SFB 823

On the method of probability weighted moments in regional frequency analysis

Jona Lilienthal, Paul Kinsvater,
Roland Fried

Discussion Paper

Nr. 63/2016

On the method of probability weighted moments in regional frequency analysis

Jona Lilienthal*, Paul Kinsvater and Roland Fried

Abstract

In regional flood frequency analysis it is of interest to estimate high quantiles of a local river flow distribution by gathering information from similar stations in the neighborhood. E. g., the popular Index Flood (IF) approach is based on an assumption termed regional homogeneity, which states that the quantile curves of those stations only differ by a site-specific factor, the so-called index flood, and it is assumed that the station's distribution is known up to some finite-dimensional parameter. In this context the method of probability weighted moments (or equivalently L-moments) is most popular for parameter estimation. While the observations often can be regarded as independent in time, a challenge arises from the fact that river flows from nearby stations are strongly dependent in space. To the best of our knowledge, none of the approaches from the literature based on the IF-model and on L-moments is able to take spatial dependence adequately into account. Our goal is to fill this gap. We present asymptotic theory that does not ignore inter-site dependence, which, for instance, allows to evaluate estimation uncertainty. As an application of this theory, a test procedure to check for regional homogeneity under index-flood assumptions is given and reviewed in a simulation study.

1 Introduction

Probability weighted moments have been introduced by Greenwood et al. (1979). Since then they have attracted a lot of attention in environmental science, for instance, in flood frequency analysis, where it is of interest to estimate high quantiles of river flow distributions.
Let F be a continuous distribution function on $\mathbb{R}$ with finite mean and let $X_1, \ldots, X_n$ denote a sample of i.i.d. observations from F. The k-th probability weighted moment (PWM) β_k of F, $k \in \mathbb{N}_0$, and its sample version $\hat{\beta}_{k,n}$ are defined by

$$\beta_k = \int_{\mathbb{R}} x \cdot F^k(x)\ dF(x) \quad \text{and} \quad \hat{\beta}_{k,n} = \int_{\mathbb{R}} x \cdot F_n^k(x)\ dF_n(x) = \frac{1}{n}\sum_{i=1}^{n} X_i \cdot F_n^k(X_i), \tag{1}$$

respectively, with F_n denoting the empirical distribution of the sample. Hosking (1990) proved that every distribution with finite first moment is uniquely determined by its sequence of probability weighted moments $(\beta_k)_{k\in\mathbb{N}_0}$. In case of a parametric family $F = F_{\boldsymbol{\vartheta}}$, $\boldsymbol{\vartheta} \in \Theta \subset \mathbb{R}^p$, some finite number of PWMs is enough in order to determine the parameter θ. As a typical example we consider the family of generalized extreme value (GEV) distribution functions

$$G_{\mu,\sigma,\xi}(x) = \exp\left(-\left[1 + \xi\frac{x-\mu}{\sigma}\right]^{-1/\xi}\right),\ 1 + \xi\frac{x-\mu}{\sigma} > 0,$$

*Corresponding author. Department of Statistics, TU Dortmund University, Vogelpothsweg 87, 44221 Dortmund, Germany. E-mail: lilenthal@statistik.tu-dortmund.de.

with parameters $\boldsymbol{\vartheta} = (\mu, \sigma, \xi)' \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}$ called location, scale and shape, respectively. If we assume that $\xi < 1$, we can apply the so-called method of PWMs: Hosking et al. (1985) showed that the parameter vector $\boldsymbol{\vartheta} = \phi(\boldsymbol{\beta})$ of the GEV is uniquely determined by the first three PWMs $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$ of $G_{\boldsymbol{\vartheta}}$, where ϕ is implicitly defined through an equation system. Even more, if $\xi < 1/2$ holds, Hosking et al. (1985) proved asymptotic normality of the canonical estimator $\hat{\theta}_n = \phi(\hat{\boldsymbol{\beta}}_n)$ computed from sample PWMs $\hat{\boldsymbol{\beta}}_n = \left(\hat{\beta}_{0,n}, \hat{\beta}_{1,n}, \hat{\beta}_{2,n}\right)'$.

In some applications, where we observe variables at many sites $j \in \{1, \ldots, d\}$ of a region with site-specific distributions F_j, it is of interest to combine information in order to estimate a target distribution, say, $F = F_1$. These pooling methods are based on certain assumptions called *regional homogeneity*. As an important example, the so-called Index Flood (IF) method (Dalrymple, 1960) considers the homogeneity hypothesis

$$\mathcal{H}_{0,IF}: \ F_j^{-1} = s_j \cdot G_{\boldsymbol{\vartheta}}^{-1} \ \text{ for all } \ j = 1, \ldots, d, \tag{2}$$

where $s_j = s(F_j)$ for some factor s (e.g. population mean or any location parameter) and where $G_{\boldsymbol{\vartheta}}$ is a given parametric distribution with $\boldsymbol{\vartheta}$ unknown (e.g. the GEV distribution).

Nowadays the most popular estimation method in regional flood frequency analysis considers assumption (2) and applies the method of PWMs for parameter estimation (Hosking and Wallis, 2005). However, satisfactory results proving asymptotic normality of such regional estimators based on PWMs and consistency of related tests of homogeneity (Hosking and Wallis, 2005, Chap. 4.3) have not been available so far. We are going to present a new limit theorem that allows us to fill these gaps. Our limit theorem enables us to estimate the variability of regional PWM estimators consistently, without relying on parametric dependence models or re-sampling schemes.

The remainder of this article is organized as follows. Section 2 presents a new central limit theorem for sample PWMs in a regional setting. As an immediate consequence, asymptotic theory for regional estimation by the method of TL-moments is provided in Section 3. We particularly focus on a new test of regional homogeneity and study its finite-sample properties by simulation in Section 4. All technical details are deferred to an appendix.

2 Limit theorem for sample PWMs

Let $\boldsymbol{X} = (X_1, \ldots, X_d)'$ be a d-dimensional random vector whose continuous marginal distribution functions are denoted by $F_j(x) = \mathbb{P}(X_j \leqslant x)$, $j = 1, \ldots, d$. In the applications we will consider river flow observations from d different measurement stations, where each margin F_j represents a station's local flow distribution. We stress out that we do not assume the components to be independent. Let $K \in \mathbb{N}$ be fixed. The first K PWMs of F_j are denoted by $\boldsymbol{\beta}_j = (\beta_{0,j}, \beta_{1,j}, \ldots, \beta_{K-1,j})'$, where

$$\beta_{k,j} = \int_{\mathbb{R}} x \cdot F_j^k(x) \, dF_j(x), \qquad k = 0, 1, \ldots, K-1 \ \text{ and } \ j = 1, \ldots, d.$$

All these local PWM vectors are summarized in $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_d') \in \mathbb{R}^{dK}$.

Suppose that $\boldsymbol{X}_i = (X_{i,1}, \ldots, X_{i,d})'$, $i = 1, \ldots, n$, denote independent copies of $\boldsymbol{X}$, where i is interpreted as a time index and with $\{1, \ldots, n\}$ covering the observation period. However, when considering observations from different river stations, it is unlikely that the observation period is (almost) the same for all d sites. Let $n = n_1 \geqslant n_2 \geqslant \ldots \geqslant n_d$ denote the local sample lengths, which are rearranged by length for ease of representation. A more appealing scenario is that we observe a scheme

$$\begin{array}{rrrrrrr} X_{1,1}, & X_{2,1}, & X_{3,1}, & X_{4,1}, & X_{5,1}, & \ldots, & X_{n,1}, \\ & X_{a_2+1,2}, & X_{a_2+2,2}, & X_{a_2+3,2}, & & \ldots, & X_{n,2}, \\ & & & \ddots & & \vdots & \\ & & & X_{a_d+1,d}, & X_{a_d+2,d}, & \ldots, & X_{n,d}, \end{array} \tag{3}$$

with $a_j = n - n_j$ and where each row contains only observations from the same station. It is important to account for the structure of the scheme in order to be able to capture properly the dependence between local estimates of probability weighted moments. For the asymptotic results we let $n \to \infty$ and we assume that $n_j/n \to r_j \in (0,1)$ in order to account for possibly very different local sample lengths, i.e., we set $n_j = \lfloor nr_j \rfloor$.
The sample version of $\beta_{k,j}$ computed from those observations is given by

$$\hat{\beta}_{k,j} = \hat{\beta}_{k,j,r_j,n} = \int_{\mathbb{R}} x \cdot F^k_{j,a_j+1:n}(x)\, dF_{j,a_j+1:n}(x) = \frac{1}{n_j}\sum_{i=1}^{n_j} X_{a_j+i} \cdot F^k_{j,a_j+1:n}(X_{a_j+i}),$$

where $F_{j,\ell:m}$ is the empirical distribution function of $X_{j,\ell}, X_{j,\ell+1}, \ldots, X_{j,m}$. Sample counterparts of $\boldsymbol{\beta}_j \in \mathbb{R}^K$ and $\boldsymbol{\beta} \in \mathbb{R}^{d\cdot K}$ are denoted by

$$\hat{\boldsymbol{\beta}}_{j,r_j,n} = \left(\hat{\beta}_{0,j}, \ldots, \hat{\beta}_{K-1,j}\right)' \quad \text{and} \quad \hat{\boldsymbol{\beta}}_{\boldsymbol{r},n} = \left(\hat{\boldsymbol{\beta}}'_{1,r_1,n}, \ldots, \hat{\boldsymbol{\beta}}'_{d,r_d,n}\right)', \tag{4}$$

respectively, where $\boldsymbol{r} = (r_1, \ldots, r_d)$ highlights the dependency on scheme (3).

Theorem 1. *Suppose that $\boldsymbol{X}_i$, $i \geqslant 1$, is a sequence of independent copies of $\boldsymbol{X} = (X_1, \ldots, X_d)'$, whose PWMs are summarized in the vector $\boldsymbol{\beta} \in \mathbb{R}^{d\cdot K}$ and with*

$$\mathbb{E}\left[X_j F_j^k(X_j) X_\ell F_\ell^m(X_\ell)\right] < \infty \quad \textit{for all} \quad 1 \leqslant j, \ell \leqslant d \quad \textit{and} \quad 0 \leqslant k, m < K.$$

Suppose further that $\sup_{x\in\mathbb{R}} |x\{F_j(x)(1 - F_j(x))\}^w| < \infty$ for all $j = 1, \ldots, d$ and some $w \in [0, 1/2)$. Then, for fixed $\boldsymbol{r} \in (0,1)^d$ and $n \to \infty$, we have that

$$\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{\boldsymbol{r},n} - \boldsymbol{\beta}\right) \xrightarrow{D} \mathcal{N}\left(0,\ \Sigma_{\boldsymbol{r}}\right),$$

where the limiting variance matrix $\Sigma_{\boldsymbol{r}} \in \mathbb{R}^{dK\times dK}$ is provided in Appendix A.

Theorem 1 and a consistent estimator $\hat{\Sigma}_{\boldsymbol{r},n}$ of $\Sigma_{\boldsymbol{r}}$ (see Appendix A) allow us to develop asymptotically consistent methods for regional frequency analysis, which is summarized in the next two sections.

3 Limit theorems for sample TL-moments and estimation of GEV parameters

L-moments $\lambda_k = \lambda_k(F)$, $k \in \mathbb{N}$, as defined by Hosking (1990) turn out to be useful summary statistics of heavy-tailed distributions F, since their existence requires only a finite first moment and because they are interpretable analogously to summary statistics based on classical product moments $\mu_k = \int x^k\, dF(x)$, for instance, with λ_1, λ_2, and $\tau_3 = \lambda_3/\lambda_2$ representing location, dispersion, and skewness of F, respectively. More generally, practitioners from hydrology nowadays consider so-called Trimmed L-moments (TL-moments)

$$\lambda_k^{(s,t)} = \frac{1}{k}\sum_{i=0}^{k-1}(-1)^i\binom{k-1}{i}\mathbb{E}\left(X_{k+s-i:k+s+t}\right)$$

of F, with $s, t \in \mathbb{N}_0$ interpreted as left- and right-trimming parameters, respectively, and $\lambda_k^{(0,0)} = \lambda_k$. $X_{1:n} \leqslant \ldots \leqslant X_{n:n}$ denote order statistics of a random sample of size n drawn from F . It is known that every TL-moment can be represented as a linear combination of a finite number of PWMs, provided F has finite mean. This fact, by referring to Theorem 1, allows us toderive central limit theorems for sample TL-moments and related methods easily.

3.1 At-site statistics

Throughout this paper we will assume that F has finite mean. The TL-moment of F of order $k \in \mathbb{N}$ with trimming $s, t \in \mathbb{N}_0$ is known to satisfy

$$\lambda_k^{(s,t)} = \sum_{i=0}^{k+s+t-1} z_{k-1,i}^{(s,t)} \, \beta_i = (\boldsymbol{z}_{k-1}^{(s,t)})' \boldsymbol{\beta},$$

with $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_{k+s+t-1})'$ being the vector of the first $k+s+t$ PWMs of F and $\boldsymbol{z}_{k-1}^{(s,t)} = (z_{k-1,0}^{(s,t)}, \ldots, z_{k-1,k+s+t-1}^{(s,t)})'$ being a coefficient vector with components

$$z_{k,i}^{(s,t)} = \frac{k!(k+s+t+1)!}{(k+1)(k+s)!(k+t)!}(-1)^{s+k+i}\binom{k+t}{i+s}\binom{k+i}{k}.$$

Let $Z^{(s,t)} = (\boldsymbol{z}_0^{(s,t)}, \ldots, \boldsymbol{z}_{m-1}^{(s,t)})'$ denote the linear mapping such that $\boldsymbol{\lambda}^{(s,t)} = Z^{(s,t)}\boldsymbol{\beta}$ with $\boldsymbol{\lambda}^{(s,t)} = (\lambda_1^{(s,t)}, \ldots, \lambda_m^{(s,t)})'$. For ease of notation we suppress the sample length n in the notation of the estimators, i.e., $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_n$. The first m sample TL(s,t)-moments $\hat{\boldsymbol{\lambda}}^{(s,t)} = (\hat{\lambda}_1^{(s,t)}, \ldots, \hat{\lambda}_m^{(s,t)})'$ and the corresponding covariance matrix are given by

$$\hat{\boldsymbol{\lambda}}^{(s,t)} = Z^{(s,t)}\hat{\boldsymbol{\beta}}, \quad \mathrm{Var}(\hat{\boldsymbol{\lambda}}^{(s,t)}) = Z^{(s,t)}\mathrm{Var}(\hat{\boldsymbol{\beta}})(Z^{(s,t)})'.$$

Recall that $n\mathrm{Var}(\hat{\boldsymbol{\beta}}) \xrightarrow{\mathbb{P}} \Sigma$ for $n \to \infty$ and some matrix Σ. From Theorem 1 and the delta method we obtain for $n \to \infty$

$$\sqrt{n}\left(\hat{\boldsymbol{\lambda}}^{(s,t)} - \boldsymbol{\lambda}^{(s,t)}\right) \xrightarrow{D} \mathcal{N}\left(0,\ Z^{(s,t)}\Sigma(Z^{(s,t)})'\right).$$

So far we have introduced TL-moments as summary statistics of distributions without restricting to any parametric family. In practice, however, one usually assumes that $F = F_{\boldsymbol{\vartheta}}$ for some unknown parameter vector $\boldsymbol{\vartheta} \in \Theta \subset \mathbb{R}^p$. Relationships between TL-moments and the distribution parameters are employed, which allows us to estimate these parameters by plugging in sample TL-moments into the formulas.

More specifically, let $g^{(s,t)} : \mathbb{R}^m \mapsto \mathbb{R}^p$ be a differentiable function that maps the first m TL(s,t)-moments of $F_{\boldsymbol{\vartheta}}$ onto its parameter vector $\boldsymbol{\vartheta}$. From the delta method, for $\hat{\boldsymbol{\vartheta}} = g^{(s,t)}(\hat{\boldsymbol{\lambda}}^{(s,t)})$ and $n \to \infty$, we immediately obtain that

$$\sqrt{n}\left(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}\right) \xrightarrow{D} \mathcal{N}\left(0,\ A_{\boldsymbol{\lambda}^{(s,t)}}^{(s,t)} Z^{(s,t)}\Sigma(Z^{(s,t)})'(A_{\boldsymbol{\lambda}^{(s,t)}}^{(s,t)})'\right), \tag{5}$$

where $A_{\boldsymbol{\lambda}}^{(s,t)} = \frac{\partial}{\partial \boldsymbol{\lambda}} g^{(s,t)}(\boldsymbol{\lambda}) \in \mathbb{R}^{p \times m}$ denotes the Jacobi matrix of $g^{(s,t)}$ evaluated at $\boldsymbol{\lambda} \in \mathbb{R}^m$. Relationships between GEV parameters and TL(0,0)-moments (resp. TL(0,1)-moments) with corresponding matrices $A_{\boldsymbol{\lambda}^{(s,t)}}^{(s,t)}$ are summarized in Appendix C.

In flood frequency analysis we are usually not interested in the estimation of parameters but in quantiles $\hat{\boldsymbol{q}} = (q_1, \ldots, q_k)'$. Suppose that $h(\boldsymbol{q}) = (F_{\boldsymbol{\vartheta}}^{-1}(q_1), \ldots, F_{\boldsymbol{\vartheta}}^{-1}(q_k))'$ is differentiable in

$\boldsymbol{\vartheta}$ and let $B_{\boldsymbol{\vartheta}} \in \mathbb{R}^{pk \times p}$ denote the corresponding Jacobi matrix. Again, from the delta method and for $n \to \infty$, we obtain

$$\sqrt{n}\,(\hat{\boldsymbol{q}} - \boldsymbol{q}) \xrightarrow{D} \mathcal{N}\left(0,\ B_{\boldsymbol{\vartheta}} A^{(s,t)}_{\boldsymbol{\lambda}^{(s,t)}} Z^{(s,t)} \Sigma (Z^{(s,t)})' (A^{(s,t)}_{\boldsymbol{\lambda}^{(s,t)}})' B'_{\boldsymbol{\vartheta}}\right). \tag{6}$$

Considering again the GEV(μ, σ, ξ) family with quantile function $h(q_i) = F^{-1}_{\boldsymbol{\vartheta}}(q_i) = \mu - \frac{\sigma}{\xi}(1 - (-\log(q_i))^{-\xi})$, the matrix $B_\vartheta = (B^{(q_1)}_\vartheta, \ldots, B^{(q_k)}_\vartheta)'$ is given row-wise by

$$B^{(q_i)}_{\mu,\sigma,\xi} = \begin{pmatrix} 1 \\ \frac{(-\log(q_i))^{-\xi} - 1}{\xi} \\ \frac{\sigma(\xi^{-1} - (-\log(q_i))^{-\xi}(\log(-\log(q_i)) + \xi^{-1}))}{\xi} \end{pmatrix}.$$

3.2 Joint estimation at multiple stations

We switch to a regional scale by considering multivariate observations as given in scheme (3). Recall that $\hat{\boldsymbol{\beta}}_{\boldsymbol{r}} = \hat{\boldsymbol{\beta}}_{\boldsymbol{r},n}$ from (4) contains sample PWMs of all d marginal distributions F_j involved in scheme (3). In analogy to (4), the vector of all sample TL(s,t)-moments is denoted by

$$\hat{\boldsymbol{\lambda}}^{(s,t)}_{\boldsymbol{r}} = \hat{\boldsymbol{\lambda}}^{(s,t)}_{\boldsymbol{r},n} = \left((\hat{\boldsymbol{\lambda}}^{(s,t)}_{1,r_1,n})', \ldots, (\hat{\boldsymbol{\lambda}}^{(s,t)}_{d,r_d,n})'\right)'$$

with population counterpart $\boldsymbol{\lambda}^{(s,t)} = \left((\boldsymbol{\lambda}^{(s,t)}_1)', \ldots, (\boldsymbol{\lambda}^{(s,t)}_d)'\right)' \in \mathbb{R}^{md}$. By Theorem 1, the delta method and for $n \to \infty$ we obtain that

$$\sqrt{n}\left(\hat{\boldsymbol{\lambda}}^{(s,t)}_{\boldsymbol{r}} - \boldsymbol{\lambda}^{(s,t)}\right) \xrightarrow{D} \mathcal{N}\left(0,\ \tilde{Z}^{(s,t)} \Sigma_{\boldsymbol{r}} (\tilde{Z}^{(s,t)})'\right),$$

with $\Sigma_{\boldsymbol{r}}$ being defined in Appendix A and with block-diagonal matrix

$$\tilde{Z}^{(s,t)} = \operatorname{diag}(Z^{(s,t)}, Z^{(s,t)}, \ldots, Z^{(s,t)}) = \begin{pmatrix} Z^{(s,t)} & 0 & \ldots & 0 \\ 0 & Z^{(s,t)} & & \vdots \\ \vdots & & \ddots & \\ 0 & \ldots & & Z^{(s,t)} \end{pmatrix}.$$

Similarly, under the assumption that $F_j = F_{\boldsymbol{\vartheta}_j}$ for $j = 1, \ldots, d$, with block-diagonal matrices $\tilde{A}^{(s,t)} = \operatorname{diag}(A^{(s,t)}_{\boldsymbol{\lambda}_1}, \ldots, A^{(s,t)}_{\boldsymbol{\lambda}_d})$ and $\tilde{B} = \operatorname{diag}(B_{\boldsymbol{\vartheta}_1}, \ldots, B_{\boldsymbol{\vartheta}_d})$ taken into account, one can easily obtain the joint limiting distribution of parameter and quantile estimators for all d stations.

4 Test of regional homogeneity

When considering observations from multiple stations, e.g. scheme (3), in flood frequency analysis mostly the Index Flood assumption $\mathcal{H}_{0,IF}$ stated in (2) is applied in order to decrease the estimation variability. However, while a moderate amount of heterogeneity of the group may still lead to an overall improvement compared to local estimation (Lettenmaier et al., 1987), strong heterogeneity typically leads to a severe bias, which again increases the overall estimation error. It is thus important to be able to identify serious sources of heterogeneity. We are going to introduce a statistical test that proves to be advantageous in several aspects to competitive procedures from the literature.

4.1 Test statistic

Suppose that we have observed scheme (3) with site-specific distribution functions $F_j = F_{\boldsymbol{\vartheta}_j}$ and that $F_{\boldsymbol{\vartheta}_j} = G_{\mu_j,\sigma_j,\xi_j}$ is the GEV distribution function with parameters $\boldsymbol{\vartheta}_j = (\mu_j, \sigma_j, \xi_j)'$. In this case hypothesis (2) is equivalent to

$$\delta_1 = \ldots = \delta_d \text{ with } \delta_i = \frac{\sigma_i}{\mu_i} \quad \text{and} \quad \xi_1 = \ldots = \xi_d. \tag{7}$$

Let $\hat{\boldsymbol{\vartheta}}_{\boldsymbol{r}} = \hat{\boldsymbol{\vartheta}}_{\boldsymbol{r},n} = (\hat{\mu}_1, \hat{\sigma}_1, \hat{\xi}_1, \ldots, \hat{\mu}_d, \hat{\sigma}_d, \hat{\xi}_d)'$ denote an estimator of local parameters obtained from scheme (3). We apply the TL(s,t)-moment estimator $\hat{\boldsymbol{\vartheta}}_{\boldsymbol{r}}$ of $\boldsymbol{\vartheta}$ from Section 3. Let g denote the map $\boldsymbol{\vartheta} \mapsto (\delta_1, \xi_1, \ldots, \delta_d, \xi_d)'$ with corresponding Jacobi-matrix $C = \frac{\partial}{\partial \boldsymbol{\vartheta}} g(\boldsymbol{\vartheta})$. Again, from the delta method, we immediately obtain that

$$\sqrt{n}\left(g(\hat{\boldsymbol{\vartheta}}_{\boldsymbol{r}}) - g(\boldsymbol{\vartheta})\right) \xrightarrow{D} \mathcal{N}(0,\ \Gamma_{\boldsymbol{r}}) \quad \text{with} \quad \Gamma_{\boldsymbol{r}} = \ C\tilde{A}^{(s,t)}\tilde{Z}^{(s,t)}\Sigma_{\boldsymbol{r}}(C\tilde{A}^{(s,t)}\tilde{Z}^{(s,t)})'$$

as $n \to \infty$. In order to evaluate hypothesis $\mathcal{H}_{0,IF}$, which is equivalent to $R \cdot g(\boldsymbol{\vartheta}) = 0$ with

$$R = \begin{pmatrix} 1 & 0 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & & & \ddots & & & & \vdots \\ 0 & \cdots & & & & 1 & 0 & -1 \end{pmatrix},$$

we propose a Wald-type test statistic

$$T_n = n\left(R\, g(\hat{\boldsymbol{\vartheta}}_{\boldsymbol{r}})\right)' (R\, \hat{\Gamma}_{\boldsymbol{r}}\, R')^{-1}(R\, g(\hat{\boldsymbol{\vartheta}}_{\boldsymbol{r}})).$$

Under $\mathcal{H}_{0,IF}$, for $n \longrightarrow \infty$ and under the assumptions of Theorem 1, we have that $T_n \xrightarrow{D} \chi^2_{2(d-1)}$, while under fixed alternatives we have $T_n \xrightarrow{\mathbb{P}} \infty$.

4.2 Simulation study

To check the capability of the proposed homogeneity test a small simulation study is conducted at a nominal level of $\alpha = 5\%$. The data is generated from $d = 6$ dependent stations with different local sample lengths (we refer to scheme (3)), with margins $F_j = GEV(\mu_j, \sigma_j, \xi_j)$ and, for simplicity, with Gumbel-Hougaard copula C_β and dependence parameter $\beta = 1.5$.
Table 1 summarizes the particular choice of GEV parameters and local sample lengths. Note that assumption $\mathcal{H}_{0,IF}$ from (2) is satisfied only if $\tilde{\sigma} = 30$ and $\tilde{\xi} = 0.3$. We conducted 10 000 independent replications of the experiment for each scenario on the grid $(\tilde{\sigma},\ \tilde{\xi}) \in \{24.0, 25.5, 27.0, \ldots, 36.0\} \times \{0.10, 0.15, 0.20, \ldots, 0.50\}$. These values are consistent with our experience from real data applications. Corresponding rejection rates of the Wald type test statistic with $\Sigma_{\boldsymbol{r}}$ estimated by the check-version $\check{\Sigma}_{\boldsymbol{r},n}$ from Appendix A.2 are summarized in Figure 1.
The left panel of Figure 1 depicts the test's rejection rate for a maximal sample length of $n = 100$. The type-I-error of the test, i.e. when $\tilde{\sigma} = 30$ and $\tilde{\xi} = 0.3$, is 7.39% with a standard error of roughly 0.26%. We observe that our proposed method captures deviations from $\mathcal{H}_{0,IF}$ in all possible directions and therefore, our test seems to be a suitable procedure for testing the Index Flood assumption $\mathcal{H}_{0,IF}$ from 7.
Lastly, a closer look at the type-I-error rate is taken in Figure 1, right panel. There the rejection rate under the null is depicted as a function of n. The plot indicates that the empirical level approaches the nominal level of 5% with increasing sample length n.

station	μ	σ	ξ	length
1	10	5	0.3	$1.0\,n$
2	20	10	0.3	$0.85\,n$
3	30	15	0.3	$0.70\,n$
4	40	20	0.3	$0.70\,n$
5	50	25	0.3	$0.85\,n$
6	60	$\tilde{\sigma}$	$\tilde{\xi}$	$1.0\,n$

Table 1: Parameters of the marginal distributions used in the simulation study. The $\tilde{\sigma}$ and $\tilde{\xi}$ parameters are allowed to vary in order to simulate different grades of deviation.

Acknowledgements

We are grateful to Professor Andreas Schumann from the Department of Civil Engineering, Ruhr-University Bochum, Germany, for providing us hydrological data and for helpful discussions. The financial support of the Deutsche Forschungsgemeinschaft (SFB 823, "Statistical modelling of nonlinear dynamic processes") is gratefully acknowledged.

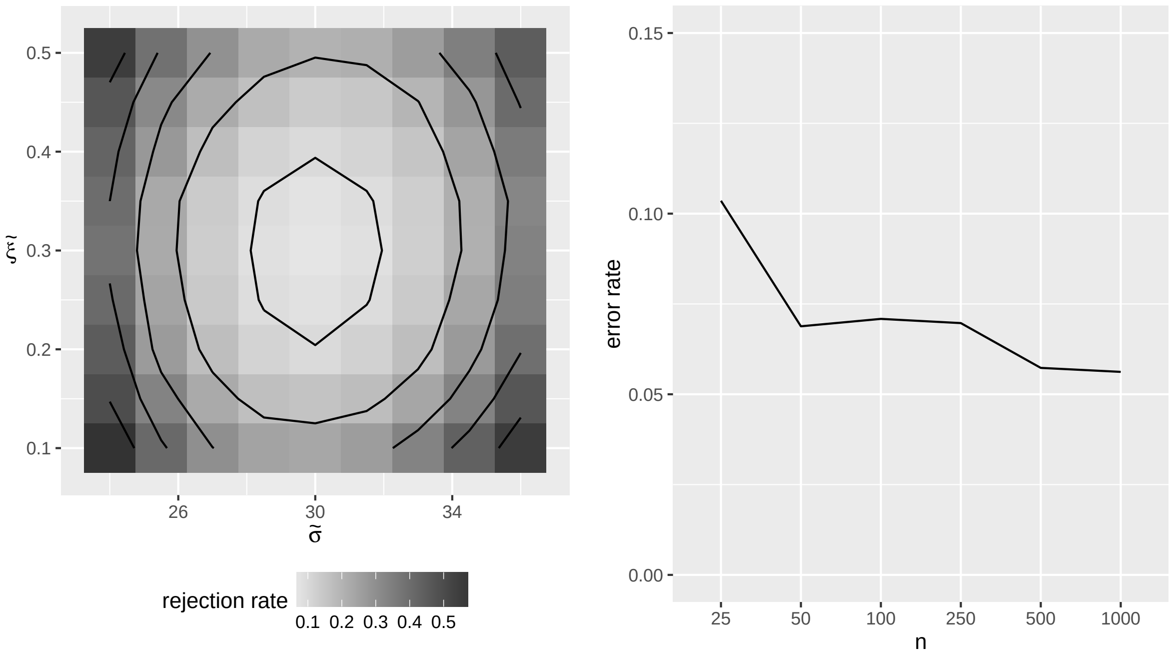

Figure 1: Left: Rejection rates of the proposed test for $n = 100$. The position on the grid defines the values of the varying parameters δ and ξ. The contour lines are printed in 0.1-steps. Right: Error rates of the proposed procedure under the null as a function of the observation length n.

A Estimation of the variance matrix $\Sigma_{\mathbf{r}}$

The limiting variance matrix $\Sigma_{\boldsymbol{r}} = \lim_{n\to\infty} \operatorname{Var}\left(\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{\boldsymbol{r},n} - \boldsymbol{\beta}\right)\right)$ from Theorem 1 is defined block-wise by

$$\lim_{n\to\infty} \operatorname{Cov}\left(\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{j,r_j,n} - \boldsymbol{\beta}_j\right),\ \sqrt{n}\left(\hat{\boldsymbol{\beta}}_{\ell,r_\ell,n} - \boldsymbol{\beta}_\ell\right)\right) = \frac{\min(r_j, r_\ell)}{r_j \cdot r_\ell} \cdot \operatorname{Cov}(\boldsymbol{Z}_j,\ \boldsymbol{Z}_\ell) \in \mathbb{R}^{K\times K}$$

and where $\boldsymbol{Z}_j = (Z_{0,j}, Z_{1,j}, \ldots, Z_{K-1}, j)'$, $j = 1, \ldots, d$, are random vectors defined through

$$Z_{k,j} = X_j \cdot F_j^k(X_j) + \int_{\mathbb{R}} x \cdot k \cdot F_j^{k-1}(x) \cdot \mathbb{1}(X_j \leqslant x)\, dF_j(x). \tag{8}$$

In words, empirical probability weighted moments are asymptotically jointly normal with limiting variance matrix obtained from that of the variables defined in (8).

A.1 Empirical estimator of $\Sigma_{\mathbf{r}}$

Suppose that we have collected an observation scheme given in (3). In practice the variance matrices $\operatorname{Cov}(\boldsymbol{Z}_j, \boldsymbol{Z}_\ell)$ can be consistently estimated by their sample analogues: Let

$$\hat{Z}_{i,k,j} = X_{i,j} \cdot F^k_{j,a_j+1:n}(X_{i,j}) + \frac{1}{n_j}\sum_{\ell=1}^{n_j} X_{\ell,j} \cdot k \cdot F^{k-1}_{j,a_j+1:n}(X_{\ell,j}) \cdot \mathbb{1}(X_{i,j} \leqslant X_{\ell,j}) \tag{9}$$

and $\hat{\boldsymbol{Z}}_{i,j} = (Z_{i,0,j}, Z_{i,1,j}, \ldots, Z_{i,K-1,j})'$, $i = a_j + 1, \ldots, n$. For $1 \leqslant j, \ell, \leqslant d$, the covariance matrix $\operatorname{Cov}(\boldsymbol{Z}_j, \boldsymbol{Z}_\ell)$ is estimated by the empirical covariance matrix of the sample

$$\left\{\left(\hat{\boldsymbol{Z}}_{\max(a_j,a_\ell)+1,j}, \hat{\boldsymbol{Z}}_{\max(a_j,a_\ell)+1,\ell}\right), \ldots, \left(\hat{\boldsymbol{Z}}_{n,j}, \hat{\boldsymbol{Z}}_{n,\ell}\right)\right\}.$$

The resulting estimator of $\Sigma_{\boldsymbol{r}}$ is denoted by $\hat{\Sigma}_{\boldsymbol{r},n}$ and is called empirical estimator.

Corollary 2. *Under the assumptions of Theorem 1 and for $n \to \infty$ we have that $\hat{\Sigma}_{\boldsymbol{r},n} \xrightarrow{\mathbb{P}} \Sigma_{\boldsymbol{r}}$.*

A.2 A parametric modification on the block diagonal

In typical applications we will assume that the margins $F_j = G_{\boldsymbol{\vartheta}_j}$ are known up to some finite dimensional parameters $\boldsymbol{\vartheta}_j$. For instance, considering the GEV family $G_{\boldsymbol{\vartheta}}$, Hosking et al. (1985) derived a parametric expression for the local covariance matrices $\operatorname{Cov}(\mathbf{Z}_j, \mathbf{Z}_j) = \operatorname{Var}(\mathbf{Z}_j) = \Sigma(\boldsymbol{\vartheta}_j)$, $j = 1, \ldots, d$, involved in $\Sigma_{\boldsymbol{r}}$. We thus may wish to replace the local part of $\hat{\Sigma}_{\boldsymbol{r}}$ by parametric estimates $\hat{\operatorname{Var}}(\mathbf{Z}_j) = \Sigma(\hat{\boldsymbol{\vartheta}}_j)$, where $\hat{\boldsymbol{\vartheta}}_j$ are consistent estimates of $\boldsymbol{\vartheta}_j$, e.g., TL-moment estimators of GEV parameters. The modified estimator of $\Sigma_{\boldsymbol{r}}$ is denoted by $\check{\Sigma}_{\boldsymbol{r},n}$.
Unsurprisingly, the check-version, which is also a consistent estimator of $\Sigma_{\boldsymbol{r}}$, is way more efficient than the empirical estimator, especially when the sample length $n \leqslant 100$ is small. However, $\check{\Sigma}_{\boldsymbol{r},n}$ is not necessarily a valid covariance matrix, contrary to $\hat{\Sigma}_{\boldsymbol{r},n}$. The mixture of non-parametric and parametric parts involved in the check-version produces negative eigenvalues in some cases. In the simulation study reported in Section 4.2 we observed negative eigenvalues in about 1% of the repetitions.

B Proofs of Theorem 1 and Corollary 2

For sake of readability the proofs are given for $d = 2$. The derivation for arbitrary dimensions $d \geqslant 2$ can be established at the cost of a more complex notation but without additional technical difficulties. Even more, we assume the same beginnings and different end points, that is, we compute the statistics purely from the variables $X_1, \ldots, X_{\lfloor nr_1 \rfloor}$ and $Y_1, \ldots, Y_{\lfloor nr_2 \rfloor}$, where (X_i, Y_i), $i \geqslant 1$, is a sequence of independent and identically distributed bivariate vectors with margins $F(x) = \mathbb{P}(X_i \leqslant x)$ and $G(y) = \mathbb{P}(Y \leqslant y)$, respectively. The corresponding first K probability weighted moments of F and G are denoted by $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_{K-1})'$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{K-1})'$, respectively, and we let $\boldsymbol{\gamma} = (\boldsymbol{\alpha}', \boldsymbol{\beta}')' \in \mathbb{R}^{2K}$. We set

$$\hat{\alpha}_{k,r_1,n} = \frac{1}{\lfloor nr_1 \rfloor} \sum_{i=1}^{\lfloor nr_1 \rfloor} X_i \cdot F_{\lfloor nr_1 \rfloor}^k(X_i) \quad \text{and} \quad \hat{\beta}_{k,r_2,n} = \frac{1}{\lfloor nr_2 \rfloor} \sum_{i=1}^{\lfloor nr_2 \rfloor} Y_i \cdot G_{\lfloor nr_2 \rfloor}^k(Y_i)$$

with F_{n_1} (resp. G_{n_2}) denoting the empirical distribution function of the sample $X_1, \ldots, X_{n_1}$ (resp. $Y_1, \ldots, Y_{n_2}$). All these components are collected in $\hat{\boldsymbol{\alpha}}_{r_1,n}$, $\hat{\boldsymbol{\beta}}_{r_2,n}$ and $\hat{\boldsymbol{\gamma}}_{\boldsymbol{r},n} = (\hat{\boldsymbol{\alpha}}'_{r_1,n}, \hat{\boldsymbol{\beta}}'_{r_2,n})'$.

Proof of Theorem 1.

Let $\tilde{\boldsymbol{\alpha}}_{r_1,n}$, $\tilde{\boldsymbol{\beta}}_{r_2,n}$ and $\tilde{\boldsymbol{\gamma}}_{\boldsymbol{r},n} = (\tilde{\boldsymbol{\alpha}}'_{r_1,n}, \tilde{\boldsymbol{\beta}}'_{r_2,n})'$ be defined analogously to the hat-versions but with $F_{\lfloor nr_1 \rfloor}$ and $G_{\lfloor nr_2 \rfloor}$ replaced by their true counterparts F and G, respectively. We write

$$\sqrt{n}\,(\hat{\boldsymbol{\gamma}}_{\boldsymbol{r},n} - \boldsymbol{\gamma}) = \boldsymbol{Q}_{\boldsymbol{r},n} + \boldsymbol{\Delta}_{\boldsymbol{r},n}, \tag{10}$$

where $\boldsymbol{Q}_{\boldsymbol{r},n} = \sqrt{n}\,(\tilde{\boldsymbol{\gamma}}_{\boldsymbol{r},n} - \boldsymbol{\gamma})$ and $\boldsymbol{\Delta}_{\boldsymbol{r},n} = \sqrt{n}\,(\hat{\boldsymbol{\gamma}}_{\boldsymbol{r},n} - \tilde{\boldsymbol{\gamma}}_{\boldsymbol{r},n})$. The remainder of the proof is organized in the following three steps:

a) Verify that $\boldsymbol{Q}_{\boldsymbol{r},n} \xrightarrow{D} \boldsymbol{Q}_{\boldsymbol{r}}$, where the limit is a zero mean normally distributed random vector and show that the convergence holds jointly with that of the weighted empirical processes $\mathbb{U}_{r_1,n}$ and $\mathbb{V}_{r_2,n}$ defined below.

b) Show that $\boldsymbol{\Delta}_{\boldsymbol{r},n} = \boldsymbol{R}_{\boldsymbol{r},n} + o_{\mathbb{P}}(1)$ for $n \to \infty$, where all components of $\boldsymbol{R}_{\boldsymbol{r},n}$ can be represented as continuous functionals of either $\mathbb{U}_{r_1,n}$ or $\mathbb{V}_{r_2,n}$. Verify that $\boldsymbol{R}_{\boldsymbol{r},n}$ converges weakly towards a zero mean normally distributed random vector $\boldsymbol{R}_{\boldsymbol{r}}$.

c) Conclude that (10) is asymptotically normal with mean zero and compute the limiting variance matrix $\Sigma_{\boldsymbol{r}} = \operatorname{Var}(\boldsymbol{Q}_{\boldsymbol{r}} + \boldsymbol{R}_{\boldsymbol{r}})$.

Step a) Let $\mathbb{U}_{r_1,n}$ and $\mathbb{V}_{r_2,n}$ be $\ell^\infty[0,1]$-valued processes defined by

$$\mathbb{U}_{r_1,n}(u) = \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nr_1 \rfloor} \{\mathbb{1}(F(X_i) \leqslant u) - u\}}{\{u(1-u)\}^w} \quad \text{and} \quad \mathbb{V}_{r_2,n}(v) = \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nr_2 \rfloor} \{\mathbb{1}(G(Y_i) \leqslant v) - v\}}{\{v(1-v)\}^w}$$

for $u, v \in [0,1]$. These are called weighted empirical processes and their weak convergence is studied, e.g., in Genest and Segers (2009, Appendix G) and Kojadinovic and Naveau (2015, Appendix B) in a more general context. The weighting is needed for step b) of the proof in order to be able to express the components of $\boldsymbol{R}_{\boldsymbol{r},n}$ as continuous functionals of the empirical processes. Without loss of generality let $r_1 \leqslant r_2$ and note that

$$\mathbb{W}_{\boldsymbol{r},n} = (\mathbb{U}_{r_1,n}, \mathbb{V}_{r_2,n}) = (\mathbb{U}_{r_1,n}, \mathbb{V}_{r_1,n}) + (0, \mathbb{V}_{r_2,n} - \mathbb{V}_{r_1,n})$$

is a sum of two independent processes with $\mathbb{V}_{r_2,n}-\mathbb{V}_{r_1,n} \stackrel{D}{=} \mathbb{V}_{r_2-r_1,n}$. By the continuous mapping theorem and by Genest and Segers (2009, Th. G.1), both summands on the right-hand side of the previous equation converge weakly in $(\ell^\infty[0,1])^2$ towards centered Gaussian processes and, by independence of the summands, also does $\mathbb{W}_{\boldsymbol{r},n}$. Let $\mathbb{W}_{\boldsymbol{r}}$ denote the limiting process.
In almost the same manner we can write $\sqrt{n}\,(\tilde{\boldsymbol{\gamma}}_{\boldsymbol{r},n}-\boldsymbol{\gamma})$ as a sum of two independent random vectors, where weak convergence of both summands towards centered normal distributions easily follows from the central limit theorem for sums of i.i.d. random vectors. The limit is denoted by $\sqrt{n}\,(\tilde{\boldsymbol{\gamma}}_{\boldsymbol{r},n}-\boldsymbol{\gamma}) \stackrel{D}{\to} \boldsymbol{Q}_{\boldsymbol{r}}$. In fact, weak convergence of $\mathbb{W}_{\boldsymbol{r},n}$ and that of $\sqrt{n}\,(\tilde{\boldsymbol{\gamma}}_{\boldsymbol{r},n}-\boldsymbol{\gamma})$ holds jointly as a random element in $(\ell^\infty[0,1]^2)\times\mathbb{R}^{2K}$. The only thing left to verify is that the finite dimensional convergence holds, which again follows from the central limit theorem for sums of i.i.d. random vectors.

Step b) Let $\boldsymbol{R}_{\boldsymbol{r},n}=(\boldsymbol{S}'_{r_1,n},\boldsymbol{T}'_{r_2,n})'$ with $\boldsymbol{S}_{r_1,n}=(S_{0,r_1,n},\dots,S_{K-1,r_1,n})'$,

$$S_{k,r_1,n}=\frac{1}{r_1}\int_{\mathbb{R}} x\cdot k\cdot F^{k-1}(x)\cdot\mathbb{U}_{r_1,n}(F(x))\cdot\{F(x)(1-F(x))\}^w\,dF(x)$$

and analogously define $\boldsymbol{T}_{r_2,n}$ but with $(r_1,F,\mathbb{U})$ replaced by $(r_2,G,\mathbb{V})$. In order to show that $\boldsymbol{\Delta}_{\boldsymbol{r},n}=\boldsymbol{R}_{\boldsymbol{r},n}+o_{\mathbb{P}}(1)$ for $n\to\infty$, it suffices to consider each component separately by proving

$$\sqrt{n}\,(\hat{\alpha}_{k,r_1,n}-\tilde{\alpha}_{k,r_1,n})=S_{k,r_1,n}+o_{\mathbb{P}}(1)$$

for each $k=0,\dots,K-1$ and analogously for the β-components. But this follows from (C.9) in the proof of Proposition C.2 in Kojadinovic and Naveau (2015).
Let $\varphi_k:\ell^\infty[0,1]\to\mathbb{R}$, $k=0,\dots,K-1$, be defined by

$$\varphi_k(g)=\int_{\mathbb{R}} x\cdot k\cdot F^{k-1}(x)\{F(x)(1-F(x))\}^w\cdot g(F(x))\,dF(x)$$

and note that $S_{k,r_1,n}=\varphi_k(\mathbb{U}_{r_1,n})$. Since $\sup_{x\in\mathbb{R}}\left|x\cdot k\cdot F^{k-1}(x)\{F(x)(1-F(x))\}^w\right|<\infty$ by assumption, it follows that φ_k is a continuous map. Similarly we can define continuous maps ψ_k, $k=0,\dots,K-1$, such that $T_{k,r_2,n}=\psi_k(\mathbb{V}_{r_2,n})$. Bringing things together we conclude that $\boldsymbol{R}_{\boldsymbol{r},n}=\Psi(\mathbb{W}_{\boldsymbol{r},n})\stackrel{D}{\to}\Psi(\mathbb{W}_{\boldsymbol{r},n})=\boldsymbol{R}_{\boldsymbol{r}}$, where $\Psi:(\ell^\infty[0,1])^2\to\mathbb{R}^{2K}$ with

$$\Psi(f,g)=(\varphi_0(f),\dots,\varphi_{K-1}(f),\psi_0(g),\dots,\psi_{K-1}(g))'$$

is continuous. Since each component of $\boldsymbol{R}_{\boldsymbol{r},n}$ is a sum of i.i.d. zero-mean random variables with existing second moments, we conclude that the limit is a zero-mean normal distribution.

Step c) From steps a) and b) we obviously obtain the joint asymptotic normality of $\mathbf{Q}_{\boldsymbol{r},n}$ and $\mathbf{R}_{\boldsymbol{r},n}$. By the continuous mapping theorem we conclude that

$$\sqrt{n}\,(\hat{\boldsymbol{\gamma}}_{\boldsymbol{r},n}-\boldsymbol{\gamma})\xrightarrow{D}\mathcal{N}(0,\ \Sigma_{\boldsymbol{r}})\quad\text{for}\quad n\to\infty,$$

where $\Sigma_{\boldsymbol{r}}=\operatorname{Var}(\boldsymbol{Q}_{\boldsymbol{r}}+\boldsymbol{R}_{\boldsymbol{r}})$. The calculation of the variance matrix is a simple exercise since each component of the random vector $\boldsymbol{Q}_{\boldsymbol{r},n}+\boldsymbol{R}_{\boldsymbol{r},n}$ is a sum of i.i.d. random variables and

$\Sigma_{\boldsymbol{r}} = \lim_{n\to\infty} \mathrm{Var}(\boldsymbol{Q}_{\boldsymbol{r},n} + \boldsymbol{R}_{\boldsymbol{r},n})$. E.g., we have that

$$\begin{aligned}
&\lim_{n\to\infty} \mathrm{Cov}\left(\sqrt{n}(\hat{\alpha}_{k,r_1,n} - \alpha_k), \sqrt{n}(\hat{\beta}_{\ell,r_2,n} - \beta_\ell)\right) \\
&= \lim_{n\to\infty} \mathrm{Cov}\Bigg(\frac{\sqrt{n}}{\lfloor nr_1 \rfloor}\sum_{i=1}^{\lfloor nr_1 \rfloor} X_i F^k(X_i) + \frac{1}{r_1\sqrt{n}}\int xkF^{k-1}(x)\mathbb{1}(X_i \leqslant x)\, dF(x), \\
&\qquad\qquad \frac{\sqrt{n}}{\lfloor nr_2 \rfloor}\sum_{i=1}^{\lfloor nr_2 \rfloor} Y_i G^\ell(Y_i) + \frac{1}{r_2\sqrt{n}}\int y\ell G^{\ell-1}(y)\mathbb{1}(Y_i \leqslant y)\, dG(y)\Bigg) \\
&= \frac{\min(r_1, r_2)}{r_1 \cdot r_2} \cdot \mathrm{Cov}\Bigg(X_1F^k(X_1) + \int xkF^{k-1}(x)\mathbb{1}(X_1 \leqslant x)\, dF(x), \\
&\qquad\qquad Y_1G^\ell(Y_1) + \int y\ell G^{\ell-1}(y)\mathbb{1}(Y_1 \leqslant y)\, dG(y)\Bigg)
\end{aligned}$$

□

Proof of Corollary 2.

Let

$$\begin{aligned}
Z_{i,k,x} &= X_i \cdot F^k(X_i) + \int_{\mathbb{R}} xkF^{k-1}(x)\mathbb{1}(X_i \leqslant x)\, dF(x), \\
Z_{i,\ell,y} &= Y_i \cdot G^\ell(Y_i) + \int_{\mathbb{R}} y\ell G^{\ell-1}(y)\mathbb{1}(Y_i \leqslant y)\, dG(y)
\end{aligned}$$

for $k, \ell \in \mathbb{N}_0$, $i = 1, \ldots, m$ and $m = \min\{nr_1, nr_2\}$. We further let $\hat{Z}_{i,k,x}$ (resp. $\hat{Z}_{i,\ell,y}$) be defined analogue with F (resp. G) replaced by its empirical counterpart $F_{\lfloor nr_1 \rfloor}$ (resp. $G_{\lfloor nr_2 \rfloor}$). We denote by $\tilde{\sigma}_{k,\ell,m}$ (resp. $\hat{\sigma}_{k,\ell,m}$) the empirical covariance of the bivariate sample $(Z_{i,k,x}, Z_{i,\ell,y})$, $i = 1, \ldots, m$ (resp. $(\hat{Z}_{i,k,x}, \hat{Z}_{i,\ell,y})$, $i = 1, \ldots, m$). From the strong law of large numbers we immediately obtain that $\tilde{\sigma}_{k,\ell,m} \overset{a.s.}{\to} \mathrm{Cov}(Z_{1,k,x}, Z_{1,\ell,y})$ for $n \to \infty$. It thus remains to show that

$$|\hat{\sigma}_{k,\ell,m} - \tilde{\sigma}_{k,\ell,m}| \xrightarrow{\mathbb{P}} 0 \quad \text{for } n \to \infty. \tag{11}$$

To make a long story short, (11) follows from the consistency of probability weighted moments proven in Theorem 1, (C.12) in Kojadinovic and Naveau (2015) and from the consistency of the empirical process $\mathbb{W}_{\boldsymbol{r},n}$ defined in the proof of Theorem 1. A detailed presentation is omitted for the sake of brevity. □

C Re-parametrization of the GEV distribution by TL-moments

This section recaps the equation systems used to calculate GEV parameters from TL(0,0)- and TL(0,1)-moments, respectively. We also present the corresponding Jacobi matrices involved in formulas (5) and (6).

TL(0,0)

Let $\boldsymbol{\vartheta} = (\mu, \sigma, \xi)'$ with $\xi < 1$ and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)'$ denote parameters and untrimmed L-moments of a GEV distribution, respectively. Hosking et al. (1985) proved that $\boldsymbol{\vartheta} = \phi(\boldsymbol{\lambda})$, where ϕ is a bijective function implicitly defined by equation system

$$\begin{cases} \frac{2\cdot 3^\xi - 3\cdot 2^\xi + 1}{2^\xi - 1} &= \frac{\lambda_3}{\lambda_2} \\ \sigma &= \frac{\lambda_2 \xi}{\Gamma(1-\xi)(2^\xi - 1)} \\ \mu &= \lambda_1 + \frac{\sigma}{\xi}(1 - \Gamma(1-\xi)) \end{cases}$$

and with Γ denoting the gamma function. However, there is no explicit expression for ϕ as a function of $\boldsymbol{\lambda}$. Practitioners thus commonly replace the first line by

$$\xi = -7.859z - 2.9554z^2, \; z = \frac{2}{3 + \lambda_3/\lambda_2} - \frac{\log 2}{\log 3}$$

based on a second order polynomial approximation in order to obtain an explicit solution. Accordingly the Jacobi matrix $A^* = \frac{\partial}{\partial \boldsymbol{\lambda}} \phi(\boldsymbol{\lambda})$ involved in the asymptotic distribution of L-moment estimators is approximated by that of the explicit solution. For the latter we obtain

$$A = \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix}$$

with

$$\begin{aligned}
a_{12} =& \frac{\log(2)\,\lambda_2\,(\Gamma(1-\pi)-1)\,2^\pi\,\rho\,\theta}{\Gamma(1-\pi)\,(1-2^\pi)^2} + \frac{\lambda_2\,\psi_0(1-\pi)\,(\Gamma(1-\pi)-1)\,\rho\,\theta}{\Gamma(1-\pi)\,(1-2^\pi)} - \\
& \frac{\lambda_2\,\psi_0(1-\pi)\,\rho\,\theta}{1-2^\pi} + \frac{\Gamma(1-\pi)-1}{\Gamma(1-\pi)\,(1-2^\pi)} \\
a_{13} =& -\frac{\log(2)\,{\lambda_2}^2\,(\Gamma(1-\pi)-1)\,2^{\pi+1}\,\rho\,\zeta^2}{\Gamma(1-\pi)\,(1-2^\pi)^2} - \frac{2\,{\lambda_2}^2\,\psi_0(1-\pi)\,(\Gamma(1-\pi)-1)\,\rho\,\zeta^2}{\Gamma(1-\pi)\,(1-2^\pi)} + \\
& \frac{2\,{\lambda_2}^2\,\psi_0(1-\pi)\,\rho\,\zeta^2}{1-2^\pi} \\
a_{22} =& -\frac{\log(2)\,\lambda_2\,\pi\,2^\pi\,\rho\,\theta}{\Gamma(1-\pi)\,(1-2^\pi)^2} - \frac{\lambda_2\,\rho\,\theta\,(\psi_0(1-\pi)\,\pi+1)}{\Gamma(1-\pi)\,(1-2^\pi)} - \frac{\pi}{\Gamma(1-\pi)\,(1-2^\pi)} \\
a_{23} =& \frac{\log(2)\,{\lambda_2}^2\,\pi\,2^{\pi+1}\,\rho\,\zeta^2}{\Gamma(1-\pi)\,(1-2^\pi)^2} + \frac{2\,{\lambda_2}^2\,\rho\,\zeta^2\,(\psi_0(1-\pi)\,\pi+1)}{\Gamma(1-\pi)\,(1-2^\pi)} \\
a_{32} =& -2\,\lambda_3\,(2\,b\,\kappa\,\lambda_3 - a\,\lambda_3 + 6\,b\,\kappa\,\lambda_2 - 4\,b\,\lambda_2 - 3\,a\,\lambda_2)\,\zeta^3 \\
a_{33} =& 2\,\lambda_2\,(2\,b\,\kappa\,\lambda_3 - a\,\lambda_3 + 6\,b\,\kappa\,\lambda_2 - 4\,b\,\lambda_2 - 3\,a\,\lambda_2)\,\zeta^3
\end{aligned}$$

and with $a = -7.859, \;\; b = -2.9554, \;\; \kappa = \frac{\log 2}{\log 3},$

$\zeta = 1/(\lambda_3 + 3\lambda_2), \;\; \theta = (2\zeta - 6\lambda_2\zeta^2), \;\; \eta = (2\lambda_2\zeta - \kappa), \;\; \pi = b\eta^2 + a\eta, \;\; \rho = 2b\eta + a,$
$\psi_0(x) = \Gamma'(x)/\Gamma(x)$

TL(0,1)

Considering trimmed L-moments $\boldsymbol{\lambda}^{(0,1)}$ of a GEV distribution with parameters $\boldsymbol{\vartheta}$ it is also known that $\boldsymbol{\vartheta} = \psi(\boldsymbol{\lambda}^{(0,1)})$, where again ψ is implicitly defined by

$$\begin{cases} \frac{5\cdot 4^\xi - 12\cdot 3^\xi + 9\cdot 2^\xi - 2}{3^\xi - 2^{\xi+1} + 1} &= \frac{9\lambda_3^{(0,1)}}{4\lambda_2^{(0,1)}} \\ \sigma &= \frac{2\cdot\lambda_2^{(0,1)}}{3\Gamma(-\xi)\cdot(3^\xi - 2^{\xi+1} + 1)} \\ \mu &= \lambda_1^{(0,1)} + \frac{\sigma}{\xi} - \frac{\sigma\cdot\Gamma(-\xi)}{(2^\xi - 2)^{-1}} \end{cases}.$$

In order to obtain an explicit solution, the first line can by replacement by a second order polynomial approximation

$$\xi = -8.5674z + 0.6760z^2, \; z = \frac{10}{9}\frac{\lambda_2^{(0,1)}}{2\lambda_2^{(0,1)} + \lambda_3^{(0,1)}} - \frac{2\log 2 - \log 3}{3\log 3 - 2\log 4}.$$

The Jacobi matrix of ψ is approximated by

$$A = \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix},$$

where

$$\begin{aligned}
a_{12} =& -\frac{2\,\lambda_2^{(0,1)}\,\psi_0\,(-\pi)\,\left(-2\,b\,(\zeta-\eta)\,\left(\lambda_2^{(0,1)}\,\zeta-\kappa\right)-a\,(\zeta-\eta)\right)+2}{3\,\pi\,\left(-2^{\pi+1}+3^{\pi}+1\right)\,\gamma\,(-\pi)} \\
&- \frac{2\,\lambda_2^{(0,1)}\,\left(\log(3)\,\iota\,3^{\pi}-\log(2)\,\iota\,2^{\pi+1}\right)(1-(2^{\pi}-2)\,\pi\,\gamma(-\pi))}{3\,\pi\,(-2^{\pi+1}+3^{\pi}+1)^2\,\gamma\,(-\pi)} \\
&- \frac{2\,\iota\,\lambda_2^{(0,1)}}{3\,\pi^2\,\left(-2^{\pi+1}+3^{\pi}+1\right)\,\gamma\,(-\pi)} - \frac{\log(2)\,\iota\,\lambda_2^{(0,1)}\,2^{\pi+1}+2\,\left(2^{\pi}-2\right)}{3\,\left(-2^{\pi+1}+3^{\pi}+1\right)} \\
a_{13} =& -\frac{2\,\lambda_2^{(0,1)}\,\left(\log(2)\,\,2^{\pi+1}\,\rho-\log(3)\,\,3^{\pi}\,\rho\right)\,\left(1-\left(2^{\pi}-2\right)\,\pi\,\gamma(-\pi)\right)}{3\,\pi\,(-2^{\pi+1}+3^{\pi}+1)^2\,\gamma\,(-\pi)} \\
&- \frac{2\,\lambda_2^{(0,1)}\,\rho\,(\psi_0(-\pi)\,\pi-2^{\pi}\log(2)-1)}{3\,\pi^2\,\left(-2^{\pi+1}+3^{\pi}+1\right)\,\gamma\,(-\pi)} \\
a_{22} =& -\frac{2\,\lambda_2^{(0,1)}\,\psi_0\,(-\pi)\,\left(-2\,b\,(\zeta-\eta)\,\left(\lambda_2^{(0,1)}\,\zeta-\kappa\right)-a\,(\zeta-\eta)\right)+2}{3\,\left(-2^{\pi+1}+3^{\pi}+1\right)\,\gamma\,(-\pi)} \\
&- \frac{2\,\lambda_2^{(0,1)}\,\left(\log(3)\,\iota\,3^{\pi}-\log(2)\,\iota\,2^{\pi+1}\right)}{3\,(-2^{\pi+1}+3^{\pi}+1)^2\,\gamma\,(-\pi)} \\
a_{23} =& -\frac{2\,\lambda_2^{(0,1)}\,\left(\log(2)\,\,2^{\pi+1}\,\rho-\log(3)\,\,3^{\pi}\,\rho\right)}{3\,(-2^{\pi+1}+3^{\pi}+1)^2\,\gamma\,(-\pi)} - \frac{2\,\lambda_2^{(0,1)}\,\psi_0\,(-\pi)\,\rho}{3\,\left(-2^{\pi+1}+3^{\pi}+1\right)\,\gamma\,(-\pi)} \\
a_{32} =& \iota \\
a_{33} =& -\rho
\end{aligned}$$

and with $a = -8.5674, \;\; b = 0.6760, \;\; \kappa = \dfrac{2\log 2 - \log 3}{3\log 3 - 2\log 4},$

$$\begin{aligned}
&\theta = 3(\lambda_3^{(0,1)} + 2\lambda_2^{(0,1)}), \;\; \zeta = 10/(3\theta), \;\; \eta = 20\lambda_2^{(0,1)}/\theta^2, \;\; \pi = b(\lambda_2^{(0,1)}\zeta-\kappa)^2 + a(\lambda_2^{(0,1)}\zeta-\kappa) \\
&\rho = -b\eta(\lambda_2^{(0,1)}\zeta-\kappa), \;\; \iota = 2b(\zeta-\eta)(\lambda_2^{(0,1)}\zeta-\kappa) + a(\zeta-\eta) \\
&\psi_0(x) = \Gamma'(x)/\Gamma(x)
\end{aligned}$$

References

Dalrymple, T. (1960). Flood-frequency analyses, manual of hydrology: Part 3. Technical report, USGPO.

Genest, C. and Segers, J. (2009). Rank-based inference for bivariate extreme-value copulas. *Ann. Statist.*, 37(5B):2990–3022.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J. R. (1979). Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, 15(5):1049–1054.

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1):105–124.

Hosking, J. R. M. and Wallis, J. R. (2005). *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3):pp. 251–261.

Kojadinovic, I. and Naveau, P. (2015). Nonparametric tests for change-point detection in the distribution of block maxima based on probability weighted moments. *ArXiv:1507.06121*.

Lettenmaier, D. P., Wallis, J. R., and Wood, E. F. (1987). Effect of regional heterogeneity on flood frequency estimation. *Water Resources Research*, 23(2):313–323.