On tail trend detection: modeling relative risk[*]

Laurens de Haan

University of Lisbon

Erasmus University Rotterdam

Albert Klein Tank

KNMI, Royal Netherlands

Meteorological Institute

Cláudia Neves

CEAUL and University of Aveiro

Abstract

The climate change dispute is about changes over time of environmental characteristics (such as rainfall). Some people say that a possible change is not so much in the mean but rather in the extreme phenomena (that is, the average rainfall may not change much but heavy storms may become more or less frequent). The paper studies changes over time in the probability that some high threshold is exceeded. The model is such that the threshold does not need to be specified, the results hold for any high threshold. For simplicity a certain linear trend is studied depending on one real parameter. Estimation and testing procedures (is there a trend?) are developed. Simulation results are presented. The method is applied to trends in heavy rainfall at 18 gauging stations across Germany and The Netherlands. A tentative conclusion is that the trend seems to depend on whether or not a station is close to the sea.

KEY WORDS AND PHRASES: extreme value distribution, regular variation, extreme rainfall

1 Introduction

In the climate change dispute some people suggest (Klein Tank and Können (2003); Groisman et al. (2005); Alexander et al. (2006); Zolina et al. (2009)) that perhaps there is no or little change in the mean of the probability distribution of daily rainfall over time but there is a change in the tail that is, more extreme events occur more frequently. The present paper – like Smith (1989); Hall and Tajvidi (2000); Hanel et al. (2009) – considers a trend in extremes from the point of view of extreme value theory.

If one wants to concentrate on a trend connected with extreme events rather than with the central part of the probability distribution function F, one should look at a high quantile $F^{\leftarrow}(p)$ (i.e. the inverse of F) for p close to one or at the exceedance probability $1 - F(x)$ at a high level x. Hence we consider the limit behavior of $F^{\leftarrow}(p)$ as $p \uparrow 1$ or $F(x)$ as $x \uparrow x^*$, which is the right end point of the probability distribution ($x^* := \sup\{x : F(x) < 1\}$). Since the limit relation for F is simpler than for $F^{\leftarrow}$, we concentrate on the behavior of $F(x)$ as $x \uparrow x^*$.

Two models for trends connected with extremes spring forward: a trend in the extreme value index γ or a situation with constant extreme value index but with heteroscedastic observations (in the tail). We consider the latter option here. In that case probabilities of extreme events are not assumed asymptotically equal (which would mean tail equivalence cf. Resnick, 1971) but asymptotically comparable as follows.

Consider random variables $X(s)$ where $s \geq 0$ is time. Write $F_s(x) := P\{X(s) \leq x\}$ for $x \in \mathbb{R}$. We assume that for all $s > 0$

$$\frac{1 - F_s(x)}{1 - F_0(x)}$$

tends to a positive constant $d(s)$ for all $s > 0$ when x tends to the right endpoint x^* of F_0 where d is a positive continuous function. Hence the exceedance probability at time s is systematically a factor times the exceedance probability at time zero. We consider a simple model for relative risk and assume that for some real trend constant c and all $s \geq 0$

$$\lim_{x \uparrow x^*} \frac{1 - F_s(x)}{1 - F_0(x)} = e^{cs}. \tag{1}$$

This means that for example (with $s = 1$ and $c = 1$) that the probability of any extreme event taking place at time s is e times the corresponding probability at time zero. For s small the limit function is approximately linear.

For our analysis we shall need the context of extreme value theory that is, we assume that the distribution function F_0 is in the domain of attraction of some extreme value distribution G_γ i.e., there exist sequences of constants $a_n > 0$ and b_n ($n = 1, 2, \ldots$) such that the normalized maximum of a sample from F_0 converges to G_γ for some $\gamma \in \mathbb{R}$:

$$\lim_{n \to \infty} F_0^n(a_n x + b_n) = G_\gamma(x) := \exp\{-(1 + \gamma x)^{-1/\gamma}\} \tag{2}$$

for all x for which $1 + \gamma x > 0$ (notation $F_0 \in \mathcal{D}(G_\gamma)$). This condition is really a condition on the tail $1 - F_0$ of the probability distribution since it is equivalent to

$$\lim_{t \to \infty} t\{1 - F_0\big(a_0(t)x + b_0(t)\big)\} = (1 + \gamma x)^{-1/\gamma} \tag{3}$$

for $x > 0$ where $a_0(t) := a_{[t]}$ and $b_0(t) := b_{[t]}$ (and $[t]$ is the integral part of t) (cf. e.g. Coles, 2001;

de Haan and Ferreira, 2006).

We return to condition (1). Let us first translate this condition into a condition for high quantiles as we discussed before. Proofs follow later.

Condition (1) is equivalent to the following condition on the quantile function $F^{\leftarrow}$:

$$\lim_{p\uparrow 1} \frac{F_s^{\leftarrow}(p) - F_0^{\leftarrow}(p)}{a_0\left(\frac{1}{1-p}\right)} = \frac{e^{c\gamma s} - 1}{\gamma}. \tag{4}$$

It is convenient (and sometimes usual) to change the notation a bit at this point. Consider the functions U_s defined by

$$U_s(t) = F_s^{\leftarrow}\left(1 - \frac{1}{t}\right) = \left(\frac{1}{1 - F_s}\right)^{\leftarrow}(t) \quad \text{for } t > 1.$$

Relation (1) is seen to be equivalent to

$$\lim_{t\to\infty} \frac{U_s(t) - U_0(t)}{a_0(t)} = \frac{e^{c\gamma s} - 1}{\gamma}. \tag{5}$$

For $\gamma > 0$ this relation can be simplified. In this case it is equivalent to

$$\lim_{t\to\infty} \frac{U_s(t)}{U_0(t)} = e^{c\gamma s}. \tag{6}$$

The condition $F_0 \in \mathcal{D}(G_\gamma)$ in conjunction with (1) implies that the extreme value condition holds for any F_s with the same limit distribution i.e. (cf. (3))

$$\lim_{t\to\infty} t\left\{1 - F_s\big(a_s(t)x + b_s(t)\big)\right\} = (1 + \gamma x)^{-1/\gamma} \tag{7}$$

where $a_s > 0$ and b_s are appropriately chosen functions. We shall use the well-known facts (cf. de Haan and Ferreira, 2006, section 1.1.2) that (7) is equivalent to

$$\lim_{t\to\infty} \frac{U_s(tx) - U_s(t)}{a_s(t)} = \frac{x^\gamma - 1}{\gamma} \quad \text{for } x > 0 \tag{8}$$

and that

$$b_s(t) - U_s(t) = o\big(a_s(t)\big),\ t \to \infty\,; \quad \lim_{t\to\infty} \frac{a_s(tx)}{a_s(t)} = x^\gamma \ \text{for } x > 0. \tag{9}$$

Moreover relation (1) implies for $s > 0$

$$\lim_{t\to\infty} \frac{a_s(t)}{a_0(t)} = e^{c\gamma s}. \tag{10}$$

All the mentioned implications will be proved in Appendix A.

In the future we shall study more general (not monotone) changes in a similar manner, possibly with

adjustments enabling other appropriate estimation procedures.

The aim of the paper is to develop estimators and testing procedures for the parameter c. This will be done in a semi-parametric way on the basis of the limit relations (1), (5) and (6).

The problem of defining and estimating a trend in extreme value theory has been considered by a number of authors including Smith (1989); Hall and Tajvidi (2000); Coles (2001); Yee and Stephenson (2007) and more recently addressed by Mannshardt-Shamseldin et al. (2010). What distinguishes our approach from the traditional ones is (among others):

- The results are directly interpretable (it is about how probabilities of extreme events change over time).

- Asymptotic justification: we prove that our estimators are not only valid when the observations come from an extreme value distribution but also under the more realistic assumption that they come from a distribution in the domain of attraction.

- Some existing proposals are not completely satisfactory. A review and discussion of existing results is given in Appendix B.

Recently, smoothing techniques were combined with models for extreme values and resulted in flexible exploratory techniques. We refer the paper by Gaetan and Grigoletto (2004) in this respect. Davison and Ramesh (2000) and Hall and Tajvidi (2000) propose to use local polynomial likelihood of the Generalized Extreme Value (GEV) distribution (cf. right hand-side of (2)). Chavez-Demoulin (1999) and Pauli and Coles (2001) suggest an approach based on penalized likelihood.

The outline of this paper is as follows. Section 2 contains our proposals for estimating the trend. In section 3 consistency and asymptotic normality of the three estimators introduced in section 2 is discussed. Proofs are postponed to section 6. In section 4 we collect some simulation results for illustrating and assessing finite sample performance of the various estimators for the trend. In section 5 we apply the methods to daily rainfall at 18 stations across Germany and The Netherlands and give a tentative interpretation of the results. Indeed for some stations the probability of extreme rainfall has increased by about 2% in each decade.

2   Estimators

In this section we develop three estimators of c.

Suppose that we have repeated observations at discrete time points $0 = s_0 < s_1 < s_2 < \ldots < s_m$. It is assumed that $\left\{ X_i(s_j) \right\}_{i=1\ j=1}^{n\ \ m}$ are all independent and that $X_1(s_j), X_2(s_j), \ldots, X_n(s_j)$ have the same distribution function F_{s_j} for all j. Let $X_{1,n}(s_j) \leq X_{2,n}(s_j) \leq \ldots \leq X_{n,n}(s_j)$ be their order statistics.

Since in the limit relations (e.g. relation (4)) only high quantiles play a role, we can expect that the impact of such a condition can be detected only among the higher order statistics. Our estimators will be based on the set of k highest order statistics $\left(X_{n-k,n}(s_j), X_{n-k+1,n}(s_j), \ldots, X_{n,n}(s_j)\right)$ for each j. In order to be able to apply asymptotic theory we need to let k depend on n, $k = k_n$ and $\lim_{n\to\infty} k_n = \infty$. On the other hand we want to determine k in such a way that we deal with the tail of the distribution only. That leads to the condition $k_n = o(n)$, $n \to \infty$. For the asymptotic normality of the estimators a further restriction will be imposed on the sequence k_n.

Now we are ready to construct estimators for c.

(i) Consider first the simplest case, $\gamma > 0$. Relation (6) implies $(s \geq 0)$

$$\lim_{n\to\infty} \log U_s\left(\frac{n}{k}\right) - \log U_0\left(\frac{n}{k}\right) = c\gamma s. \tag{11}$$

We are going to replace the quantities at the left hand side by their sample analogs. It will be proved (Appendix C) that for $j = 0, 1, 2, \ldots, m$

$$\log X_{n-k,n}(s_j) - \log U_{s_j}\left(\frac{n}{k}\right) \xrightarrow[n\to\infty]{P} 0. \tag{12}$$

as $n \to \infty$. Hence under condition (11)

$$\sum_{j=1}^{m} \left(\log X_{n-k,n}(s_j) - \log X_{n-k,n}(0) - c\gamma s_j\right)^2$$

should be small. This leads to the least squares estimator

$$\hat{c}^{(1)} := \frac{\sum_{j=1}^{m} s_j \left(\log X_{n-k,n}(s_j) - \log X_{n-k,n}(0)\right)}{\hat{\gamma}_{n,k}^{+} \sum_{j=1}^{m} s_j^2}, \tag{13}$$

where $\hat{\gamma}_{n,k}^{+}$ is some estimator of $\gamma^{+} := \max(\gamma, 0)$ (see section 3). We shall discuss specific estimators of γ^{+} in section 3 and the simulations section 4.

(ii) For γ not restricted to be positive (5) leads to an estimator for c. Intuitively relation (5) means that

$$\left(1 + \hat{\gamma}_{n,k}\frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)}\right)^{\frac{1}{\hat{\gamma}_{n,k}}} \approx e^{cs_j},$$

where $\hat{\gamma}_{n,k}$ is an estimator for γ and $\hat{a}_0(n/k)$ is an estimator for $a_0(n/k)$. Define $\hat{c}^{(2)}$ by

$$\arg\min_c \sum_{j=1}^m \left\{ \log\left(1 + \hat{\gamma}_{n,k} \frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)}\right)^{\frac{1}{\hat{\gamma}_{n,k}}} - cs_j \right\}^2$$

i.e.,

$$\hat{c}^{(2)} := \frac{\sum\limits_{j=1}^m s_j \log\left(1 + \hat{\gamma}_{n,k} \frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0(n/k)}\right)^{\frac{1}{\hat{\gamma}_{n,k}}}}{\sum\limits_{j=1}^m s_j^2}. \tag{14}$$

For $\hat{\gamma}_{n,k} = 0$, the estimator is defined by continuity. Again, specific (well-known) estimators $\hat{\gamma}_{n,k}$ and $\hat{a}_0(n/k)$ will be discussed in section 3.

(iii) Finally relation (1) also leads to an estimator for c. Intuitively relation (1) means that

$$\log \frac{1 - \widehat{F}_s\big(X_{n-k,n}(0)\big)}{1 - \widehat{F}_0\big(X_{n-k,n}(0)\big)} \approx cs$$

where $\widehat{F}_s$ is the empirical distribution function at time s. Note that $1 - \widehat{F}_0\big(X_{n-k,n}(0)\big) \approx k/n$. Hence the following estimator:

$$\hat{c}^{(3)} := \frac{\sum\limits_{j=1}^m \log\left(\frac{1}{k}\sum\limits_{i=1}^n I_{\{X_i(s_j) > X_{n-k,n}(0)\}}\right)}{\sum\limits_{j=1}^m s_j}. \tag{15}$$

3 Results

(i) Let us consider $\hat{c}^{(1)}$ first and suppose that $\gamma > 0$. For part of our results we need a second order strengthening of conditions (1) and (2).

CONDITION A   Suppose there exists a positive or negative function β with $\lim_{t\to\infty}\beta(t) = 0$ such that for $x > 0$

$$\lim_{t\to\infty} \frac{\frac{U_0(tx)}{U_0(t)} - x^{\gamma^+}}{\beta(t)} = x^{\gamma^+} \frac{x^{\widetilde{\rho}} - 1}{\widetilde{\rho}} \tag{16}$$

with $\widetilde{\rho}$ a non-positive parameter. Further we need a second order strengthening of condition (6): suppose that for all j

$$\lim_{t\to\infty} \frac{\frac{U_{s_j}(t)}{U_0(t)} - e^{c\gamma^+ s_j}}{\beta(t)} = e^{c\gamma^+ s_j} \frac{e^{c\widetilde{\rho}s_j} - 1}{\widetilde{\rho}}.$$

Equivalently

$$\lim_{t\to\infty} \frac{\log U_{s_j}(t) - \log U_0(t) - c\gamma^+ s_j}{\beta(t)} = \frac{e^{c\widetilde{\rho}s_j} - 1}{\widetilde{\rho}}. \tag{17}$$

We consider an estimator $\hat{\gamma}_{n,k}^+(s_j)$ of γ^+ that is consistent i.e., $\hat{\gamma}_{n,k}^+(s_j) \xrightarrow[n\to\infty]{P} \gamma^+$ provided $k = k_n \to \infty$, $k_n/n \to 0$. Furthermore we require that under condition A

$$\sqrt{k}\left(\hat{\gamma}_{n,k}^+(s_j) - \gamma^+, \ \log X_{n-k,n}(s_j) - \log U_{s_j}\left(\frac{n}{k}\right)\right) \xrightarrow[n\to\infty]{d} \left(\Gamma^+(s_j), \ B^+(s_j)\right), \tag{18}$$

say, for all j, where $\left(\Gamma^+(s_j), \ B^+(s_j)\right)$ has a multivariate normal distribution provided k (the number of upper order statistics used in $\hat{\gamma}_{n,k}^+(s_j)$ for all j) satisfies $k = k_n \to \infty$ and

$$\lim_{n\to\infty} \sqrt{k_n}\, \beta\left(\frac{n}{k_n}\right) =: \lambda \tag{19}$$

exists finite. By setting $\lambda = 0$, we are imposing an upper bound to the intermediate sequence k_n, whereas $\lambda \neq 0$ imposes a lower and an upper bound to k_n and the asymptotic bias component then intervenes. Various estimators $\hat{\gamma}_{n,k}^+(s_j)$ are known with this property, notably Hill's estimator (Hill (1975)) as explained now:

Remark 1 *In sections 4 and 5 we shall choose*

$$\hat{\gamma}_{n,k}(s_j) := \frac{1}{k} \sum_{i=0}^{k-1} \log X_{n-i,n}(s_j) - \log X_{n-k,n}(s_j)$$

(Hill's estimator). In that case

$$\left(\Gamma^+(s_j), \ B^+(s_j)\right) \stackrel{d}{=} \left(\gamma \int_0^1 \left(s^{-1}W(s) - W(1)\right) ds + \frac{\lambda}{1-\rho}, \ W(1)\right)$$

with W Brownian motion, hence $\Gamma^+(s_j)$ and $B^+(s_j)$ are independent, $Var\left(\Gamma^+(s_j)\right) = \gamma^2$, $Var\left(B^+(s_j)\right) = 1$ (de Haan and Ferreira, 2006, pages 52 and 76).

(ii) Next we consider $\hat{c}^{(2)}$ (and $\hat{c}^{(3)}$). Again we need a second order strengthening of conditions (1) and (2) for part of the results.

CONDITION B   Suppose there exists a positive or negative function α_0 with $\lim_{t\to\infty} \alpha_0(t) = 0$ such that for each $x > 0$

$$\lim_{t\to\infty} \frac{\frac{U_0(tx)-U_0(t)}{a_0(t)} - \frac{x^\gamma-1}{\gamma}}{\alpha_0(t)} = \frac{1}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma-1}{\gamma}\right) =: H_{\gamma,\rho}(x) \tag{20}$$

where ρ is a non-positive parameter. For $\gamma = 0$ and/or $\rho = 0$ the limit function is defined by continuity. Further we need a strengthening of condition (1) or rather (5) for all j:

$$\lim_{t\to\infty} \frac{\frac{U_{s_j}(t)-U_0(t)}{a_0(t)} - \frac{e^{c\gamma s_j}-1}{\gamma}}{\alpha_0(t)} = H_{\gamma,\rho}(e^{cs_j}). \tag{21}$$

Relations (20) and (21) imply that all functions U_{s_j} satisfy a second order relation (cf. Lemma 8 below). We consider estimators $\hat{\gamma}_{n,k}(s_j)$ and $\hat{a}_{s_j}(n/k)$ that are consistent i.e.,

$$\hat{\gamma}_{n,k}(s_j) \xrightarrow[n\to\infty]{P} \gamma, \quad \frac{\hat{a}_{s_j}\left(\frac{n}{k}\right)}{a_{s_j}\left(\frac{n}{k}\right)} \xrightarrow[n\to\infty]{P} 1 \tag{22}$$

provided $k = k_n \to \infty$, $k_n/n \to 0$, $n \to \infty$. Furthermore we require that under condition B

$$\sqrt{k}\left(\hat{\gamma}_{n,k}(s_j) - \gamma, \ \frac{\hat{a}_{s_j}\left(\frac{n}{k}\right)}{a_{s_j}\left(\frac{n}{k}\right)} - 1, \ \frac{X_{n-k,n}(s_j) - U_{s_j}(n/k)}{a_{s_j}\left(\frac{n}{k}\right)}\right)$$
$$\xrightarrow[n\to\infty]{d} \quad \big(\Gamma(s_j), \ A(s_j), \ B(s_j)\big),$$

say, where $\big(\Gamma(s_j), A(s_j), \ B(s_j)\big)$, $j = 1, 2, \ldots, m$, are independent random vectors and have a multivariate normal distribution for each j provided $k = k_n \to \infty$, and

$$\lim_{n\to\infty} \sqrt{k_n}\, \alpha_0\left(\frac{n}{k_n}\right) =: \lambda \tag{23}$$

exists finite. Various estimators are known with these properties.

Remark 2 *Condition (23) results in a bound on the intermediate sequence k_n: for large values of λ the sequence k_n increases more rapidly, hence the variance decreases and the bias grows. We did not attempt to use bias-corrected estimators since these have been developed only for a limited number of estimators. A bias arises not only for the marginal parameters as in (22) but also from the c-estimator itself.*

Remark 3 *In sections 4 and 5 we shall choose the moment estimator for $\hat{\gamma}_{n,k}(s_j)$ and the associated scale estimator ((3.5.9) p.102 and (4.2.4) p.130 de Haan and Ferreira, 2006) for $\hat{a}_{s_j}(n/k)$. In this case, if $\lim_{n\to\infty} \sqrt{k}\, \beta(n/k) = 0$, $B(s_j)$ and $\big(\Gamma(s_j), A(s_j)\big)$ are independent, $Var\big(B(s_j)\big) = 1$,*

$$Var\big(\Gamma(s_j)\big) = \begin{cases} \gamma^2 + 1, & \gamma \geq 0 \\ \frac{(1-\gamma)^2(1-2\gamma)(1-\gamma+6\gamma^2)}{(1-3\gamma)(1-4\gamma)}, & \gamma < 0 \end{cases}$$

$$Var\big(A(s_j)\big) = \begin{cases} \gamma^2 + 2, & \gamma \geq 0 \\ \frac{2-16\gamma+51\gamma^2-69\gamma^3+50\gamma^4-24\gamma^5}{(1-2\gamma)(1-3\gamma)(1-4\gamma)}, & \gamma < 0 \end{cases}$$

and

$$Cov\big(\Gamma(s_j), A(s_j)\big) = \begin{cases} \gamma - 1, & \gamma \geq 0 \\ \frac{(1-\gamma)^2(-1+4\gamma-12\gamma^2)}{(1-3\gamma)(1-4\gamma)}, & \gamma < 0 \end{cases}$$

(pages 104, 131, 133 respectively de Haan and Ferreira (2006); the asymptotic biases – in case $\lambda \neq 0$ in (19) – can be found on the same pages).

We have the following results.

Theorem 4 *1. Under conditions* (1) *and* (2)

$$\hat{c}^{(1)} \xrightarrow[n\to\infty]{P} c.$$

Under condition A

$$\sqrt{k}\left(\hat{c}^{(1)} - c\right) \xrightarrow[n\to\infty]{d} \frac{\sum\limits_{j=1}^{m} s_j\left(\frac{B^+(s_j)-B^+(0)}{\gamma^+} + \frac{e^{c\tilde{\rho}s_j}-1}{\tilde{\rho}\gamma^+}\lambda\right)}{\sum\limits_{j=1}^{m} s_j^2} - \frac{c}{\gamma^+}\frac{1}{m}\sum\limits_{j=1}^{m}\Gamma^+(s_j). \qquad (24)$$

2. Under conditions (1) *and* (2)

$$\hat{c}^{(r)} \xrightarrow[n\to\infty]{P} c \quad for \ r = 2, 3.$$

Under condition B

$$\sqrt{k}\left(\hat{c}^{(2)} - c\right) \xrightarrow[n\to\infty]{d} \sum\limits_{j=1}^{m} s_j\left\{\frac{1 - e^{-c\gamma s_j} - c\gamma s_j}{\gamma^2}\frac{1}{m}\sum\limits_{i=1}^{m}\Gamma(s_i)\right.$$

$$\left. + B(s_j) - e^{-c\gamma s_j}B(0) - \frac{1-e^{-c\gamma s_j}}{\gamma}A(0) + \lambda\,e^{-c\gamma s_j}H_{\gamma,\rho}(e^{cs_j})\right\}\Big/\sum\limits_{j=1}^{m} s_j^2$$

and

$$\sqrt{k}\left(\hat{c}^{(3)} - c\right) \xrightarrow[n\to\infty]{d} \sum\limits_{j=1}^{m}\left\{e^{-cs_j}W^{(s_j)}(e^{cs_j}) - W^{(0)}(1) + \lambda\,b_3(s_j)\right\}\Big/\sum\limits_{j=1}^{m} s_j,$$

where $\{W^{(s_j)}(t)\}_{t\geq 0}$ *are independent standard Brownian motions* $(j = 1, 2, \ldots, m)$ *and*

$$b_3(s_j) = \begin{cases} -\frac{e^{-cs_j\rho}+1}{\rho(\gamma+\rho)}, & \gamma+\rho \neq 0,\ \rho < 0, \\ -cs_j\frac{e^{cs_j\gamma}+1}{\gamma}, & \gamma+\rho = 0,\ \rho < 0, \\ \frac{2cs_j}{\gamma}, & \rho = 0 \neq \gamma, \\ -(cs_j)^2, & \gamma = \rho = 0. \end{cases}$$

Remark 5 *When choosing the estimators of the extreme value index and scale according to Remarks 1 and 3 we get:*

- *the variance of the limit distribution of* $\sqrt{k}\left(\hat{c}^{(1)} - c\right)$ *is*

$$\frac{1}{\left(\sum\limits_{j=1}^{m} s_j^2\right)^2}\frac{1}{(\gamma^+)^2}\left\{\sum\limits_{j=1}^{m} s_j^2 + \left(\sum\limits_{j=1}^{m} s_j\right)^2\right\} + \frac{c^2}{m};$$

- *the variance of the limit distribution of $\sqrt{k}\left(\hat{c}^{(2)} - c\right)$ is*

$$\frac{\left(\sum\limits_{j=1}^{m} s_j \frac{1-e^{-c\gamma s_j}-c\gamma s_j}{\gamma^2}\right)^2 \frac{Var\left(\Gamma(0)\right)}{m} + \sum\limits_{j=1}^{m} s_j^2 + \left(\sum\limits_{j=1}^{m} s_j e^{-c\gamma s_j}\right)^2 + \left(\sum\limits_{j=1}^{m} s_j \frac{1-e^{-c\gamma s_j}}{\gamma}\right)^2 Var\left(A(0)\right)}{\left(\sum\limits_{j=1}^{m} s_j^2\right)^2}$$

  with the variances of $\Gamma(0)$ and $A(0)$ as in Remark 3.

- *The variance of the limit distribution of $\sqrt{k}\left(\hat{c}^{(3)} - c\right)$ is free from estimators of γ and for the scale, and is equal to*

$$\frac{\sum\limits_{j=1}^{m}\left(m + e^{-cs_j}\right)}{\left(\sum\limits_{j=1}^{m} s_j\right)^2}.$$

Figure 1 offers a comparison of the variances of $\hat{c}^{(2)}$ and $\hat{c}^{(3)}$ at the value $\gamma = 0.1$ which is close to the value of γ that plays a role in the application section 5. We avoid depicting the asymptotic variance of $\hat{c}^{(1)}$ in Figure 1 because we are setting γ at a small positive value. Hence a high variance of $\hat{c}^{(1)}$ is bound to occur. In fact, for the range of values considered here the variance is always bigger than 200 but smaller than 250. Figure 2 shows the three variances for $\gamma = 0.5$.

Figure 1: Asymptotic variances from Remark 5 with equal lengths, i.e. $s_j = j/m$, $j = 1, 2, \ldots, m$, and $m = 17$ for several values of c and constant $\gamma = 0.1$.
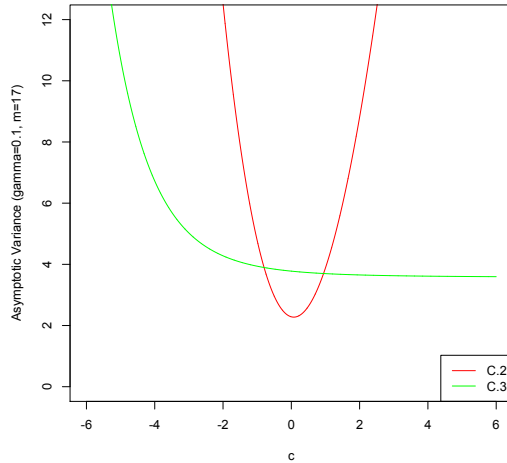
Corollary 6 *Assume $c = 0$. Under the conditions of the Theorem,*

Figure 2: Asymptotic variances from Remark 5 with equal lengths, i.e. $s_j = j/m$, $j = 1, 2, \ldots, m$, and $m = 17$ for several values of c and constant $\gamma = 0.5$.
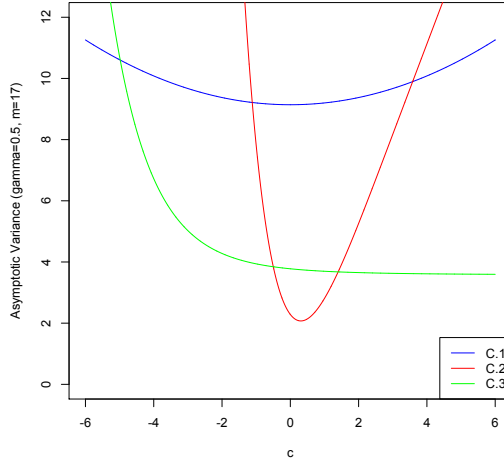
1. *if $k = k_n$ is such that $\sqrt{k}\,\beta(n/k) \to 0$, as $n \to \infty$ then*

$$Q_{m,n}^{(1)} := \sum_{j=1}^{m} \frac{k}{2} \left\{ \frac{\log X_{n-k,n}(s_j) - \log X_{n-k,n}(0)}{\hat{\gamma}_{n,k}^{+}} \right\}^2 \xrightarrow[n\to\infty]{d} \chi^2(m); \tag{25}$$

2. *if $k = k_n$ is such that $\sqrt{k}\,\alpha_0(n/k) \to 0$, as $n \to \infty$ then*

$$Q_{m,n}^{(2)} := \sum_{j=1}^{m} \frac{k}{2} \left\{ \frac{1}{k} \sum_{i=1}^{n} I_{\{X_i(s_j) > X_{n-k,n}(0)\}} - 1 \right\}^2 \xrightarrow[n\to\infty]{d} \chi^2(m). \tag{26}$$

Here $\chi^2(m)$ is a standard chi-squared distributed random variable with m degrees of freedom.

Corollary 6 gives rise to a testing procedure for detecting the presence of a trend in the tail of the underlying distribution functions F_s all lying in the same domain of attraction. That is, $Q_{m,n}^{(r)}$, $r = 1, 2$, defined above can be used as test statistics to evaluate the null hypothesis $H_0 : c = 0$ against the alternative $H_1 : c \neq 0$. Whence H_0 should be rejected at a significance level $\alpha \in (0,1)$ for any observed value of $Q_{m,n}^{(r)}$ verifying $Q_{obs}^{(r)} > q_{1-\alpha}(m)$, the latter being the $(1-\alpha)$-quantile pertaining to the chi-squared distribution with m degrees of freedom.

Remark 7 *A test for the extreme relative risk model* (1) *is provided by the statistic (cf.* (15))

$$\sup_{1 \leq j \leq m} \left| \log \Big(\frac{1}{k} \sum_{j=1}^{m} I_{\{X_i(s_j) > X_{n-k,n}(0)\}} \Big) \right|. \tag{27}$$

Then, the null hypothesis which states (1) *as our main model should be rejected for large values of* (27). *Large sample properties can be obtained using the methods of this paper. We do not pursue the matter here.*

4   Simulations

Simulations have been carried out for three distributions: (i) the generalized Pareto distribution; (ii) the ordinary Pareto distribution and (iii) the Burr$(1, \tau, \lambda)$ distribution with distribution function $F(x) = 1 - (1 + x^\tau)^{-\lambda}$, $x > 0$ and $\lambda, \tau > 0$. The number of time points is 100 (i.e. $m = 100$) with $s_j = j/m$, $j = 1, 2, \ldots, m$. At each location there are 300 i.i.d. observations ($n = 300$). Then there are 1000 replications which serve to obtain the means of $\hat{c}^{(r)}$, $r = 1, 2, 3$ as a function of the number $(k + 1)$ of tail related observations.

By definition, the finite sample behavior of $\hat{c}^{(1)}$ and $\hat{c}^{(2)}$ is inexorably attached to the estimation of the extreme value index γ. The parameter γ, which can be seen as a gauge of tail heaviness of the underlying distribution function F_s is thus an important design factor in the present numerical study. Since γ does not depend on s, we use combined estimators

$$\widehat{\gamma}^+ = \hat{\gamma}_{n,k}^+ := \frac{1}{m} \sum_{j=1}^{m} \hat{\gamma}_{n,k}^+(s_j) \quad \text{and} \quad \hat{\gamma} = \hat{\gamma}_{n,k} := \frac{1}{m} \sum_{j=1}^{m} \hat{\gamma}_{n,k}(s_j) \tag{28}$$

in accordance with Remarks 1 and 3. Estimator $\hat{c}^{(3)}$ is a shift invariant estimator not depending on γ. But the second order parameter $\rho \leq 0$ also plays a relevant role in the performance evaluation of the three estimators since it contributes for the (second order) dominant component of the asymptotic bias. In the present framework, providing a full array of combinations of parameter values c, γ and ρ, for a various number of time points m, would make a simulation study quite cumbersome and ultimately of unenthusiastic reading. For the sake of brevity, we have settled with $\gamma = 0.1$ and 0.5 (since 0.1 is a typical value for rainfall, the application topic). Different rates of convergence are addressed via several values of ρ. In particular:

(i) The generalized Pareto distribution (GPD) with distribution function $1 - (1 + \gamma x)^{-1/\gamma}$ for those x for which $1 + \gamma x > 0$ has been considered. Relation (20) holds with limit zero since the left hand-side is zero (exact fit). In this case, the values $c = -0.1$ and $c = 0.1$ have been considered.

The starting point is a r.v. X from the GPD distribution. For each location s_j we then take $X(s_j) \overset{d}{=} e^{cs_j\gamma} X + (e^{cs_j\gamma} - 1)/\gamma$. That way the relations (1) and (21) hold.

Figure 3: Estimated means of $\hat{c}^{(r)}$, $r = 1, 2, 3$, plotted against the same number k of top observations on each location $s_j = j/m$, $j = 1, 2, \ldots, m$, with underlying Generalized Pareto distribution, in case of true value $c = 0.1$.
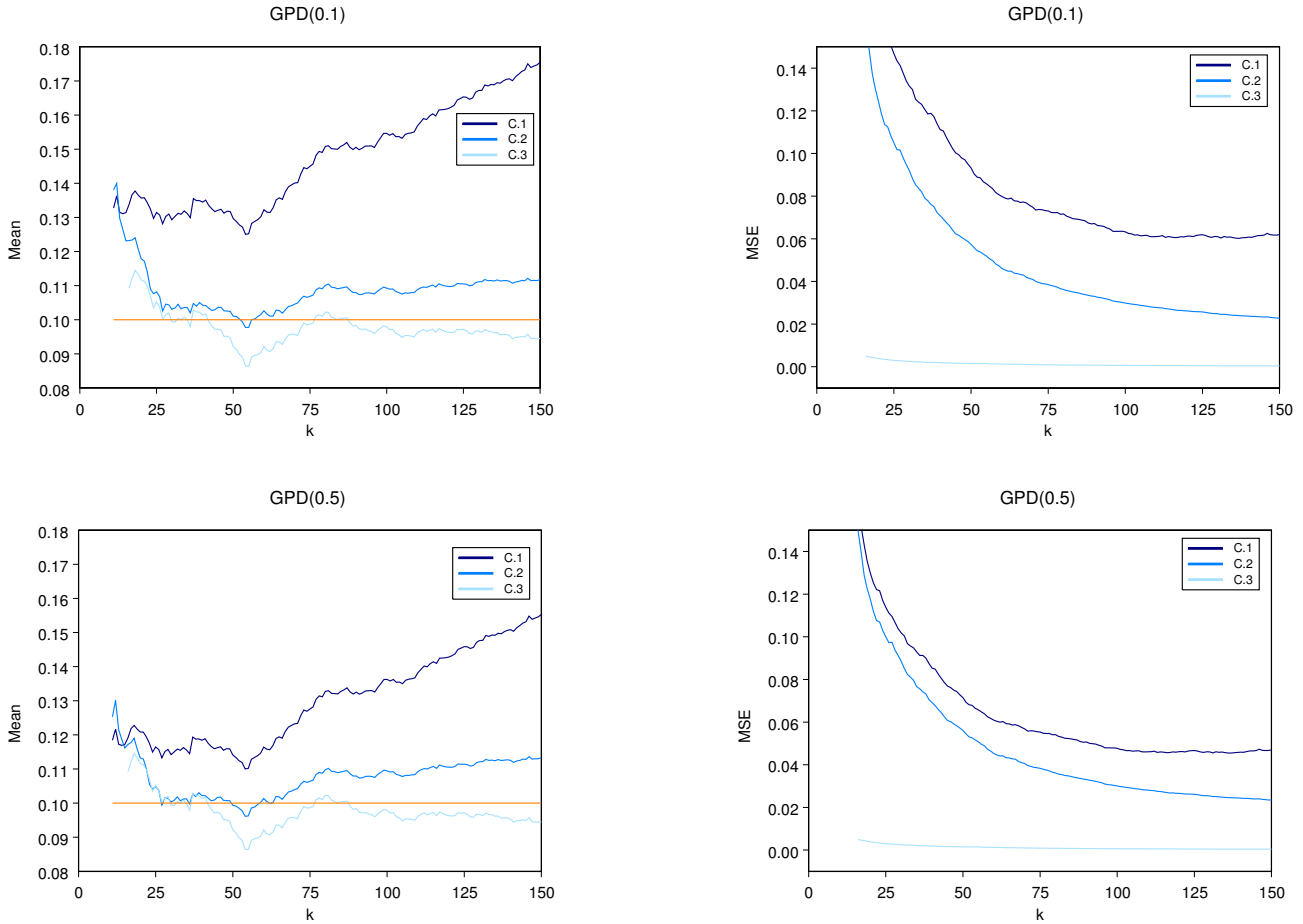

Figure 3 displays the average (over 1000 replications) values of the estimators $\hat{c}^{(1)}$ (only for positive γ), $\hat{c}^{(2)}$ and $\hat{c}^{(3)}$ as functions of the number k of upper order statistics above $X_{n-k,n}(s_j)$ for all s_j. As usual in graphs of this type there is a stretch of the graph that is more or less straight; the idea is that in that part both the variance and bias are not too high. We note that if, by the one hand, only a very tiny sample fraction (k/n) is selected then huge variance arises; on the other hand, if we get further into the original sample by selecting a very large number k of upper order statistics then bias increases. This sort of bias/variance trade-off is a common requirement in extreme value statistics.

In both Figures 3 and 4, the extreme value index γ and scale a_0 have been estimated by the estimators prescribed in Remarks 1 and 3. Note that in these and later graphs a realistic choice of the number of upper order statistics seems to be in the range between $k = 20$ and $k = 50$. This is the main reason why

Figure 4: Estimated means of $\hat{c}^{(r)}$, $r = 1, 2, 3$, plotted against the same number k of top observations on each location $s_j = j/m$, $j = 1, 2, \ldots, m$, with underlying Generalized Pareto distribution, in case of true value $c = -0.1$.
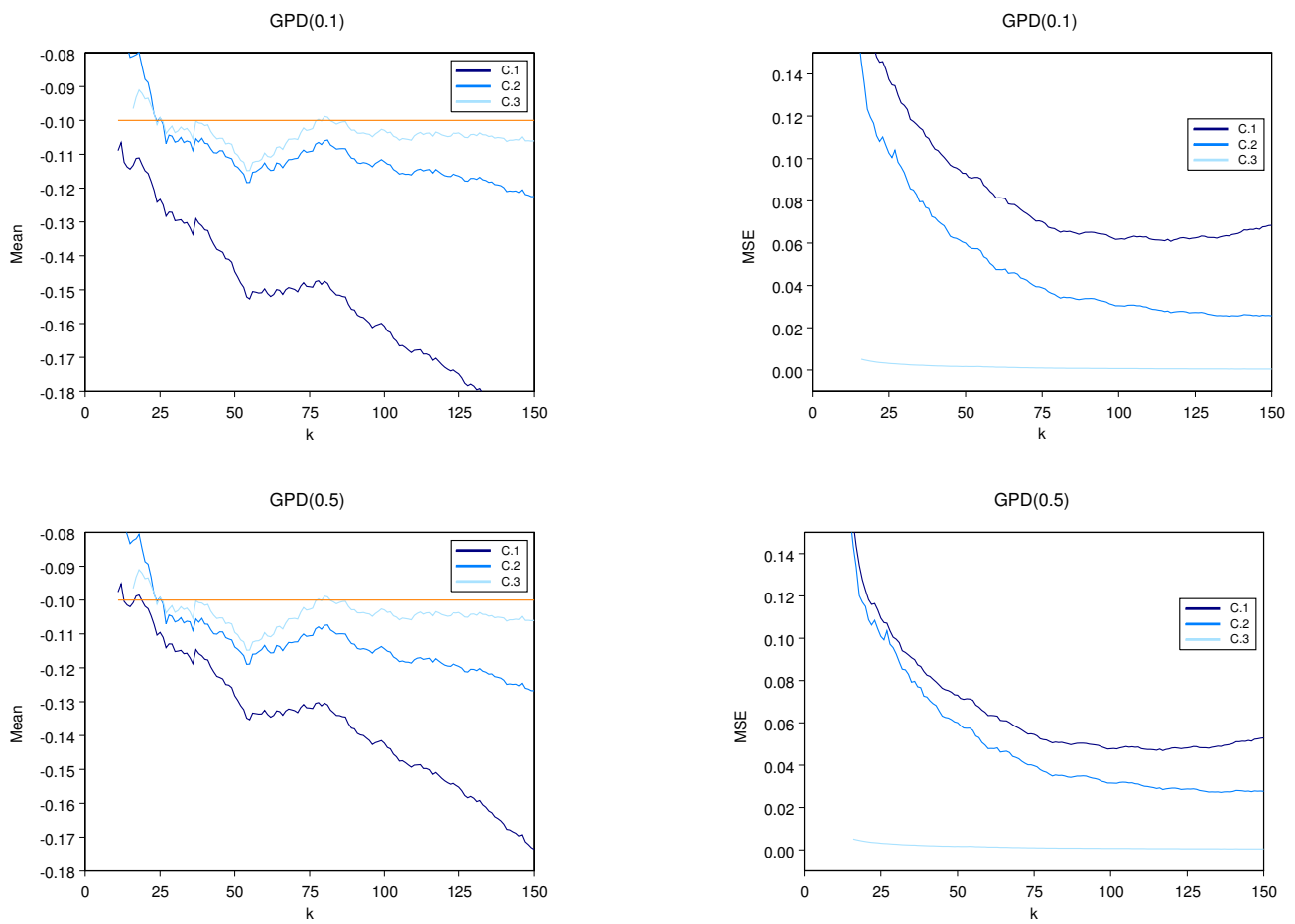
we say that the third estimator, $\hat{c}^{(3)}$, has the best performance among the three estimators. The estimator $\hat{c}^{(3)}$ begins to return estimates in a close vicinity of the true value c already for small k and confirms the trajectory of small bias as k gets closer to the whole sample size n. This occurs because the estimator $\hat{c}^{(3)}$ is the most direct empirical counterpart of relation (1) and the underlying Generalized Pareto distribution is the precise limit distribution in condition (3) regarding tail distribution.

The corresponding Mean Squared Errors (MSE) are also displayed on the right hand-side panels of Figures 3 and 4. In this respect, the third estimator surpasses all the other, which is mostly due to the low variance.

Overall, not the estimator $\hat{c}^{(1)}$, but the other two estimators of c seem to return stable trajectories in a close vicinity of the actual c-value quite often. Because the estimator $\hat{c}^{(1)}$ is subject to γ positive, one should expect that $\hat{c}^{(1)}$ is more prone to bias and/or variance inflations due to the presence of a true γ near zero. This is verified by the simulations: Figures 3 and 4 show that setting $\gamma = 0.1$ results in considerable bias displayed by $\hat{c}^{(1)}$. Estimators $\hat{c}^{(2)}$ and $\hat{c}^{(3)}$ come out with the best performance for intermediate values of k.

Figure 5 gives a comparison with the maximum likelihood estimators, valid for $\gamma \geq -1$. The simulations in this case however have been carried out with $m = 200$ and $n = 500$ mainly for facilitating the numerical convergence attached to the choice of $\gamma = 0.1$ (small positive value). The plot on the left panel of Figure 5 epitomizes the behavior of $\hat{c}^{(2)}$, either with the maximum likelihood estimator or with the moment estimator, because in other simulations we have conducted the moment estimator has been recognized so as to instill less bias in $\hat{c}^{(2)}$ while pertaining to moderate values of k, which are the most adequate in the context of extreme value theory. Moreover, the fact that the maximum likelihood estimates for shape and scale, γ and a_0, have to be numerically obtained can pose a practical difficulty to our trend estimation procedure. The convergence of appropriate numerical procedures may be rather poor when the true value of γ is close to zero. For $\gamma = 0.1$, the number of times the algorithm has converged thus returning feasible estimates of γ and a_0 is depicted on the right hand-side of Figure 5. Other estimators for the extreme value index, including bias-corrected estimators (see e.g. Caeiro et al. (2005) and Gomes et al. (2008) for $\gamma > 0$ and Cai et al. (2013) for general γ) could have been considered. Here, we have restricted ourselves to the case which poses the most proximity to the exact GPD model in the POT approach (see point 2. in Appendix B).

(ii) For the ordinary Pareto distribution with distribution function $1 - x^{-1/\gamma}$, $x \geq 1$, $\gamma > 0$, we simulate the trend by taking $X(s_j) \overset{d}{=} e^{cs_j\gamma}X$ where X follows the Pareto distribution. Again we have an exact fit in view of condition (6) for $\gamma > 0$.

(iii) For the Burr$(1, \tau, \lambda)$ distribution again the trend is simulated by taking $X(s_j) \overset{d}{=} e^{cs_j/2}X$ (since we are taking $\gamma = (\lambda\,\tau)^{-1}$ equal to $1/2$). Relations (20) and (21) hold. In this case $|\alpha_0|$ is a regularly varying function with index $\rho = -1/\lambda$. We set $\rho = -1, -1/2$, ranging from fast to a moderate converge rate.

Figure 5: Estimated means of $\hat{c}^{(2)}$ either with Moment estimator and ML estimator for the Generalized Pareto distribution with $\gamma = 0.1$ and $c = 0.1$. The number of samples amongst the 1000 replicates that have produced valid ML-estimates is presented on the right hand-side.
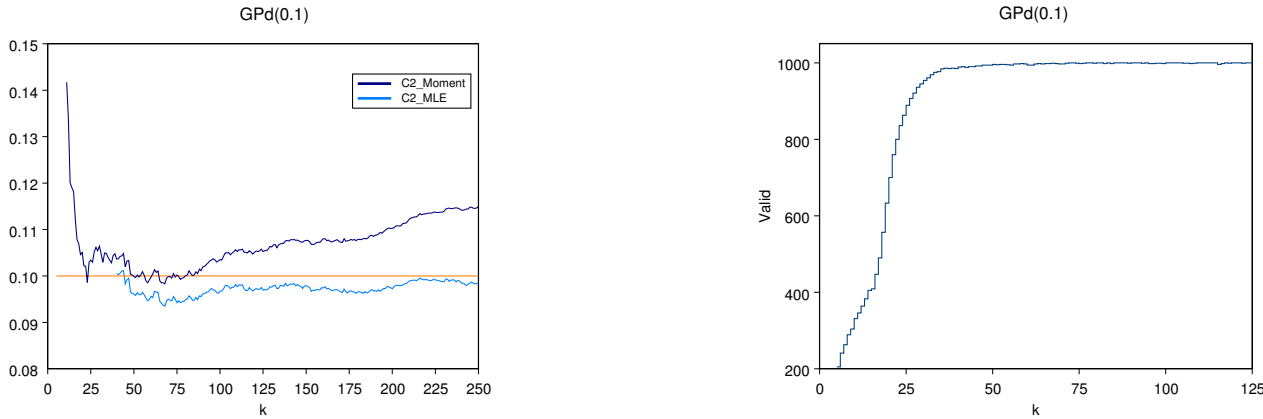
Figures 6 and 7 display the simulation results for the Pareto and Burr distributions, always with $c = 0.1$. Again, the extreme value index γ and scale a_0 have been estimated by the moment estimator. In the particular case of the Pareto distribution, the estimates process of $\hat{c}^{(3)}$ is a repeat of the previous exact model described in (i), due to its invariance towards a shift in location and/or changes to in the scale of the observed data. Although $\hat{c}^{(1)}$ is only valid for $\gamma > 0$, it remains a matching competitor against $\hat{c}^{(2)}$ and $\hat{c}^{(3)}$ under the three parent distribution functions considered in Figures 6 and 7. The estimators $\hat{c}^{(r)}$, $r = 1, 2$ and 3, are quite close to the real value if one chooses k close to 30; the graph in that area is relatively flat. In Figures 6 and 7, the MSE is also depicted. We have similar findings in case of Pareto and Burr parent distributions: the empirical MSE is decreasing fast until the end of the plateau of stability observed in the mean estimates plot, then it slows down. The stability range for the mean is quite clear in Figures 6 and 7, with the values $k = 50$ and $k = 75$, respectively, marking the end of the stability zone. From this point on, the variance becomes less dominant and bias overpowers. Hence, these frontier values of k provide a reasonable indicator for the number of upper order statistics to be used in the estimation of the trend.

We emphasize that a much more comprehensive simulation study has been performed. From the described simulations and the other ones we conclude that the estimators perform reasonably well. The main conclusion is that estimators $\hat{c}^{(2)}$ and $\hat{c}^{(3)}$ seem to behave better than $\hat{c}^{(1)}$. In the next (application) section we shall adopt estimator $\hat{c}^{(3)}$. In general applications, a possible focus standing on $\hat{c}^{(3)}$ alone could be justified by its implicit detachment from the extreme value index γ thus making $\hat{c}^{(3)}$ more versatile in the estimation of c for a wider range of underlying models pertaining to diverse values of γ.

Figure 6: Estimated means of $\hat{c}^{(r)}$, $r = 1, 2, 3$, plotted against the same number k of top observations on each location $s_j = j/m$, $j = 1, 2, \ldots, m$, in case of true value $c = 0.1$, with Pareto parent distribution.
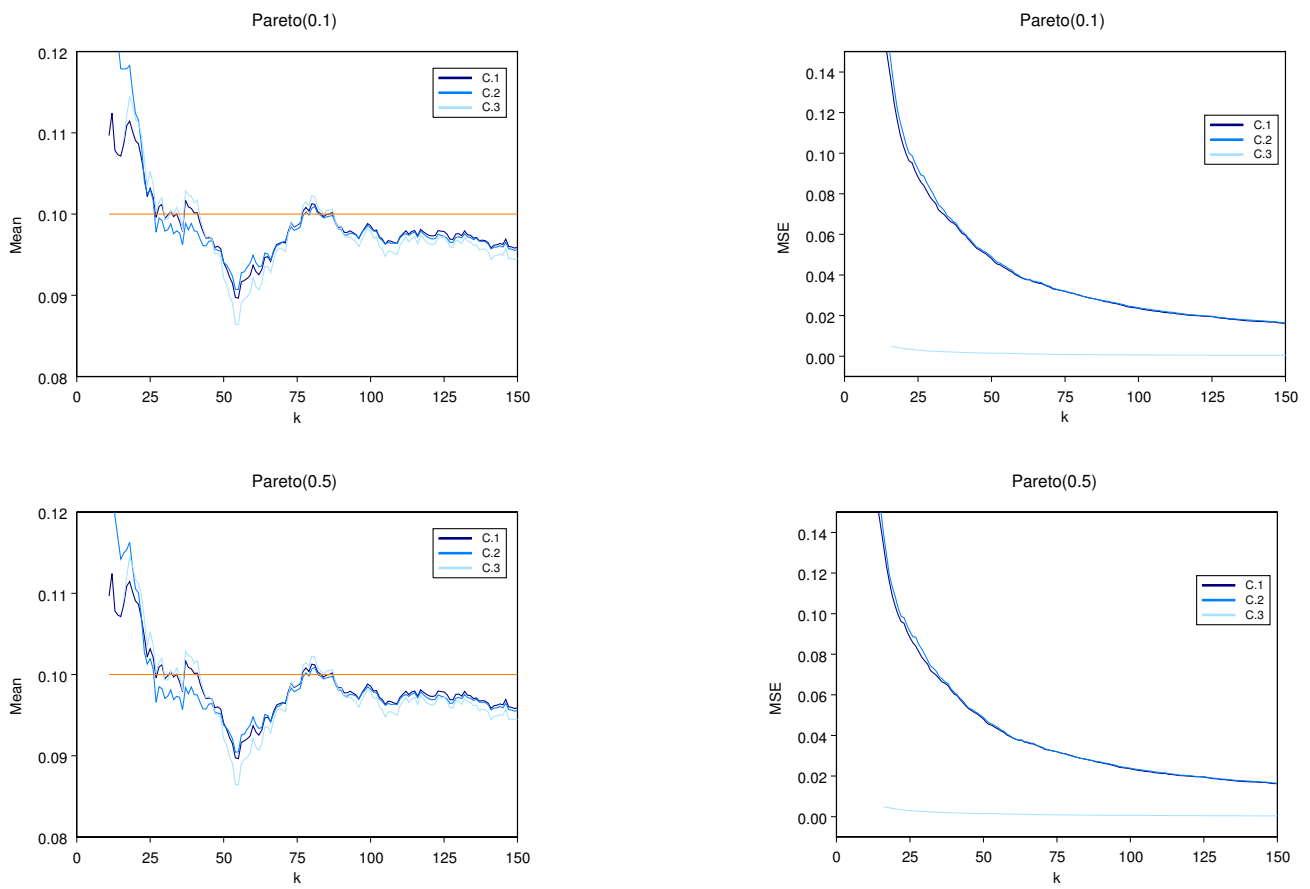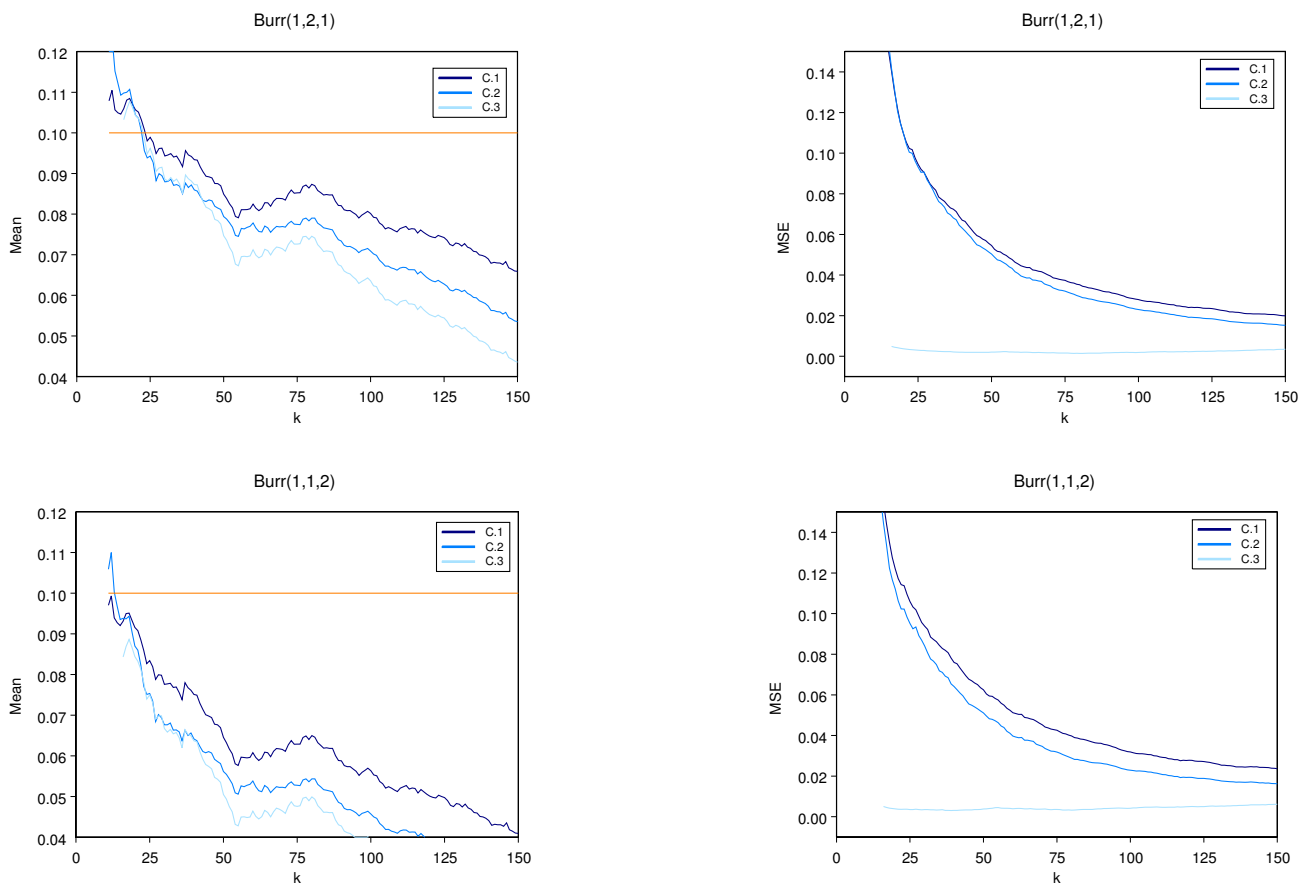
Figure 7: Estimated means of $\hat{c}^{(r)}$, $r = 1, 2, 3$, plotted against the same number k of top observations on each location $s_j = j/m$, $j = 1, 2, \ldots, m$, in case of true value $c = 0.1$, with Burr parent distribution.

In view of the testing procedures for detecting the presence of a trend, built on corollary 6, we present some more simulation results in order to assess power of the test statistics $Q^{(r)} = Q_{m,n}^{(r)}$, $r = 1, 2$. Figure 8 illustrates the simulated power and type I error of the test that rejects the null hypothesis of no trend, i.e. $H_0 : c = 0$, for large values of the test statistics $Q^{(1)}$ or $Q^{(2)}$ (cf. text below corollary 6). We focus on the special cases of the GPD and Burr distributions with $\gamma = 0.1$ and $\gamma = 0.5$. For assessing power, we have simulated from these parent distributions with $c = 0.1$ and $c = 0.5$ in points (i) and (iii). The significance level is $\alpha = 0.05$ and the number of observations ($n = 300$) and time points ($m = 100$) are the same as in the case of c-estimation. Again, there are 1000 replicates. Despite the small frequency of rejections under the Burr(1,1,2) model, the test based on $Q^{(1)}$ (solid lines in Figure 8) seems to perform much better than the test adopting $Q^{(2)}$ (dashes lines in Figure 8). By construction, the estimated type I error of $Q^{(2)}$ obtained under the Burr(1,1,2) coincides with the case of GPD(0.5). Hence the former is omitted from the last panel in Figure 8. The major drawback in both tests pertains in fact to the type I error since the actual size of the tests tends to be higher than the nominal significance level 0.05. This leads to many undesirable rejections of the null hypothesis of no trend for small values of k in particular. Proceeding with this admonition to the next section 5, we find useful to apply the two testing procedures to rainfall data.

5   Data Analysis

As an application of the tail trend assessment methodology developed in this paper, we will look at daily rainfall totals collected in 18 gauging stations across Germany and The Netherlands, comprising latitude 47N-53N and longitude 5E-13E. The geographic location of the stations is displayed in Figure 9. Rainfall data are from the European Climate Assessment and Dataset (ECA&D). We note that different stations suffer from different coverage in time in the sense that not all stations have started regular recording of data at the same year. Moreover there are some stations with missing values. All of them however meet the basic criterium for completeness that there is less than 10 days missing per year which leaves us with 90 years of complete data from 1918 up to 2007.
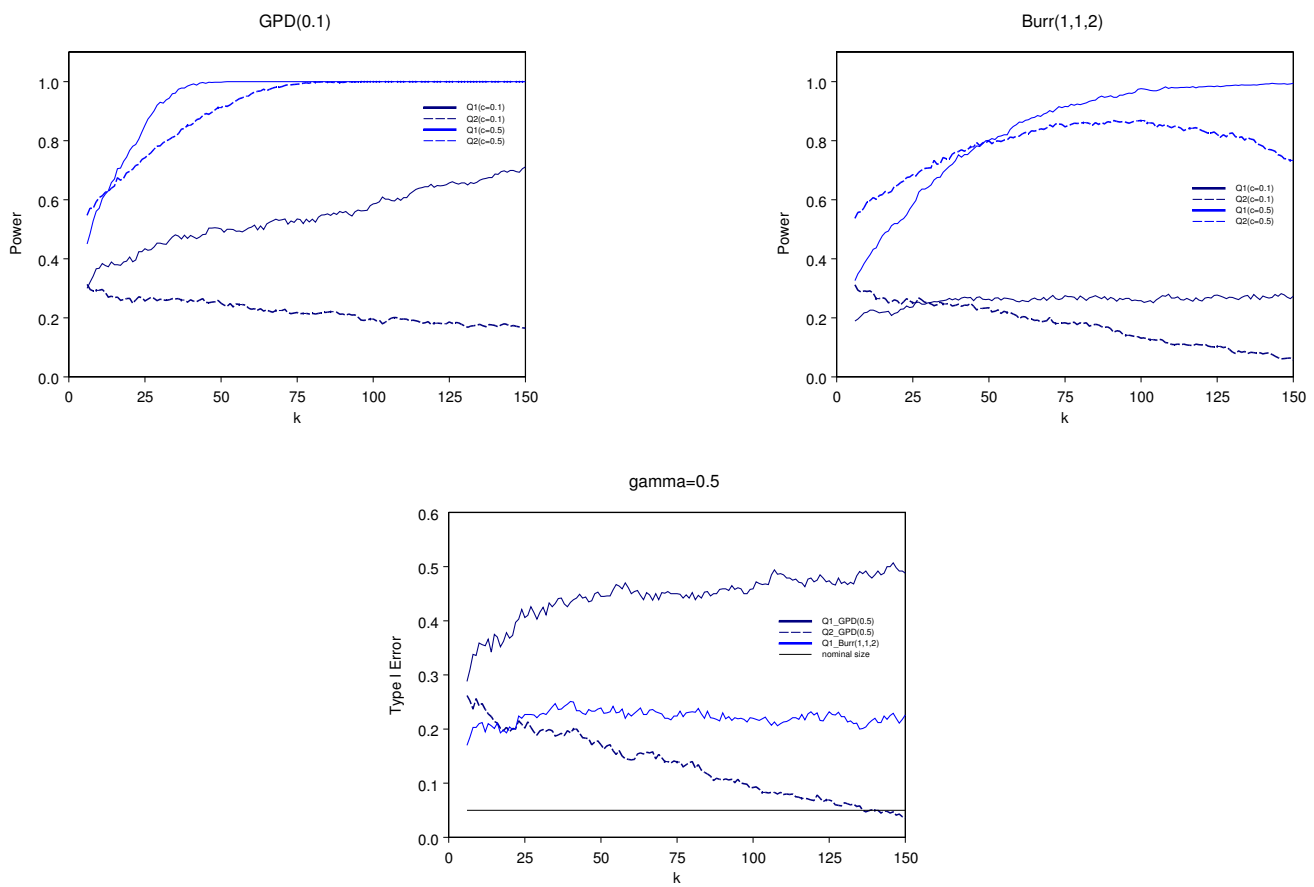
5.1   Trend estimation in the extreme relative risk model

Figure 10 displays yearly maxima plots for several stations on the basis of available data. We get a mixed picture. In *STN41-Halle*, for instance, precipitation does not seem to be as severe now as in the first half of the 20th century anymore, whereas *STN39-Dresden* shows increasingly annual maxima with the largest peak of $158\,mm$ of rain, spot on the catastrophic event of 12 August 2002.

In what follows we shall assume that as long as there is at least one day in between, there is not much dependence in the amount of rainfall on two different days. For each station we select first the highest
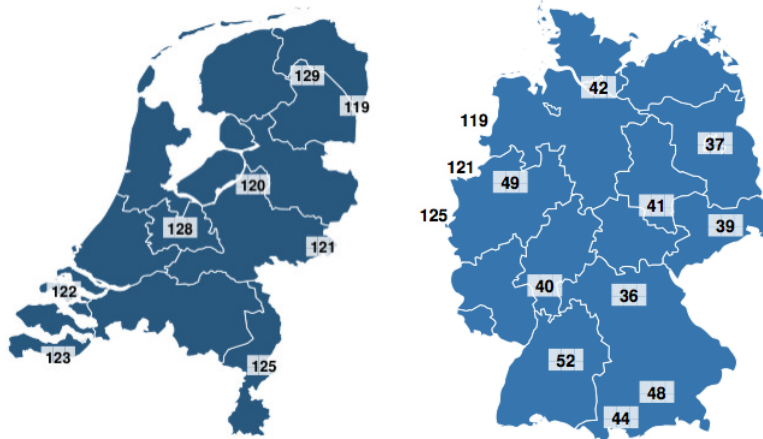
Figure 8: Mean number of rejections of $H_0 : c = 0$ by the test statistics $Q^{(r)}$, $r = 1, 2$, plotted against the same number k of top observations on each location $s_j = j/m$, $j = 1, 2, \ldots, m$.

Figure 9: Selected gauging stations in The Netherlands (*left*); selected gauging stations in Germany and Dutch stations near the borderline (*right*).
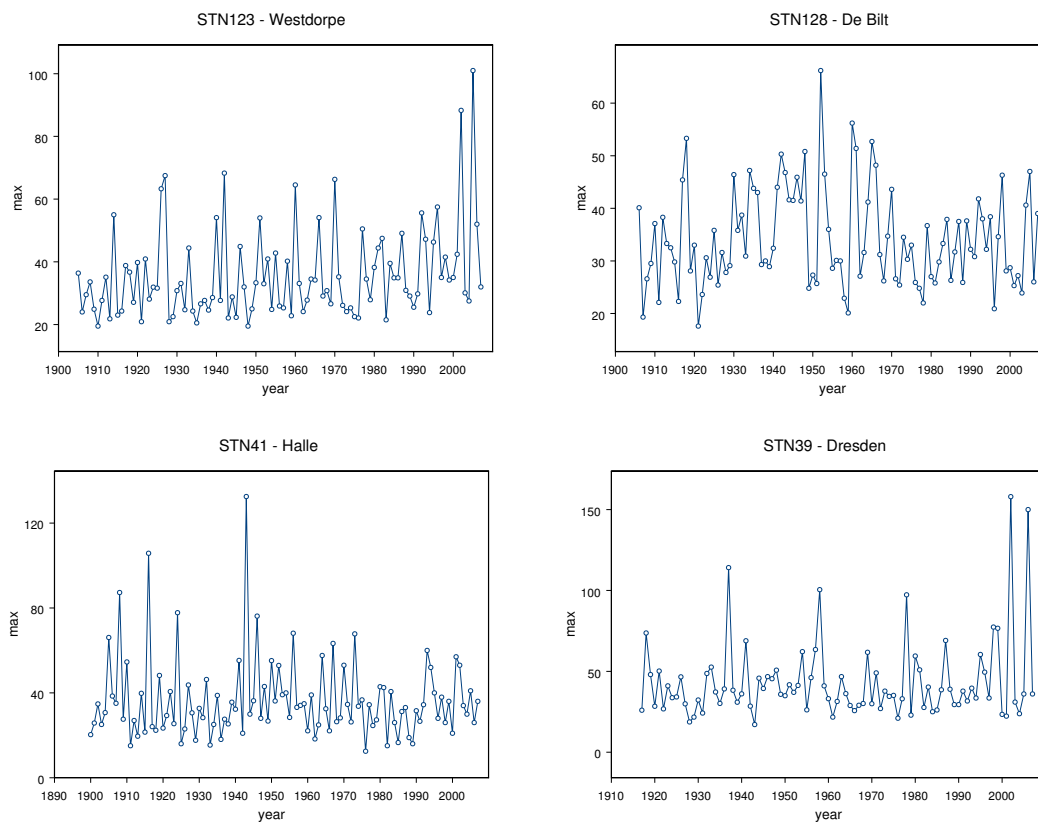

observation. Then we remove the observations on the day before and after. Next we select the highest observation from the remaining data, etc. This goes on until we have selected 70 days or the threshold of $1\,mm$ is reached. That way we get a sequence of higher order statistics from i.i.d. data. Table 1 displays the number of rain days (i.e. with at least $1\,mm$ of rain) per station. Our approach for *declustering* large rainfall observations follows identical procedure to AghaKouchak and Nasrollahi (2010) in the sense that only the one largest value from consecutive days of rainfall enters in the calculations. A similar argument can be found in Tomassini and Jacob (2009). The authors claim that dependent extreme precipitation occurs within two days at most and conclude that rainfall observations rarely compose clusters of more than three days. Hence, the approach described in section 6.(iij) of de Haan (1990) for selecting extreme observations, with fixed $i = 1$ therein, is a plausible approach in the present context. Furthermore, Burauskaite-Harju et al. (2012) study the ECA&D in the Netherlands by taking into account the entire year of rainfall records as opposed to taking data from a specific season. Rainfall patterns tend to differ in summer and winter. The heaviest rainfall typically occurs in winter. Since our statistical analysis is based on the k largest observations in a 5-year period, they are much more likely to occur in the winter time hence the season does not play a big role. With a similar bearing, Yee and Stephenson (2007) introduced an *ad hoc* procedure for selecting the largest observations from an entire year of daily rainfall records. Yee and Stephenson (2007) also give accounts of only a slight tendency to cluster at high levels of daily rainfall.

Since we are not looking for a spatial trend now we shall make a study of the highest daily rainfall amounts in the 90 year period for each station separately. At each location, for $\hat{\gamma}^+$ (in connection with

Figure 10: Some yearly maxima of daily rainfall.

Table 1: Total number of rain days in the period from 1918 to 2007 for each selected station.

STN	Name	Country	Lat.	Lon.	Rain days
36	Bamberg	D	48°49'N	11°33'E	6276
37	Berlin	D	52°31'N	13°20'E	6270
39	Dresden	D	51°31'N	13°44'E	6293
40	Frankfurt	D	50°6'N	08°40'E	6069
41	Halle	D	51°28'N	11°57'E	6230
42	Hamburg	D	53°33'N	09°59'E	6125
44	Hohenpeissenberg	D	47°48'N	10°59'E	6230
48	München	D	48°08'N	11°34'E	6153
49	Münster	D	52°59'N	07°41'E	5904
52	Stuttgart	D	48°46'N	09°46'E	6289
119	Ter Apel	NL	52°53'N	07°04'E	6069
120	Heerde	NL	52°24'N	06°03'E	6300
121	Winterswijk	NL	51°59'N	06°42'E	6298
122	Kerkwerve	NL	51°40'N	03°51'E	6276
123	Westdorpe	NL	51°13'N	03°52'E	6300
125	Roermond	NL	51°11'N	05°58'E	6300
128	De Bilt	NL	51°06'N	05°11'E	6299
129	Eelde	NL	53°08'N	06°35'E	6300

$\hat{c}^{(1)}$) we use Hill's estimator and for $\hat{\gamma}$ and $\hat{a}_0$ (for $\hat{c}^{(2)}$) we use the moment estimator (cf. sections 3.5 and 4.2 of de Haan and Ferreira (2006))

The point estimation of the extreme value index γ and trend estimation is conducted with the same number of upper order statistics k, just as prescribed in each definition of $\hat{c}^{(r)}$, $r = 1, 2, 3$, introduced in (13), (14), and (15), respectively.

In order to have enough tail related rain measurements per time point we found reasonable to take consecutive intervals of 5 years over the 90-year span. The disjoint intervals serve as our time points indexed by $j = 0, 1, 2, \ldots, m$, with $m = 17$.

For the purpose of data analysis, the gauging stations have been divided into two groups, determined by their alignments in the general climate characteristics (according to the Köppen-Geiger climate classification system, see e.g. Kottek et al. (2006) and Peel et al. (2007)). Each selected station in Germany was classified as either *humid oceanic* or *humid continental*. All stations across The Netherlands are classified as *humid oceanic*. Although we are not looking for spatial coherence we hope to benefit from this information to get a more systematic presentation of our results.

Figure 11 includes sample paths of the three proposals for estimating the tail trend parameter $c \in \mathbb{R}$ for some typical gauging stations. As already discussed, we shall handle estimation of c by screening plots as in Figure 11 for plateaus of stability in the early part of the graphs pertaining to the estimator $\hat{c}^{(3)}$, coherent with the path patterns of $\hat{c}^{(1)}$ and $\hat{c}^{(2)}$ whenever possible.

Table 2 contains the estimated values of c for each station by their increasing order of magnitude. Standard errors are also provided. Bearing in mind the simulation results from section 4, here we shall

Figure 11: Sample path of the overall moment estimator for γ and sample trajectories of $\hat{c}^{(r)}$, $r = 1, 2, 3$, as a function of the same number k of top observations for each 5-year interval between 1918 and 2007 for several stations.
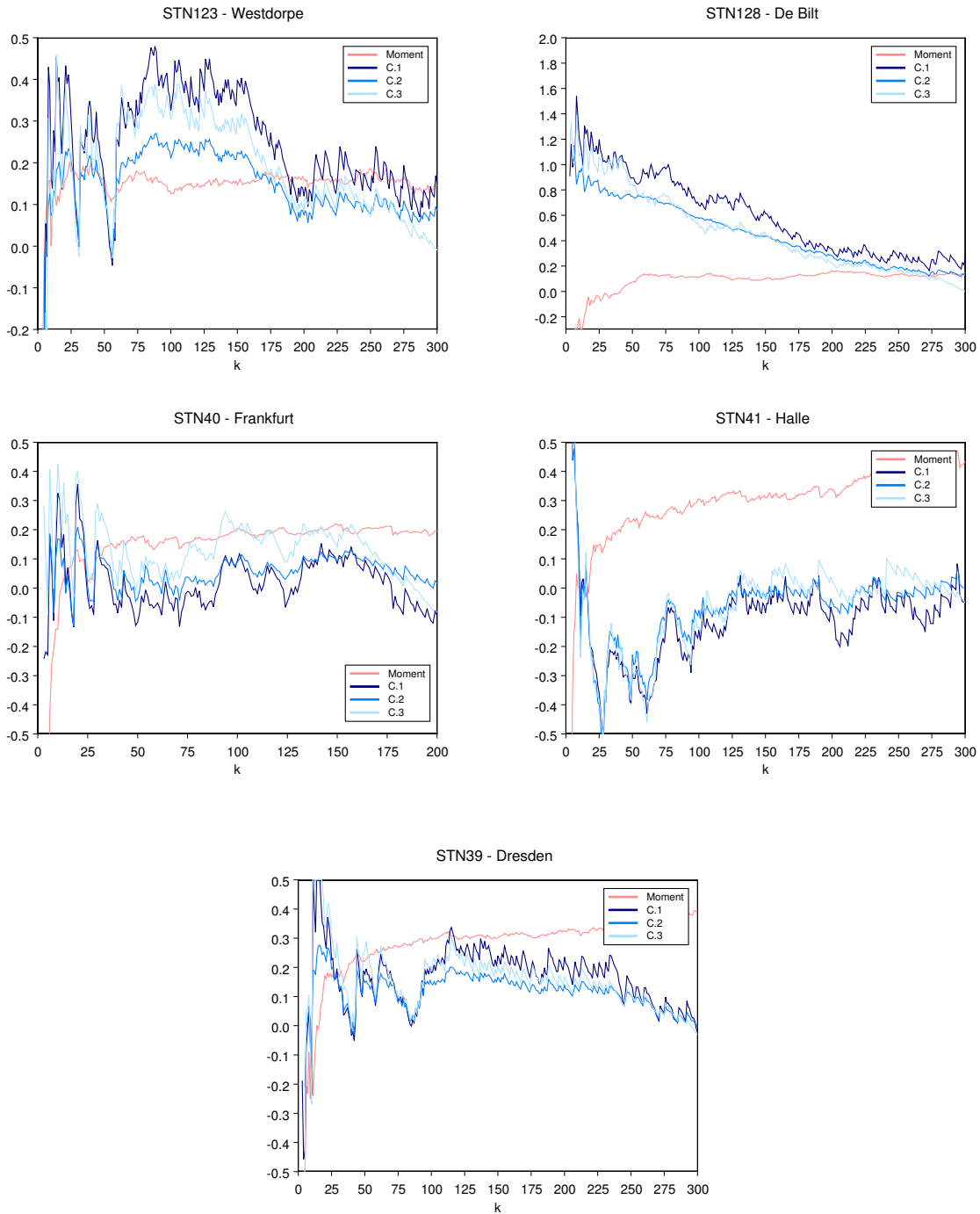
Table 2: Station-wise estimates for the trend parameter c by adopting estimator $\hat{c} = \hat{c}^{(3)}$ with accompanying Moment-estimates for the extreme value index γ.

Oceanic Stations

Station	$\hat{c}$	$\widehat{s.e.}(\hat{c})$	$\hat{\gamma}$
STN 49	0.11	0.388	0.12
STN 40	0.22	0.404	0.1
STN 123	0.23	0.396	0.17
STN 52	0.28	0.380	0.08
STN 122	0.53	0.403	0.16
STN 120	0.69	0.394	0.15
STN 119	0.80	0.358	0.1
STN 42	1.0	0.430	0.05
STN 121	1.02	0.384	0.05
STN 128	1.06	0.384	−0.05
STN 129	1.16	0.357	0.07
STN 125	1.61	0.428	0.08

Continental Stations

Station	$\hat{c}$	$\widehat{s.e.}(\hat{c})$	$\hat{\gamma}$
STN 44	−0.4	0.326	0.11
STN 41	−0.2	0.316	0.2
STN 48	−0.1	0.382	0.13
STN 37	−0.06	0.415	0.2
STN 39	0.26	0.289	0.24
STN 36	0.29	0.289	0.18

confine attention to the leading estimator $\hat{c}^{(3)}$. The asymptotic standard error of $\hat{c}^{(3)} = \hat{c}$ is estimated by

$$\widehat{s.e.}(\hat{c}^{(3)}) \quad := \quad \frac{1}{\sqrt{k}} \left\{ \frac{\sum\limits_{j=1}^{m} \left(m + e^{-c\,j/m} \right)}{\left(\frac{1+m}{2} \right)^2} \right\}^{\frac{1}{2}}$$

(cf. Remark 5 in section 3). Although the estimation of the extreme value index γ is not an actual requirement here because we are using estimator $\hat{c}^{(3)}$ for the trend, Table 2 also displays the Moment estimates for γ. Estimates of γ in a vicinity of 0.1 often emerge in connection with the extremal behavior of distributions underlying rainfall records (see e.g. Buishand et al. (2008), p.239; also Mannshardt-Shamseldin et al. (2010), p.492). AghaKouchak and Nasrollahi (2010) apply several estimators for the extreme value index, namely Pickands, Hill, Moment, Probability Weighted Moment and Maximum Likelihood estimators, in the context of rainfall. The obtained estimates for γ are again of similar magnitude, i.e. values of γ approximately equal to 0.1. This also seems to hold for most of the stations we are considering in this study although there is a lot of variation. Stations from the Continental group seem to yield heavier tails than the Oceanic group, which is in agreement with Figure 4(a) in Tomassini and Jacob (2009) regarding rainfall amounts for several stations in Germany between 1961 and 2004.

5.2   Detecting a trend in extreme rainfall

It remains to assess whether the stations with near zero estimates in fact have a null trend. On the basis of Table 2, examples are *STN37–Berlin*, *STN48–München* and *STN49–Münster*. The last site we refer to is *STN40–Frankfurt*, where testing for the presence of a trend is also of practical importance given the poor circumstances involving the estimation of the parameter c: the erratic sample paths displayed by the three

Figure 12: Sample paths of $Q^{(1)}_{18,n}$ and $Q^{(2)}_{18,n}$ test statistics and corresponding critical values for the two-sided test at a significance level $\alpha = 0.05$ ($\chi^{-1}_{0.95}(17) = 8.67$) for several stations across Germany.
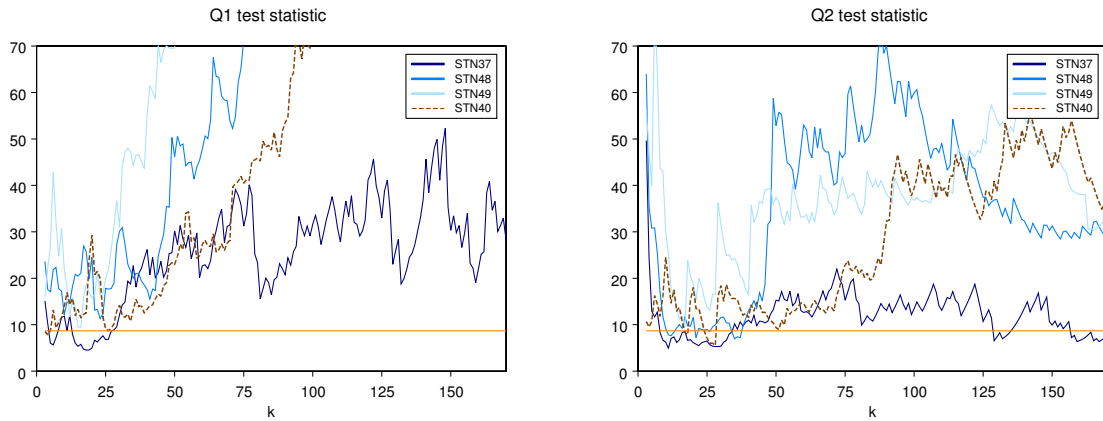
estimators often cross the $c = 0$ line (cf. Figure 11). In the case of *STN40–Frankfurt* it seems difficult to find a "plateau of stability" in Figure 11; the estimate $\hat{c} = 0.22$ is rather uncertain. Therefore, we aim at a more definite decision on the value of c by means of a testing procedure upon *STN40–Frankfurt* in particular.

In order to tackle the problem of testing the presence of a trend in time, i.e. the problem of testing hypothesis

$$H_0 : c = 0 \quad versus \quad H_1 : c \neq 0, \tag{29}$$

we shall use $Q^{(r)}_{m,n}$ from corollary 6 as our test statistics. Hence, for $r = 1,\ 2$, the null hypothesis $H_0 : c = 0$ is rejected in favor of the two-sided alternative $H_1 : c \neq 0$ if $Q^{(r)}_{m,n} > q_{m,1-\alpha}$, where $q_{m,1-\alpha}$ denotes the $(1 - \alpha)$-quantile of the chi-square distribution with m degrees of freedom.

Figure 12 depicts the sample trajectories pertaining to the two-sided test statistics $Q^{(1)}_{m,n}$ and $Q^{(2)}_{m,n}$ in companion with critical values at the nominal size $\alpha = 0.05$ and with respect to the referenced stations *STN37–Berlin*, *STN48–München*, *STN49–Münster* and *STN40–Frankfurt*. It seems that the hypothesis of no trend has to be rejected in case of *STN40–Frankfurt*. The two tests also ascertain a non-null trend for *STN48–München* and *STN49–Münster*. There is no evidence of a particular trend at *STN37–Berlin*.

5.3 Discussion

Broadly speaking, the stations in the Oceanic group present higher positive values of c (but not always smaller values of γ) than the Continental stations. Overall, fairly positive values of c may be interpreted as being influenced by the ocean.

Allen and Ingram (2002) describe how the intensity of extreme rainfall events depends on the availability of moisture. Because moisture availability is constraint on temperature (through the Clausius-Clapeyron relationship), an increase in rainfall extremes is expected in a warming climate. Lenderink et al. (2009) show that higher increases can be expected at locations that are under the influence of the sea. Increasing sea surface temperatures contribute to higher rainfall amounts. The results obtained in this study for the behavior of extreme rainfall at locations in Germany and The Netherlands are consistent with these findings. Overall, the Oceanic group of stations shows a stronger increase in extreme rainfall than the Continental group of stations.

Next we discuss the interpretation of c. There is no evidence of a significant trend in the extreme relative risk at *STN37–Berlin*, meaning that the value $c = 0$ can be assigned to this gauging station. If $c = 1.16$, like the estimated value at *STN129–Eelde* (see Table 2), then in view of the fact that s is measured in periods of 5 years, the probability of really heavy rainfall increases approximately by 15% during each decade. Similarly if $c = 0.2$, which is approximately the estimate for the trend at *STN40–Frankfurt* and *STN123–Westdorpe*, then the probability of heavy rainfall increases in the same period approximately by 2%. These results are in good agreement with the positive trend of 3% per decade found by Zolina et al. (2009) for the second half of the century (1950-2000) using a different metric. Figure 13 may help to clarify the contrast in these values by plotting the empirical log-relative risk

$$\log \frac{1 - \widehat{F}_s\big(X_{n-k,n}(0)\big)}{1 - \widehat{F}_0\big(X_{n-k,n}(0)\big)}$$

against several values of k and for $s_j = j/m \in [0,1]$. At *STN37–Berlin*, there is no clear evidence of a trend, whereas at *STN129–Eelde* the estimated log-relative risk seems to increase as s_j approaches 1.

6 Proofs

We shall use the following representation:

$$\big\{X_{n-i,n}(s_j)\big\}_{i=0}^{n-1} \overset{d}{=} \big\{U_{s_j}\big(Y_{n-i,n}(s_j)\big)\big\}_{i=0}^{n-1}$$

where $\{Y_{n-i,n}(s_j)\}_{i=0}^{n-1}$ are the $n-$th order statistics from the distribution function $1 - x^{-1}$, $x \geq 1$, independently for $j = 1, 2, \ldots, m$.

Figure 13: Estimated $\log$-relative risk for $s_j = j/17$, with $j = 1, 2, \ldots, 17$ marking periods of 5 years. Equal distances between s_j are chosen to reflect time periods of equal lenght.

STN37 - Berlin

STN129 - Eelde
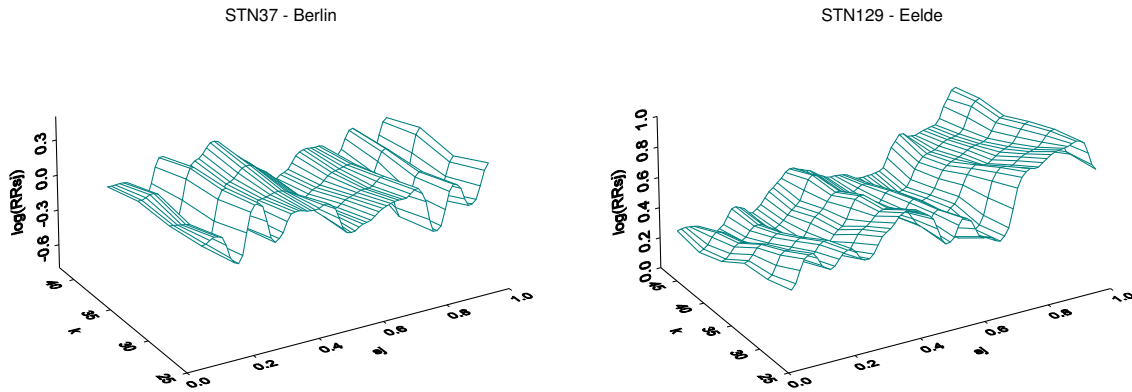
Proof of consistency

For the consistency of $\hat{c}^{(1)}$ note that $(k/n)\, Y_{n-k,n}(s_j) \xrightarrow{P} 1$, $n \to \infty$ (cf. de Haan and Ferreira (2006), Corollary 2.2.2) and that $\lim_{t \to \infty} U_s(tx)/U_s(t) = x^{\gamma^+}$ locally uniformly for $x > 0$. Hence for $j = 1, 2, \ldots, m$

$$\log X_{n-k,n}(s_j) - \log U_{s_j}\Big(\frac{n}{k}\Big) \overset{d}{=} \log U_{s_j}\Big(\frac{n}{k}\big\{\frac{k}{n} Y_{n-k,n}(s_j)\big\}\Big) - \log U_{s_j}\Big(\frac{n}{k}\Big) \xrightarrow[n\to\infty]{P} 0.$$

The rest is easy.

Similarly with respect to $\hat{c}^{(2)}$ we get that

$$\frac{X_{n-k,n}(s_j) - U_{s_j}\big(\frac{n}{k}\big)}{a_{s_j}\big(\frac{n}{k}\big)} \xrightarrow[n\to\infty]{P} 0. \tag{30}$$

This limit relation combined with relation (5) leads directly to the consistency of $\hat{c}^{(2)}$.

With respect to the consistency of $\hat{c}^{(3)}$, we begin by noting that the domain of attraction condition

$$\lim_{n\to\infty} \frac{n}{k}\, P\Big\{X_l(0) > U_0\Big(\frac{n}{k}\Big) + x\, a_0\Big(\frac{n}{k}\Big)\Big\} = (1 + \gamma x)^{-1/\gamma}$$

for all $l = 1, 2, \ldots$, combined with (1) implies

$$\lim_{n \to \infty} e^{-cs_j} \frac{n}{k} P\Big\{ X_l(s_j) > U_0\big(\tfrac{n}{k}\big) + x\, a_0\big(\tfrac{n}{k}\big) \Big\} = (1 + \gamma x)^{-1/\gamma},$$

$j = 1, 2, \ldots, m$. Hence the characteristic functions converge:

$$E \exp\Big\{ i \frac{t}{k} \sum_{l=1}^n I_{\{X_l(s_j) > U_0(n/k) + x\, a_0(n/k)\}} \Big\}$$

$$= \Big(E \exp\Big\{ i \frac{t}{k} I_{\{X_1(s_j) > U_0(n/k) + x\, a_0(n/k)\}} \Big\} \Big)^n$$

$$= \Big(e^{i \frac{t}{k}} P\Big\{ X_1(s_j) > U_0\big(\tfrac{n}{k}\big) + x\, a_0\big(\tfrac{n}{k}\big) \Big\}$$

$$+ 1 - P\Big\{ X_1(s_j) > U_0\big(\tfrac{n}{k}\big) + x\, a_0\big(\tfrac{n}{k}\big) \Big\} \Big)^n$$

$$= \Big(1 + e^{cs_j} k \big(e^{it/k} - 1 \big) \frac{e^{-cs_j}(n/k)\big[1 - F_{s_j}(U_0(n/k) + x a_0(n/k)) \big]}{n} \Big)^n$$

$$\xrightarrow[n \to \infty]{} \exp\Big\{ it\, e^{cs_j} (1 + \gamma x)^{-1/\gamma} \Big\},$$

for every $t \in \mathbb{R}$. Owing to Lévy's continuity theorem, the latter implies

$$\frac{1}{k} \sum_{i=1}^n I_{\{X_i(s_j) > U_0(n/k) + x\, a_0(n/k)\}} \xrightarrow[n \to \infty]{P} e^{cs_j}(1 + \gamma\, x)^{-1/\gamma}.$$

Next use (30). ❏

For the proof of the asymptotic normality we need an auxiliary result.

Lemma 8 *Assume conditions* (20) *and* (21). *Define*

$$\alpha_s(t) := e^{cs\rho} \alpha_0(t),$$

$$a_s(t) := e^{cs\gamma} a_0(t) \Big(1 + \alpha_0(t) \frac{e^{cs\rho} - 1}{\rho} \Big). \tag{31}$$

Then for $s \in \mathbb{R}$ and $x > 0$

$$\lim_{t \to \infty} \frac{\frac{U_s(tx) - U_s(t)}{a_s(t)} - \frac{x^\gamma - 1}{\gamma}}{\alpha_s(t)} = H_{\gamma, \rho}(x). \tag{32}$$

Proof: For simplicity we write d for e^{cs} in this proof.

Relation (20) implies (cf. de Haan and Ferreira, 2006, p.44) that for $x > 0$

$$\lim_{t \to \infty} \frac{\frac{a_0(tx)}{a_0(t)} - x^\gamma}{\alpha_0(t)} = x^\gamma \frac{x^\rho - 1}{\rho}$$

and

$$\lim_{t\to\infty} \frac{\alpha_0(tx)}{\alpha_0(t)} = x^\rho.$$

First note that (20) and (21) imply that

$$
\frac{\frac{U_s(tx)-U_s(t)}{a_0(t)} - d^\gamma \frac{x^\gamma-1}{\gamma}}{\alpha_0(t)} = \frac{\frac{U_0(tx)-U_0(t)}{a_0(t)} - \frac{x^\gamma-1}{\gamma}}{\alpha_0(t)} + \frac{\frac{a_0(tx)}{a_0(t)} - x^\gamma}{\alpha_0(t)} \frac{d^\gamma-1}{\gamma}
$$
$$
+ \left(\frac{\alpha_0(tx)\, a_0(tx)}{\alpha_0(t)\, a_0(t)} - 1 \right) H_{\gamma,\rho}(d)\big(1+o(1)\big)
$$

converges to

$$H_{\gamma,\rho}(x) + x^\gamma \frac{x^\rho-1}{\rho}\frac{d^\gamma-1}{\gamma} + (x^{\gamma+\rho}-1)H_{\gamma,\rho}(d).$$

Next write

$$
\frac{\frac{U_s(tx)-U_s(t)}{a_s(t)} - \frac{x^\gamma-1}{\gamma}}{\alpha_s(t)} = \frac{\alpha_0(t)}{\alpha_s(t)}\left(\frac{\frac{U_s(tx)-U_s(t)}{a_0(t)} - d^\gamma\frac{x^\gamma-1}{\gamma}}{\alpha_0(t)}\frac{a_0(t)}{a_s(t)} + \frac{x^\gamma-1}{\gamma}\frac{\frac{d^\gamma a_0(t)}{a_s(t)}-1}{\alpha_0(t)} \right).
$$

This converges to

$$
d^{-(\gamma+\rho)}\left(H_{\gamma,\rho}(x) + x^\gamma \frac{x^\rho-1}{\rho}\frac{d^\gamma-1}{\gamma} + (x^{\gamma+\rho}-1)H_{\gamma,\rho}(d) \right) - \frac{x^\gamma-1}{\gamma}\frac{1-d^{-\rho}}{\rho}
$$

which is equal to $H_{\gamma,\rho}(x)$. ❏

Remark 9 *The analogue of Lemma 8 stemming from conditions* (16) *and* (17) *– i.e.* $\gamma > 0$ *– holds with the auxiliary function* $\beta_s(t) := e^{cs\widetilde{\rho}}\beta(t)$. *This leads to the following relation for every* $s \in \mathbb{R}$,

$$\lim_{t\to\infty} \frac{\frac{U_s(tx)}{U_s(t)} - x^{\gamma^+}}{\beta_s(t)} = x^{\gamma^+}\frac{x^{\widetilde{\rho}}-1}{\widetilde{\rho}}, \quad x > 0.$$

Proof of asymptotic normality

We write

$$
\sqrt{k}\left(\hat{c}^{(1)} - c\right) = \sqrt{k}\left(\frac{1}{\hat{\gamma}_{n,k}^+} - \frac{1}{\gamma^+} \right)\hat{\gamma}_{n,k}^+\,\hat{c}^{(1)} + \left(\gamma^+ \sum_{j=1}^m s_j^2 \right)^{-1}
$$
$$
\times \Bigg\{ \sum_{j=1}^m s_j\Big[\sqrt{k}\Big(\log X_{n-k,n}(s_j) - \log U_{s_j}\big(\tfrac{n}{k}\big) \Big)
$$
$$
- \sqrt{k}\Big(\log X_{n-k,n}(0) - \log U_0\big(\tfrac{n}{k}\big) \Big)
$$
$$
+ \sqrt{k}\,\beta\big(\tfrac{n}{k}\big)\frac{\log U_{s_j}(n/k) - \log U_0(n/k) - c\gamma^+ s_j}{\beta(n/k)} \Big] \Bigg\}.
$$

The result follows from (18), (17) and (19).

For $\hat{c}^{(2)}$ it is sufficient to consider

$$
\sqrt{k}\left\{\log\left(1 + \hat{\gamma}_{n,k}\frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)}\right)^{\frac{1}{\hat{\gamma}_{n,k}}} - cs_j\right\}
$$

for $j = 1, 2, \ldots, m$ where $\hat{\gamma}_{n,k} = 1/m \sum_{j=1}^m \hat{\gamma}_{n,k}(s_j)$. We use Cramér's delta method.

$$
\frac{\partial}{\partial\gamma}\log(1 + \gamma x)^{\frac{1}{\gamma}} = \frac{1}{\gamma^2}\left(\frac{\gamma x}{1 + \gamma x} - \log(1 + \gamma x)\right)
$$

(which is $x^2/2$ for $\gamma = 0$) and

$$
\frac{\partial}{\partial x}\log(1 + \gamma x)^{\frac{1}{\gamma}} = \frac{1}{1 + \gamma x}
$$

($= 1$ for $x = 0$). Further we write

$$
\frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)}
$$
$$
= \frac{a_0\left(\frac{n}{k}\right)}{\hat{a}_0\left(\frac{n}{k}\right)}\left\{\frac{a_{s_j}\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)}\frac{X_{n-k,n}(s_j) - U_{s_j}\left(\frac{n}{k}\right)}{a_{s_j}\left(\frac{n}{k}\right)}\right.
$$
$$
\left. - \frac{X_{n-k,n}(0) - U_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} + \frac{U_{s_j}\left(\frac{n}{k}\right) - U_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)}\right\},
$$

implying (cf. (10))

$$
\sqrt{k}\left(\frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)} - \frac{e^{c\gamma s_j} - 1}{\gamma}\right)
$$
$$
= \frac{e^{c\gamma s_j} - 1}{\gamma}\sqrt{k}\left(\frac{a_0\left(\frac{n}{k}\right)}{\hat{a}_0\left(\frac{n}{k}\right)} - 1\right) + \frac{a_0\left(\frac{n}{k}\right)}{\hat{a}_0\left(\frac{n}{k}\right)}\left\{\frac{a_{s_j}\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)}\sqrt{k}\frac{X_{n-k,n}(s_j) - U_{s_j}\left(\frac{n}{k}\right)}{a_{s_j}\left(\frac{n}{k}\right)}\right.
$$
$$
\left. - \sqrt{k}\frac{X_{n-k,n}(0) - U_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} + \sqrt{k}\left(\frac{U_{s_j}\left(\frac{n}{k}\right) - U_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} - \frac{e^{c\gamma s_j} - 1}{\gamma}\right)\right\}.
$$

Hence by (10)

$$
\sqrt{k}\left(\frac{X_{n-k,n}(s_j) - X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)} - \frac{e^{c\gamma s_j} - 1}{\gamma}\right)
$$
$$
\xrightarrow[n\to\infty]{d} -\frac{e^{c\gamma s_j} - 1}{\gamma}A(0) + e^{c\gamma s_j}B(s_j) - B(0) + \lambda H_{\gamma,\rho}(e^{cs_j}).
$$

Next we apply the delta method:

$$\sqrt{k}\Big\{\log\Big(1+\hat{\gamma}_{n,k}\frac{X_{n-k,n}(s_j)-X_{n-k,n}(0)}{\hat{a}_0\left(\frac{n}{k}\right)}\Big)^{\frac{1}{\hat{\gamma}_{n,k}}}$$
$$-\log\Big(1+\gamma\frac{e^{c\gamma s_j}-1}{\gamma}\Big)^{\frac{1}{\gamma}}\Big\}$$

$$\xrightarrow[n\to\infty]{d}\quad \Big[\frac{\partial}{\partial\gamma}\,\log(1+\gamma x)^{\frac{1}{\gamma}}\Big]_{x=\frac{e^{c\gamma s_j}-1}{\gamma}}\cdot\frac{1}{m}\sum_{i=1}^{m}\Gamma(s_i)$$
$$+\Big[\frac{\partial}{\partial x}\,\log(1+\gamma x)^{\frac{1}{\gamma}}\Big]_{x=\frac{e^{c\gamma s_j}-1}{\gamma}}\Big\{-\frac{e^{c\gamma s_j}-1}{\gamma}\,A(0)$$
$$+e^{c\gamma s_j}B(s_j)-B(0)+\lambda\,H_{\gamma,\rho}(e^{cs_j})\Big\}$$

$$=\quad \frac{1-e^{-c\gamma s_j}-c\gamma s_j}{\gamma^2}\frac{1}{m}\sum_{i=1}^{m}\Gamma(s_i)+e^{-c\gamma s_j}\Big\{e^{c\gamma s_j}B(s_j)$$
$$-B(0)-\frac{e^{c\gamma s_j}-1}{\gamma}A(0)+\lambda\,H_{\gamma,\rho}(e^{cs_j})\Big\}.$$

The result follows.

Finally for $\hat{c}^{(3)}$ consider first

$$\frac{1}{k}\sum_{i=1}^{n}I_{\left\{X_i(s)>X_{n-k,n}(0)\right\}}=\frac{1}{k}\sum_{i=1}^{n}I_{\left\{\frac{X_i(s)-U_s^\star(n/k)}{a_s^\star(n/k)}>\frac{X_{n-k,n}(0)-U_s^\star(n/k)}{a_s^\star(n/k)}\right\}}. \tag{33}$$

We write

$$x_n(s):=\frac{X_{n-k,n}(0)-U_s^\star\left(\frac{n}{k}\right)}{a_s^\star\left(\frac{n}{k}\right)},$$

with $U^\star$ and $a^\star$ from Corollary 2.3.7 (but with a different notation since we use the subscript 0 for another purpose here) of de Haan and Ferreira (2006), coupled with Lemma 8,

$$\alpha_s^\star(t)\quad:=\quad\begin{cases}\alpha_0(t)\frac{e^{cs\rho}}{\rho}, & \rho<0,\\ \alpha_0(t), & \rho=0,\end{cases}$$

$$a_s^\star(t)\quad:=\quad\begin{cases}e^{cs\gamma}a_0(t)\big(1-\alpha_0(t)\frac{1}{\rho}\big), & \rho<0,\\ e^{cs\gamma}a_0(t)\big(1+\alpha_0(t)\{cs-\frac{1}{\gamma}\}\big), & \rho=0\neq\gamma,\\ a_0(t)\big(1+\alpha_0(t)cs\big), & \gamma=\rho=0,\end{cases}$$

$$U_s^\star(t)\quad:=\quad\begin{cases}U_s(t)-\frac{e^{cs(\gamma+\rho)}}{\gamma+\rho}\,a_0^\star(t)\alpha_0^\star(t), & \gamma+\rho\neq0,\,\rho<0,\\ U_s(t), & \text{otherwise},\end{cases}$$

and write (33) as

$$\frac{n}{k}\left\{1 - F_n^{(s)}\left(U_s^\star\left(\frac{n}{k}\right) + x_n(s)\,a_s^\star\left(\frac{n}{k}\right)\right)\right\}.$$

where $F_n^{(s)}$ is the empirical distribution function of the random sample $X_1(s), X_2(s), \ldots, X_n(s)$. We consider $x_n(s)$ first. We use (21) and Theorem 2.4.2 of de Haan and Ferreira (2006).

$$
\begin{aligned}
&\sqrt{k}\Big(x_n(s) - \frac{e^{-cs\gamma}-1}{\gamma}\Big)\\
=\ &\sqrt{k}\Big(\frac{X_{n-k,n}(0) - U_s^\star\left(\frac{n}{k}\right)}{a_s^\star\left(\frac{n}{k}\right)} - \frac{e^{-cs\gamma}-1}{\gamma}\Big)\\
=\ &\frac{a_0^\star\left(\frac{n}{k}\right)}{a_s^\star\left(\frac{n}{k}\right)}\sqrt{k}\,\frac{X_{n-k,n}(0) - U_0^\star\left(\frac{n}{k}\right)}{a_0^\star\left(\frac{n}{k}\right)} - \frac{a_0^\star\left(\frac{n}{k}\right)}{a_s^\star\left(\frac{n}{k}\right)}\sqrt{k}\Big(\frac{U_s^\star\left(\frac{n}{k}\right) - U_0^\star\left(\frac{n}{k}\right)}{a_0^\star\left(\frac{n}{k}\right)} - \frac{e^{cs\gamma}-1}{\gamma}\Big)\\
&-\sqrt{k}\Big(\frac{a_0^\star\left(\frac{n}{k}\right)}{a_s^\star\left(\frac{n}{k}\right)} - e^{-cs\gamma}\Big)\frac{e^{cs\gamma}-1}{\gamma}\\
=\ &\frac{a_0^\star\left(\frac{n}{k}\right)}{a_s^\star\left(\frac{n}{k}\right)}\sqrt{k}\,\frac{X_{n-k,n}(0) - U_0\left(\frac{n}{k}\right)}{a_0^\star\left(\frac{n}{k}\right)} - \frac{a_0^\star\left(\frac{n}{k}\right)}{a_s^\star\left(\frac{n}{k}\right)}\sqrt{k}\Big(\frac{U_0^\star\left(\frac{n}{k}\right) - U_0\left(\frac{n}{k}\right)}{a_0^\star\left(\frac{n}{k}\right)} + \frac{U_s^\star\left(\frac{n}{k}\right) - U_0^\star\left(\frac{n}{k}\right)}{a_0^\star\left(\frac{n}{k}\right)} - \frac{e^{cs\gamma}-1}{\gamma}\Big)\\
&-\sqrt{k}\Big(\frac{a_0^\star\left(\frac{n}{k}\right)}{a_s^\star\left(\frac{n}{k}\right)} - e^{-cs\gamma}\Big)\frac{e^{cs\gamma}-1}{\gamma}. \qquad (34)
\end{aligned}
$$

The first term converges in distribution to $e^{-cs\gamma}W^{(0)}(1)$. The second term converges to $\lambda^\star b(s)$ defined by

$$
b(s) := \begin{cases}
\frac{e^{-cs\gamma}}{\gamma+\rho}, & \gamma+\rho \neq 0,\ \rho < 0,\\
-e^{-cs\gamma}cs, & \gamma+\rho = 0,\ \rho < 0,\\
-\frac{cs}{\gamma}, & \rho = 0 \neq \gamma,\\
-\frac{1}{2}(cs)^2, & \gamma = \rho = 0
\end{cases}
$$

and the third term converges to

$$\lambda^\star e^{-cs\gamma}\frac{e^{cs\gamma}-1}{\gamma}cs\,I_{\{\rho=0\}}.$$

Here

$$\lambda^\star := \lim_{n\to\infty}\sqrt{k}\,\alpha_0^\star(n/k) = \begin{cases} \lambda/\rho, & \rho < 0\\ \lambda, & \rho = 0. \end{cases}$$

Further by Theorem 5.1.2 of de Haan and Ferreira (2006), since $x = x_n$ is asymptotically constant,

$$
\begin{aligned}
\sqrt{k}\Big\{&\frac{n}{k}\left(1 - F_n^{(s)}\left(U_s^\star\left(\frac{n}{k}\right) + x_n(s)\,a_s^\star\left(\frac{n}{k}\right)\right)\right) - (1 + \gamma\,x_n(s))^{-\frac{1}{\gamma}}\Big\}\\
&- W_n^{(s)}\Big((1 + \gamma\,x_n(s))^{-\frac{1}{\gamma}}\Big)\\
&- \sqrt{k}\,\alpha_s^\star\left(\frac{n}{k}\right)(1 + \gamma\,x_n(s))^{-\frac{1}{\gamma}-1}\overline{\Psi}_{\gamma,\rho}\Big((1 + \gamma\,x_n(s))^{\frac{1}{\gamma}}\Big) \xrightarrow[n\to\infty]{P} 0
\end{aligned} \qquad (35)
$$

for a sequence of standard Brownian motions $\{W_n^{(s)}(t)\}_{t \geq 0}$ and with

$$
\overline{\Psi}_{\gamma,\rho}(x) := \begin{cases}
\frac{x^{\gamma+\rho}}{\gamma+\rho}, & \gamma+\rho \neq 0,\ \rho < 0, \\
\log x, & \gamma+\rho = 0,\ \rho < 0, \\
\frac{1}{\gamma}\, x^\gamma \log x, & \rho = 0 \neq \gamma, \\
\frac{1}{2}\,(\log x)^2, & \gamma = \rho = 0.
\end{cases}
$$

Finally we write

$$
\sqrt{k}\left\{ \frac{n}{k}\left(1 - F_n^{(s)}\left(U_s^\star\left(\frac{n}{k}\right) + x_n(s)\, a_s^\star\left(\frac{n}{k}\right)\right)\right) - e^{cs} \right\}
$$

$$
= \sqrt{k}\left\{ \frac{n}{k}\left(1 - F_n^{(s)}\left(U_s^\star\left(\frac{n}{k}\right) + x_n(s)\, a_s^\star\left(\frac{n}{k}\right)\right)\right) - \left(1 + \gamma\, x_n(s)\right)^{-\frac{1}{\gamma}} \right\}
$$

$$
+ \sqrt{k}\left\{ \left(1 + \gamma\, x_n(s)\right)^{-\frac{1}{\gamma}} - \left(1 + \gamma\, \frac{e^{-cs\gamma}-1}{\gamma}\right)^{-\frac{1}{\gamma}} \right\}.
$$

Since $x_n(s) \xrightarrow[n\to\infty]{P} (e^{-cs\gamma} - 1)/\gamma$, by (35) the first term converges in distribution to

$$
W^{(s)}(e^{cs}) - \lambda^\star e^{cs(\gamma+1)}\overline{\Psi}_{\gamma,\rho}(e^{-cs}).
$$

By (34) and the delta method the second term converges to

$$
-e^{cs} W^{(0)}(1) - \lambda^\star e^{cs(\gamma+1)}\left(b(s) + \frac{1 - e^{-cs\gamma}}{\gamma}\, cs\, I_{\{\rho=0\}}\right).
$$

The result follows by Carmér's delta method again,

$$
\sqrt{k}\left\{ \log\left[\frac{n}{k}\left(1 - F_n^{(s)}\left(U_s^\star\left(\frac{n}{k}\right) + x_n(s)\, a_s^\star\left(\frac{n}{k}\right)\right)\right)\right] - cs \right\}
$$

$$
\xrightarrow[n\to\infty]{d} e^{-cs} W^{(s)}(e^{cs}) - W^{(0)}(1) - \lambda^\star e^{cs\gamma}\left(\overline{\Psi}(e^{-cs}) + b(s) + \frac{1 - e^{-cs\gamma}}{\gamma}\, cs\, I_{\{\rho=0\}}\right).
$$

❏

Proof:[of Corollary 6] By virtue of Rao's theorem (Rao (1973), Section 3.b.4; see also Serfling (2002), p.128) pertaining to quadratic forms of asymptotically normal random vectors, statements (25) and (26) follow immediately from the theorem.

❏

A Proofs concerning relations (1), (5), (6), (10)

<u>Proof</u> of (1) $\Leftrightarrow$ (5).

We use the following result of S.I. Resnick (Resnick, 1971; de Haan and Ferreira, 2006, Lemma 1.2.12 p.23). Let F_1 and F_2 be two probability distribution functions and let $F_1 \in \mathcal{D}(G_\gamma)$ for some $\gamma \in \mathbb{R}$. The following two statements are equivalent (with x^* the right endpoint of F_1).

(i) $\qquad \lim_{x\uparrow x^*} \dfrac{1 - F_2(x)}{1 - F_1(x)} = 1$

(ii) $\qquad \lim_{t\to\infty} \dfrac{U_2(t) - U_1(t)}{a_1(t)} = 0$

where $U_i := \left(1/(1 - F_i)\right)^{\leftarrow}$, $i = 1, 2$. Define the distribution function F_s^* by

$$1 - F_s^*(x) = \min\left(1,\ e^{-cs}\left(1 - F_s(x)\right)\right).$$

Then (using (i), (ii))

$$(1) \ \Leftrightarrow \ \lim_{x\uparrow x^*} \frac{1 - F_s^*(x)}{1 - F_0(x)} = 1 \ \Leftrightarrow \ \lim_{t\to\infty} \frac{U_s^*(t) - U_0(t)}{a_0(t)} = 0$$

which holds if and only if (since $U_s^*(t) = \left(1/(e^{-sc}(1 - F_s))\right)^{\leftarrow}(t) = U_s(t\,e^{-cs})$)

$$\lim_{t\to\infty} \frac{U_s(t\,e^{-sc}) - U_0(t)}{a_0(t)} = 0.$$

Hence (1) holds if and only if

$$\frac{U_s(t) - U_0(t)}{a_0(t)} = \frac{U_s(t) - U_0(t\,e^{cs})}{a_0(t\,e^{cs})} \frac{a_0(t\,e^{cs})}{a_0(t)} + \frac{U_0(t\,e^{cs}) - U_0(t)}{a_0(t)} \longrightarrow 0 + \frac{e^{\gamma cs} - 1}{\gamma}$$

as $t \to \infty$. Here we have used that $F_0 \in \mathcal{D}(G_\gamma)$ implies

$$\lim_{t\to\infty} \frac{U_0(tx) - U_0(t)}{a_0(t)} = \frac{x^\gamma - 1}{\gamma} \qquad \text{and} \qquad \lim_{t\to\infty} \frac{a_0(tx)}{a_0(t)} = x^\gamma$$

for $x > 0$.

<u>Proof</u> of (1) $\Leftrightarrow$ (6) for $\gamma > 0$.

The proof of Lemma 1.2.12 in de Haan and Ferreira (2006) shows that for $\gamma > 0$ the following statements are equivalent.

(i) $\qquad \lim_{x\to\infty} \dfrac{1 - F_2(x)}{1 - F_1(x)} = 1$

(ii) $\quad \lim\limits_{t \to \infty} \dfrac{U_2(t)}{U_1(t)} = 1.$

The rest of the proof is very similar to the previous one and is omitted.

<u>Proof</u> of (1) $\Rightarrow$ (10).

$\quad F_s \in \mathcal{D}(G_\gamma)$ implies

$$\lim_{t \to \infty} t \left\{ 1 - F_s \big(U_s(t) + x a_s(t) \big) \right\} = (1 + \gamma x)^{-1/\gamma}. \tag{36}$$

Relation (1) implies

$$\lim_{t \to \infty} t \left\{ 1 - F_s \big(U_0(t) + x a_0(t) \big) \right\} = e^{cs} (1 + \gamma x)^{-1/\gamma}. \tag{37}$$

We combine this with relation (8) and

$$\lim_{t \to \infty} \frac{a_0(t\, e^{sc})}{a_0(t)} = e^{sc\gamma}. \tag{38}$$

We then get

$$
\begin{aligned}
& t \left\{ 1 - F_s \big(U_0(t\, e^{sc}) - x a_0(t\, e^{sc}) \big) \right\} \\
=\ & t \left\{ 1 - F_s \Big(U_0(t) + \frac{e^{sc\gamma - 1}}{\gamma} a_0(t) \big(1 + o(1)\big) + x\, a_0(t) e^{sc\gamma} \big(1 + o(1)\big) \Big) \right\} \\
=\ & t \left\{ 1 - F_s \Big(U_0(t) + a_0(t) \big(\frac{e^{cs\gamma} - 1}{\gamma} + x\, e^{sc\gamma} \big) \big(1 + o(1)\big) \Big) \right\} \\
\longrightarrow\ & e^{cs} \Big(1 + \gamma \big\{ \frac{e^{sc\gamma} - 1}{\gamma} + x\, e^{sc\gamma} \big\} \Big)^{-1/\gamma} = (1 + \gamma x)^{-1/\gamma}
\end{aligned}
$$

as $t \to \infty$. Comparing this with (36) we conclude by Khinchine's convergence-to-types theorem

$$\lim_{t \to \infty} \frac{U_s(t) - U_0(t)}{a_0(t)} = 0 \qquad \text{and} \qquad \lim_{t \to \infty} \frac{a_s(t)}{a_0(t\, e^{sc})} = 1.$$

The latter relation gives the result by (38).

B   Sketch of alternative approaches

The subject of extreme value theory (EVT) is the study of the right (or left) tail of a probability distribution near the endpoint. Hence by nature EVT is an asymptotic theory. The basic assumption is

$$\lim_{n \to \infty} P \Big\{ \frac{\max(X_1, X_2, \dots, X_n) - b_n}{a_n} \le x \Big\} = \exp \big\{ -(1 + \gamma\, x)^{-1/\gamma} \big\}, \tag{39}$$

where $X_1, X_2, \ldots$ are i.i.d. random variables. It follows that the limit distribution has only one parameter, the shift $b_n \in \mathbb{R}$ and scale $a_n > 0$ are not parameters of the limit distribution. They depend essentially on the distribution of X_1.

When it comes to statistics there are three basic methods:

1. <u>Yearly maxima</u> (or block maxima). Over a number of years one takes the yearly maximum. Since the yearly maximum is taken over many underlying random variables (albeit not i.i.d.) the assumption is that the yearly maximum M_j can be considered the maximum over a large number n of i.i.d random variables so that

$$P\{M_j \leq x\} \approx \exp\left\{-\left(1 + \gamma \frac{x - b_n}{a_n}\right)^{-\frac{1}{\gamma}}\right\}$$

where n is unknown. The random variables M_j are i.i.d.. The right hand-side can then be interpreted as a parametric model (GEV: Generalized Extreme Value distribution) so that e.g. the method of maximum likelihood can be applied.

The interpretation of b_n is: the level that has a return period (the mean time between consecutive exceedances of the level) of $e/(e-1) \approx 1.58$ years. Hence there is no direct intuitive meaning for b_n. Also the behavior of b_n as $n \to \infty$ can not be found. This method carries a bias stemming from replacing an approximate equality with a firm equality. In contrast to the next case it seems difficult to control that bias.

2. <u>Peaks over threshold</u>. The basic assumption (39) implies that with $b(t) = b_{[t]}$, $a(t) = a_{[t]}$ and $[t]$ the integer part of t for $x > 0$

$$\lim_{t \to \infty} P\left\{\frac{X_1 - b(t)}{a(t)} > x | X_1 > b(t)\right\} = (1 + \gamma x)^{-1/\gamma}. \tag{40}$$

Select out of n i.i.d. observations the ones that are larger than $b(t)$. These are approximately i.i.d. and (when normalized) follow approximately the GPD distribution $1 - (1 + \gamma x)^{-1/\gamma}$ (Pickands (1975)).

One can take for $b(t)$ one of the order statistics $X_{n-k,n}$. In order to get meaningful results we need to have $k = k_n$ and $k_n \to \infty$, $k(n)/n \to 0$ as $n \to \infty$. Then $X_{n-k,n}$ is close to $b(n/k)$, i.e., the quantile $F^{\leftarrow}(1 - k/n)$.

Again, since for $x > b(t)$

$$P\{X_1 > x | X_1 > b(t)\} \approx \left(1 + \frac{x - b(t)}{a(t)}\right)^{-1/\gamma},$$

one can consider the right hand-side as a parametric model so that e.g. the method of maximum likelihood can be applied.

Next one can prove that the obtained estimators are consistent and asymptotically normal as n, the

number of observations, tends to infinity. That is, the vector

$$\sqrt{k}\left(\frac{\hat{a}\left(\frac{n}{k}\right)}{a\left(\frac{n}{k}\right)} - 1,\ \hat{\gamma}_{n,k} - \gamma\right)$$

has asymptotically a normal distribution(Smith (1987); Drees et al. (2004); Zhou (2009)). There are also methods to minimize the bias by choosing the number k appropriately.

3. <u>Point process convergence</u>. Suppose that the basic assumption holds. Take the point process on $\mathbb{R}^2$ with points

$$\left\{\left(\frac{i}{n},\ \frac{X_i - b_n}{a_n}\right)\right\}_{i=1}^{n}. \tag{41}$$

This point process converges in distribution to a Poisson point process on $(0,1) \times \mathbb{R}$ with intensity measure $dt \cdot (1 + \gamma x)^{-1/\gamma - 1} dx$ (cf. Pickands (1971)). Note that the intensity measure is unbounded. Those points in (41) for which $(X_i - b_n)/a_n$ exceeds some threshold u are approximately points from a Poisson point process with (finite) parametric intensity measure so that the method of maximum likelihood can be applied supplying estimators for γ, b_n and a_n (see Smith (1989), cf. Coles (2001), Chapter 7). No asymptotic behavior ($n \to \infty$, $u = u_n$ decreasing) seems to be known for these estimators.

The three methods have been explained in detail in the book of Coles (2001). A trend in the EVT analysis has been considered in all three methods.

1$'$. Chapter 6 of Coles (2001) book treats trends in the block maxima / GEV setup. One considers time points $j = 0, 1, 2, \ldots$ and assumes (for example)

$$b_n(j) = b_n(0) + j\,c$$

or/and

$$\log a_n(j) = \log a_n(0) + j\,c'.$$

As we saw before, b_n is the level that has a return period of just $e/(e-1)$ years. The scale a_n can be interpreted with some liberty as a derivative, i.e., speed of change of location. The interpretation of both seems less straightforward than that of (1).

There is also another complication. If one is interested in the location parameter over a longer period, say, of 2 years i.e. n replaced with $2n$, the relation is

$$
\begin{aligned}
b_{2n}(j) &\approx b_n(j) + a_n(j)\frac{2^\gamma - 1}{\gamma} \\
&\approx b_{2n}(0) + j\,c + (a_n(j) - a_n(0))\frac{2^\gamma - 1}{\gamma}.
\end{aligned}
$$

This is no longer a linear trend in general. Note that our framework, combining trends in location and scale, is not bound to a certain period.

2′.<u>Peaks over threshold</u>. Davison and Smith (1990) consider a linear trend in both γ and $a(n/k)$. Coles (2001), p.119, considers a linear trend in $\log a(n/k)$. Estimation is done by maximum likelihood. No asymptotic analysis of the quality of the estimators as the number of observations tends to infinity is made. Again the interpretation of a trend in $a(n/k)$ or $\log a(n/k)$ seems difficult.

Hall and Tajvidi (2000) consider a nonlinear trend in γ. The method is likelihood based. No large sample results are given.

3′. Smith (1989) (c.f. Coles (2001) p.133 sqq.) considers a trend in the location

$$b_n(j) = b_n(0) + j\,c$$

(simplified) in the point process model. Estimation is by maximum likelihood. No asymptotic analysis ($n \to \infty$) is made. Another possible problem is that the trend for $b_n(j)$ could get out of range if $\gamma < 0$.

In short: the present paper looks at changes in (tail) probabilities whereas in the literature changes in various quantiles have been considered. The two viewpoints are not equivalent.

C   An auxiliary result

Lemma 10 *Let $X_1, X_2, X_3, \ldots$ be i.i.d. random variables with distribution function F. Let $X_{1,n} \leq X_{2,n} \leq \ldots \leq X_{n,n}$ be the n-th order statistics. Define $U := \big(1/(1-F)\big)^{\leftarrow}$. If $F \in \mathcal{D}(G_\gamma)$ for some $\gamma \in \mathbb{R}$, then*

$$\frac{X_{n-k,n}}{U\left(\frac{n}{k}\right)} \xrightarrow[n\to\infty]{P} 1,$$

as $n \to \infty$ where $k = k_n$, $k_n \to \infty$, $k_n/n \to 0$ as $n \to \infty$.

Proof: First note that $\{X_1, X_2, X_3, \ldots\} \overset{d}{=} \{U(Y_1), U(Y_2), U(Y_3), \ldots\}$ where $Y_1, Y_2, Y_3, \ldots$ are i.i.d. with distribution function $1 - 1/x$, $x \geq 1$. Hence $X_{n-k,n} \overset{d}{=} U(Y_{n-k,n})$ for all n. Next note that

$$\frac{k}{n} Y_{n-k,n} \xrightarrow[n\to\infty]{P} 1 \tag{42}$$

(cf. e.g. Corollary 2.2.2 p.41 de Haan and Ferreira, 2006).

The regular variation of U (cf. Lemma 1.2.9 p.22 de Haan and Ferreira, 2006) implies

$$\lim_{t\to\infty} \frac{U\big(t\,x(t)\big)}{U(t)} = 1 \tag{43}$$

provided $\lim_{t \to \infty} x(t) = 1$. Hence

$$\frac{X_{n-k,n}}{U\left(\frac{n}{k}\right)} \stackrel{d}{=} \frac{U(Y_{n-k,n})}{U\left(\frac{n}{k}\right)} \stackrel{d}{=} \frac{U\left(\frac{n}{k}\left(\frac{k}{n}Y_{n-k,n}\right)\right)}{U\left(\frac{n}{k}\right)} \xrightarrow[n\to\infty]{P} 1$$

where in the last step we have used (42) and (43).

❏

References

AghaKouchak, A. and Nasrollahi, N. (2010). Semi-parametric and parametric inference of extreme value models for rainfall data. *Water Resour. Manage.*, 24:1229–1249.

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa Kumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D. B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., and Vazquez-Aguirre, J. L. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.*, 111:(D05109).

Allen, M. R. and Ingram, W. J. (2002). Constraints on future changes in climate and the hydrologic cycle. *Nature*, 419:224–232.

Buishand, T. A., de Haan, L., and Zhou, C. (2008). On spatial extremes: With application to a rainfall problem. *Ann. Appl. Statist.*, 2:624–642.

Burauskaite-Harju, A., Grimvall, A., and von Brömssen, C. (2012). Analysing trends in precipitation extremes for a network of stations. *International Journal of Climatology*, 32:86–94.

Caeiro, F., Gomes, M. I., and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat*, 3:113–136.

Cai, J., de Haan, L., and Zhou, C. (2013). Direct reduction of bias of the classical Hill estimator. *Extremes*, 16:173–201.

Chavez-Demoulin, V. (1999). *Two Problems in Environmental Statistics: Capture-recapture Analysis and Smooth Extremal Models, PhD Thesis*. Department of Mathematics, Swiss Federal Institute of Technology, Lausanne.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer Verlag.

Davison, A. C. and Ramesh, N. (2000). Local likelihood smoothing of sample extremes. *J.R. Statist. Soc. B*, 62:191–208.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds (with discussion). *J.R. Statist. Soc. B*, 52:393–442.

de Haan, L. (1990). Fighting the arch-enemy with mathematics. *Statist. Neerlandica*, 44:45–68.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer.

Drees, H., Ferreira, A., and de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *Ann. Appl. Probab.*, 14:1179–1201.

Gaetan, C. and Grigoletto, M. (2004). Smoothing sample extremes with dynamic models. *Extremes*, 7:221–236.

Gomes, M. I., de Haan, L., and Henriques Rodrigues, L. (2008). Tail index estimation through accommodation of bias in the weighted log-excesses. *J. Roy. Statist. Soc. Ser. B*, 70:31–52.

Groisman, P. Y., Knight, R. W., Easterling, D. R., Karl, T. R., Hegerl, G. C., and Razuvaev, V. N. (2005). Trends in intense precipitation in the climate record. *J. Climate*, 18:1326–1350.

Hall, P. and Tajvidi, N. (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statist. Sci.*, 15:153–167.

Hanel, M., Buishand, T. A., and Ferro, C. A. T. (2009). A nonstationary index flood model for precipitation extremes in transient regional climate model simulations. *J. Geophys. Res.*, 114:(D15107).

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

Klein Tank, A. M. G. and Können, G. P. (2003). Trends in indices of daily temperature and precipitation extremes in Europe, 1946-99. *J. Climate*, 16:3665–3680.

Kottek, M., Grieser, J., Christoph, B., Rudolph, B., and Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorol. Z.*, 15:259–263.

Lenderink, G., van Meijgaard, E., and Selten, F. (2009). Intense coastal rainfall in The Netherlands in response to high sea surface temperatures: analysis of the event of august 2006 from the perspective of a changing climate. *Clim. Dyn.*, 32:19–33.

Mannshardt-Shamseldin, E. C., Smith, R. L., Sain, S. R., Mearns, L. O., and Cooley, D. (2010). Downscaling extremes: A comparison of extreme value distributions in point-source and gridded precipitation data. *Ann. Appl. Statist.*, 4:484–502.

Pauli, F. and Coles, S. G. (2001). Penalized likelihood inference in extreme value analyses. *J. Appl. Stat.*, 28:547–560.

Peel, M., Finlayson, B. L., and McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.*, 11:1633–1644.

Pickands, J. (1971). The two-dimensional Poisson process and extremal processes. *J. Appl. Probab.*, 8:745–756.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.*, 3:119–131.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York.

Resnick, S. I. (1971). Tail equivalence and its applications. *J. Appl. Prob.*, 8:135–156.

Serfling, R. J. (2002). *Approximation Theorems of Mathematical Statistics*. Wiley Series in Probability and Statistics, Wiley, New York.

Smith, R. L. (1987). Estimating tails of probability distributions. *Ann. Statist.*, 15:1174–1207.

Smith, R. L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, 4:367–393.

Tomassini, L. and Jacob, D. (2009). Spatial analysis of trends in extreme precipitation events in high-resolution climate model results and observations for Germany. *Journal of Geophysical Research*, 3:D1211.

Yee, T. W. and Stephenson, A. G. (2007). Vector generalized linear and additive extreme value models. *Extremes*, 10:1–19.

Zhou, C. (2009). Existence and consistency of the maximum likelihood estimator for the extreme value index. *J. Multivariate Anal.*, 100:794–815.

Zolina, O., Simmer, C., Belyaev, K., Kapala, A., and Gulev, S. (2009). Improving estimates of heavy and extreme precipitation using daily records from European rain gauges. *J. Hydrometeorol.*, 10:701–716.