to 1 as x goes to infinity, which means that (4) is satisfied with $\theta = 1/2$. As for condition (5) note that

$$x \frac{\ell'_{|X|}(x)}{\ell_{|X|}(x)} = \frac{xh(x)}{H(x)} - 2,$$

which gives for x large enough that

$$x \frac{|\ell'_{|X|}(x)|}{\ell_{|X|}(x)} \leq C \frac{\log x}{x^2},$$

for some positive constant C . Hence, (5) is equivalent to the condition $\frac{\sqrt{k}}{\log n/k} \rightarrow 0$, and (6) is satisfied if $\sqrt{k} \frac{\log \log \frac{n}{k}}{\log \frac{n}{k}} \rightarrow 0$. For the $|\mathcal{N}(\mu, \sigma^2)|$ distribution with $\mu \neq 0$ one can show, after tedious calculations, that (4) is satisfied with $\theta = 1/2$, and that (5) and (6) hold if $k/\log(n/k) \rightarrow 0$.

Compared to Goegebeur et al. (2022) where the estimation of the conditional tail moment has also been studied but for Pareto-type distributions, the assumptions required to show the convergence in distribution of the estimator in the extrapolation case are not the same as those of Theorem 4.1, nor the speed of convergence. These differences can be explained from the expression of the estimators in the case of extrapolation which are, in Pareto and Weibull cases, based on a Weissman-type construction as follows

$$\left(\frac{U_X(1/p)}{U_X(n/k)}\right)^\beta \times \text{estimator in the intermediate case,}$$

but with a different ratio

$$\frac{U_X(1/p)}{U_X(n/k)} \sim \begin{cases} \left(\frac{k}{np}\right)^\gamma & \text{in case of Pareto-type distributions,} \\ \left(\frac{\log(1/p)}{\log(n/k)}\right)^\theta & \text{in case of Weibull-type distributions,} \end{cases}$$

where γ is the extreme value index. Thus, if $\hat{\gamma}_k$ is an estimator of γ such that $\sqrt{k}(\hat{\gamma}_k - \gamma) \xrightarrow{d} \Gamma$, the dominant term in the expansion of the ratio between the conditional tail moment estimator and the true value is given by

$$\begin{cases} 1 + \beta (\hat{\gamma}_k - \gamma) \log \frac{k}{np} & \text{in case of Pareto-type distributions,} \\ 1 + \beta (\hat{\theta}_k - \theta) \log \frac{\log 1/p}{\log n/k} & \text{in case of Weibull-type distributions,} \end{cases}$$

from which different conditions (involving p) follows.

Note also that it is quite usual to have differences between two domains of attraction, see, e.g., Gardes et al. (2011), where some different asymptotic properties are highlighted between the Weibull tail-coefficient estimator and the Hill estimator (see Hill, 1975), in terms of bias and rate of convergence.

5 Simulation experiment

In this section we illustrate the finite sample behavior of the proposed estimator by means of a simulation experiment. The considered distributions are: *Weibull*(1), $|\mathcal{N}(1.2, 1)|$, $\Gamma(2, 1)$ and *EWeibull*(1.5, -1). From each distribution we simulate 500 datasets of sizes $n \in \{500, 2000\}$, and we consider estimation of $CTM_{\beta,p}$ with $p \in \{1/n, 1/(5n)\}$, and $\beta \in \{1, 2\}$.

In Figure 1 we show for the *Weibull*(1) distribution the boxplots of $\widehat{CTM}_{1,p}/CTM_{1,p}$, computed over the 500 replications, as a function of k for $n = 500$ (left) and $n = 2000$ (right), and where the top row corresponds with $p = 1/n$ and the bottom row with $p = 1/(5n)$. Figures 2 – 4 are constructed analogously for the other distributions. From the simulations we observe the following:

- For all the distributions considered the estimation is good when k is small compared to n , which is in agreement with the theoretical condition $k/n \rightarrow 0$. With increasing k the estimators show typically an increase in bias. As for the bias we comment that estimation for Weibull-type distributions is practically challenging, with a bias that comes typically rather fast with increasing k . This was also already observed in the simpler context of the estimation of the Weibull-tail coefficient θ , see, e.g., Girard (2004), Gardes and Girard (2008), Goegebeur et al. (2010), to mention only a few, and obviously also manifests itself here. Note however that the bias for the Extended Weibull distribution in Figure 4 shows a different pattern compared to the other distributions shown in Figures 1 – 3. In fact, the bias of our estimator $\widehat{CTM}_{\beta,p}$ comes from different terms and thus it is difficult to theoretically explain their interaction, see the proof of our Theorem 4.1. Nevertheless, the estimator for θ , which is one of these terms from which the bias comes (see term $T_{11,n}$ in the proof of our Theorem 4.1), has for the Extended Weibull distribution an asymptotic normality which requires less assumptions (controlling the bias) compared to the same result for general Weibull-type distributions: see Corollary 2 for the Extended Weibull distribution and Theorem 1 for general Weibull-type distributions in Girard (2004).
- The estimation results improve with increasing sample size n .
- Decreasing p leads to increased variability of the estimates, which can be expected as smaller values of p lead to more severe extrapolations.
- Larger values of θ lead to more variable results. This is also in line with the expectations as θ determines the heaviness of the upper tail of the Weibull-type model.

In Figure 5 we illustrate the estimation of the second conditional tail moment, $CTM_{2,p}$, for the $|\mathcal{N}(1.2, 1)|$ distribution. Similar figures can be found in the online Supporting Information for

the three other distributions considered here. As expected, estimation of the second conditional tail moment is more challenging than that of the conditional tail expectation ($\beta = 1$), with larger bias and variability of the estimates.

Next, using Theorem 4.1 combined with Lemma 4.1, we can construct a confidence interval for the conditional tail moment. To this aim and in order to improve the coverage probabilities, as suggested by Drees (2003), we use a log-scale version of Theorem 4.1, i.e.,

$$\sqrt{k} \log \frac{\widehat{CTM}_{\beta,p}}{CTM_{\beta,p}} \xrightarrow{d} \beta \log(\lambda)\Gamma,$$

from which we deduce the approximate $100(1 - \alpha)\%$ confidence interval

$$\left[\widehat{CTM}_{\beta,p} \exp \left\{ -\Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\hat{\theta}_k \beta}{\sqrt{k}} \log \frac{\log \frac{1}{p}}{\log \frac{n}{k}} \right\}; \widehat{CTM}_{\beta,p} \exp \left\{ \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\hat{\theta}_k \beta}{\sqrt{k}} \log \frac{\log \frac{1}{p}}{\log \frac{n}{k}} \right\} \right],$$

where Φ^{-1} denotes the quantile function of the standard normal distribution. To evaluate the accuracy of the approximate confidence intervals, we show in Figure 6 for the four considered distributions the empirical coverage probabilities of the 95% confidence intervals for $CTM_{\beta,p}$ as a function of k , for $n \in \{500, 2000\}$, $p \in \{1/n, 1/(5n)\}$ and $\beta = 1$. As is clear from the results, the confidence intervals only reach the 95% level for the smaller values of k , except for the extended Weibull distribution which has a longer stable pattern. Remark that in the construction of the confidence intervals the bias is assumed to be zero. In reality this is not the case, which explains the deviation from the nominal level 0.95 as k increases. Concerning the better coverage probabilities in case of the Extended Weibull distribution, it might be due to the fact that the estimates for $CTM_{\beta,p}$ have a better bias behavior compared to the other distributions, and also that the Weibull tail coefficient θ was for the Extended Weibull distribution set at 1.5, which is the largest among the considered cases, leading to wider confidence intervals. For all the distributions, the coverage probabilities increase with n and they are also larger for $p = 1/(5n)$ compared to $p = 1/n$. The latter can be explained by the fact that the confidence intervals become wider for decreasing values of p .

Finally, we discuss the choice of the parameter k on which the estimator $\widehat{CTM}_{\beta,p}$ is based. Unlike the asymptotic variance, the asymptotic bias comes from several sources and cannot

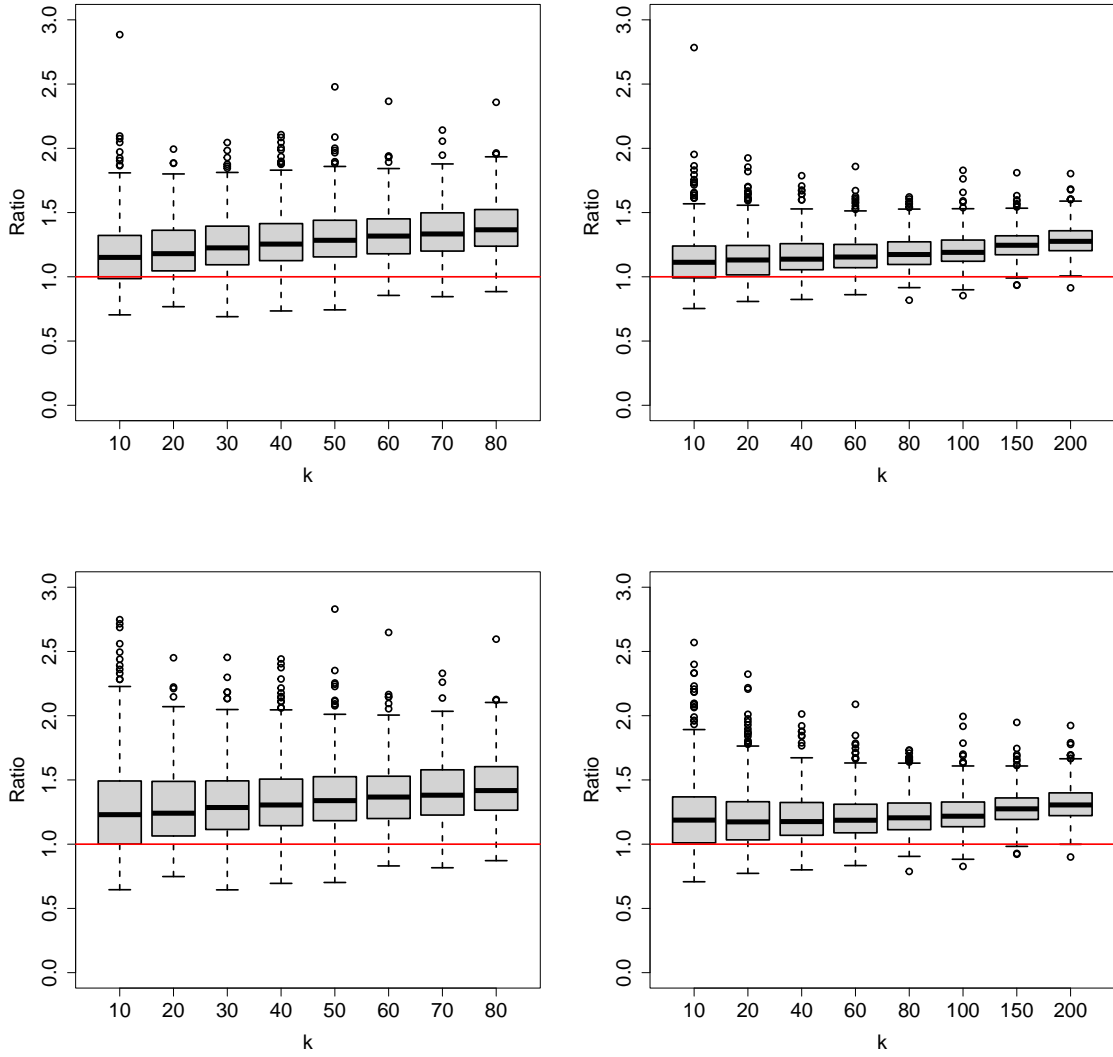


Figure 1: *Weibull*(1) distribution. Boxplots of $\widehat{CTM}_{1,p}/CTM_{1,p}$ as a function of k for $n = 500$ (left) and $n = 2000$ (right) with $p = 1/n$ (top) and $p = 1/(5n)$ (bottom).

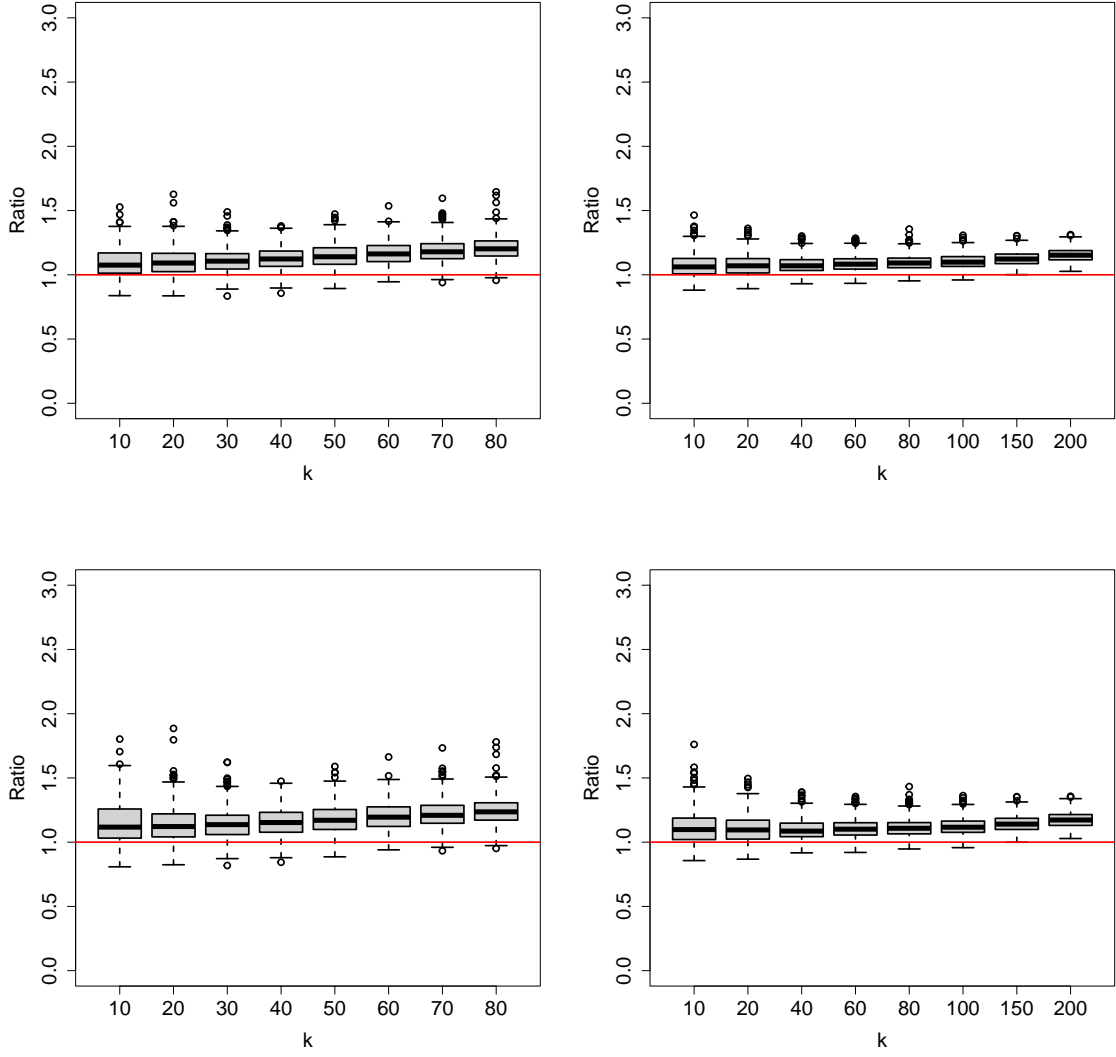


Figure 2: $|\mathcal{N}(1.2, 1)|$ distribution. Boxplots of $\widehat{CTM}_{1,p} / CTM_{1,p}$ as a function of k for $n = 500$ (left) and $n = 2000$ (right) with $p = 1/n$ (top) and $p = 1/(5n)$ (bottom).

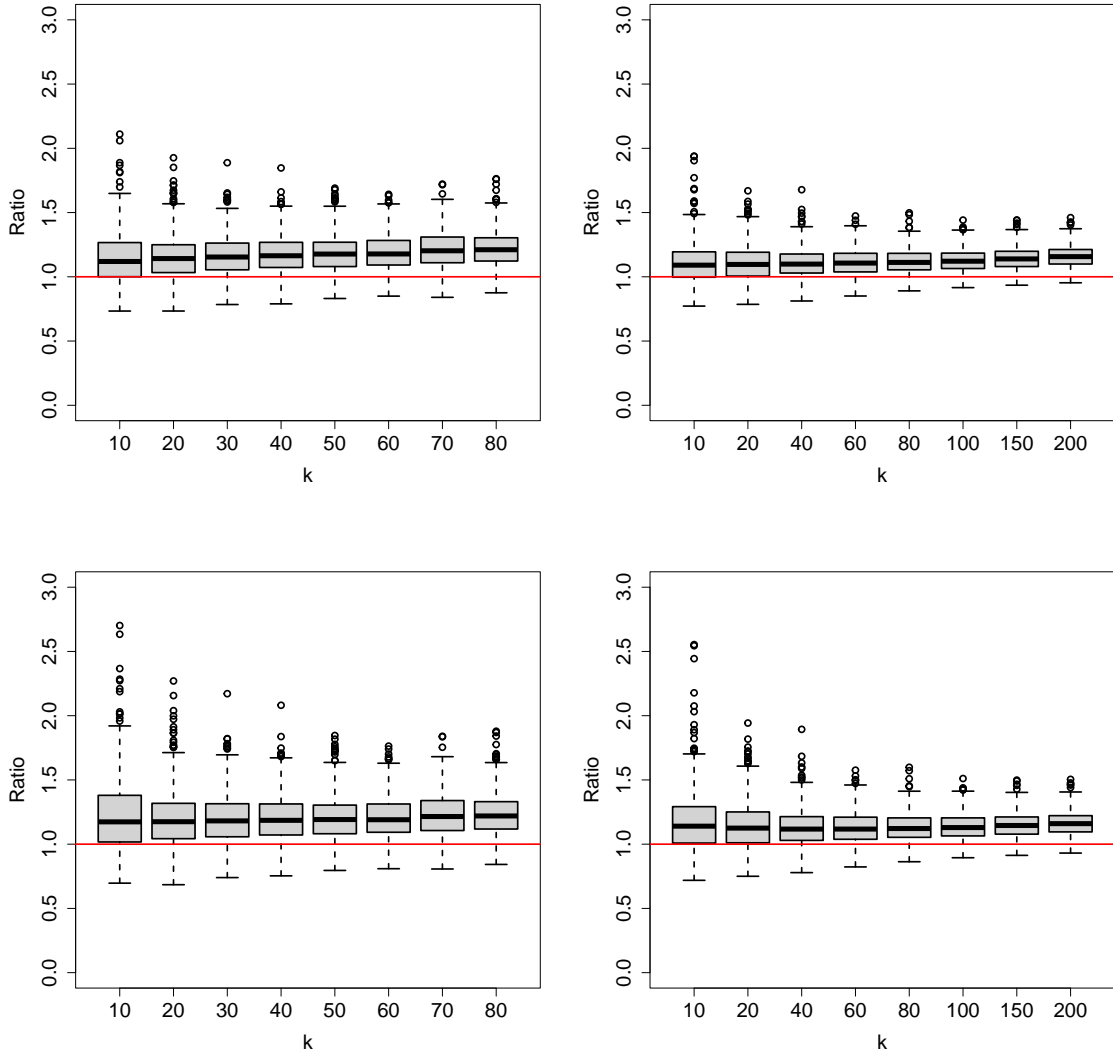


Figure 3: $\Gamma(2, 1)$ distribution. Boxplots of $\widehat{CTM}_{1,p}/CTM_{1,p}$ as a function of k for $n = 500$ (left) and $n = 2000$ (right) with $p = 1/n$ (top) and $p = 1/(5n)$ (bottom).

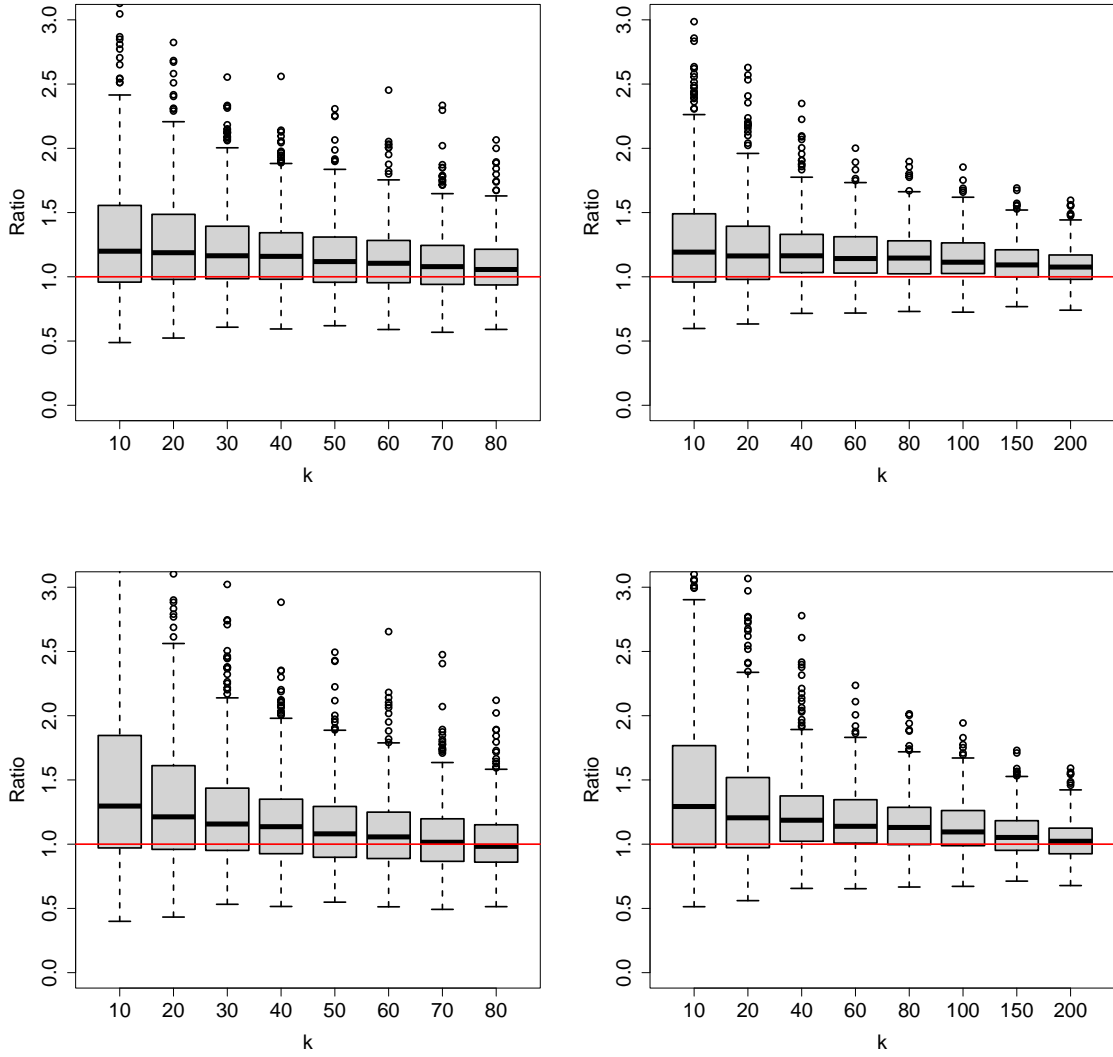


Figure 4: $EWeibull(1.5, -1)$ distribution. Boxplots of $\widehat{CTM}_{1,p}/CTM_{1,p}$ as a function of k for $n = 500$ (left) and $n = 2000$ (right) with $p = 1/n$ (top) and $p = 1/(5n)$ (bottom).

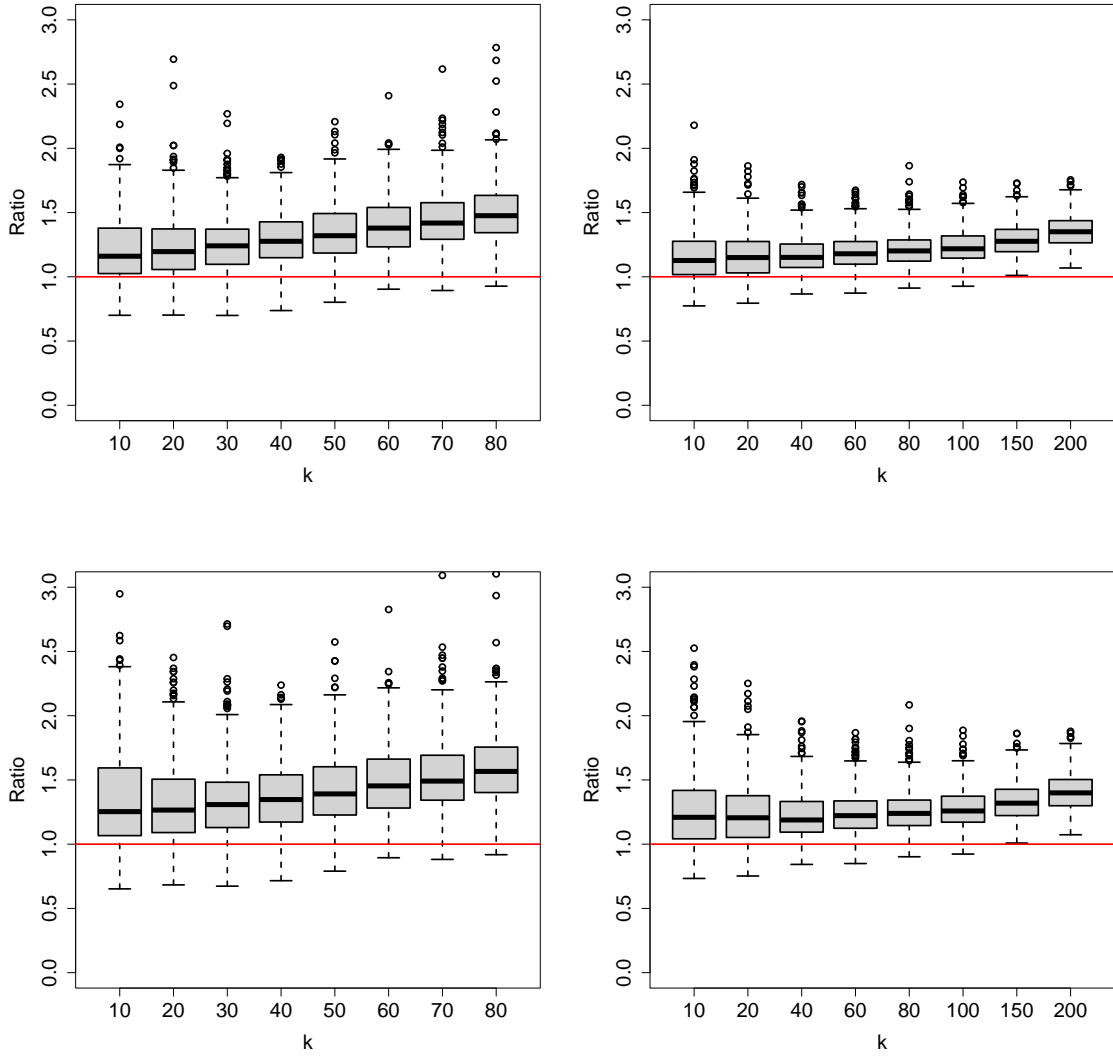


Figure 5: $|\mathcal{N}(1.2, 1)|$ distribution. Boxplots of $\widehat{CTM}_{2,p} / CTM_{2,p}$ as a function of k for $n = 500$ (left) and $n = 2000$ (right) with $p = 1/n$ (top) and $p = 1/(5n)$ (bottom).

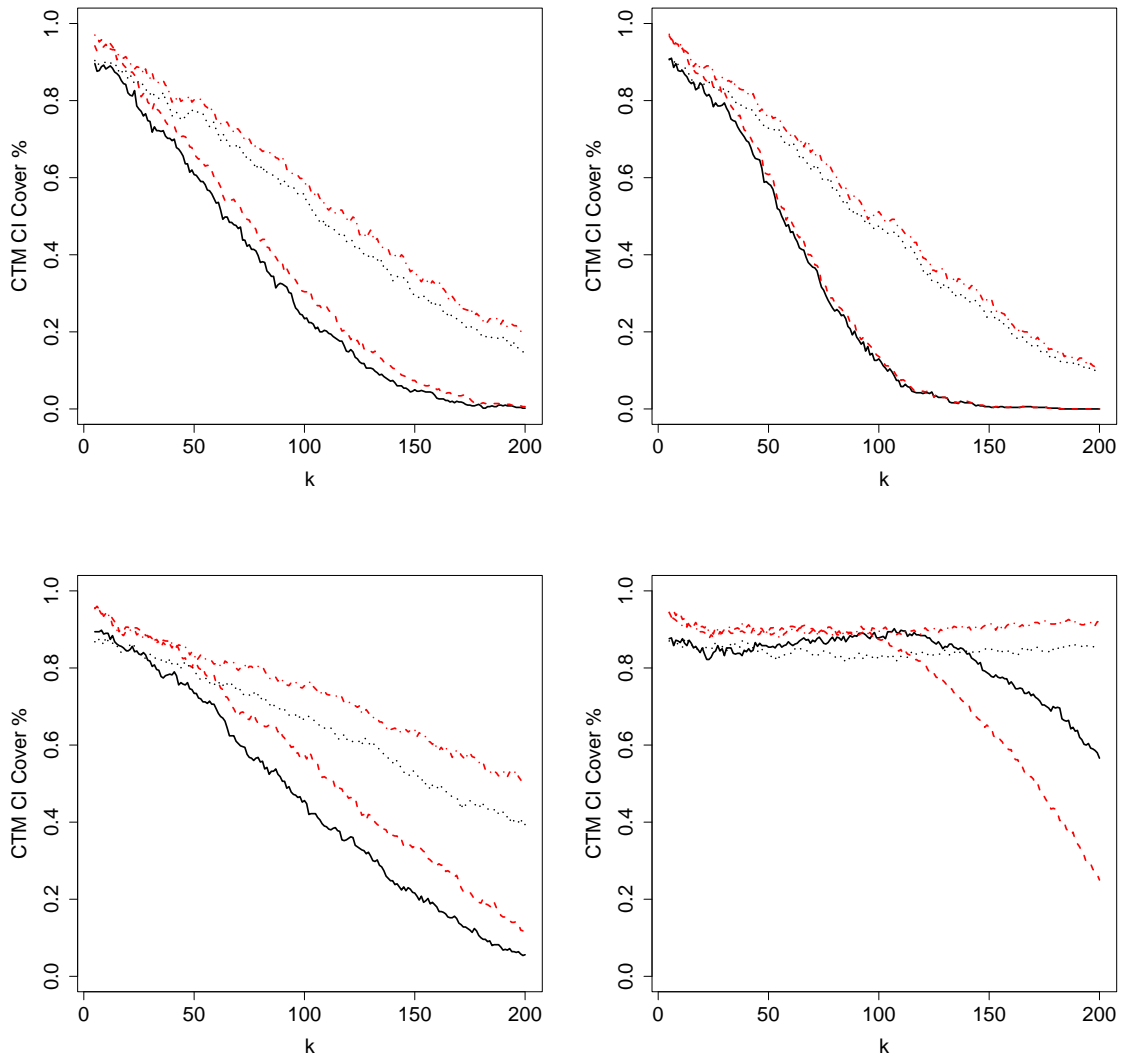


Figure 6: Coverage probabilities of approximate 95% confidence intervals as a function of k for $(n, p) = (500, 0.002)$ (black solid line), $(500, 0.0004)$ (red dashed line), $(2000, 0.0005)$ (black dotted line) and $(2000, 0.0001)$ (red dashed-dotted line). The distributions considered are $Weibull(1)$ (top left), $|\mathcal{N}(1.2, 1)|$ (top right), $\Gamma(2, 1)$ (bottom left) and $EWeibull(1.5, -1)$ (bottom right).

be made explicit. Thus, this makes the actual implementation of a direct criterion based on the AMSE nearly impossible. In such a case, a common approach in extreme value statistics is to look at the plot $(k, \widehat{CTM}_{\beta,p})$, and to try to identify a stable part in this plot. Rather than looking for this stability with the naked eye, we propose a selection procedure like the one suggested by Goegebeur et al. (2019) making this 'visual detection' of the stable part automatic, namely:

1. we compute $\widehat{CTM}_{\beta,p}$ for $k = 5, \dots, \min(0.1n, 50)$,
2. we compute the standard deviation of $\widehat{CTM}_{\beta,p}$ in a moving block of 10 successive k -values,
3. we select the block with the smallest standard deviation,
4. within the block selected in step 3, we repeat the steps 2 and 3, now with block size 5,
5. the estimate for $CTM_{\beta,p}$ is the median of the $\widehat{CTM}_{\beta,p}$ in the finally selected block.

As is clear from Figure 7, for fixed n , smaller values of p lead to more variable results while increasing n leads to less variability of the estimates, as expected. After the data-driven selection of k some estimation bias remains, which is in line with the boxplots shown in Figures 1 till 5.

In Table 1, we report the empirical coverage probabilities of the approximate 95% confidence intervals based on $\widehat{CTM}_{1,p}$, at the optimally selected k -value. As observed before, decreasing p for a fixed n leads to increased coverage probabilities, and increasing n has also a positive impact on the coverage probabilities. Overall these coverage probabilities are reasonably close to the nominal level 95% for all the distributions, sample sizes and values for p .

6 Real data analysis

6.1 Wind speeds at Anholt offshore wind farm

Following the worldwide economic development and population growth, global energy demand has also shown an important upward trend during the last decades. Traditional fossil fuels have many disadvantages: their reserves are limited and their use contributes to increased carbon

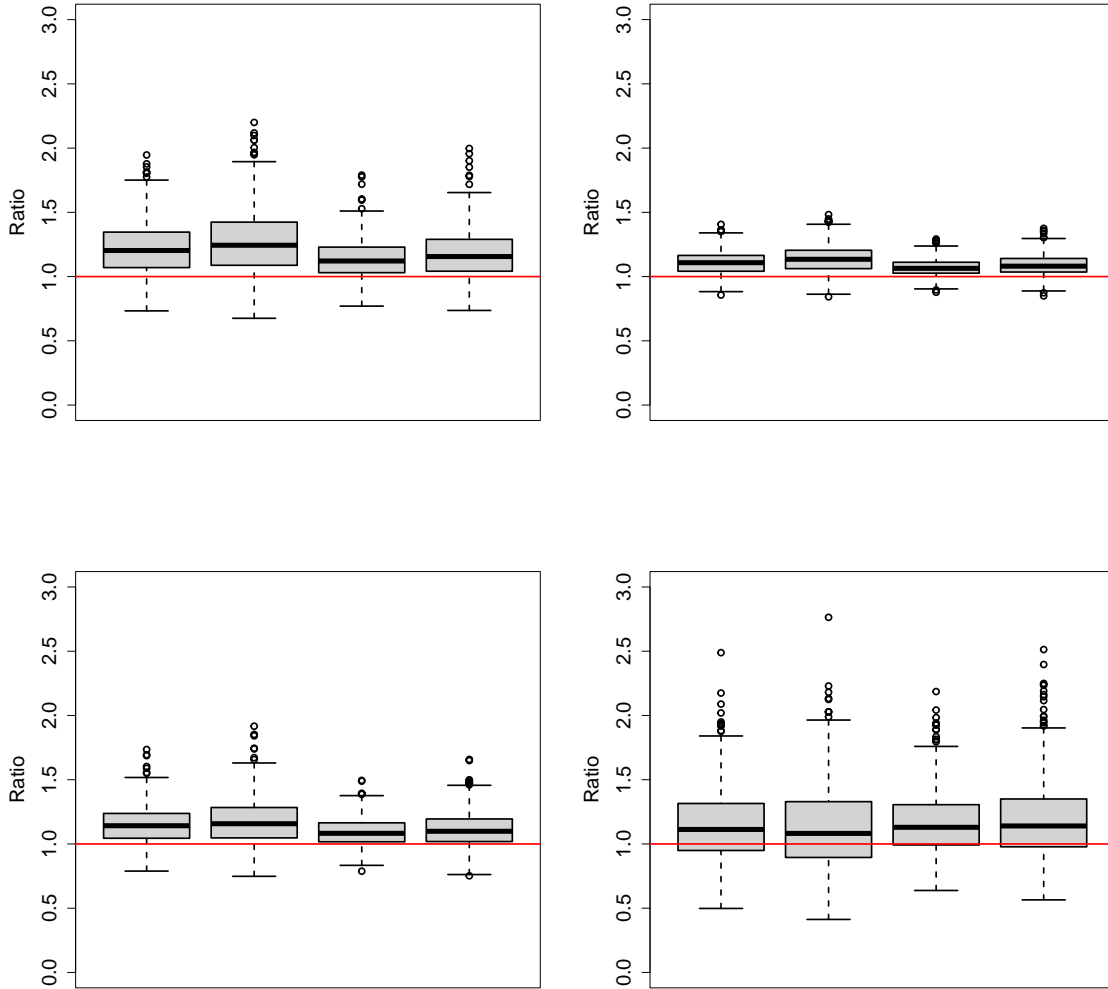


Figure 7: Boxplots of $\widehat{CTM}_{1,p}/CTM_{1,p}$ calculated at the optimal k -value, with from left to right $(n, p) = (500, 0.002), (500, 0.0004), (2000, 0.0005)$ and $(2000, 0.0001)$. The distributions considered are $Weibull(1)$ (top left), $|\mathcal{N}(1.2, 1)|$ (top right), $\Gamma(2, 1)$ (bottom left) and $EWeibull(1.5, -1)$ (bottom right).

	$n = 500$		$n = 2000$	
	$p = 1/n$	$p = 1/(5n)$	$p = 1/n$	$p = 1/(5n)$
<i>Weibull</i> (1)	79.4	83.8	85.4	88.6
$ \mathcal{N}(1.2, 1) $	80.0	82.4	82.4	86.6
$\Gamma(2, 1)$	86.2	91.4	85.8	88.6
<i>EWeibull</i> (1.5, -1)	88.8	92.2	88.0	92.6

Table 1: Empirical coverage probabilities of 95% confidence intervals for $CTM_{1,p}$ based on 500 simulated data sets.

dioxide levels in the atmosphere. On the other hand, renewable energy like wind power is inexhaustible and causes little pollution. Accurate modelling of the wind speed distribution is of crucial importance for wind energy output estimation, e.g., at an offshore wind farm. For a review of wind speed modelling approaches we refer to Shi et al. (2021).

Our interest is in an analysis of the upper tail of the wind speed distribution at the Anholt offshore wind farm, which is a Danish offshore wind power farm in the Kattegat, between Djursland and the island Anholt. We estimate the Weibull-tail index θ , the first conditional tail moment $CTM_{1,p}$ (also called the conditional tail expectation) and the conditional tail variance V_p , using a dataset that is publicly available at <https://orsted.com/en/what-we-do/renewable-energy-solutions/offshore-wind/offshore-wind-data>. In particular we use LIDAR measured wind speeds (in meters per second) at 100 meters altitude. The original dataset consists of a time series of 10 minutes measurements during the years 2013-2014, but in order to reduce temporal dependencies only one observation every three days is kept for the analysis, leading to $n = 242$. In order to evaluate the appropriateness of the Weibull-type model (1) we construct the Weibull QQ plot, which has coordinates

$$\left(\log \left(-\log \left(1 - \frac{i}{n+1} \right) \right), \log X_{i,n} \right), \quad i = 1, \dots, n,$$

see Figure 8, top left panel. Clearly, the Weibull QQ plot becomes approximately linear in the largest observations, which supports the Weibull-type model. We refer to Goegebeur and

Guillou (2010) for a discussion of the Weibull QQ plot. Next we estimate the Weibull-tail index θ , see Figure 8, top right, where we show $\hat{\theta}_k$ along with approximate 95% confidence intervals as a function of k . This plot shows a stable horizontal part for k between, approximately, 30 and 60, indicating an estimate for θ of about 0.4. Finally, the bottom row of Figure 8 shows the estimates for $CTM_{1,1/n}$ (left) and $CTM_{1,1/(5n)}$ (right) with approximate 95% confidence intervals as a function of k . Applying the data-driven selection method to determine k , as described in the simulation section, yields $\widehat{CTM}_{1,1/n} = 27.3$ and $\widehat{CTM}_{1,1/(5n)} = 30.9$, obtained at $k = 21$ and $k = 19$, respectively. These k values are clearly selected in the stable part of the plots of $\widehat{CTM}_{1,p}$ as function of k . Besides $CTM_{1,p}$, being an average above a high quantile, it can also be relevant to estimate the variability in the upper tail, V_p , see Figure 9. Focusing on the stable k -region we have an estimate of V_p between 30-40 for $p = 1/n$ and $p = 1/(5n)$.

6.2 PM_{2.5} air pollution

Air pollution is among the greatest environmental risks to human health. Of particular concern is the concentration of PM_{2.5}, referring to atmospheric particulate matter with a diameter of less than 2.5 micrometers. Since these particles are so small and light, they tend to stay longer in the air than heavier ones, increasing the chances of humans and animals inhaling them. PM_{2.5} can travel deep in the respiratory tract, reaching the lungs and entering the blood stream. The adverse health effects of PM_{2.5} are well studied, see, e.g., Feng et al. (2016) and Li et al. (2017), and the references therein.

We apply the developed methodology to the PM_{2.5} air pollution in the city of Slavonski Brod, Croatia, which is with an average of 28.0 micrograms per cubic meter ($\mu\text{g}/\text{m}^3$), measured during the last two years, amongst the most polluted cities in Europe. As a reference, note that the 2021 World Health Organisation's updated health-based guidelines for air quality recommend a maximum level of 5 $\mu\text{g}/\text{m}^3$ for long term exposure, in order to protect health. We use a dataset of daily average concentrations of PM_{2.5} over the period 2015-2021, publicly available on the website of the European Environment Agency at <https://discomap.eea.europa.eu/map/fme/AirQualityExport.htm>. We base the analysis on the data from the winter season (December-

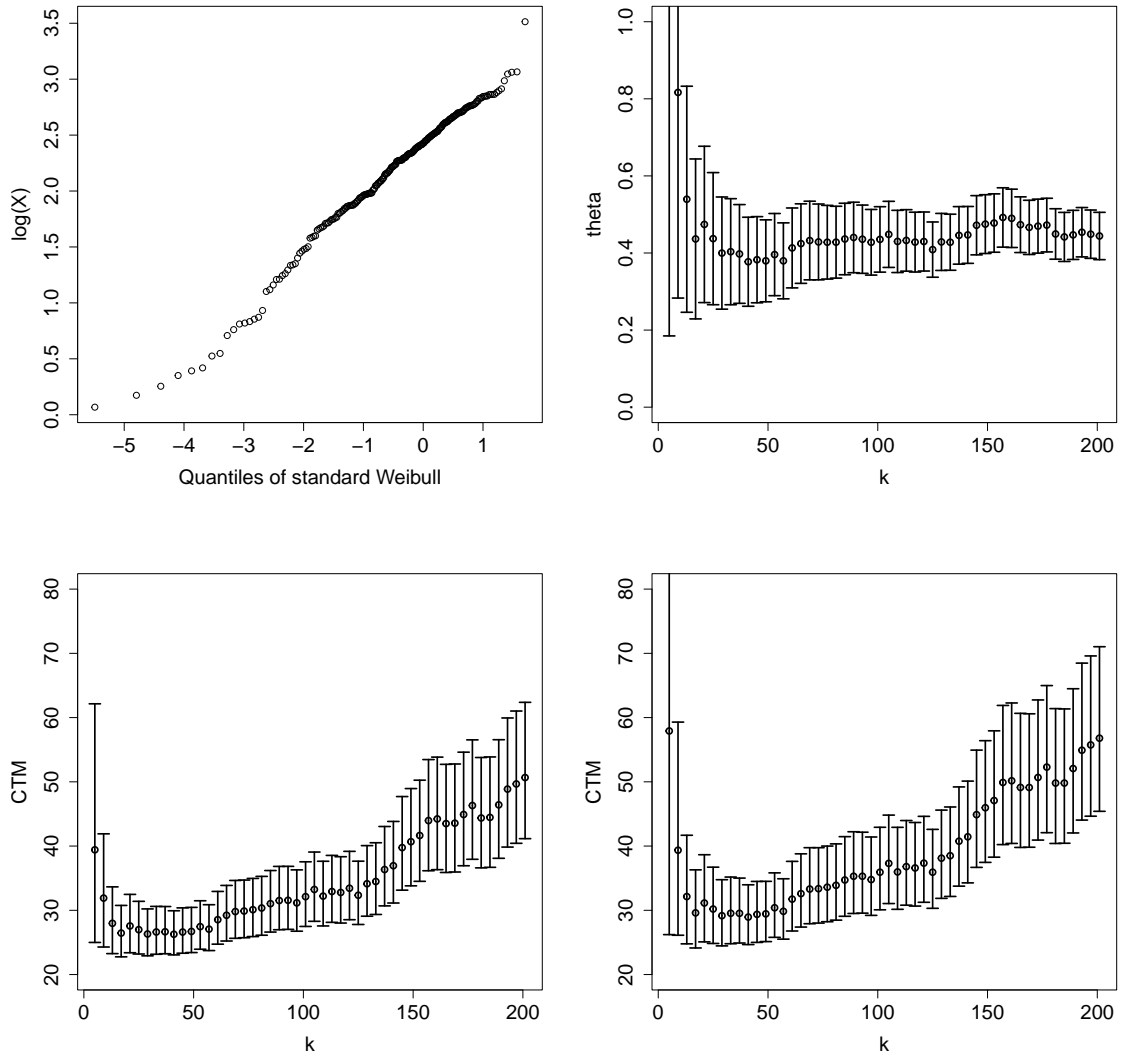


Figure 8: Anholt wind speed data. Weibull QQ plot (top left), $\hat{\theta}_k$ with approximate 95% confidence intervals as a function of k (top right), $\widehat{CTM}_{1,1/n}$ (bottom left) and $\widehat{CTM}_{1,1/(5n)}$ (bottom right) with approximate 95% confidence intervals as a function of k .

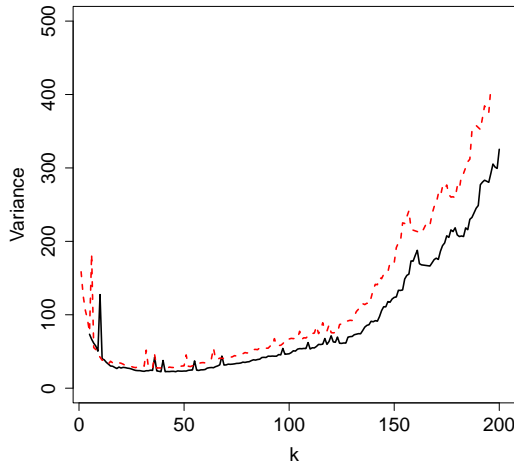


Figure 9: Anholt wind speed data. \widehat{V}_p with $p = 1/n$ (solid line) and $p = 1/(5n)$ (dashed line) as a function of k .

February), and in order to reduce the temporal dependence we only consider measurements that are three days spaced apart, leading to $n = 209$. The Weibull QQ plot shown in Figure 10, top left, becomes ultimately linear, supporting the assumption of an underlying Weibull-type distribution, and the plot of the Weibull-tail index estimates $\widehat{\theta}_k$ (Figure 10, top right) along with approximate 95% confidence intervals as a function of k is stable for k between approximately 20 and 80, indicating an estimate of about 0.7-0.8. In the bottom row of Figure 10 we show $\widehat{CTM}_{1,1/n}$ (left panel) and $\widehat{CTM}_{1,1/(5n)}$ (right panel) along with approximate 95% confidence intervals as a function of k . The data-driven selection method gives $\widehat{CTM}_{1,1/n} = 312.8$ and $\widehat{CTM}_{1,1/(5n)} = 386.6$, both obtained at $k = 15$. In Figure 11 we show \widehat{V}_p for $p = 1/n$ and $p = 1/(5n)$ as a function of k , indicating estimates around 10 000 and 15 000 for $p = 1/n$ and $p = 1/(5n)$, respectively.

7 Concluding remarks

Below we list some topics that will be investigated in future projects.

- In the present paper the estimation problem was considered in the context of i.i.d. random

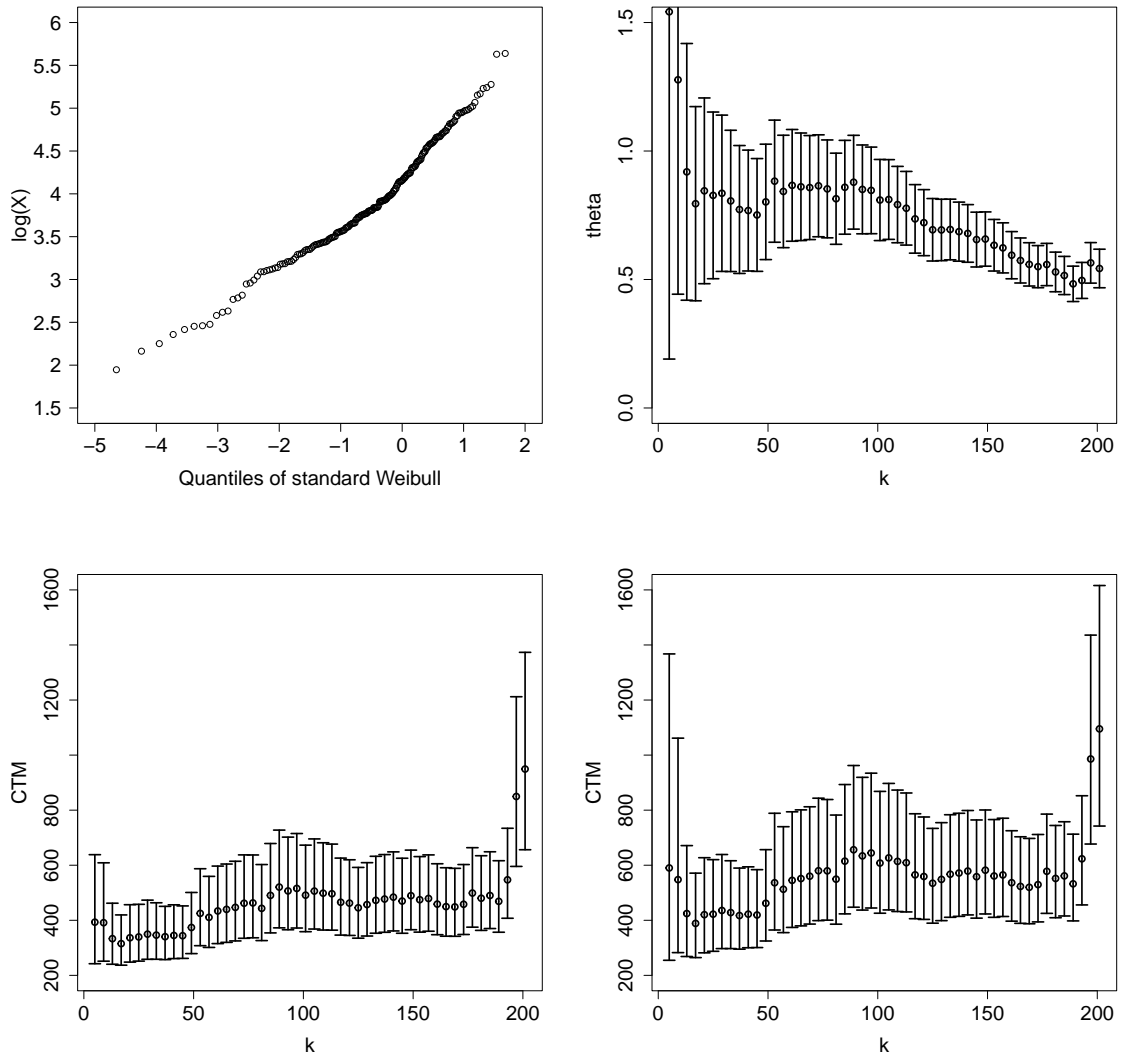


Figure 10: Air pollution data. Weibull QQ plot (top left), $\hat{\theta}_k$ with approximate 95% confidence intervals as a function of k (top right), $\widehat{CTM}_{1,1/n}$ (bottom left) and $\widehat{CTM}_{1,1/(5n)}$ (bottom right) with approximate 95% confidence intervals as a function of k .

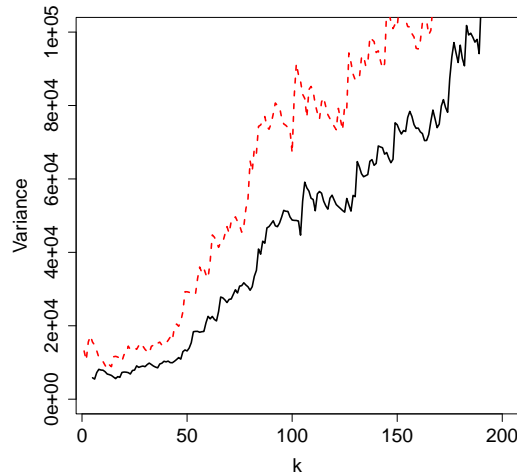


Figure 11: Air pollution data. \widehat{V}_p with $p = 1/n$ (solid line) and $p = 1/(5n)$ (dashed line) as a function of k .

variables. A natural extension is then to derive the asymptotic behavior of the proposed estimator when the data are of the form of a stationary time series. In the recent literature there are several contributions concerning the estimation of risk measures based on time series data, satisfying some mixing condition, though most of these are in the context of multivariate regular variation/heavy-tailed distributions. For instance, Linton and Xiao (2013) consider the estimation of the conditional tail expectation based on a stationary regularly varying time series while Davison et al. (2023) study the estimation of the marginal expected shortfall for a heavy-tailed multivariate time series. See also Davis and Mikosch (2009), and Padoan et al. (2023) for other examples of estimation problems with multivariate regularly varying time series. Drees (2003) contains results for estimation of the extreme value index and extreme quantiles based on strictly stationary β -mixing time series where the marginal distribution belongs to the general max-domain of attraction. Although the result of Theorem 6.1 and Theorem 6.2 of the latter paper could possibly also be used to address the estimation problem considered in our paper, as the Weibull-type distributions are a subclass of the general max-domain of attraction, it would be desirable to have a result that is specific for the Weibull-type class. Indeed, as mentioned

in the introduction to our paper, by considering the Weibull-type distributions one can differentiate the tail behaviour by the parameter θ , while from a more general perspective, i.e., by considering the Gumbel class, one has an extreme value index $\gamma = 0$. Focusing on our estimation problem, we would need a process convergence result similar to (9) below for the case of stationary β -mixing time series. Obtaining such a result is not trivial. A possible approach could be to use the theory on the weak convergence of empirical processes of cluster functionals, developed in Drees and Rootzén (2010), but this will require completely different proofs and is outside the scope of the present paper. Note also that in the time series context the asymptotic variances are typically more complicated as they are of the form of a series; see, e.g., Theorem 6.1 in Drees (2003) and Davison et al. (2023), making the results more difficult to use in practice. The development and study of a bootstrap procedure to estimate the asymptotic variance is another topic of future investigations.

- The developed methodology could also be extended to the full max-domain of attraction. This generalisation would then include the Weibull-type models considered in the present paper, but also the Pareto-type and the light-tailed distributions which have a finite right endpoint.
- Padoan and Stupfler (2022), see also Gomes and Pestana (2007), proposed correction methods in the estimate of the asymptotic variance for the construction of confidence intervals with better coverage probabilities. It would be thus interesting to develop a similar approach and to compare it with the method used in this paper based on an automatic selection of k .

8 Proofs

Remark that

$$\begin{aligned}
& \overline{CTM}_{\beta,k/n}(x) \\
&= \beta \frac{1}{k} \sum_{i=1}^n \int_0^{X_i} u^{\beta-1} du \mathbb{1}_{\{X_i > U_X(\frac{n}{k}) + a_X(\frac{n}{k})x\}} \\
&= \beta \int_0^\infty u^{\beta-1} \frac{1}{k} \sum_{i=1}^n \mathbb{1}_{\{X_i > \max(u, U_X(\frac{n}{k}) + a_X(\frac{n}{k})x)\}} du \\
&= \beta \int_0^{U_X(\frac{n}{k}) + a_X(\frac{n}{k})x} u^{\beta-1} du \frac{1}{k} \sum_{i=1}^n \mathbb{1}_{\{X_i > U_X(\frac{n}{k}) + a_X(\frac{n}{k})x\}} \\
&\quad + \beta \int_{U_X(\frac{n}{k}) + a_X(\frac{n}{k})x}^\infty u^{\beta-1} \frac{1}{k} \sum_{i=1}^n \mathbb{1}_{\{X_i > u\}} du \\
&= \left[U_X\left(\frac{n}{k}\right) + a_X\left(\frac{n}{k}\right)x \right]^\beta \frac{1}{k} \sum_{i=1}^n \mathbb{1}_{\{X_i > U_X(\frac{n}{k}) + a_X(\frac{n}{k})x\}} \\
&\quad + \beta a_X\left(\frac{n}{k}\right) \int_0^\infty \left[U_X\left(\frac{n}{k}\right) + a_X\left(\frac{n}{k}\right)(x+v) \right]^{\beta-1} \frac{1}{k} \sum_{i=1}^n \mathbb{1}_{\{X_i > U_X(\frac{n}{k}) + a_X(\frac{n}{k})(x+v)\}} dv \\
&= \left[U_X\left(\frac{n}{k}\right) \right]^\beta \left\{ \left[1 + \frac{a_X\left(\frac{n}{k}\right)}{U_X\left(\frac{n}{k}\right)}x \right]^\beta T_n\left(\frac{n}{k}\overline{F}_X\left(U_X\left(\frac{n}{k}\right) + a_X\left(\frac{n}{k}\right)x\right)\right) \right. \\
&\quad \left. + \beta \frac{a_X\left(\frac{n}{k}\right)}{U_X\left(\frac{n}{k}\right)} \int_0^\infty \left[1 + \frac{a_X\left(\frac{n}{k}\right)}{U_X\left(\frac{n}{k}\right)}(x+v) \right]^{\beta-1} T_n\left(\frac{n}{k}\overline{F}_X\left(U_X\left(\frac{n}{k}\right) + a_X\left(\frac{n}{k}\right)(x+v)\right)\right) dv \right\} \text{ a.s.},
\end{aligned}$$

where

$$T_n(x) := \frac{1}{k} \sum_{i=1}^n \mathbb{1}_{\{\overline{F}_X(X_i) < \frac{k}{n}x\}}.$$

The asymptotic behavior of $T_n(x)$, after proper normalisation, is given in Lemma 1 of Cai et al. (2015), according which, for any $\eta \in [0, 1/2)$ and $T > 0$ we have, for $n \rightarrow \infty$,

$$\sup_{x \in (0, T]} \left| \frac{\sqrt{k}[T_n(x) - x] - W(x)}{x^\eta} \right| \rightarrow 0 \text{ a.s.} \quad (9)$$

Note that in the above result, all the involved processes are defined on the same probability space via the Skorohod construction. See also Einmahl et al. (2006) for convergence properties of related stochastic processes.

For completeness, we also include below three lemmas from Goegebeur et al. (2023), useful for our proofs. Note that the last one is actually a consequence of the proof of Lemma 5.4 in the aforementioned paper.

Lemma 8.1 Let F_X satisfy (1), H be differentiable with derivative $h \in RV_{1/\theta-1}$, then, as $t \rightarrow \infty$

$$\frac{\overline{F}_X\left(t + \frac{z}{h(t)}\right)}{\overline{F}_X(t)} \rightarrow e^{-z},$$

uniformly for $z \geq -\delta$, for any $\delta > 0$.

Lemma 8.2 If F_X satisfies (1) and H is differentiable with derivative $h \in RV_{1/\theta-1}$, then, for a sequence k such that $k \rightarrow \infty$ with $k/n \rightarrow 0$, we have

$$\sqrt{k} \frac{X_{n-k,n} - U_X\left(\frac{n}{k}\right)}{a_X\left(\frac{n}{k}\right)} \xrightarrow{d} W(1).$$

Lemma 8.3 Assume F_X satisfies (1), H is differentiable with derivative h which is ultimately monotone, and h is differentiable with derivative h' satisfying $xh'(x)/h(x) \rightarrow 1/\theta - 1$ as $x \rightarrow \infty$. Then, if $k \rightarrow \infty$ as $n \rightarrow \infty$ such that $k/n \rightarrow 0$ and $\sqrt{k}a_X(n/k)/U_X(n/k) \rightarrow 0$, we have, for any $\delta > 0$,

$$\sup_{x \in [-\delta, \delta]} \sqrt{k} \int_0^\infty \left| \frac{n}{k} \overline{F}_X\left(U_X\left(\frac{n}{k}\right) + a_X\left(\frac{n}{k}\right)(x+v)\right) - e^{-(x+v)} \right| dv^\beta \rightarrow 0.$$

Proof of Lemma 2.1. We have

$$\begin{aligned} CTM_{\beta,p} &= - \int_{U_X(1/p)}^\infty \frac{x^\beta}{\overline{F}_X(U_X(1/p))} d\overline{F}_X(x) \\ &= [U_X(1/p)]^\beta + \beta \int_{U_X(1/p)}^\infty x^{\beta-1} \frac{\overline{F}_X(x)}{\overline{F}_X(U_X(1/p))} dx \\ &= [U_X(1/p)]^\beta \left\{ 1 + \beta \frac{a_X(1/p)}{U_X(1/p)} \int_0^\infty \left[1 + z \frac{a_X(1/p)}{U_X(1/p)} \right]^{\beta-1} \frac{\overline{F}_X(U_X(1/p) + z a_X(1/p))}{\overline{F}_X(U_X(1/p))} dz \right\}. \end{aligned} \quad (10)$$

Since $\frac{a_X(1/p)}{U_X(1/p)} \rightarrow 0$, to prove Lemma 2.1, it is sufficient to show that the integral in the right-hand side of (10) is finite. To this aim, by the mean value theorem, with $\xi \in (0, 1)$, we have

$$\begin{aligned} \frac{\overline{F}_X(U_X(1/p) + z a_X(1/p))}{\overline{F}_X(U_X(1/p))} &= e^{-H(U_X(1/p) + z a_X(1/p)) + H(U_X(1/p))} \\ &= e^{-\frac{h(U_X(1/p) + \xi z a_X(1/p))}{h(U_X(1/p))} z}, \end{aligned}$$

from which, by the use of the Potter bounds (see Proposition B.1.9.5 in de Haan and Ferreira, 2006), for $0 < \delta_1 < 1$, $\delta_2 > 0$ and n large we have

$$\begin{aligned} & \int_0^\infty \left[1 + z \frac{a_X(1/p)}{U_X(1/p)} \right]^{\beta-1} \frac{\overline{F}_X(U_X(1/p) + z a_X(1/p))}{\overline{F}_X(U_X(1/p))} dz \\ & \leq \int_0^\infty \left[1 + z \frac{a_X(1/p)}{U_X(1/p)} \right]^{\beta-1} e^{-(1-\delta_1) [1+\xi \frac{a_X(1/p)}{U_X(1/p)} z]^{\frac{1}{\theta}} - 1 - \delta_2} z dz. \end{aligned}$$

If $\theta < 1$, take $0 < \delta_2 < 1/\theta - 1$, and obtain, for n large

$$\begin{aligned} & \int_0^\infty \left[1 + z \frac{a_X(1/p)}{U_X(1/p)} \right]^{\beta-1} \frac{\overline{F}_X(U_X(1/p) + z a_X(1/p))}{\overline{F}_X(U_X(1/p))} dz \\ & \leq \int_0^\infty \left[1 + z \frac{a_X(1/p)}{U_X(1/p)} \right]^{\beta-1} e^{-(1-\delta_1)z} dz \\ & = \frac{U_X(1/p)}{a_X(1/p)} e^{(1-\delta_1) \frac{U_X(1/p)}{a_X(1/p)}} \int_1^\infty u^{\beta-1} e^{-(1-\delta_1) \frac{U_X(1/p)}{a_X(1/p)} u} du \\ & = \frac{1}{1-\delta_1} \frac{\Gamma_I \left(\beta, (1-\delta_1) \frac{U_X(1/p)}{a_X(1/p)} \right)}{\left[(1-\delta_1) \frac{U_X(1/p)}{a_X(1/p)} \right]^{\beta-1} e^{-(1-\delta_1) \frac{U_X(1/p)}{a_X(1/p)}}}. \end{aligned} \tag{11}$$

For $p \downarrow 0$, the right-hand side in (11) is bounded.

Now, if $\theta \geq 1$, take $0 < \delta_2 < 1/\theta$, and for $\varepsilon, \Delta > 0$, we obtain for n large

$$\begin{aligned}
& \int_0^\infty \left[1 + z \frac{a_X(1/p)}{U_X(1/p)} \right]^{\beta-1} \frac{\overline{F}_X(U_X(1/p) + z a_X(1/p))}{\overline{F}_X(U_X(1/p))} dz \\
& \leq \int_0^\infty \left[\mathbb{1}_{\{\beta < 1\}} + (1 + \varepsilon z)^{\beta-1} \mathbb{1}_{\{\beta \geq 1\}} \right] e^{-(1-\delta_1)[1+\varepsilon z]^{\frac{1}{\theta}-1-\delta_2} z} dz \\
& = \int_0^\Delta \left[\mathbb{1}_{\{\beta < 1\}} + (1 + \varepsilon z)^{\beta-1} \mathbb{1}_{\{\beta \geq 1\}} \right] e^{-(1-\delta_1)[1+\varepsilon z]^{\frac{1}{\theta}-1-\delta_2} z} dz \\
& \quad + \int_\Delta^\infty \left[\mathbb{1}_{\{\beta < 1\}} + (1 + \varepsilon z)^{\beta-1} \mathbb{1}_{\{\beta \geq 1\}} \right] e^{-(1-\delta_1)[\varepsilon + \frac{1}{z}]^{\frac{1}{\theta}-1-\delta_2} z^{\frac{1}{\theta}-\delta_2}} dz \\
& \leq \left[\mathbb{1}_{\{\beta < 1\}} + (1 + \varepsilon \Delta)^{\beta-1} \mathbb{1}_{\{\beta \geq 1\}} \right] \int_0^\Delta e^{-(1-\delta_1)[1+\varepsilon \Delta]^{\frac{1}{\theta}-1-\delta_2} z} dz \\
& \quad + \int_\Delta^\infty \left[\mathbb{1}_{\{\beta < 1\}} + (1 + \varepsilon z)^{\beta-1} \mathbb{1}_{\{\beta \geq 1\}} \right] e^{-(1-\delta_1)[\varepsilon + \frac{1}{\Delta}]^{\frac{1}{\theta}-1-\delta_2} z^{\frac{1}{\theta}-\delta_2}} dz \\
& \leq \Delta \left[\mathbb{1}_{\{\beta < 1\}} + (1 + \varepsilon \Delta)^{\beta-1} \mathbb{1}_{\{\beta \geq 1\}} \right] \\
& \quad + \mathbb{1}_{\{\beta < 1\}} \int_\Delta^\infty e^{-(1-\delta_1)[\varepsilon + \frac{1}{\Delta}]^{\frac{1}{\theta}-1-\delta_2} z^{\frac{1}{\theta}-\delta_2}} dz \\
& \quad + \mathbb{1}_{\{\beta \geq 1\}} \left(\varepsilon + \frac{1}{\Delta} \right)^{\beta-1} \int_\Delta^\infty z^{\beta-1} e^{-(1-\delta_1)[\varepsilon + \frac{1}{\Delta}]^{\frac{1}{\theta}-1-\delta_2} z^{\frac{1}{\theta}-\delta_2}} dz \\
& \leq \Delta \left[\mathbb{1}_{\{\beta < 1\}} + (1 + \varepsilon \Delta)^{\beta-1} \mathbb{1}_{\{\beta \geq 1\}} \right] \\
& \quad + \mathbb{1}_{\{\beta < 1\}} \frac{\theta}{1 - \theta \delta_2} \left[\frac{1}{(1 - \delta_1)(\varepsilon + \frac{1}{\Delta})^{\frac{1}{\theta}-1-\delta_2}} \right]^{\frac{\theta}{1-\theta \delta_2}} \Gamma \left(\frac{\theta}{1 - \theta \delta_2} \right) \\
& \quad + \mathbb{1}_{\{\beta \geq 1\}} \left(\varepsilon + \frac{1}{\Delta} \right)^{\beta-1} \frac{\theta}{1 - \theta \delta_2} \left[\frac{1}{(1 - \delta_1)(\varepsilon + \frac{1}{\Delta})^{\frac{1}{\theta}-1-\delta_2}} \right]^{\frac{\beta \theta}{1-\theta \delta_2}} \Gamma \left(\frac{\beta \theta}{1 - \theta \delta_2} \right)
\end{aligned}$$

which is also bounded. This achieves the proof of Lemma 2.1.

Proof of Theorem 3.1. Remark that

$$\frac{\overline{CTM}_{\beta,k/n}}{CTM_{\beta,k/n}} - 1 = \frac{[U_X(n/k)]^\beta}{CTM_{\beta,k/n}} \left\{ \frac{\overline{CTM}_{\beta,k/n}(\hat{u}_{n,k})}{[U_X(n/k)]^\beta} - \frac{CTM_{\beta,k/n}}{[U_X(n/k)]^\beta} \right\}. \quad (12)$$

The leading factor in the right-hand side of (12) tends to 1 by Lemma 2.1. Thus we only need to look at the term in the braces. First, we note that

$$T_n \left(\frac{n}{k} \overline{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) \hat{u}_{n,k} \right) \right) = 1 \text{ a.s.},$$

from which we deduce that

$$\begin{aligned}
& \frac{\overline{CTM}_{\beta,k/n}(\hat{u}_{n,k})}{[U_X(n/k)]^\beta} - \frac{CTM_{\beta,k/n}}{[U_X(n/k)]^\beta} \\
= & \left\{ \left[\frac{X_{n-k,n}}{U_X(n/k)} \right]^\beta - 1 \right\} + \beta \frac{1}{\sqrt{k}} \frac{a_X(n/k)}{U_X(n/k)} \int_0^\infty W(e^{-v}) dv \\
& + \beta \frac{a_X(n/k)}{U_X(n/k)} \int_0^\infty \left[1 + \frac{a_X(n/k)}{U_X(n/k)} (\hat{u}_{n,k} + v) \right]^{\beta-1} \\
& \quad \times \left[T_n \left(\frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) \right) - \frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) \right. \\
& \quad \left. - \frac{1}{\sqrt{k}} W \left(\frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) \right) \right] dv \\
& + \beta \frac{a_X(n/k)}{U_X(n/k)} \int_0^\infty \left[1 + \frac{a_X(n/k)}{U_X(n/k)} (\hat{u}_{n,k} + v) \right]^{\beta-1} \\
& \quad \times \left[\frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) - e^{-(\hat{u}_{n,k}+v)} \right] dv \\
& + \beta \frac{a_X(n/k)}{U_X(n/k)} e^{-\hat{u}_{n,k}} \int_0^\infty \left[\left(1 + \frac{a_X(n/k)}{U_X(n/k)} (\hat{u}_{n,k} + v) \right)^{\beta-1} - \left(1 + \frac{a_X(n/k)}{U_X(n/k)} v \right)^{\beta-1} \right] e^{-v} dv \\
& + \beta \frac{a_X(n/k)}{U_X(n/k)} \left(e^{-\hat{u}_{n,k}} - 1 \right) \int_0^\infty \left(1 + \frac{a_X(n/k)}{U_X(n/k)} v \right)^{\beta-1} e^{-v} dv \\
& + \beta \frac{a_X(n/k)}{U_X(n/k)} \int_0^\infty \left[1 + \frac{a_X(n/k)}{U_X(n/k)} v \right]^{\beta-1} \left[e^{-v} - \frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) v \right) \right] dv \\
& + \beta \frac{1}{\sqrt{k}} \frac{a_X(n/k)}{U_X(n/k)} \int_0^\infty \left[1 + \frac{a_X(n/k)}{U_X(n/k)} (\hat{u}_{n,k} + v) \right]^{\beta-1} \\
& \quad \times \left[W \left(\frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) \right) - W \left(e^{-(\hat{u}_{n,k}+v)} \right) \right] dv \\
& + \beta \frac{1}{\sqrt{k}} \frac{a_X(n/k)}{U_X(n/k)} \int_0^\infty \left[1 + \frac{a_X(n/k)}{U_X(n/k)} (\hat{u}_{n,k} + v) \right]^{\beta-1} \left[W \left(e^{-(\hat{u}_{n,k}+v)} \right) - W \left(e^{-v} \right) \right] dv \\
& + \beta \frac{1}{\sqrt{k}} \frac{a_X(n/k)}{U_X(n/k)} \int_0^\infty \left[\left(1 + \frac{a_X(n/k)}{U_X(n/k)} (\hat{u}_{n,k} + v) \right)^{\beta-1} - 1 \right] W \left(e^{-v} \right) dv \\
= & \sum_{i=1}^{10} T_{i,n}.
\end{aligned}$$

We will study all the terms separately.

Term $T_{1,n}$. According to Lemma 8.2, we have

$$\sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} T_{1,n} = \beta \sqrt{k} \frac{X_{n-k,n} - U_X(n/k)}{a_X(n/k)} (1 + o_{\mathbb{P}}(1)) = \beta W(1) + o_{\mathbb{P}}(1). \quad (13)$$

Term $T_{2,n}$. Clearly

$$\sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} T_{2,n} = \beta \int_0^\infty W(e^{-v}) dv. \quad (14)$$

Term $T_{3,n}$. Using Lemmas 8.1 and 8.2, with arbitrary large probability, for n large, $c > 0$ and $0 < \eta < 1/2$, we have

$$\begin{aligned} \sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} |T_{3,n}| &\leq \sup_{0 < y \leq e^\delta + c} \frac{|\sqrt{k}[T_n(y) - y] - W(y)|}{y^\eta} \\ &\times \beta \int_0^\infty \left[1 + \frac{a_X(n/k)}{U_X(n/k)} (\hat{u}_{n,k} + v) \right]^{\beta-1} \left(\frac{\bar{F}_X(U_X(\frac{n}{k}) + a_X(\frac{n}{k}) (\hat{u}_{n,k} + v))}{\bar{F}_X(U_X(\frac{n}{k}))} \right)^\eta dv. \end{aligned}$$

The supremum term is negligible a.s. according to (9) and a treatment similar to that of the proof of Lemma 2.1 with Potter bounds ensures that the integral is bounded. This yields

$$\sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} T_{3,n} = o_{\mathbb{P}}(1). \quad (15)$$

Term $T_{4,n}$. Let us define

$$I_n(x; \beta) = \sqrt{k} \int_0^\infty \left| \frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (x + v) \right) - e^{-(x+v)} \right| dv^\beta.$$

Using Lemma 8.2, with arbitrary large probability, for n large, $0 < \varepsilon < 1/\delta$ and any $\Delta > 0$, we have

$$\begin{aligned} &\sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} |T_{4,n}| \\ &\leq \mathbb{1}_{\{\beta < 1\}} \beta (1 - \delta\varepsilon)^{\beta-1} \sqrt{k} \int_0^\infty \left| \frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) - e^{-(\hat{u}_{n,k} + v)} \right| dv \\ &\quad + \mathbb{1}_{\{\beta \geq 1\}} \beta (1 + \varepsilon(\delta + \Delta))^{\beta-1} \sqrt{k} \int_0^\Delta \left| \frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) - e^{-(\hat{u}_{n,k} + v)} \right| dv \\ &\quad + \mathbb{1}_{\{\beta \geq 1\}} \left(\varepsilon + \frac{1 + \varepsilon\delta}{\Delta} \right)^{\beta-1} \sqrt{k} \int_\Delta^\infty \left| \frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) - e^{-(\hat{u}_{n,k} + v)} \right| dv^\beta \\ &\leq \left(\sup_{x \in [-\delta, \delta]} I_n(x; 1) \right) \left\{ \mathbb{1}_{\{\beta < 1\}} \beta (1 - \delta\varepsilon)^{\beta-1} + \mathbb{1}_{\{\beta \geq 1\}} \beta (1 + \varepsilon(\delta + \Delta))^{\beta-1} \right\} \\ &\quad + \left(\sup_{x \in [-\delta, \delta]} I_n(x; \beta) \right) \mathbb{1}_{\{\beta \geq 1\}} \left(\varepsilon + \frac{1 + \varepsilon\delta}{\Delta} \right)^{\beta-1}. \end{aligned}$$

Thus, Lemma 8.3 implies that

$$\sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} T_{4,n} = o_{\mathbb{P}}(1). \quad (16)$$

Term $T_{5,n}$. By Lemma 8.2, using the mean value theorem, with $\xi \in (0, 1)$, we have with large probability, for n large and $0 < \varepsilon < 1/\delta$

$$\begin{aligned} \sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} |T_{5,n}| &\leq \beta |\beta - 1| \frac{a_X(n/k)}{U_X(n/k)} \sqrt{k} \hat{u}_{n,k} \int_0^\infty \left[1 + \frac{a_X(n/k)}{U_X(n/k)} (v + \xi \hat{u}_{n,k}) \right]^{\beta-2} e^{-v} dv (1 + o_{\mathbb{P}}(1)) \\ &\leq \beta |\beta - 1| \frac{a_X(n/k)}{U_X(n/k)} \sqrt{k} \hat{u}_{n,k} (1 + o_{\mathbb{P}}(1)) \\ &\quad \times \left\{ \mathbb{1}_{\{\beta < 2\}} (1 - \varepsilon \delta)^{\beta-2} \int_0^\infty e^{-v} dv + \mathbb{1}_{\{\beta \geq 2\}} \int_0^\infty [1 + \varepsilon(v + \delta)]^{\beta-2} e^{-v} dv \right\} \\ &= o_{\mathbb{P}}(1), \end{aligned} \tag{17}$$

again by Lemma 8.2 and since $\frac{a_X(n/k)}{U_X(n/k)} \rightarrow 0$.

Term $T_{6,n}$. According to Lemma 8.2, we have

$$\sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} T_{6,n} = -\beta \sqrt{k} \left(\frac{X_{n-k,n} - U_X(n/k)}{a_X(n/k)} \right) + o_{\mathbb{P}}(1) = -\beta W(1) + o_{\mathbb{P}}(1). \tag{18}$$

Term $T_{7,n}$. It can be handled similarly as $T_{4,n}$. This yields

$$\sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} T_{7,n} = o(1). \tag{19}$$

Term $T_{8,n}$. With arbitrary large probability, for n large, $0 < \varepsilon < 1/\delta$ and any $\Delta > 0$, we have

$$\begin{aligned} \sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} |T_{8,n}| &\leq \left\{ \mathbb{1}_{\{\beta < 1\}} \beta (1 - \delta \varepsilon)^{\beta-1} + \mathbb{1}_{\{\beta \geq 1\}} \beta (1 + \varepsilon(\delta + \Delta))^{\beta-1} \right\} \mathbb{W}_n(1) \\ &\quad + \mathbb{1}_{\{\beta \geq 1\}} \left(\varepsilon + \frac{1 + \varepsilon \delta}{\Delta} \right)^{\beta-1} \mathbb{W}_n(\beta), \end{aligned} \tag{20}$$

where

$$\mathbb{W}_n(\beta) := \int_0^\infty \left| W \left(\frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) \right) - W \left(e^{-(\hat{u}_{n,k} + v)} \right) \right| dv^\beta.$$

For $T > 0$, we have

$$\begin{aligned} \mathbb{W}_n(\beta) &\leq \int_T^\infty \left| W \left(e^{-(\hat{u}_{n,k} + v)} \right) \right| dv^\beta + \int_T^\infty \left| W \left(\frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) \right) \right| dv^\beta \\ &\quad + \int_0^T \left| W \left(\frac{n}{k} \bar{F}_X \left(U_X \left(\frac{n}{k} \right) + a_X \left(\frac{n}{k} \right) (\hat{u}_{n,k} + v) \right) \right) - W \left(e^{-(\hat{u}_{n,k} + v)} \right) \right| dv^\beta \\ &=: \mathbb{W}_n^{(1)}(\beta) + \mathbb{W}_n^{(2)}(\beta) + \mathbb{W}_n^{(3)}(\beta). \end{aligned}$$

Then, for $a > 0$ arbitrary, we have

$$\mathbb{P}(\mathbb{W}_n(\beta) > a) \leq \mathbb{P} \left(\mathbb{W}_n^{(1)}(\beta) > \frac{a}{4} \right) + \mathbb{P} \left(\mathbb{W}_n^{(2)}(\beta) > \frac{a}{4} \right) + \mathbb{P} \left(\mathbb{W}_n^{(3)}(\beta) > \frac{a}{2} \right).$$

Clearly, for $\eta \in (0, 1/2)$, we have

$$\mathbb{P}\left(\mathbb{W}_n^{(1)}(\beta) > \frac{a}{4}\right) \leq \mathbb{P}\left(\sup_{0 < y \leq e^\delta} \frac{|W(y)|}{y^\eta} > \frac{a\eta^\beta}{4\beta} \frac{e^{-\eta\delta}}{\Gamma_T(\beta, T\eta)}\right) + \mathbb{P}(|\hat{u}_{n,k}| > \delta),$$

and similarly, by Lemma 8.1, $c > 0$ and n large

$$\begin{aligned} \mathbb{P}\left(\mathbb{W}_n^{(2)}(\beta) > \frac{a}{4}\right) &\leq \mathbb{P}\left(\sup_{0 < y \leq e^{\delta+c}} \frac{|W(y)|}{y^\eta} \int_T^\infty \left[\frac{n}{k} \bar{F}_X\left(U_X\left(\frac{n}{k}\right) + a_X\left(\frac{n}{k}\right)(-\delta+v)\right)\right]^\eta dv^\beta > \frac{a}{4}\right) \\ &\quad + \mathbb{P}(|\hat{u}_{n,k}| > \delta). \end{aligned}$$

Following the lines of proof of Lemma 2.1 with the Potter bounds, the integral from the above display can be made arbitrary small by taking T large. Thus, combining Lemma 8.2 with Lemma 2 in Cai et al. (2015), and by choosing T large enough, for $\varepsilon > 0$, there exists n_1 such that for $n > n_1$ we have

$$\mathbb{P}\left(\mathbb{W}_n^{(1)}(\beta) > \frac{a}{4}\right) + \mathbb{P}\left(\mathbb{W}_n^{(2)}(\beta) > \frac{a}{4}\right) \leq \frac{\varepsilon}{2}.$$

With this specific value of T , we have

$$\begin{aligned} \mathbb{P}\left(\mathbb{W}_n^{(3)}(\beta) > \frac{a}{2}\right) &\leq \mathbb{P}\left(\sup_{y \in [-\delta, \delta+T]} \left|W\left(\frac{n}{k} \bar{F}_X\left(U_X\left(\frac{n}{k}\right) + a_X\left(\frac{n}{k}\right)y\right)\right) - W(e^{-y})\right| > \frac{a}{2T^\beta}\right) \\ &\quad + \mathbb{P}(|\hat{u}_{n,k}| > \delta). \end{aligned}$$

Lemmas 8.1 and 8.2 combined with the uniform continuity of W on compact sets yield, for $n > n_2$

$$\mathbb{P}\left(\mathbb{W}_n^{(3)}(\beta) > \frac{a}{2}\right) \leq \frac{\varepsilon}{2},$$

from which we deduce that, for $n > \max(n_1, n_2)$

$$\mathbb{P}(\mathbb{W}_n(\beta) > a) \leq \varepsilon. \tag{21}$$

Combining (20) and (21), we have

$$\sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} T_{8,n} = o_{\mathbb{P}}(1). \tag{22}$$

Term $T_{9,n}$. It can be handled similarly as $T_{8,n}$, except for

$$\widetilde{\mathbb{W}}_n^{(3)}(\beta) := \int_0^T \left|W\left(e^{-(\hat{u}_{n,k}+v)}\right) - W(e^{-v})\right| dv^\beta,$$

where we can use the bound

$$\mathbb{P}\left(\widetilde{\mathbb{W}}_n^{(3)}(\beta) > \frac{a}{2}\right) \leq \mathbb{P}\left(\sup_{v \in [0, T], |x| \leq k^{-1/4}} \left|W\left(e^{-(v+x)}\right) - W\left(e^{-v}\right)\right| > \frac{a}{2T^\beta}\right) + \mathbb{P}\left(|\widehat{u}_{n,k}| > k^{-1/4}\right).$$

By the uniform continuity of W on compact sets and Lemma 8.2, we have for $n > n_2$

$$\mathbb{P}\left(\widetilde{\mathbb{W}}_n^{(3)}(\beta) > \frac{a}{2}\right) \leq \frac{\varepsilon}{2},$$

from which we deduce that

$$\sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} T_{9,n} = o_{\mathbb{P}}(1). \quad (23)$$

Term $T_{10,n}$. By the mean value theorem, with $\xi \in (0, 1)$, we have with large probability, for n large, and $0 < \varepsilon < 1/\delta$ and $0 < \eta < 1/2$,

$$\begin{aligned} \sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} |T_{10,n}| &= \left| \beta(\beta - 1) \frac{a_X(n/k)}{U_X(n/k)} \int_0^\infty [\widehat{u}_{n,k} + v] \left[1 + \xi \frac{a_X(n/k)}{U_X(n/k)} (\widehat{u}_{n,k} + v)\right]^{\beta-2} W(e^{-v}) dv \right| \\ &\leq |\beta| |\beta - 1| \frac{a_X(n/k)}{U_X(n/k)} \\ &\quad \times \left\{ \mathbb{1}_{\{\beta < 2\}} [1 - \delta\varepsilon]^{\beta-2} \int_0^\infty [\delta + v] |W(e^{-v})| dv \right. \\ &\quad \left. + \mathbb{1}_{\{\beta \geq 2\}} \int_0^\infty [\delta + v] [1 + \varepsilon(\delta + v)]^{\beta-2} |W(e^{-v})| dv \right\} \\ &\leq |\beta| |\beta - 1| \frac{a_X(n/k)}{U_X(n/k)} \\ &\quad \times \left\{ \mathbb{1}_{\{\beta < 2\}} [1 - \delta\varepsilon]^{\beta-2} \sup_{0 < y \leq 1} \frac{|W(y)|}{y^\eta} \int_0^\infty [\delta + v] e^{-\eta v} dv \right. \\ &\quad \left. + \mathbb{1}_{\{\beta \geq 2\}} \sup_{0 < y \leq 1} \frac{|W(y)|}{y^\eta} \int_0^\infty [\delta + v] [1 + \varepsilon(\delta + v)]^{\beta-2} e^{-\eta v} dv \right\} \\ &\leq C \frac{a_X(n/k)}{U_X(n/k)}, \end{aligned} \quad (24)$$

by Lemma 2 in Cai et al. (2015), and where C is a positive constant.

Combining (13)-(19) with (22)-(24), Theorem 3.1 follows.

Proof of Theorem 4.1. Consider the following decomposition

$$\begin{aligned}
& \sqrt{k} \left(\frac{\widehat{CTM}_{\beta,p}}{CTM_{\beta,p}} - 1 \right) \\
&= \sqrt{k} \left(\underbrace{\left\{ \left(\frac{\log 1/p}{\log n/k} \right)^{\beta[\widehat{\theta}_k - \theta]} \right\}}_{T_{11,n}} \underbrace{\left\{ \frac{\overline{CTM}_{\beta, \frac{k}{n}}}{CTM_{\beta, \frac{k}{n}}} \right\}}_{T_{12,n}} \underbrace{\left\{ \frac{(\log n/k)^{-\beta\theta} CTM_{\beta, \frac{k}{n}}}{(\log 1/p)^{-\beta\theta} CTM_{\beta,p}} \right\}}_{T_{13,n}} - 1 \right) \\
&= \sqrt{k}(T_{11,n} - 1)T_{12,n}T_{13,n} + \sqrt{k}(T_{12,n} - 1)T_{13,n} + \sqrt{k}(T_{13,n} - 1).
\end{aligned}$$

By a Taylor series expansion we obtain

$$\sqrt{k}(T_{11,n} - 1) = \beta \log(\lambda) \sqrt{k} [\widehat{\theta}_k - \theta] (1 + o_{\mathbb{P}}(1)),$$

and hence $\sqrt{k}(T_{11,n} - 1) \xrightarrow{d} \beta \log(\lambda) \Gamma$.

As for $T_{12,n}$, we have

$$\sqrt{k}(T_{12,n} - 1) = \sqrt{k} \frac{U_X(n/k)}{a_X(n/k)} \left\{ \frac{\overline{CTM}_{\beta, \frac{k}{n}}}{CTM_{\beta, \frac{k}{n}}} - 1 \right\} \frac{a_X(n/k)}{U_X(n/k)},$$

yielding $\sqrt{k}(T_{12,n} - 1) = o_{\mathbb{P}}(1)$ by Theorem 3.1.

Finally, we analyse $T_{13,n}$, for which we use the decomposition

$$\begin{aligned}
\sqrt{k}(T_{13,n} - 1) &= \sqrt{k} \left\{ \frac{CTM_{\beta, \frac{k}{n}}/[U_X(n/k)]^{\beta}}{CTM_{\beta,p}/[U_X(1/p)]^{\beta}} - 1 \right\} \left(\frac{U_X(n/k)}{U_X(1/p)} \right)^{\beta} \left(\frac{\log n/k}{\log 1/p} \right)^{-\beta\theta} \\
&\quad + \sqrt{k} \left(\left(\frac{U_X(n/k)}{U_X(1/p)} \right)^{\beta} \left(\frac{\log n/k}{\log 1/p} \right)^{-\beta\theta} - 1 \right). \tag{25}
\end{aligned}$$

To handle the first term of (25), we remark that, according to (10)

$$\begin{aligned}
\frac{CTM_{\beta, \frac{k}{n}}}{[U_X(n/k)]^{\beta}} &= 1 + \beta \frac{a_X(n/k)}{U_X(n/k)} \left\{ 1 + \int_0^{\infty} \left[\left(1 + \frac{a_X(n/k)}{U_X(n/k)} z \right)^{\beta-1} - 1 \right] e^{-z} dz \right. \\
&\quad \left. + \int_0^{\infty} \left(1 + \frac{a_X(n/k)}{U_X(n/k)} z \right)^{\beta-1} \left(\frac{\overline{F}_X(U_X(n/k) + a_X(n/k)z)}{\overline{F}_X(U_X(n/k))} - e^{-z} \right) dz \right\} \\
&= 1 + o\left(\frac{1}{\sqrt{k}}\right),
\end{aligned}$$

and also we have a similar result for $CTM_{\beta, \frac{1}{p}}/[U_X(1/p)]^\beta$. This yields

$$\frac{CTM_{\beta, \frac{k}{n}}/[U_X(n/k)]^\beta}{CTM_{\beta, p}/[U_X(1/p)]^\beta} - 1 = o\left(\frac{1}{\sqrt{k}}\right).$$

Now, for the last term of (25), we have

$$\left(\frac{U_X(n/k)}{U_X(1/p)}\right)^\beta \left(\frac{\log n/k}{\log 1/p}\right)^{-\beta\theta} - 1 = \left(\frac{\ell_U(\log n/k)}{\ell_U(\log 1/p)}\right)^\beta - 1 = \left(\frac{\ell_X(U_X(1/p))}{\ell_X(U_X(n/k))}\right)^{\beta\theta} - 1,$$

since $\ell_U(\log x) = [\ell_X(U_X(x))]^{-\theta}$. Now, using the mean value theorem with $u_n \in [U_X(n/k); U_X(1/p)]$, we have, for $\delta_1, \delta_2 > 0$ and n large,

$$\begin{aligned} \left| \frac{\ell_X(U_X(1/p))}{\ell_X(U_X(n/k))} - 1 \right| &= \frac{|\ell'_X(u_n)|}{\ell_X(U_X(n/k))} \left[U_X\left(\frac{1}{p}\right) - U_X\left(\frac{n}{k}\right) \right] \\ &= \frac{u_n |\ell'_X(u_n)|}{\ell_X(u_n)} \frac{\ell_X(u_n)}{\ell_X(U_X(n/k))} \frac{U_X(n/k)}{u_n} \left(\frac{U_X(1/p)}{U_X(n/k)} - 1 \right) \\ &\leq (1 + \delta_1) \frac{u_n |\ell'_X(u_n)|}{\ell_X(u_n)} \left(\frac{u_n}{U_X(n/k)} \right)^{\delta_2} \left(\frac{U_X(1/p)}{U_X(n/k)} - 1 \right) \\ &\leq (1 + \delta_1) \sup_{x \geq U_X(n/k)} \frac{x |\ell'_X(x)|}{\ell_X(x)} \left(\frac{U_X(1/p)}{U_X(n/k)} \right)^{\delta_2} \left(\frac{U_X(1/p)}{U_X(n/k)} - 1 \right) \\ &= o\left(\frac{1}{\sqrt{k}}\right), \end{aligned}$$

since

$$\frac{U_X(1/p)}{U_X(n/k)} - 1 = \left[\left(\frac{\log 1/p}{\log n/k} \right)^\theta - 1 \right] \frac{\ell_U(\log 1/p)}{\ell_U(\log n/k)} + \left[\frac{\ell_U(\log 1/p)}{\ell_U(\log n/k)} - 1 \right] = O(1).$$

Combining the above gives $\sqrt{k}(T_{13,n} - 1) = o(1)$.

Acknowledgement

The authors sincerely thank the referees, the Associate Editor and Editor for their helpful comments on our paper. The research of Armelle Guillou was supported by the French National Research Agency under the grant ANR-19-CE40-0013-01/ExtremReg project and by the CNRS International Research Network MaDeF.

Supporting Information

Additional information for this article is available online: additional simulation results concerning the estimation of $CTM_{2,p}$.

References

- Abramowitz, M., & Stegun, I. (1983). Handbook of mathematical functions with formulas, graphs, and mathematical tables, applied mathematics series, New York.
- Beirlant, J., Broniatowski, M., Teugels, J.L., & Vynckier, P. (1995). The mean residual life function at great age: applications to tail estimation. *Journal of Statistical Planning and Inference*, *45*, 21–48.
- Bingham, N.H., Goldie, C.M., & Teugels, J.L. (1987). *Regular Variation*. Cambridge University Press.
- Broniatowski, M. (1993). On the estimation of the Weibull tail coefficient. *Journal of Statistical Planning and Inference*, *35*, 349–366.
- Cai, J.J., Einmahl, J.H.J., de Haan, L., & Zhou, C. (2015). Estimation of the marginal expected shortfall: the mean when a related variable is extreme. *Journal of the Royal Statistical Society Series B*, *77*, 417–442.
- Coles, S. (2013). *An introduction to the statistical modeling of extreme events*. Springer.
- Cooley, D. (2013). Return periods and return levels under climate change. *Extremes in a changing climate: Detection, analysis and uncertainty, Chapter 4*, A. AghaKouchak, D. Easterling, and K. Hsu, eds., 65, Springer, New York.
- Davis, R.A., & Mikosch, T. (2009). The extremogram: A correlogram for extreme events. *Bernoulli*, *15*, 977–1009.
- Davison, A.C., Padoan, S.A., & Stupfler, G. (2023). Tail risk inference via expectiles in heavy-tailed time series. *Journal of Business & Economic Statistics*, *41*, 876–889.
- de Haan, L., & Ferreira, A. (2006). *Extreme value theory, an introduction*. Springer.

- de Wet, T., Goegebeur, Y., Guillou, A., & Osmann, M. (2016). Kernel regression with Weibull-type tails. *Annals of the Institute of Statistical Mathematics*, *68*, 1135–1162.
- Diebolt, J., Gardes, L., Girard, S., & Guillou, A. (2008). Bias-reduced estimators of the Weibull tail-coefficient. *TEST*, *17*, 311–331.
- Dierckx, G., Beirlant, J., de Waal, D., & Guillou, A. (2009). A new estimation method for Weibull-type tails based on the mean excess function. *Journal of Statistical Planning and Inference*, *139*, 1905–1920.
- Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli*, *9*, 617–657.
- Drees, H., & Rootzén, H. (2010). Limit theorems for empirical processes of cluster functionals. *The Annals of Statistics*, *38*, 2145–2186.
- Eini, E.J., & Khaloozadeh, H. (2021). Tail conditional moment for generalized skew-elliptical distributions. *Journal of Applied Statistics*, *48*, 2285–2305.
- Einmahl, J. H. J., de Haan, L., & Li, D. (2006). Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. *The Annals of Statistics*, *34*, 1987–2014.
- El Methni, J., Gardes, L., & Girard, S. (2014). Non-parametric estimation of extreme risk measures from conditional heavy-tailed distributions. *Scandinavian Journal of Statistics*, *41*, 988–1012.
- Feng, S., Gao, D., Liao, F., Zhou, F., & Wang, X. (2016). The health effects of ambient PM_{2.5} and potential mechanisms. *Ecotoxicology and Environmental Safety*, *128*, 67–74.
- Gardes, L., & Girard, S. (2005). Estimating extreme quantiles of Weibull tail distributions. *Communications in Statistics – Theory and Methods*, *34*, 1065–1080.
- Gardes, L., & Girard, S. (2008). Estimation of the Weibull-tail coefficient with linear combination of upper order statistics. *Journal of Statistical Planning and Inference*, *138*, 1416–1427.
- Gardes, L., & Girard, S. (2016). On the estimation of the functional Weibull tail-coefficient. *Journal of Multivariate Analysis*, *146*, 29–45.

- Gardes, L., Girard, S., & Guillou, A. (2011). Weibull tail-distributions revisited: a new-look at some tail estimators. *Journal of Statistical Planning and Inference*, *141*, 429–444.
- Girard, S. (2004). A Hill type estimator of the Weibull tail coefficient. *Communications in Statistics – Theory and Methods*, *33*, 205–234.
- Goegebeur, Y., Beirlant, J., & de Wet, T. (2010). Generalized kernel estimators for the Weibull tail coefficient. *Communications in Statistics – Theory and Methods*, *39*, 3695–3716.
- Goegebeur, Y., & Guillou, A. (2010). Goodness-of-fit testing for Weibull-type behavior. *Journal of Statistical Planning and Inference*, *140*, 1417–1436.
- Goegebeur, Y., & Guillou, A. (2011). A weighted mean excess function approach to the estimation of Weibull-type tails. *TEST*, *20*, 138–162.
- Goegebeur, Y., Guillou, A., Pedersen, T., & Qin, J. (2022). Extreme-value based estimation of the conditional tail moment with application to reinsurance rating. *Insurance: Mathematics and Economics*, *107*, 102–122.
- Goegebeur, Y., Guillou, A., & Qin, J. (2019). Bias-corrected estimation for conditional Pareto-type distributions with random right censoring. *Extremes*, *22*, 459–498.
- Goegebeur, Y., Guillou, A., & Qin, J. (2023). Estimation of marginal excess moments for Weibull-type distributions, <https://hal.science/hal-04012275>.
- Goegebeur, Y., Guillou, A., & Qin, J. (2024). Conditional tail moment and reinsurance premium estimation under random right censoring. *TEST*, *33*, 230–250.
- Goegebeur, Y., Guillou, A., & Rietsch, T. (2015). Robust conditional Weibull-type estimation. *Annals of the Institute of Statistical Mathematics*, *67*, 479–514.
- Gomes, M.I., & Pestana, D. (2007). A sturdy reduced-bias extreme quantile (var) estimator. *Journal of the American Statistical Association*, *102*, 280–292.
- Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, *5*, 1163–1174.
- Hong, J., & Elshahat, A. (2010). Conditional tail variance and conditional tail skewness. *Journal*

of Financial and Economic Practice, 10, 147–156.

Jorion, P. (2007). *Value at risk: the new benchmark for managing financial risk*. McGraw-Hill, New York.

Klüppelberg, C., & Villaseñor, J.A. (1993). Estimation of distribution tails – a semiparametric approach. *Deutschen Gesellschaft für Versicherungsmathematik*, 21, 213–235.

Li, Z., Wen, Q., & Zhang, R. (2017). Sources, health effects and control strategies of indoor fine particulate matter (PM_{2.5}): a review. *Science of The Total Environment*, 586, 610–622.

Linton, O., & Xiao, Z. (2013). Estimation of and inference about the expected shortfall for time series with infinite variance. *Econometric Theory*, 29, 771–807.

Padoan, S., Rizzelli, S., & Schiavone, M. (2023). Marginal expected shortfall inference under multivariate regular variation. arXiv:2304.07578.

Padoan, S.A., & Stupfler, G. (2022). Joint inference on extreme expectiles for multivariate heavy-tailed distributions. *Bernoulli*, 28, 1021–1048.

Salas, J. D., & Obeysekera, J. (2014). Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. *Journal of Hydrologic Engineering*, 19, 554–568.

Shi, H., Dong, Z., Xiao, N., & Huang, Q. (2021). Wind speed distributions used in wind energy assessment: a review. *Frontiers in Energy Research*, 9, 769920.

Singh, V.P., Jain, S.K., & Tyagi, A. (2007). *Risk and reliability analysis: A handbook for civil and environmental engineers*. Reston, VA: American Society of Civil Engineers.

Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark.

Email address: Yuri.Goegebeur@imada.sdu.dk; qin@imada.sdu.dk

Institut Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67084 Strasbourg cedex, France.

Email address: armelle.guillou@math.unistra.fr