CrossMark

Detecting distributional changes in samples of independent block maxima using probability weighted moments

Ivan Kojadinovic[1] · Philippe Naveau[2]

Abstract The analysis of seasonal or annual block maxima is of interest in fields such as hydrology, climatology or meteorology. In connection with the celebrated method of block maxima, we study several tests that can be used to assess whether the available series of maxima is identically distributed. It is assumed that block maxima are independent but not necessarily generalized extreme value distributed. The asymptotic null distributions of the test statistics are investigated and the practical computation of approximate p-values is addressed. Extensive Monte-Carlo simulations show the adequate finite-sample behavior of the studied tests for a large number of realistic data generating scenarios. Illustrations on several environmental datasets conclude the work.

Keywords Asymptotic statistics · Cumulative sum statistics · Generalized extreme value distribution · Partial-sum empirical processes · Weighted empirical processes

AMS 2000 Subject Classifications 62G32 · 62G10 · 62G20

✉ Ivan Kojadinovic
ivan.kojadinovic@univ-pau.fr

Philippe Naveau
philippe.naveau@lsce.ipsl.fr

[1] Laboratoire de Mathématiques et Applications, Université de Pau et des Pays de l'Adour, UMR CNRS IPSL UVSQ CEA 5142, B.P. 1155, 64013 Pau Cedex, France

[2] Laboratoire des Sciences du Climat et de l'Environnement, CNRS, Orme des Merisiers, Bât. 701 C.E.A Saclay, 91191 Gif-sur-Yvette, France

1 Introduction

The *block maximum method* put forward in the seminal monograph of Gumbel (1958)
is frequently applied in environmental sciences to analyse extremes of a given series
of observations. It consists of dividing the initial high frequency observations into
blocks of equal size and forming a sample of block maxima. A typical example is
the study of annual or seasonal maxima of temperatures or precipitation from daily
observations. The distribution of the block maxima is classically modeled using the
generalized extreme value (GEV) distribution. Under rather broad conditions on the
underlying initial observations, the extremal types theorem (see, e.g., de Haan and
Ferreira, 2006) states that, as the block size increases to infinity, the only possi-
ble non-degenerated limit of affinely normalized block maxima is the GEV. The
cumulative distribution function (c.d.f.) of this three-parameter distribution is

$$
G_{\mu,\sigma,\xi}(x) =
\begin{cases}
\exp\left\{ -\left(1 + \xi \cdot \frac{x-\mu}{\sigma}\right)_+^{-1/\xi} \right\}, & \text{if } \xi \neq 0, \\[2ex]
\exp\left\{ -\exp\left(-\frac{x-\mu}{\sigma}\right) \right\}, & \text{if } \xi = 0,
\end{cases}
\tag{1.1}
$$

for $x \in \mathbb{R}$, where $\mu \in \mathbb{R}$, $\sigma > 0$ and $\xi \in \mathbb{R}$ are the so-called location, scale and shape
parameters, respectively, and with the notation $(\cdot)_+ = \max(\cdot, 0)$.

   The adequacy of the GEV distribution as a model for a sample of block maxima
is predicted by the theory when the underlying high frequency observations are iden-
tically distributed and independent, or satisfy some suitable mixing conditions (see,
e.g., Leadbetter et al. 1983), and the block size is "sufficiently large". Sometimes
maxima are directly observed, making it impossible to question the stationarity of
the initial series. Furthermore, the block maxima method may be applied even when
the initial series is clearly not identically distributed. For instance, when the original
observations exhibit seasonality, it suffices to take the block size to be a multiple of
the season length.

   When the block size is large, one of the least questionable assumptions might be
that of independence among block maxima. Before fitting a GEV distribution to the
resulting extremes, two questions then arise naturally: (i) Is the series of independent
maxima identically distributed? (ii) Is the block size large enough to warrant a good
fit by the GEV distribution? The aim of this work is to focus on the first question
without assuming that the answer to the second question is positive. Specifically, we
consider that we have at hand a sample of independent block maxima $X_1, \ldots, X_n$
and our objective is to develop tests for

$$
H_0 : \exists F \text{ such that } X_1, \ldots, X_n \text{ have c.d.f. } F
\tag{1.2}
$$

against alternatives involving the non-constancy of the c.d.f. In other words, we aim
at developing tests for detecting changes in the distribution of independent block
maxima.

   Under the additional assumption that each X_i is GEV distributed, a natural
approach for testing H_0 is to apply the maximum likelihood based theory developed
in the first chapter of the well-known monograph of Csörgő and Horváth (1997) on
change-point analysis. The latter approach was considered for instance in Jaruskova

and Rencová ([2008](#)) but not without significant practical difficulties related to maximum likelihood estimation of the GEV parameters from small samples (Hosking et al. [1985](#), see also for instance the discussions in; Diebolt et al. [2008](#)).

The approach that we consider is based on *probability weighted moments* (PWM) (see, e.g., Greenwood et al. [1979](#); Hosking et al. [1985](#); Hosking and Wallis [1987](#); Diebolt et al. [2008](#); Guillou et al. [2009](#); Ferreira and de Haan [2015](#)) and has the advantage of giving rise to tests that can hold their level asymptotically even if the distributions of the block maxima X_i are not GEV. To illustrate that point, in our simulations, some scenarios involved block maxima obtained from blocks of size 1, 5, 10 and 50. The later sentence actually implies that the tests can be applied outside of the block maxima setting, although, by construction, this is less natural. These aspects will be clarified in the forthcoming sections.

The rest of the paper is organized as follows. In Section [2](#), we recall the PWM method for estimating the parameters of the GEV due to Hosking et al. ([1985](#)), propose related *cumulative sum* (CUSUM) statistics for change-point detection, establish the limiting null distributions of the latter and address the computation of approximate p-values. The third section proposes alternative tests in the framework of the generalized PWM method for estimating the GEV due to Diebolt et al. ([2008](#)). Simulation results are partially reported in the fourth section and show the adequate behavior of the tests for many realistic data generating scenarios. The last section presents several fully reproducible illustrations on environmental datasets.

All proofs are deferred to a sequence of appendices. Appendices [B](#) and [C](#) in particular establish necessary intermediate results that might be of independent interest, namely the weak convergence of the sequential weighted uniform empirical process and of certain sequential empirical processes constructed from PWM estimators. The paper is also accompanied by supplementary material available upon request from the authors and mostly consisting of simulation results that are not reported in Section [4](#). Finally, note that the studied tests for change-point detection are implemented in the package npcp (Kojadinovic [2015](#)) for the R statistical environment (R Core Team [2016](#)).

2 The PWM method and related statistics

The statistics for testing H_0 in ([1.2](#)) that we propose are related to the method of PWM for estimating the GEV parameters (Hosking et al. [1985](#)). After briefly recalling the method and suggesting related statistics for change-point detection, we study the limiting null distributions of the latter and address the practical computation of approximate p-values.

2.1 Estimation of the GEV using PWM

A general definition of PWM can be found for instance in Hosking et al. ([1985](#), Section [2](#)) (see also Appendix [C](#)). Let X be a random variable with c.d.f. F and consider the functions $v_1(x) = 1$, $v_2(x) = x$ and $v_3(x) = x^2$, $x \in [0, 1]$. The quantities $\beta_i = \mathrm{E}[Xv_i\{F(X)\}]$, $i \in \{1, 2, 3\}$, are the three (probability weighted) moments used in the PWM method for estimating the GEV parameters.

Assume that H_0 in (1.2) holds with F the c.d.f. of the GEV given in (1.1). The method exploits the following relationship between β_1, β_2 and β_3 and the parameters μ, σ and ξ of the GEV. For $\sigma > 0$, $\xi < 1$ and $\xi \neq 0$, there holds

$$\begin{cases} \beta_1 = \mu - \frac{\sigma}{\xi}\{1 - \Gamma(1 - \xi)\}, \\ 2\beta_2 - \beta_1 = \frac{\sigma}{\xi}\Gamma(1 - \xi)(2^\xi - 1), \\ \frac{3\beta_3 - \beta_1}{2\beta_2 - \beta_1} = \frac{3^\xi - 1}{2^\xi - 1}, \end{cases} \tag{2.1}$$

where Γ is the Gamma function. Using continuity arguments, some thought reveals that the constraint $\xi \neq 0$ is actually unnecessary. Estimators of μ, σ and ξ are then the solution of the above system of equations in which β_1, β_2 and β_3 are replaced by their estimators.

Two ways to estimate β_1, β_2 and β_3 were considered by Hosking et al. (1985). The first approach consists of first estimating F by a slight modification of the empirical c.d.f. of the sample, and then estimating the PWM by suitable arithmetic means. For reasons that shall become clear later in this section, we present the resulting estimators when based on a subsample $X_k, \ldots, X_l$, $1 \leq k \leq l \leq n$, of the available data. The c.d.f. F is estimated by

$$F_{k:l}(x) = \frac{1}{l - k + 1}\left\{\sum_{j=k}^{l} \mathbf{1}(X_j \leq x) + \gamma\right\}, \qquad x \in \mathbb{R}, \tag{2.2}$$

with the convention that $F_{k:l} = 0$ for $k > l$ and where γ is a constant whose value suggested by Hosking et al. (1985) is -0.35. A sensible estimator of β_i is then

$$\hat{\beta}_{i,k:l} = \frac{1}{l - k + 1}\sum_{j=k}^{l} X_j v_i\{F_{k:l}(X_j)\} \tag{2.3}$$

with the convention that $\hat{\beta}_{i,k:l} = 0$ for $k > l$.

A second equally natural possibility to estimate β_1, β_2 and β_3 is to consider the asymptotically equivalent estimators $\hat{b}_{1,1:n}$, $\hat{b}_{2,1:n}$, $\hat{b}_{3,1:n}$ proposed by Landwehr et al. (1979) and defined, for a subsample $X_k, \ldots, X_l$, $1 \leq k \leq l \leq n$, by

$$\hat{b}_{i,k:l} = \frac{1}{l - k + 1}\sum_{j=k}^{l} \frac{\prod_{m=1}^{i-1}\{(j - k + 1) - m\}}{\prod_{m=1}^{i-1}\{(l - k + 1) - m\}} X_{(j),k:l}, \tag{2.4}$$

where $X_{(k),k:l} \leq \cdots \leq X_{(l),k:l}$ are the order statistics computed from $X_k, \ldots, X_l$, and with the convention that $\hat{b}_{i,k:l} = 0$ if $l - k < i$. The latter estimators are unbiased (they are related to U-statistics).

Hosking et al. (1985, Section 4) claimed that the $\hat{\beta}_{i,k:l}$ are slightly superior to the $\hat{b}_{i,k:l}$ in terms of finite-sample performance. While this seems true when $X_1, \ldots, X_n$ are i.i.d. from a GEV with location zero and scale one, it does not seem to necessarily hold for arbitrary location and scale parameters.

Let g_μ, g_σ and g_ξ be the components of the map from $\mathbb{R}^3$ to $\mathbb{R}^3$ implicitly defined by (2.1) that transforms $\beta = (\beta_1, \beta_2, \beta_3)$ into (μ, σ, ξ) (that is, $\mu = g_\mu(\beta)$, $\sigma = g_\sigma(\beta)$ and $\xi = g_\xi(\beta)$), and let us first focus on the function g_ξ without assuming anymore that F is the c.d.f. of the GEV but only that F is continuous and

that β_1, β_2 and β_3 exist. Under the constraint $\xi < 1$, it can be easily verified that the solution $g_\xi(\beta)$ of the third equation in (2.1) exists and is unique if and only if $1 < (3\beta_3 - \beta_1)/(2\beta_2 - \beta_1) < 2$. The latter is guaranteed by the following result proved in Appendix A.

Proposition 2.1 *Assume that F is continuous and that β_1, β_2 and β_3 exist. Then, $2\beta_2 - \beta_1 > 0$, $3\beta_3 - 2\beta_2 > 0$ and $-\beta_1 + 4\beta_2 - 3\beta_3 > 0$.*

Given the equivalence

$$\begin{cases} 2\beta_2 - \beta_1 > 0, \\ 3\beta_3 - 2\beta_2 > 0, \\ -\beta_1 + 4\beta_2 - 3\beta_3 > 0, \end{cases} \iff \begin{cases} 2\beta_2 - \beta_1 > 0, \\ 3\beta_3 - \beta_1 > 0, \\ 1 < (3\beta_3 - \beta_1)/(2\beta_2 - \beta_1) < 2, \end{cases}$$

it is natural to consider g_ξ as a function defined on the convex subset

$$\mathcal{D}_\xi = \{x \in \mathbb{R}^3 : 2x_2 - x_1 > 0, 3x_3 - 2x_2 > 0, -x_1 + 4x_2 - 3x_3 > 0\}. \quad (2.5)$$

In addition, it is easy to see that, since $\beta \in \mathcal{D}_\xi$, necessarily $g_\xi(\beta) < 1$ and $g_\sigma(\beta) > 0$, which is fully in accordance with the constraints under which the system (2.1) was derived. Using the fact that $x \mapsto x/(2^x - 1)$ can be extended by continuity at 0 and is strictly positive, it is also natural to consider g_σ and g_μ as functions defined on $\mathcal{D}_\xi$.

When β_1, β_2 and β_3 are estimated using the unbiased estimators $\hat{b}_{1,1:n}, \hat{b}_{2,1:n}, \hat{b}_{3,1:n}$ defined analogously to (2.4), Hosking et al. (1985, Appendix B) showed that $\hat{b}_{1:n} = (\hat{b}_{1,1:n}, \hat{b}_{2,1:n}, \hat{b}_{3,1:n}) \in \mathcal{D}_\xi$ for all $n \geq 3$, where $\mathcal{D}_\xi$ is defined in (2.5), which implies that $g_\xi(\hat{b}_{1:n}) < 1$ and $g_\sigma(\hat{b}_{1:n}) > 0$ for all $n \geq 3$. The latter desirable property is referred to as the *feasibility criterion* of the PWM method. When $\hat{\beta}_{1:n} = (\hat{\beta}_{1,1:n}, \hat{\beta}_{2,1:n}, \hat{\beta}_{3,1:n})$ is used instead of $\hat{b}_{1:n}$, an analogue result is not available. However, provided $E(|X_1|) < \infty$, a consequence of the strong law of large numbers and the Glivenko–Cantelli lemma is that $\hat{\beta}_{1:n}$ converges almost surely to β (see Appendix D). The latter immediately implies that $P(\hat{\beta}_{1:n} \in \mathcal{D}_\xi) \to 1$.

2.2 Test statistics

The above developments suggest to study tests for change-point detection based on the three CUSUM like statistics

$$S_{g,n} = \max_{1 \leq k \leq n-1} \frac{k(n-k)}{n^{3/2}} \mathbf{1}(\hat{\beta}_{1:k} \in \mathcal{D}_\xi, \hat{\beta}_{k+1:n} \in \mathcal{D}_\xi) \left| g(\hat{\beta}_{1:k}) - g(\hat{\beta}_{k+1:n}) \right|, \ g \in \{g_\mu, g_\sigma, g_\xi\}, \quad (2.6)$$

or on the three alternative ones

$$T_{g,n} = \max_{3 \leq k \leq n-3} \frac{k(n-k)}{n^{3/2}} \left| g(\hat{b}_{1:k}) - g(\hat{b}_{k+1:n}) \right|, \qquad g \in \{g_\mu, g_\sigma, g_\xi\}. \quad (2.7)$$

In (2.6), $\hat{\beta}_{1:k} = (\hat{\beta}_{1,1:k}, \hat{\beta}_{2,1:k}, \hat{\beta}_{3,1:k})$ and $\hat{\beta}_{k+1:n} = (\hat{\beta}_{1,k+1:n}, \hat{\beta}_{2,k+1:n}, \hat{\beta}_{3,k+1:n})$ are defined according to (2.3), and $\mathcal{D}_\xi$ is given in (2.5), while in (2.7), $\hat{b}_{1:k} =$

$(\hat{b}_{1,1:k}, \hat{b}_{2,1:k}, \hat{b}_{3,1:k})$ and $\hat{b}_{k+1:n} = (\hat{b}_{1,k+1:n}, \hat{b}_{2,k+1:n}, \hat{b}_{3,k+1:n})$ are defined according to (2.4).

For instance, under H_0 in (1.2) with F the c.d.f. of the GEV, $S_{g_\xi,n}$ is the maximum over $k \in \{1, \ldots, n-1\}$ of the normalized absolute difference between $g_\xi(\hat{\beta}_{1:k})$, the estimator of the shape parameter ξ computed from $X_1, \ldots, X_k$, and $g_\xi(\hat{\beta}_{k+1:n})$, the same estimator computed from $X_{k+1}, \ldots, X_n$. The role of the indicator $\mathbf{1}(\hat{\beta}_{1:k} \in \mathcal{D}_\xi, \hat{\beta}_{k+1:n} \in \mathcal{D}_\xi)$ is to allow the previous evaluations of g_ξ only if $\hat{\beta}_{1:k} \in \mathcal{D}_\xi$ and $\hat{\beta}_{k+1:n} \in \mathcal{D}_\xi$. The coefficient $k(n-k)/n^{3/2}$ in front of each absolute difference in (2.6) is the classical normalizing term in the CUSUM approach ensuring that, under additional suitable assumptions, $S_{g_\xi,n}$ converges in distribution under the null. The interpretation of $T_{g_\xi,n}$ is similar. The reason for the disappearance of the indicators in its expression is that, as discussed in Section 2.1, $\hat{b}_{1:k} \in \mathcal{D}_\xi$ for all $k \geq 3$ (Hosking et al. 1985, Appendix B).

If each block maxima in the sample $X_1, \ldots, X_n$ is GEV distributed, $S_{g_\xi,n}$ and $T_{g_\xi,n}$ (resp. $S_{g_\sigma,n}$ and $T_{g_\sigma,n}$, $S_{g_\mu,n}$ and $T_{g_\mu,n}$) are thus test statistics particularly sensitive to changes in the shape parameter ξ (resp. the scale parameter σ, the location parameter μ) of the GEV. It is however important to keep in mind that the proposed tests are actually nonparametric in the sense that they remain potentially meaningful even if the block maxima are not GEV distributed. The latter follows from the fact $\hat{\beta}_{1,1:n} = \hat{b}_{1,1:n}$ is simply the sample mean, while, for instance from Hosking et al. (1985, page 252), we have that $2\hat{b}_{2,1:n} - \hat{b}_{1,1:n}$ is actually half of the Gini's mean difference and, under H_0 in (1.2), is thus an estimator of scale as measured by $2\beta_2 - \beta_1 = \mathrm{E}(|X_1 - X_2|)/2$. Consequently, the tests based on $T_{g_\mu,n}$ and $S_{g_\mu,n}$ (resp. $T_{g_\sigma,n}$ and $S_{g_\sigma,n}$) can in general be interpreted as tests for change-point detection particularly sensitive to changes in the location (resp. scale) of the distribution. Similarly, the tests based on $T_{g_\xi,n}$ and $S_{g_\xi,n}$ can be interpreted in general as tests for change-point detection particularly sensitive to changes in the upper tail of F. Indeed, the ratio $(3\beta_3 - \beta_1)/(2\beta_2 - \beta_1)$ appearing in the expression of g_ξ can be rewritten as

$$\frac{\mathrm{E}\{\max(X_1, X_2, X_3) - (X_1 + X_2 + X_3)/3\}}{\mathrm{E}\{\max(X_1, X_2) - (X_1 + X_2)/2\}}$$

and is thus clearly invariant to the replacement of the X_i by $cX_i + d$ if $c > 0$. Some thought reveals that this ratio captures some information on the upper tail of F.

The statement that the tests are actually nonparametric will also become clearer in view of the theoretical results established in the next two subsections and will be illustrated in the Monte Carlo experiments partially reported in Section 4.

2.3 Limiting null distributions of $S_{g,n}$ and $T_{g,n}$

We shall now state conditions under which $S_{g,n}$ in (2.6) and $T_{g,n}$ in (2.7) converge in distribution under the null.

Condition 2.2 The random variables $X_1, \ldots, X_n$ are i.i.d. from a continuous distribution with c.d.f. F such that $\mathrm{E}(X_1^2) < \infty$ and there exists $\alpha \in [0, 1/2]$ such that $\sup_{x \in \mathbb{R}} |H_\alpha(x)| < \infty$, where $H_\alpha(x) = x[F(x)\{1 - F(x)\}]^\alpha$, $x \in \mathbb{R}$.

Condition 2.2 is for instance satisfied when F is the c.d.f. of the GEV in (1.1) with $\xi < 1/2$, that is, when it has finite variance. In addition, it can be verified that when both $\max(X_1, 0)$ and $-\min(X_1, 0)$ have an ultimately monotone density and are in the maximum domain of attraction of the GEV with shape parameter $\xi < 1/2$, then $\sup_{x \in \mathbb{R}} |H_\alpha(x)| < \infty$ for any $\alpha \in (\xi, 1/2)$ (see Section 3 in the supplementary material).

In the sequel, the first-order partial derivatives of $g \in \{g_\mu, g_\sigma, g_\xi\}$ will be denoted by $\partial_1 g$, $\partial_2 g$ and $\partial_3 g$, respectively. Recall furthermore that $v_1(x) = 1$, $v_2(x) = x$ and $v_3(x) = x^2$, $x \in [0, 1]$.

The following result is proved in Appendix D.

Proposition 2.3 (Limiting null distributions of $S_{g,n}$ and $T_{g,n}$) *Under Condition 2.2, for any $g \in \{g_\mu, g_\sigma, g_\xi\}$, $S_{g,n} - T_{g,n} = o_P(1)$ and $S_{g,n}$ converges in distribution to $\sigma_g \sup_{s \in [0,1]} |\mathbb{U}(s) - s\mathbb{U}(1)|$, where $\mathbb{U}$ is a standard Brownian motion on $[0, 1]$ and*

$$\sigma_g^2 = \sum_{i,j=1}^{3} \partial_i g(\beta) \partial_j g(\beta) \mathrm{Cov}(Y_{v_i}, Y_{v_j}), \tag{2.8}$$

with, for any $v \in \{v_1, v_2, v_3\}$,

$$Y_v = X_1 v\{F(X_1)\} + \int_{\mathbb{R}} x v'\{F(x)\} \mathbf{1}(X_1 \leq x) \mathrm{d}F(x). \tag{2.9}$$

2.4 Computation of approximate p-values

Proposition 2.3 suggests to base the computation of approximate p-values for $S_{g,n}$ and $T_{g,n}$, $g \in \{g_\mu, g_\sigma, g_\xi\}$, on their estimated asymptotic null distributions. Indeed, the distribution of the supremum of a Brownian bridge, known as the Kolmogorov distribution, can be approximated very well in practice. It therefore simply remains to estimate the unknown variance σ_g^2 in (2.8).

To do so, it is necessary to estimate the first-order partial derivatives of g at β and the covariances $\mathrm{Cov}(Y_{v_i}, Y_{v_j})$, $i, j \in \{1, 2, 3\}$, where Y_{v_i} is defined in (2.9). Natural estimators of the former are $\partial_1 g(\hat{\beta}_{1:n}), \ldots, \partial_p g(\hat{\beta}_{1:n})$ or $\partial_1 g(\hat{b}_{1:n}), \ldots, \partial_p g(\hat{b}_{1:n})$, while estimators of the latter can be based on the pseudo-observations

$$Y_{v,i,n} = X_i v\{F_{1:n}(X_i)\} + \frac{1}{n} \sum_{j=1}^{n} X_j v'\{F_{1:n}(X_j)\} \mathbf{1}(X_i \leq X_j),$$

for $i \in \{1, \ldots, n\}$ and $v \in \{v_1, v_2, v_3\}$, where $F_{1:n}$ is defined in (2.2). Indeed, a sensible estimator $\widehat{\mathrm{Cov}}_n(Y_{v_i}, Y_{v_j})$ of $\mathrm{Cov}(Y_{v_i}, Y_{v_j})$, $i, j \in \{1, 2, 3\}$, under H_0 is then simply the sample covariance computed from $(Y_{v_i,1,n}, Y_{v_j,1,n}), \ldots, (Y_{v_i,n,n}, Y_{v_j,n,n})$.

The proof of the following result is omitted for the sake of brevity.

Proposition 2.4 *Under Condition 2.2, for any $g \in \{g_\mu, g_\sigma, g_\xi\}$,*

$$\hat{\sigma}_{g,n}^2 = \sum_{i,j=1}^{p} \partial_i g(\hat{\beta}_{1:n}) \partial_j g(\hat{\beta}_{1:n}) \widehat{\mathrm{Cov}}_n(Y_{v_i}, Y_{v_j}) \tag{2.10}$$

and

$$\check{\sigma}_{g,n}^2 = \sum_{i,j=1}^{p} \partial_i g(\hat{b}_{1:n}) \partial_j g(\hat{b}_{1:n}) \widehat{\mathrm{Cov}}_n(Y_{v_i}, Y_{v_j})$$

converge almost surely to σ_g^2 *in* (2.8).

As we continue, we shall thus compute approximate p-values for $S_{g,n}$ as $1 - F_K(S_{g,n}/\hat{\sigma}_{g,n})$ and approximate p-values for $T_{g,n}$ as $1 - F_K(T_{g,n}/\check{\sigma}_{g,n})$, where F_K is the c.d.f. of the Kolmogorov distribution. As classically done, we approached F_K by the c.d.f. of the statistic of the Kolmogorov–Smirnov goodness-of-fit test for a simple hypothesis. From a practical perspective, we used the function `pkolmogorov1x` given in the code of the R function `ks.test`.

An even more important practical issue is the computation of $S_{g,n}$ and $T_{g,n}$ for g equal to g_μ, g_σ and g_ξ. The numerical solving of the system (2.1) can actually be avoided by using the following accurate approximations proposed in Hosking et al. (1985):

$$\xi = \tilde{g}_\xi(\beta) = f \circ f_\xi(\beta) \quad \text{with} \quad \begin{cases} f(x) = -7.8590x - 2.9554x^2, \\ f_\xi(x_1, x_2, x_3) = \frac{2x_2 - x_1}{3x_3 - x_1} - \frac{\log 2}{\log 3}, \end{cases} \quad (2.11)$$

$$\sigma = \tilde{g}_\sigma(\beta) = \frac{(2\beta_{v_2} - \beta_{v_1})\xi}{\Gamma(1 - \xi)(2^\xi - 1)}, \tag{2.12}$$

$$\mu = \tilde{g}_\mu(\beta) = \beta_{v_1} + \frac{\sigma}{\xi}\{1 - \Gamma(1 - \xi)\}. \tag{2.13}$$

We also used the above approximations in combination with symbolic differentiation to obtain approximations of the first-order partial derivatives of g_μ, g_σ and g_ξ necessary for computing $\hat{\sigma}_{g,n}^2$ and $\check{\sigma}_{g,n}^2$ in Proposition 2.4.

3 The GPWM method and related statistics

Alternative test statistics could yet be based on the generalization of the PWM method for estimating the parameters of the GEV proposed by Diebolt et al. (2008). Under the assumption that H_0 in (1.2) holds with F the c.d.f. of the GEV distribution, the estimation method involves the quantities $\beta_i = \mathrm{E}[X v_i\{F(X)\}]$, $i \in \{1, 2, 3\}$, with $v_1(x) = -x \log x$, $v_2(x) = x(\log x)^2$ and $v_3(x) = -x^2 \log x$, $x \in [0, 1]$. The moments β_1, β_2 and β_3 are called *generalized* PWM in Diebolt et al. (2008) as they are not PWM in the classical sense considered for instance in Greenwood et al. (1979) because of the definitions of the functions v_1, v_2 and v_3.

The relationship between the GEV parameters μ, σ and ξ and the GPWM β_1, β_2 and β_3 is given by the following system of equations, which holds for $\sigma > 0$, $\xi < 2$ and $\xi \neq 0$:

$$\begin{cases} 4\beta_1 = \mu - \frac{\sigma}{\xi}\{1 - 2^\xi \Gamma(2 - \xi)\}, \\ \beta_1 - \beta_2 = \frac{\sigma}{2^{3-\xi}} \Gamma(2 - \xi), \\ \frac{2(\beta_1 - \beta_2)}{\beta_1 - 9/4\beta_3} = \frac{\xi}{1 - (3/2)^\xi}. \end{cases} \tag{3.1}$$

Again, the constraint $\xi \neq 0$ can be removed using continuity arguments. Estimators of μ, σ and ξ are then the solution of the above system in which β_1, β_2 and β_3 are replaced by suitable estimators. A sensible choice is to consider $\hat{\beta}_{1,1:n}$, $\hat{\beta}_{2,1:n}$ and $\hat{\beta}_{3,1:n}$, respectively, defined analogously to (2.3).

Let h_μ, h_σ and h_ξ be the components of the map from $\mathbb{R}^3$ to $\mathbb{R}^3$ implicitly defined by (3.1) that transforms $\beta = (\beta_1, \beta_2, \beta_3)$ into (μ, σ, ξ). In practice, to evaluate whether $\hat{\beta}_{1:n} = (\hat{\beta}_{1,1:n}, \hat{\beta}_{2,1:n}, \hat{\beta}_{3,1:n}) \in \mathbb{R}^3$ belongs to the domain of definition $\mathcal{D}_h$ of h_ξ, h_σ and h_μ, we shall attempt to solve numerically (3.1) and will conclude that $\hat{\beta}_{1:n} \in \mathcal{D}_h$ if and only if $h_\xi(\hat{\beta}_{1:n}) < 2$, $h_\sigma(\hat{\beta}_{1:n}) > 0$ and $h_\mu(\hat{\beta}_{1:n}) \in \mathbb{R}$.

The above ingredients lead to the following generalization of the statistics $S_{g,n}$ in (2.6):

$$S_{h,n} = \max_{1 \leq k \leq n-1} \frac{k(n-k)}{n^{3/2}} \mathbf{1}(\hat{\beta}_{1:k} \in \mathcal{D}_h, \hat{\beta}_{k+1:n} \in \mathcal{D}_h) \left| h(\hat{\beta}_{1:k}) - h(\hat{\beta}_{k+1:n}) \right|, \ h \in \{h_\mu, h_\sigma, h_\xi\}. \tag{3.2}$$

Should the block maxima X_i be GEV distributed, $S_{h_\xi,n}$ (resp. $S_{h_\sigma,n}$, $S_{h_\mu,n}$) is a test statistic particularly sensitive to changes in the shape (resp. scale, location) parameter.

We conjecture that the asymptotics of the statistics in (3.2) under H_0 in (1.2) take the same form as those of the statistics $S_{g,n}$ considered in the previous section, the only difference being the definitions of the functions ν_1, ν_2 and ν_3. Let σ_h^2 be the analogue of (2.8). Specifically, we conjecture that, under the null, $S_{h,n}$ converges in distribution to $\sigma_h \sup_{s \in [0,1]} |\mathbb{U}(s) - s\mathbb{U}(1)|$, where $\mathbb{U}$ is a standard Brownian motion on $[0, 1]$. By means of a sequential extension of the method of proof used in Diebolt et al. (2008), it might be possible to show the aforementioned weak convergence when F is the c.d.f. of the GEV and h is either h_ξ, h_σ or h_μ.

To carry out the tests, consider the analogue $\hat{\sigma}_{h,n}^2$ of (2.10). Approximate p-values for $S_{h,n}$ can be computed as $1 - F_K(S_{h,n}/\hat{\sigma}_{h,n})$, where F_K is the c.d.f. of the Kolmogorov distribution.

From a practical perspective, as in the previous section, the numerical solving of system (3.1) can be avoided. Starting from the fact that the function $x \mapsto \{x/(1-(3/2)^x)\}^{\log(3/2)}$ is almost linear on $[-1, 1]$, we propose the following accurate approximations of h_ξ, h_σ and h_μ:

$$\xi = \tilde{h}_\xi(\beta) = f \circ f_\xi(\beta) \qquad \text{with} \qquad \begin{cases} f(x) = \frac{1.442853 - (-x)^{0.4054651}}{0.1183375}, \\ f_\xi(x_1, x_2, x_3) = \frac{2(x_1 - x_2)}{x_1 - \frac{9}{4}x_3}, \end{cases} \tag{3.3}$$

$$\sigma = \tilde{h}_\sigma(\beta) = \frac{2^{3-\xi}(\beta_1 - \beta_2)}{\Gamma(2 - \xi)}, \tag{3.4}$$

$$\mu = \tilde{h}_\mu(\beta) = 4\beta_1 + \frac{\sigma}{\xi}\{1 - 2^\xi \Gamma(2 - \xi)\}. \tag{3.5}$$

As in Section 2.4, approximations of the partial derivatives of h_ξ, h_σ and h_μ which are necessary for computing $\hat{\sigma}_{h,n}^2$, are obtained by differentiating $\tilde{h}_\xi$, $\tilde{h}_\sigma$ and $\tilde{h}_\mu$.

4 Monte-Carlo experiments

Monte-Carlo experiments were carried out in order to investigate the finite-sample behavior of the tests studied in the previous sections. Nine statistics were considered: $S_{g,n}$ in (2.6) and $T_{g,n}$ in (2.7) with g being respectively $\tilde{g}_\xi$ in (2.11), $\tilde{g}_\sigma$ in (2.12) and $\tilde{g}_\mu$ in (2.13), and the statistic $S_{h,n}$ in (3.2) with h being respectively $\tilde{h}_\xi$ in (3.3), $\tilde{h}_\sigma$ in (3.4) and $\tilde{h}_\mu$ in (3.5). Recall that, under the conditions of Proposition 2.3, the statistics $S_{g,n}$ and $T_{g,n}$ are asymptotically equivalent under the null, and that the functions $\tilde{g}_\xi$, $\tilde{g}_\sigma$, $\tilde{g}_\mu$ (resp. $\tilde{h}_\xi$, $\tilde{h}_\sigma$, $\tilde{h}_\mu$) are related to the PWM (resp. GPWM) method for estimating the GEV parameters summarized in Section 2.1 (resp. Section 3). When computing the statistics, the following asymptotically negligible change was made: given a small integer $r \geq 1$, maxima in (2.6), (2.7) and (3.2) were taken over the set $\{r, \ldots, n-r\}$ instead of the set $\{1, \ldots, n-1\}$ or $\{3, \ldots, n-3\}$. The role of the integer r is to exclude subsamples of cardinality strictly smaller than r. In our simulations, we used the value $r = 10$ and, in all the considered settings, encountered hardly any sample for which the indicators in (2.6) and (3.2) turned to be zero for some value of k.

We started our experiments by investigating the empirical levels of the nine tests for samples of size $n \in \{50, 100, 200, 400\}$ generated from the GEV(μ, σ, ξ). To estimate the power of the tests, 1000 samples were generated under each data generating scenario and all the tests were carried out at the 5 % significance level. Our main finding was that all tests had a tendency of being too liberal when $|\mu/\sigma|$ was large. This prompted us to run the tests based on $S_{g,n}$ in (2.6) (resp. $T_{g,n}$ in (2.7), $S_{h,n}$ in (3.2)) on the data translated by $-\tilde{g}_\mu(\hat{\beta}_{1:n})$ (resp. $-\tilde{g}_\mu(\hat{b}_{1:n})$, $-\tilde{h}_\mu(\hat{\beta}_{1:n})$). Since $\tilde{g}_\mu(\hat{\beta}_{1:n}) \xrightarrow{\text{a.s.}} \tilde{g}_\mu(\beta)$ and $\tilde{g}_\mu(\hat{b}_{1:n}) \xrightarrow{\text{a.s.}} \tilde{g}_\mu(\beta)$ under Condition 2.2, the asymptotic results stated in Section 2.3 remain valid when the tests based on the statistics $S_{g,n}$ and $T_{g,n}$ are applied on the translated data.

The adopted translation of the data led to a significant improvement of the empirical levels. Additional improvement of the levels of the test based on $T_{\tilde{g}_\sigma,n}$ (resp. $T_{\tilde{g}_\xi,n}$) was obtained by multiplying $\hat{\sigma}^2_{\tilde{g}_\sigma,n}$ (resp. $\hat{\sigma}^2_{\tilde{g}_\xi,n}$) in (2.10) by $(n + 10)/n$ (resp. $(n + 20)/n$). A last asymptotically negligible change was carried out for the tests based on $S_{h,n}$: the statistics given in (3.2) were computed with γ in (2.2) equal to zero.

As the tests based on $S_{g,n}$ in (2.6) turned out to be consistently worse-behaved than those based on $T_{g,n}$ in (2.7), for the sake of brevity, we do not report any results for the former in the forthcoming tables.

Table 1 gives the empirical levels of the six remaining tests for samples generated from the GEV$(\mu, 1, \xi)$ for $\mu = 0$ and $\xi \in \{-1, -0.9, \ldots, 1, 1.1\}$. Analogue tables for $\mu \in \{-100, -10, 10, 100\}$ can be found in Section 1.1 of the supplementary material (see Tables 1–4 therein). For the sample sizes under consideration, the tests based on $T_{g,n}$ appear to hold their level well for $-1 \leq \xi \leq 0.4$, that is, as long as the variance of the GEV is finite. The tests based on $S_{h,n}$ are overall too conservative for $-1 \leq \xi \leq 0.3$. Their empirical levels are closer to the 5 % nominal level for $0.4 \leq \xi \leq 0.8$ and they start being too liberal for $\xi \geq 0.9$.

Table 1 Percentage of rejection of H_0 computed from 1000 samples of size $n \in \{50, 100, 200, 400\}$ generated from the GEV(0,1,ξ)

ξ	n	PWM / $T_{g,n}$			GPWM / $S_{h,n}$			ξ	n	PWM / $T_{g,n}$			GPWM / $S_{h,n}$		
		$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$			$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$
-1.0	50	3.0	0.7	4.1	7.2	0.8	0.2	0.1	50	3.8	3.2	4.0	1.2	0.6	1.4
	100	2.9	1.6	3.7	7.4	1.7	0.2		100	4.9	3.4	2.7	2.7	1.4	1.3
	200	4.0	2.9	4.1	5.4	3.2	0.6		200	4.7	4.2	2.5	3.8	2.4	2.9
	400	3.8	3.3	4.4	3.9	2.9	1.1		400	4.4	4.5	3.4	4.1	3.5	3.9
-0.9	50	2.6	0.7	3.8	4.8	0.2	0.4	0.2	50	6.2	5.4	4.4	2.3	1.5	1.5
	100	4.3	2.0	4.2	6.3	1.5	0.2		100	5.2	4.9	3.6	3.6	2.2	1.6
	200	4.8	2.8	5.1	4.8	2.4	0.7		200	5.0	4.4	3.2	4.3	2.6	4.2
	400	4.0	3.3	3.6	3.4	3.5	1.1		400	4.6	4.2	4.2	4.2	3.9	4.4
-0.8	50	2.2	0.8	4.0	4.5	0.2	0.3	0.3	50	5.7	3.7	5.0	2.2	0.8	1.4
	100	2.2	2.4	2.8	4.3	1.6	0.1		100	4.3	4.1	2.9	2.6	1.4	2.0
	200	4.1	3.8	3.8	4.7	3.6	0.5		200	6.1	6.0	2.5	4.5	3.0	3.2
	400	5.6	4.1	4.2	5.9	4.3	2.0		400	5.0	4.8	2.6	3.7	3.7	4.0
-0.7	50	2.4	1.4	3.9	3.6	0.1	0.0	0.4	50	6.6	4.4	8.1	3.1	1.5	2.1
	100	3.1	2.5	4.3	4.2	1.1	0.1		100	6.8	5.6	4.5	4.5	2.0	2.9
	200	3.9	3.7	4.1	5.0	2.7	0.6		200	6.1	5.0	3.3	4.0	3.8	3.0
	400	3.7	3.7	3.9	4.3	3.5	1.3		400	5.3	4.5	2.1	3.6	4.8	4.3
-0.6	50	2.2	1.9	3.6	2.2	0.1	0.3	0.5	50	8.0	4.2	9.5	2.7	1.3	3.5
	100	4.1	2.9	3.8	3.7	0.7	0.3		100	9.2	7.0	6.8	4.8	3.2	2.9
	200	5.1	4.1	3.5	4.3	2.5	0.7		200	5.8	4.5	3.4	4.8	3.3	3.5
	400	4.3	4.6	4.0	4.4	4.1	2.4		400	5.4	5.6	2.9	4.5	4.1	4.2
-0.5	50	3.3	3.2	4.5	2.7	0.3	0.7	0.6	50	6.7	4.8	16.6	4.2	2.5	3.4
	100	3.4	2.3	4.7	3.7	0.5	0.7		100	8.5	6.8	10.9	5.6	4.4	3.0
	200	3.4	3.8	4.9	3.8	2.2	1.3		200	6.4	5.9	7.6	4.5	3.9	3.7
	400	4.5	4.4	4.1	4.5	3.3	2.4		400	5.3	7.6	2.7	3.8	4.8	3.4
-0.4	50	2.5	3.2	4.9	2.0	0.4	0.3	0.7	50	7.8	4.0	18.5	3.9	2.1	3.2
	100	3.6	3.5	3.9	2.8	1.2	1.0		100	10.4	7.2	17.1	5.8	5.1	2.4
	200	3.0	5.0	3.9	3.2	2.7	1.5		200	7.1	6.6	8.5	5.3	4.7	3.2
	400	5.3	4.9	3.7	5.2	3.5	1.9		400	7.4	6.9	6.5	6.3	5.4	3.5
-0.3	50	2.8	4.2	4.4	1.3	0.7	0.6	0.8	50	9.6	3.1	27.4	4.6	4.3	5.9
	100	4.6	3.0	4.3	3.9	0.9	0.7		100	10.1	7.3	24.3	7.2	6.0	3.2
	200	4.6	4.2	4.6	4.4	1.9	2.1		200	9.4	7.7	18.8	6.8	6.3	2.7
	400	4.2	4.1	4.7	3.6	3.0	3.1		400	7.5	7.4	10.6	5.4	5.1	3.0
-0.2	50	2.6	2.9	3.0	1.5	0.4	0.4	0.9	50	11.9	5.5	31.1	7.0	5.3	6.1
	100	5.1	4.3	3.9	3.6	0.9	1.0		100	10.8	7.4	33.6	7.8	6.0	2.2
	200	4.7	4.6	4.3	4.0	2.7	2.0		200	10.4	9.3	25.5	7.4	6.0	3.0
	400	5.1	4.1	4.6	3.9	3.0	3.2		400	7.6	7.4	18.9	6.0	6.7	2.7
-0.1	50	4.0	3.8	3.9	1.8	0.7	0.8	1.0	50	12.0	4.6	37.8	7.6	6.4	9.5
	100	4.2	4.2	4.7	2.6	1.6	1.5		100	11.6	7.4	39.1	7.4	6.7	4.2

Table 1 (continued)

ξ	n	PWM / $T_{g,n}$			GPWM / $S_{h,n}$			ξ	n	PWM / $T_{g,n}$			GPWM / $S_{h,n}$		
		$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$			$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$
	200	4.7	4.4	4.5	4.1	2.4	2.1		200	12.3	8.5	32.4	9.6	10.1	3.0
	400	4.0	3.7	4.2	3.4	2.4	3.8		400	7.8	5.2	27.5	5.8	6.2	2.8
0.0	50	4.7	3.0	4.0	1.8	0.3	1.3	1.1	50	13.9	4.8	38.8	10.8	7.6	12.2
	100	3.6	3.3	2.8	2.6	1.3	1.5		100	15.3	8.3	44.9	12.0	9.6	5.3
	200	4.5	3.6	3.8	3.6	2.2	1.7		200	9.8	7.4	44.4	7.9	9.5	2.3
	400	4.6	3.4	3.7	4.0	2.8	4.0		400	7.5	6.4	38.0	7.0	8.3	2.4

A similar experiment was carried out for samples of independent block maxima obtained from blocks of size $b \in \{1, 5, 10, 50\}$ from the generalized Pareto distribution (GPD) with location parameter equal to 0, scale parameter equal to 1 and shape parameter equal to ξ (abbreviated as GPD(0,1,ξ) as we continue). The experiment was designed having in mind the fact that the GPD with shape parameter ξ is in the maximum domain of attraction of the GEV with shape parameter ξ. The results for $b = 1$ are reported in Table 2. The results for $b \in \{5, 10, 50\}$ are reported in Section 1.2 of the supplementary material (see Tables 5–7 therein). Another variant of this experiment was carried out by using the absolute value of the standard Student t distribution with $1/\xi$ degrees of freedom instead of the GPD(0,1,ξ). Indeed, the absolute value of the standard Student t distribution with $1/\xi$ degrees of freedom, $\xi > 0$, is in the maximum domain of attraction of the GEV with shape parameter ξ. The results are reported in Section 1.3 of the supplementary material (see Tables 8–11 therein). The conclusions in terms of ξ are overall the same as for the first experiment: the tests based on $T_{g,n}$ behave mostly adequately when $\xi \leq 0.4$ and become too liberal for $\xi \geq 0.5$; the tests based on $S_{h,n}$ are overall too conservative for $\xi \leq 0.3$ and are too liberal for $\xi \geq 0.9$.

Additional experiments under the null (whose results are reported in Table 12 of the supplementary material) were carried out to assess the empirical levels of the tests for samples generated from various normal and exponential distributions. No test turned out to be too liberal.

The next experiments focused on the power of the tests. To ease reading of the forthcoming tables, among the six tests proposed in the paper, the rejection percentages of those that are expected to be sensitive to the alternative under consideration are colored in light gray.

We started by assessing the power of the tests in the case of a change in the shape parameter of the GEV. First, for $t \in \{0.25, 0.5, 0.75\}$, samples of size $n \in \{100, 200\}$ were generated such that the $\lfloor nt \rfloor$ first observations were from a GEV(μ,1,-0.4) and the $n - \lfloor nt \rfloor$ last observations were from a GEV(μ, 1, ξ), for $\xi \in \{-0.2, 0, 0.2, 0.4\}$ and $\mu \in \{-100, 0, 100\}$. The results for $\mu = 0$ are reported in Table 3 while those for $\mu \in \{-100, 100\}$ are given in Tables 13 and 14 of the supplementary material. The column F_n gives the rejection percentages obtained using the

Table 2 Percentage of rejection of H_0 computed from 1000 samples of $n \in \{50, 100, 200, 400\}$ independent block maxima obtained from blocks of size 1 from the GPD(0,1,ξ)

ξ	n	PWM / $T_{g,n}$ $\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	GPWM / $S_{h,n}$ $\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$	ξ	n	PWM / $T_{g,n}$ $\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	GPWM / $S_{h,n}$ $\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$
-1.0	50	3.0	3.9	1.7	1.6	3.0	1.8	0.1	50	6.1	2.9	5.5	3.5	1.5	5.0
	100	3.7	4.3	2.5	2.2	2.8	1.6		100	5.7	3.3	4.3	3.9	2.2	4.2
	200	3.8	4.3	3.1	2.7	3.6	2.9		200	4.9	3.3	2.4	3.8	2.9	3.5
	400	4.2	4.7	3.7	3.9	4.3	3.1		400	5.5	3.5	4.2	4.8	3.7	4.9
-0.9	50	3.4	3.5	2.8	1.5	3.0	1.8	0.2	50	6.6	3.4	5.0	2.6	1.5	3.9
	100	4.9	3.5	3.8	2.8	1.8	2.5		100	5.4	3.6	3.7	3.6	2.4	4.3
	200	4.4	4.4	3.5	3.6	3.7	3.2		200	5.7	4.0	3.1	5.0	4.5	4.8
	400	3.5	4.3	3.5	3.3	3.6	4.2		400	5.0	5.3	3.1	3.8	4.6	4.4
-0.8	50	3.8	1.6	2.3	1.1	1.3	1.7	0.3	50	6.3	3.7	6.3	3.0	1.8	5.2
	100	3.5	2.7	3.0	2.7	1.9	2.0		100	6.5	4.8	5.5	4.4	2.9	4.6
	200	4.4	3.7	3.5	3.5	2.9	3.4		200	5.7	5.0	2.3	4.0	3.4	4.7
	400	4.8	4.2	4.3	4.4	3.5	4.5		400	5.2	4.1	2.8	4.3	3.8	4.5
-0.7	50	4.5	1.9	2.0	1.9	1.0	1.8	0.4	50	9.8	5.1	9.0	5.7	2.9	4.3
	100	5.2	2.5	3.2	3.5	2.0	1.5		100	7.8	6.1	5.0	5.2	3.0	4.1
	200	4.3	3.3	3.8	2.9	2.1	3.8		200	7.3	6.0	3.3	6.1	5.0	3.6
	400	3.4	4.4	5.3	3.4	3.9	4.6		400	5.9	4.3	2.8	5.4	3.9	4.4
-0.6	50	4.1	2.2	2.5	2.0	1.8	2.4	0.5	50	7.8	3.9	11.0	4.3	2.0	3.8
	100	4.5	3.1	3.3	3.5	2.2	3.8		100	7.0	6.8	9.2	5.1	4.9	4.5
	200	4.0	3.2	3.3	2.8	2.9	2.9		200	6.8	5.0	6.2	4.8	3.1	4.3
	400	4.1	4.5	3.4	4.1	3.8	4.0		400	6.5	5.9	3.5	4.3	4.6	3.3
-0.5	50	4.6	1.9	2.3	2.1	1.2	1.8	0.6	50	8.4	4.5	17.0	5.1	2.8	5.5
	100	4.1	2.7	2.4	2.3	1.9	2.5		100	10.1	7.1	11.1	6.7	4.9	4.9
	200	6.3	3.6	5.4	4.5	3.6	3.7		200	9.3	7.6	7.0	6.2	5.3	3.6
	400	5.7	2.9	4.9	5.0	3.0	5.6		400	5.9	6.4	4.8	5.1	5.3	3.8
-0.4	50	5.0	2.1	3.7	1.9	1.0	3.1	0.7	50	9.9	5.7	22.1	6.2	3.9	6.7
	100	4.5	1.7	3.5	2.7	1.5	3.1		100	9.9	7.0	19.8	7.2	5.2	3.5
	200	4.6	3.3	3.3	3.7	2.9	3.4		200	8.0	7.5	11.2	6.6	5.4	3.2
	400	5.2	3.8	3.8	4.6	3.8	3.3		400	7.1	5.5	5.2	6.5	4.4	4.4
-0.3	50	3.3	1.9	3.6	1.3	0.6	1.8	0.8	50	11.4	5.9	27.4	7.1	5.5	6.9
	100	5.1	3.4	3.2	2.9	2.2	3.2		100	13.2	7.1	25.2	8.8	6.6	3.8
	200	5.4	2.7	3.0	3.8	2.8	4.0		200	10.4	7.2	16.7	7.7	7.5	2.5
	400	6.0	3.6	4.1	5.8	3.4	4.1		400	6.3	6.2	12.0	5.1	4.8	3.2
-0.2	50	4.9	1.4	3.5	1.8	0.8	2.1	0.9	50	12.4	5.9	31.0	9.9	7.7	9.4
	100	4.5	2.6	3.3	3.0	1.3	2.8		100	12.4	8.6	32.0	7.9	8.9	5.2
	200	5.2	3.1	5.0	4.0	2.7	4.1		200	10.2	6.8	25.0	8.1	7.2	1.7
	400	4.2	3.9	3.2	4.7	4.1	4.0		400	8.1	7.0	18.5	6.3	5.6	2.4
-0.1	50	6.8	2.8	3.8	4.0	0.9	3.5	1.0	50	12.6	4.9	37.0	8.5	7.0	9.1
	100	5.1	2.2	3.5	3.3	1.8	4.0		100	13.6	6.8	40.4	11.1	8.8	4.8

Table 2 (continued)

		PWM / $T_{g,n}$			GPWM / $S_{h,n}$					PWM / $T_{g,n}$			GPWM / $S_{h,n}$		
ξ	n	$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$	ξ	n	$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$
	200	4.6	3.7	3.0	4.2	3.0	3.8		200	10.9	7.4	35.1	9.7	9.5	2.1
	400	5.5	4.6	5.5	5.3	3.9	4.8		400	8.0	7.4	25.6	6.7	7.3	3.1
0.0	50	6.1	3.5	2.9	2.9	1.8	3.7	1.1	50	16.3	5.6	40.0	11.7	9.2	14.3
	100	5.0	3.1	4.4	3.6	1.5	4.2		100	15.6	8.3	47.3	10.8	11.3	5.7
	200	6.2	3.7	4.3	5.1	4.0	4.1		200	11.2	8.3	41.5	11.0	9.6	3.7
	400	5.7	4.7	4.5	5.1	4.0	5.1		400	8.7	7.7	37.1	8.0	7.4	2.5

Table 3 Percentage of rejection of H_0 computed from 1000 samples of size $n \in \{100, 200\}$ such that the $\lfloor nt \rfloor$ first observations are from a GEV(0,1,-0.4) and the $n - \lfloor nt \rfloor$ last observations are from a GEV(0,1,ξ)

						PWM / $T_{g,n}$			GPWM / $S_{h,n}$		
ξ	n	t	F_n	$\bar{x}_n$	s_n^2	$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$
-0.2	100	0.25	5.7	5.1	4.3	3.2	3.8	10.2	2.3	1.7	1.2
		0.75	6.1	4.9	7.0	4.5	3.4	5.1	3.3	0.9	1.4
		0.50	5.5	6.5	8.1	4.2	3.4	12.5	2.9	1.7	2.0
	200	0.25	5.2	7.0	6.0	3.7	4.0	17.7	3.7	2.2	5.1
		0.75	7.3	8.3	8.5	5.5	3.9	15.3	4.5	1.9	3.5
		0.50	8.5	11.7	10.7	4.9	3.5	29.4	4.7	1.6	6.9
0.0	100	0.25	5.8	7.6	4.5	4.8	4.2	23.1	2.6	1.2	3.2
		0.75	7.6	12.6	13.0	5.3	4.4	11.6	4.2	1.1	4.5
		0.50	7.8	16.8	10.2	4.4	2.6	32.8	3.0	0.8	9.1
	200	0.25	8.3	16.8	7.6	4.5	2.6	54.4	3.8	1.1	22.7
		0.75	8.0	22.2	29.8	5.7	5.2	40.2	3.7	2.1	15.4
		0.50	10.3	32.1	30.7	4.6	3.1	76.1	3.2	1.5	37.9
0.2	100	0.25	8.0	9.7	2.4	6.1	4.7	42.7	3.1	1.4	14.3
		0.75	8.3	23.6	11.7	5.8	5.5	16.0	4.2	1.5	9.2
		0.50	12.0	27.7	10.3	6.8	6.7	52.8	3.8	3.0	26.8
	200	0.25	12.0	24.9	4.9	7.1	4.7	78.2	5.0	2.8	58.2
		0.75	11.5	51.3	39.4	6.8	4.4	59.7	4.2	2.2	44.6
		0.50	22.0	65.9	35.0	6.6	6.4	92.4	4.9	3.2	81.8
0.4	100	0.25	9.6	11.1	0.5	9.6	7.3	62.6	4.3	2.0	27.5
		0.75	9.3	34.4	11.7	6.4	9.5	21.7	3.5	2.5	20.9
		0.50	15.7	40.8	7.2	8.1	10.1	69.5	4.2	2.5	50.3
	200	0.25	15.5	25.1	2.3	11.1	7.0	89.2	4.6	2.9	85.1
		0.75	13.7	71.4	36.3	7.9	7.6	63.9	3.3	4.2	79.2
		0.50	29.6	80.6	26.9	12.3	8.7	95.7	6.3	3.4	97.4

Table 4 Percentage of rejection of H_0 computed from 1000 samples of size $n \in \{100, 200\}$ such that the $\lfloor nt \rfloor$ first observations are from a GEV(0,1,0) and the $n - \lfloor nt \rfloor$ last observations are from a GEV(0,1,ξ)

ξ	n	t	F_n	$\bar{x}_n$	s_n^2	PWM / $T_{g,n}$			GPWM / $S_{h,n}$		
						$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$
0.2	100	0.25	5.7	4.4	1.2	5.8	4.2	6.7	2.8	2.0	4.1
		0.75	6.7	7.4	3.2	4.6	4.6	3.9	3.5	2.2	2.9
		0.50	4.1	5.5	2.2	4.6	4.8	6.7	2.9	1.3	4.3
	200	0.25	8.2	6.7	2.2	5.9	4.4	10.0	4.8	2.4	8.8
		0.75	5.7	9.9	7.8	5.7	4.8	6.0	4.4	3.3	7.2
		0.50	6.1	9.9	7.3	5.2	4.8	14.7	4.6	2.8	15.9
0.4	100	0.25	5.9	4.6	0.4	6.2	4.4	15.5	3.3	2.1	7.0
		0.75	6.1	11.3	2.5	7.0	8.0	5.9	3.8	1.7	6.6
		0.50	6.4	13.1	4.4	6.7	8.0	19.2	2.9	2.2	17.4
	200	0.25	7.1	8.0	1.8	5.4	5.2	28.8	3.9	2.4	27.9
		0.75	7.7	25.9	15.2	6.1	7.2	13.0	4.5	3.2	24.5
		0.50	7.6	30.6	10.3	5.1	6.4	37.9	3.7	2.8	54.2
0.6	100	0.25	6.7	5.1	0.7	8.6	7.9	39.9	4.5	3.2	17.8
		0.75	7.3	19.5	5.2	7.5	13.7	13.1	3.9	3.6	15.0
		0.50	7.7	17.4	1.6	7.4	12.5	36.8	3.8	3.4	35.1
	200	0.25	8.4	9.4	1.3	9.1	8.8	57.6	5.2	3.9	57.7
		0.75	8.3	47.4	21.9	7.1	15.0	24.3	4.0	5.6	56.6
		0.50	15.4	45.7	11.5	10.6	14.7	68.3	5.1	5.7	84.1

nonparametric rank-based test for change-point detection based on empirical c.d.f.s studied in Holmes et al. (2013) and implemented in the function `cpTestFn` of the R package `npcp`. The columns $\bar{x}_n$ and s_n^2 report the empirical powers of CUSUM tests designed to be particularly sensitive to changes in the expectation (see Phillips, 1987, and the references therein) and the variance (see, e.g., Bücher and Kojadinovic, 2016), respectively. The former (resp. latter) is implemented in the function `cpTestMean` (resp. `cpTestU`) of the R package `npcp`. Tables 4 and 5 report similar results but for higher ξ values. Overall, as could have been expected, the tests based on $T_{\tilde{g}_\mu,n}$, $T_{\tilde{g}_\sigma}$, $S_{\tilde{h}_\mu,n}$ and $S_{\tilde{h}_\sigma,n}$ have hardly any power against such alternatives. Roughly speaking, the test based on $T_{\tilde{g}_\xi,n}$ is more (resp. less) powerful than the one based on $S_{\tilde{h}_\xi,n}$ when the largest shape parameter is smaller or equal than 0.2 (resp. greater or equal than 0.4). Both tests are overall more powerful than the general-purpose nonparametric test considered in Holmes et al. (2013) and the two CUSUM tests designed to be particularly sensitive to changes in the expectation and the variance, respectively.

Similar experiments were used to assess the power of the tests when the scale (resp. location) parameter increases from 0.5 to 1 (resp. from 0 to 0.5). The corresponding rejection percentages are reported in Tables 6 and 7, respectively

Table 5 Percentage of rejection of H_0 computed from 1000 samples of size $n \in \{100, 200\}$ such that the $\lfloor nt \rfloor$ first observations are from a GEV(0,1,0.2) and the $n - \lfloor nt \rfloor$ last observations are from a GEV(0,1,ξ)

						PWM / $T_{g,n}$			GPWM / $S_{h,n}$		
ξ	n	t	F_n	$\bar{x}_n$	s_n^2	$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$
0.4	100	0.25	5.7	2.8	0.5	6.8	4.7	6.1	4.0	1.2	4.5
		0.75	6.1	4.3	1.1	6.1	5.5	3.7	4.3	2.3	3.6
		0.50	5.9	6.5	0.3	6.7	6.0	6.6	3.8	1.7	5.0
	200	0.25	5.2	4.4	0.8	5.5	4.8	7.8	4.0	3.2	8.9
		0.75	6.0	7.9	4.5	5.9	6.5	3.4	4.8	4.1	8.7
		0.50	5.4	8.9	2.9	4.3	5.0	9.7	3.2	2.3	15.9
0.6	100	0.25	6.7	2.4	0.2	8.1	6.3	22.6	4.5	2.9	10.7
		0.75	7.3	8.1	2.3	8.1	9.4	7.0	5.3	3.0	6.7
		0.50	7.6	7.8	1.3	9.7	9.4	18.4	4.6	3.5	16.3
	200	0.25	6.4	6.1	0.5	6.9	7.6	27.7	4.8	4.6	24.5
		0.75	7.0	21.0	7.4	6.7	10.4	9.4	4.5	5.7	23.5
		0.50	7.5	19.2	4.5	7.7	10.0	29.5	4.0	3.7	48.9
0.8	100	0.25	6.0	3.1	0.3	10.7	8.4	49.5	4.9	4.1	16.5
		0.75	5.8	13.9	2.5	8.8	17.0	13.6	4.4	4.1	11.1
		0.50	8.3	13.1	1.0	12.9	18.4	43.9	5.3	5.8	28.3
	200	0.25	9.2	5.6	0.8	13.4	11.3	58.6	6.5	5.6	46.4
		0.75	7.4	38.3	11.2	9.7	23.0	23.6	4.6	7.1	49.1
		0.50	14.2	30.6	4.0	15.8	24.7	60.5	6.9	8.5	75.3

(additional results are available in Sections 2.2 and 2.3 of the supplementary material). For changes in the scale parameter, we see that the test based on $T_{\tilde{g}_\sigma,n}$ is more (resp. less) powerful than the test based on $S_{\tilde{h}_\sigma,n}$ for $\xi \in \{-0.4, 0\}$ (resp. $\xi \in \{0.4, 0.8\}$), the two tests becoming more powerful than the test "s_n^2" for $\xi \geq 0$. Again, as expected, the remaining PWM or GPWM tests have little power against such alternatives. For changes in the location parameter, the test based on $T_{\tilde{g}_\mu,n}$ is more (resp. less) powerful than the test based on $S_{\tilde{h}_\mu,n}$ for $\xi \in \{-0.4, 0\}$ (resp. $\xi = 0.8$). This time the rank-based test of Holmes et al. (2013) is more powerful than the two aforementioned tests, although the difference is small for $\xi \in \{0, 0.4, 0.8\}$.

In a last experiment, we investigated the power of the tests for samples generated from normal distributions with different variances or expectations. The results are reported in Tables 23 and 24 of the supplementary material. In the case of a change in variance (resp. expectation), the proposed tests designed to be sensitive to a change in scale (resp. location) appear overall slightly less powerful than the "s_n^2" (resp. "$\bar{x}_n$") tests.

Table 6 Percentage of rejection of H_0 computed from 1000 samples of size $n \in \{100, 200\}$ such that the $\lfloor nt \rfloor$ first observations are from a GEV(0,0.5,ξ) and the $n - \lfloor nt \rfloor$ last observations are from a GEV(0,1,ξ)

ξ	n	t	F_n	$\bar{x}_n$	s_n	PWM / $T_{g,n}$			GPWM / $S_{h,n}$		
						$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$
-0.4	100	0.25	17.0	7.3	54.3	4.5	62.2	9.8	4.8	38.1	1.8
		0.75	20.1	13.4	94.5	13.5	82.3	2.6	13.9	58.9	0.9
		0.50	40.7	12.4	98.1	8.5	93.6	5.7	9.8	81.9	1.0
	200	0.25	50.0	8.4	97.8	5.5	97.4	11.8	4.2	93.7	3.5
		0.75	44.7	19.6	99.9	15.5	98.7	2.8	15.6	96.5	0.4
		0.50	85.6	24.4	100.0	10.5	100.0	5.1	10.9	100.0	0.9
0.0	100	0.25	15.0	9.0	18.3	6.8	57.6	6.7	5.8	39.1	2.1
		0.75	13.7	17.7	59.6	12.5	78.6	3.1	10.3	50.6	0.7
		0.50	26.0	20.7	64.7	11.0	91.7	5.5	9.9	79.1	0.5
	200	0.25	28.1	15.7	50.6	6.3	95.6	7.5	4.8	89.9	2.7
		0.75	29.5	33.1	91.0	15.0	97.9	1.7	13.9	95.8	0.7
		0.50	63.1	41.6	94.6	10.8	99.9	4.9	10.5	99.8	0.7
0.4	100	0.25	8.2	3.9	0.6	6.9	25.2	7.1	5.0	22.0	2.6
		0.75	10.9	11.0	5.2	14.3	58.1	6.4	12.6	42.3	1.6
		0.50	17.6	14.5	5.9	12.2	70.3	7.9	10.7	63.6	1.6
	200	0.25	18.8	9.9	2.7	7.3	62.4	6.4	6.0	69.9	5.4
		0.75	20.4	22.6	11.8	12.9	85.9	3.6	13.2	85.5	1.3
		0.50	47.5	29.7	12.2	11.3	94.5	6.1	10.7	95.2	2.2
0.8	100	0.25	7.3	2.7	0.0	9.4	11.3	25.6	6.6	14.1	3.9
		0.75	10.1	5.0	0.6	18.9	34.1	23.4	14.0	36.1	2.9
		0.50	12.7	4.4	0.3	13.1	32.6	23.0	10.4	45.8	2.8
	200	0.25	16.8	3.0	0.6	8.7	23.8	21.6	6.7	40.4	3.6
		0.75	13.2	7.8	1.0	17.3	55.8	14.6	14.9	70.6	1.2
		0.50	36.8	9.5	0.9	15.7	61.5	18.6	13.4	79.9	2.1

5 Illustrations

The results of the finite-sample experiments partially reported in the previous section suggest that the overall behavior of the tests based on the statistics $T_{g,n}$ in (2.7) is better than that of the tests based on the statistics $S_{h,n}$ in (3.2) when the distributions of the block maxima have finite variances. When they have an infinite variance but a finite expectation, the tests based on the statistics $S_{h,n}$ should however clearly be preferred as those based on $T_{g,n}$ do not hold their level anymore.

When dealing with climate data, the assumption of finite variances is reasonable for temperatures, precipitation or winds, which is why we focus on the tests based on the statistics $T_{g,n}$ in the forthcoming applications.

Table 7 Percentage of rejection of H_0 computed from 1000 samples of size $n \in \{100, 200\}$ such that the $\lfloor nt \rfloor$ first observations are from a GEV(0,1,ξ) and the $n - \lfloor nt \rfloor$ last observations are from a GEV(0.5,1,ξ)

ξ	n	t	F_n	$\bar{x}_n$	s_n	PWM / $T_{g,n}$			GPWM / $S_{h,n}$		
						$\tilde{g}_\mu$	$\tilde{g}_\sigma$	$\tilde{g}_\xi$	$\tilde{h}_\mu$	$\tilde{h}_\sigma$	$\tilde{h}_\xi$
-0.4	100	0.25	39.3	39.2	2.2	27.4	3.9	3.9	27.6	1.1	0.5
		0.75	36.5	38.4	3.9	27.9	3.4	3.8	19.3	1.0	0.5
		0.50	58.0	61.9	2.1	48.0	3.5	4.1	42.7	0.6	0.5
	200	0.25	67.6	72.1	3.6	57.7	3.2	4.5	53.9	1.6	1.5
		0.75	68.6	73.3	4.1	57.1	5.1	4.2	47.8	2.5	1.0
		0.50	88.6	91.0	1.2	79.3	3.0	3.7	74.1	1.4	2.0
0.0	100	0.25	28.2	19.7	3.7	26.5	3.4	2.6	23.8	0.9	1.1
		0.75	28.2	23.3	2.2	29.0	4.7	4.2	21.2	1.8	2.0
		0.50	48.6	39.0	1.9	48.3	2.3	3.4	42.1	0.7	1.7
	200	0.25	52.8	42.2	4.0	54.6	4.1	2.5	52.8	2.3	1.8
		0.75	51.7	42.4	1.7	54.9	4.5	4.0	48.1	3.6	2.7
		0.50	76.7	66.6	3.5	78.1	3.7	3.1	76.1	2.4	2.3
0.4	100	0.25	23.1	4.7	0.5	22.1	4.7	4.7	19.8	2.0	2.3
		0.75	22.8	7.8	0.5	27.2	5.2	4.3	22.0	2.8	3.1
		0.50	43.5	10.6	0.3	45.2	3.8	3.6	42.0	1.4	3.0
	200	0.25	48.8	8.5	1.2	48.4	5.0	3.2	51.9	3.0	2.7
		0.75	48.6	8.5	1.0	54.1	5.6	2.8	49.9	3.9	4.6
		0.50	74.0	16.8	0.9	71.6	4.7	2.3	76.6	2.8	3.1
0.8	100	0.25	22.9	1.2	0.2	21.7	7.1	21.0	18.5	4.9	2.5
		0.75	24.2	2.5	0.2	27.9	6.7	19.7	26.5	5.3	4.0
		0.50	46.2	2.9	0.2	38.9	6.8	23.2	41.1	4.9	1.3
	200	0.25	48.6	2.0	0.4	40.0	7.9	15.4	45.6	6.9	2.2
		0.75	52.0	3.3	0.4	46.4	7.9	17.5	51.8	6.9	3.9
		0.50	77.5	3.4	0.4	64.7	7.8	13.6	71.6	5.1	2.0

To illustrate the pratical use of the three selected tests, we applied them on some of the environmental datasets available in the R packages evd (Stephenson 2002), extRemes (Gilleland and Katz 2011), ismev (Heffernan and Stephenson 2014) and ClusterMax (Bernard et al. 2013). The observations consist of block maxima of wind speeds, temperatures, precipitation or sea levels, and can reasonably be considered to be independent. The interest reader can easily plot the data after installing the aforementioned R packages.

The main practical difficulty that prevented a direct application of the tests is that all the datasets contain ties due to a limited precision of the underlying measurement instruments. To address this issue, we proceeded as follows. For every dataset, we computed the smallest absolute difference d between any two distinct observations. Assuming that d is a good estimate of the precision of the measurements, we added a random component uniformly distributed over $(0, d)$ to every observation. For every

Table 8 Minimum and maximum of $\tilde{g}_\mu(\hat{b}_{1:n})$, $\tilde{g}_\sigma(\hat{b}_{1:n})$ and $\tilde{g}_\xi(\hat{b}_{1:n})$ as well as minimum and maximum of the p-values of the tests based on $T_{g,n}$ in (2.7) with g respectively equal to $\tilde{g}_\mu$, $\tilde{g}_\sigma$, $\tilde{g}_\xi$ for 1000 different de-tied samples generated from each dataset as explained in Section 5

Package	Dataset	n	n'	Estimates						p-values of the test based on $T_{g,n}$ with					
				$\tilde{g}_\mu(\hat{\beta}_{v,1:n})$		$\tilde{g}_\sigma(\hat{\beta}_{v,1:n})$		$\tilde{g}_\xi(\hat{\beta}_{v,1:n})$		$g=\tilde{g}_\mu$		$g=\tilde{g}_\sigma$		$g=\tilde{g}_\xi$	
				min	max	min	max	min	max	min	max	min	max	min	max
evd	lisbon	30	21	95.79	96.22	12.62	13.07	-0.16	-0.13	0.152	0.205	0.167	0.271	0.416	0.630
evd	oxford	80	19	84.23	84.46	4.20	4.44	-0.34	-0.26	0.099	0.248	0.534	1.000	0.413	1.000
extRemes	HEAT (Tmax)	43	9	112.89	113.22	1.91	2.32	-0.41	-0.21	0.002	0.029	0.590	1.000	0.766	1.000
extRemes	HEAT (-Tmin)	43	15	-70.70	-70.37	3.67	4.07	-0.18	-0.09	0.000	0.002	0.130	0.529	0.173	0.898
extRemes	ftcanmax	100	78	135.75	135.96	55.51	55.79	0.13	0.13	0.724	0.757	0.430	0.467	1.000	1.000
ismev	fremantle	86	31	1.49	1.49	0.14	0.14	-0.21	-0.19	0.006	0.009	0.433	0.578	1.000	1.000
ismev	portpirie	65	42	3.88	3.88	0.20	0.20	-0.06	-0.04	0.537	0.603	0.788	0.949	0.782	0.928
ClusterMax	Station 8 (St-Lizier)	228	58	1.97	1.98	1.95	1.97	0.14	0.15	0.217	0.261	0.296	0.378	0.068	0.091
ClusterMax	Station 14 (Aurillac)	228	73	2.39	2.40	2.87	2.89	0.29	0.29	0.152	0.174	0.622	0.683	0.124	0.138
ClusterMax	Station 57 (Nevers)	228	56	1.82	1.84	1.90	1.92	0.18	0.18	0.287	0.349	0.155	0.202	0.034	0.046
ClusterMax	Station 78 (Niort)	228	59	1.96	1.97	2.15	2.17	0.22	0.22	0.038	0.047	0.016	0.022	0.448	0.523

The first two columns provide the names of the R packages containing the datasets and the names of the latter. The columns n and n' report the number of block maxima and the number of distinct block maxima, respectively

dataset, 1000 different de-tied samples were generated with the hope that at least one of these is close to the true unobserved sample, and the tests were applied to each such de-tied sample. The minimum and the maximum of the obtained p-values are reported in Table 8. The minimum and the maximum of the estimates $\tilde{g}_\mu(\hat{b}_{\nu,1:n})$, $\tilde{g}_\sigma(\hat{b}_{\nu,1:n})$ and $\tilde{g}_\xi(\hat{b}_{\nu,1:n})$ are provided as well. Under H_0 in (1.2) with F the c.d.f. of the GEV, the three latter quantities are estimates of the corresponding location, scale and shape parameters obtained using the PWM method summarized in Section 2.1.

Small maximal p-values provide evidence against H_0 in (1.2). Under the additional assumption of GEV distributed block maxima, the latter can be interpreted as evidence of change in the location, scale or shape parameter. Ignoring a necessary adjustment of the p-values due to multiple testing, the following conclusions could for instance be drawn from Table 8:

- there is very strong evidence of a change in the location parameter for the annual minimal temperature in the HEAT dataset, and strong evidence of the same nature for the annual maximal temperature in the HEAT dataset and for the annual maximal sea levels in the fremantle dataset,
- there is weak (resp. some) evidence of change in the shape parameter for the weekly maxima of hourly precipitation measured at Saint-Lizier (resp. Nevers), France, during the Fall season between 1993 and 2011,
- there is strong evidence of a change in the scale parameter and some evidence of a change in the location parameter for the weekly maxima of hourly precipitation measured at Niort, France, during the Fall season between 1993 and 2011.

Notice that the tests, as implemented in R package npcp, are fast. Obtaining the results reported in Table 8 took less than 30 seconds on one 2.7 GHz processor.

Let us end this section with a practical remark. To test H_0 in (1.2) from a single sample of block maxima, three tests, based on the statistics $T_{g,n}$ in (2.7) or on the statistics $S_{h,n}$ in (3.2), are to be used. As noted by a referee, such an approach, although interesting from a data analysis perspective, lacks a simple decision rule regarding the rejection or not of H_0. The later could be the subject of future research. A simple yet conservative approach would be to use a Bonferroni correction and reject H_0 as soon one p-value is below a third of the chosen significance level.

Appendix A: Proof of Proposition 2.1

Proof Some thought reveals that

$$\beta_i = \mathrm{E}[X\{F(X)\}^{i-1}] = i^{-1}\mathrm{E}\{\max(X_1, \ldots, X_i)\}, \qquad i \in \{1, 2, 3\},$$

where $X_1, \ldots, X_i$ are independent copies of X.

The first inequality is obtained by noticing that $\max(X_1, X_2) > (X_1 + X_2)/2$ almost surely since $P(X_1 = X_2) = 0$ by continuity of F, and by taking expectations on both sides. The second inequality follows from the fact that $3\max(X_1, X_2, X_3) > \max(X_1, X_2) + \max(X_1, X_3) + \max(X_2, X_3)$ almost surely. The last inequality is a consequence of the fact that

$$2\max(X_1, X_2) + 2\max(X_1, X_3) + 2\max(X_2, X_3) > 3\max(X_1, X_2, X_3) + X_1 + X_2 + X_3$$

almost surely. $\qquad\square$

Appendix B: Weak convergence of the sequential weighted uniform empirical process

The purpose of this appendix is to derive the weak convergence of a sequential version of the weighted uniform empirical process. The latter result, that may be of independent interest, is used in the proof of Proposition C.2, itself necessary for the proof, given in Appendix D, of Proposition 2.3. The forthcoming developments rely heavily on Appendix G of Genest and Segers (2009).

Let $U_1, \ldots, U_n$ be an i.i.d. sample from the standard uniform distribution. For any integers $1 \leq k \leq l \leq n$, let $G_{k:l}$ be the empirical c.d.f. computed from the sample $U_k, \ldots, U_l$ with the convention that $G_{k:l} = 0$, for all $k > l$. Then, define the *two-sided sequential uniform empirical process* by

$$\mathbb{B}_n(s, t, u) = \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} \{\mathbf{1}(U_i \leq u) - u\} = \sqrt{n}\lambda_n(s, t)\{G_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(u) - u\},$$

(B.1)

for $(s, t, u) \in \Delta \times [0, 1]$, where $\Delta = \{(s, t) \in [0, 1]^2 : s \leq t\}$ and $\lambda_n(s, t) = (\lfloor nt \rfloor - \lfloor ns \rfloor)/n$, $(s, t) \in \Delta$. The process $\mathbb{B}_n(0, \cdot, \cdot)$ is the standard sequential uniform empirical process and it is well-known (see, e.g., van der Vaart and Wellner, 2000, Theorem 2.12.1) that $\mathbb{B}_n(0, \cdot, \cdot) \rightsquigarrow \mathbb{B}(0, \cdot, \cdot)$ in $\ell^\infty([0, 1]^2)$, where $\mathbb{B}(0, \cdot, \cdot)$ is a centered Gaussian process with continuous trajectories, sometimes called a *Kieffer–Müller process*, whose covariance function is given by

$$\operatorname{Cov}\{\mathbb{B}(0, s, u), \mathbb{B}(0, t, v)\} = (s \wedge t)(u \wedge v - uv), \qquad (s, t, u, v) \in [0, 1]^4.$$

Since, for any $(s, t) \in \Delta$, $\mathbb{B}_n(s, t, \cdot) = \mathbb{B}_n(0, t, \cdot) - \mathbb{B}_n(0, s, \cdot)$, we immediately obtain from the continuous mapping that $\mathbb{B}_n \rightsquigarrow \mathbb{B}$ in $\ell^\infty(\Delta \times [0, 1])$, where, for any $(s, t) \in \Delta$, $\mathbb{B}(s, t, \cdot) = \mathbb{B}(0, t, \cdot) - \mathbb{B}(0, s, \cdot)$.

Analogously to Genest and Segers (2009), let $\alpha \geq 0$ and consider a weighted version of the processes $\mathbb{B}_n$ defined as

$$\mathbb{W}_{\alpha,n}(s, t, u) = \begin{cases} \frac{\mathbb{B}_n(s,t,u)}{\{u(1-u)\}^\alpha}, & \text{if } u \in (0, 1), \\ 0, & \text{if } u \in \{0, 1\}. \end{cases}$$

(B.2)

The following result is then a univariate sequential extension of Theorem G.1 of Genest and Segers (2009).

Proposition B.1 *For any $\alpha \in [0, 1/2)$, $\mathbb{W}_{\alpha,n} \rightsquigarrow \mathbb{W}_\alpha$ in $\ell^\infty(\Delta \times [0, 1])$, where $\mathbb{W}_\alpha$ is a centered Gaussian process with continuous trajectories such that $\mathbb{W}_\alpha(0, \cdot, u) = 0$ for $u \in \{0, 1\}$,*

$$\mathrm{Cov}\{\mathbb{W}_\alpha(0, s, u), \mathbb{W}_\alpha(0, t, v)\} = \frac{(s \wedge t)(u \wedge v - uv)}{\{u(1-u)v(1-v)\}^\alpha}, \qquad (s, t, u, v) \in [0, 1]^2 \times (0, 1)^2,$$

and $\mathbb{W}_\alpha(s, t, \cdot) = \mathbb{W}_\alpha(0, t, \cdot) - \mathbb{W}_\alpha(0, s, \cdot)$ for all $(s, t) \in \Delta$.

Proof Fix $\alpha \in [0, 1/2)$. Analogously to Genest and Segers (2009, Appendix G), for any $u \in (0, 1)$, let $f_u : (0, 1) \to \mathbb{R}$ be defined as $f_u(x) = \{u(1-u)\}^{-\alpha}\{\mathbf{1}(x \le u) - u\}$, $x \in (0, 1)$, and let $\mathcal{F} = \{f_u : u \in (0, 1)\} \cup \{0\}$, where 0 is the function vanishing everywhere on $(0, 1)$. Furthermore, let P be the uniform probability distribution on $(0, 1)$. Lemma G.2 of Genest and Segers (2009) then states that the collection $\mathcal{F}$ is P-Donsker (see, e.g., van der Vaart and Wellner, 2000; Kosorok, 2008), which is equivalent to saying that there exists a P-Brownian bridge $\mathbb{G}$ such that $\mathbb{G}_n \rightsquigarrow \mathbb{G}$ in $\ell^\infty(\mathcal{F})$, where, for any $f \in \mathcal{F}$, $\mathbb{G}_n(f) = \sqrt{n}(\mathbb{P}_n f - Pf)$ with

$$Pf = \int_{[0,1]} f(x)\mathrm{d}x, \qquad \mathbb{P}_n f = \frac{1}{n}\sum_{i=1}^n f(U_i),$$

and the convention that $\mathbb{P}_0 f = 0$. For any $s \in [0, 1]$ and $f \in \mathcal{F}$, let $\mathbb{Z}_n(s, f) = \sqrt{n}\lambda_n(0, s)(\mathbb{P}_{\lfloor ns \rfloor} f - Pf) = \sqrt{n}\lambda_n(0, s)\mathbb{P}_{\lfloor ns \rfloor} f$, where the second equality follows from the fact that $Pf = 0$ for all $f \in \mathcal{F}$. An immediate consequence of Theorem 2.12.1 of Van der Vaart and Wellner (2000) is then that $\mathbb{Z}_n \rightsquigarrow \mathbb{Z}$ in $\ell^\infty([0, 1] \times \mathcal{F})$, where $\mathbb{Z}$ is a tight centered Gaussian process with covariance function

$$\mathrm{Cov}\{\mathbb{Z}(s, f), \mathbb{Z}(t, g)\} = (s \wedge t)(Pfg - PfPg) = (s \wedge t)Pfg, \qquad s, t \in [0, 1], f, g \in \mathcal{F}. \tag{B.3}$$

To conclude the proof, we proceed analogously to Genest and Segers (2009, Appendix G). Let $\phi : [0, 1] \to \mathcal{F}$ be defined as

$$\phi(u) = \begin{cases} f_u, & \text{if } u \in (0, 1), \\ 0, & \text{if } u \in \{0, 1\}. \end{cases}$$

Then, let $T : \ell^\infty([0, 1] \times \mathcal{F}) \to \ell^\infty([0, 1]^2)$ be defined as $T(z)(s, u) = z(s, \phi(u))$ for all $z \in \ell^\infty([0, 1] \times \mathcal{F})$ and $(s, u) \in [0, 1]^2$, and notice that $T(\mathbb{Z}_n) = \mathbb{W}_{\alpha,n}(0, \cdot, \cdot)$. Indeed, for any $s \in [0, 1]$, if $u \in (0, 1)$,

$$T(\mathbb{Z}_n)(s, u) = \mathbb{Z}_n(s, f_u) = \sqrt{n}\lambda_n(0, s)\mathbb{P}_{\lfloor ns \rfloor} f_u = \mathbb{W}_{\alpha,n}(0, s, u),$$

and, if $u \in \{0, 1\}$,

$$T(\mathbb{Z}_n)(s, u) = \mathbb{Z}_n(s, 0) = \sqrt{n}\lambda_n(0, s)\mathbb{P}_{\lfloor ns \rfloor} 0 = \mathbb{W}_{\alpha,n}(0, s, u).$$

The map T being linear and bounded, it is continuous. Hence, by the continuous mapping theorem, $T(\mathbb{Z}_n) \rightsquigarrow T(\mathbb{Z})$ in $\ell^\infty([0, 1]^2)$. Let $\mathbb{W}_\alpha(0, \cdot, \cdot) = T(\mathbb{Z})$. Given the definition of T, since $\mathbb{Z}$ is centered, so is $\mathbb{W}_\alpha(0, \cdot, \cdot)$, and since the finite-dimensional

distributions of $\mathbb{Z}$ are Gaussian, so are those of $\mathbb{W}_\alpha(0, \cdot, \cdot)$. Furthermore, the covariance function of $\mathbb{W}_\alpha(0, \cdot, \cdot)$ follows immediately from (B.3). It remains to verify that the trajectories of $\mathbb{W}_\alpha(0, \cdot, \cdot)$ are continuous. From (Van der Vaart and Wellner 2000, Chapter 2.12), we know that the trajectories of $\mathbb{Z}$ are almost surely uniformly continuous with respect to the natural semimetric $|s - t| + \rho(f, g)$ on $[0, 1] \times \mathcal{F}$, where

$$\rho^2(f, g) = \mathrm{E}[\{\mathbb{G}(f) - \mathbb{G}(g)\}^2], \qquad f, g \in \mathcal{F}.$$

From Genest and Segers (2009, page 3020), we have that the map ϕ is continuous with respect to the Euclidean metric on $[0, 1]$ and the standard deviation semimetric ρ on $\mathcal{F}$. As a consequence, the trajectories of $(s, u) \mapsto \mathbb{W}_\alpha(0, s, u) = \mathbb{Z}(s, \phi(u))$ are almost surely continuous on $[0, 1]^2$.

Finally, since $\mathbb{W}_{\alpha,n}(s, t, \cdot) = \mathbb{W}_{\alpha,n}(0, t, \cdot) - \mathbb{W}_{\alpha,n}(0, s, \cdot)$ for all $(s, t) \in \Delta$, we immediately obtain from the continuous mapping theorem that $\mathbb{W}_{\alpha,n} \rightsquigarrow \mathbb{W}_\alpha$ in $\ell^\infty(\Delta \times [0, 1])$, where, for any $(s, t) \in \Delta$, $\mathbb{W}_\alpha(s, t, \cdot) = \mathbb{W}_\alpha(0, t, \cdot) - \mathbb{W}_\alpha(0, s, \cdot)$. $\qquad\square$

Appendix C: Weak convergence of sequential PWM processes

In order to study the limiting null distribution of the test statistics $S_{g,n}$ in (2.6) and $T_{g,n}$ in (2.7), it is necessary to investigate the asymptotics of sequential empirical processes constructed from PWM estimators. The main result of this section is used in the proof, given in Appendix D, of Proposition 2.3, but it might also be of independent interest.

As we consider a slightly more general theoretical framework than in the main sections of the paper, a richer notation is necessary. Let X be a random variable with c.d.f. F and let $\nu_1, \ldots, \nu_p$ be p functions on $[0, 1]$. The reals $\beta_\nu = \mathrm{E}[X\nu\{F(X)\}]$, $\nu \in \{\nu_1, \ldots, \nu_p\}$, assuming they exist, are called *generalized* PWM following Diebolt et al. (2008). When the functions ν_i are defined as $\nu_i(x) = x^{r_i}(1 - x)^{s_i}$, $x \in [0, 1]$, for some integers $r_i \geq 0$ and $s_i \geq 0$, one recovers a very usefull subset of "classical" PWM that can be used to fit a large number of distributions (see, e.g., Greenwood et al. 1979, Hosking et al. 1985; Hosking and Wallis 1987). In the sequel, we write $\nu = (\nu_1, \ldots, \nu_p)$ and $\beta_\nu = (\beta_{\nu_1}, \ldots, \beta_{\nu_p})$.

Assuming that we have at hand a random sample $X_1, \ldots, X_n$ from an unknown c.d.f. F, a natural nonparametric estimator of β_ν based on a subsample $X_k, \ldots, X_l$, $1 \leq k \leq l \leq n$, is

$$\hat{\beta}_{\nu,k:l} = (\hat{\beta}_{\nu_1,k:l}, \ldots, \hat{\beta}_{\nu_p,k:l}), \tag{C.1}$$

where $\hat{\beta}_{\nu_i,k:l}$ stands for $\hat{\beta}_{i,k:l}$ defined in (2.3) with the difference that γ in (2.2) is set to zero.

For any $x \geq 0$, let $\lfloor x \rfloor$ be the largest integer smaller or equal than x. Furthermore, recall that $\Delta = \{(s, t) \in [0, 1]^2 : s \leq t\}$ and that $\lambda_n(s, t) = (\lfloor nt \rfloor - \lfloor ns \rfloor)/n$, $(s, t) \in \Delta$. The aim of this section is to establish the weak convergence of the vector of empirical processes $\mathbf{M}_{\nu,n} = \left(\mathbb{M}_{\nu_1,n}, \ldots, \mathbb{M}_{\nu_p,n}\right)$, where, for any $\nu \in \{\nu_1, \ldots, \nu_p\}$,

$$\mathbb{M}_{\nu,n}(s, t) = \sqrt{n}\lambda_n(s, t)(\hat{\beta}_{\nu,\lfloor ns \rfloor+1:\lfloor nt \rfloor} - \beta_\nu), \qquad (s, t) \in \Delta. \tag{C.2}$$

The desired result is established under the following condition.

Condition C.1 The functions $\nu_1, \ldots, \nu_p$ are continuously differentiable on $[0, 1]$. Furthermore, $X_1, \ldots, X_n$ are i.i.d. from a continuous distribution with c.d.f. F such that:

(i) $\beta_{\nu_i} = \mathrm{E}[X_1 \nu_i \{F(X_1)\}] < \infty$ for all $i \in \{1, \ldots, p\}$,
(ii) $\mathrm{E}([X_1 \nu_i \{F(X_1)\} - \beta_{\nu_i}][X_1 \nu_j \{F(X_1)\} - \beta_{\nu_j}]) < \infty$ for all $i, j \in \{1, \ldots, p\}$,
(iii) there exists $\alpha \in [0, 1/2]$ such that $\sup_{x \in \mathbb{R}} |H_\alpha(x)| < \infty$, where

$$H_\alpha(x) = x[F(x)\{1 - F(x)\}]^\alpha, \qquad x \in \mathbb{R}.$$

It is easy to verify that if one of the functions $\nu_1, \ldots, \nu_p$ is constant on $[0, 1]$, then (i) and (ii) in Condition C.1 are equivalent to requiring that $\mathrm{E}(X_1^2) < \infty$.

Proposition C.2 (Limiting null distribution of $\mathbf{M}_{\nu,n}$) *Under Condition C.1,*

$$\mathbf{M}_{\nu,n} = \left(\mathbb{M}_{\nu_1,n}, \ldots, \mathbb{M}_{\nu_p,n}\right) \rightsquigarrow \mathbf{M}_\nu = \left(\mathbb{M}_{\nu_1}, \ldots, \mathbb{M}_{\nu_p}\right)$$

in $\{\ell^\infty(\Delta)\}^p$, where $\mathbb{M}_{\nu_1,n}, \ldots, \mathbb{M}_{\nu_p,n}$ are defined in (C.2) and $\mathbb{M}_{\nu_1}, \ldots, \mathbb{M}_{\nu_p}$ are centered Gaussian processes with continuous trajectories such that

$$\mathbb{M}_{\nu_i}(s, t) = \mathbb{M}_{\nu_i}(0, t) - \mathbb{M}_{\nu_i}(0, s), \qquad (s, t) \in \Delta, \ i \in \{1, \ldots, p\}, \qquad (\mathrm{C.3})$$

and

$$\mathrm{Cov}\{\mathbb{M}_{\nu_i}(0, s), \mathbb{M}_{\nu_j}(0, t)\} = (s \wedge t)\mathrm{Cov}(Y_{\nu_i}, Y_{\nu_j}), \qquad (s, t) \in \Delta, \ i, j \in \{1, \ldots, p\}, \qquad (\mathrm{C.4})$$

where, for any $\nu \in \{\nu_1, \ldots, \nu_p\}$, Y_ν is defined in (2.9).

Proof Fix $\nu \in \{\nu_1, \ldots, \nu_p\}$ and consider the decomposition

$$\mathbb{M}_{\nu,n}(s, t) = \mathbb{X}_{\nu,n}(s, t) + \mathbb{Y}_{\nu,n}(s, t), \qquad (s, t) \in \Delta,$$

where

$$\mathbb{X}_{\nu,n}(s, t) = \sqrt{n}\lambda_n(s, t)(\hat{\beta}_{\nu, \lfloor ns \rfloor+1:\lfloor nt \rfloor} - \tilde{\beta}_{\nu, \lfloor ns \rfloor+1:\lfloor nt \rfloor}), \qquad (\mathrm{C.5})$$
$$\mathbb{Y}_{\nu,n}(s, t) = \sqrt{n}\lambda_n(s, t)(\tilde{\beta}_{\nu, \lfloor ns \rfloor+1:\lfloor nt \rfloor} - \beta_\nu),$$

with

$$\tilde{\beta}_{\nu, \lfloor ns \rfloor+1:\lfloor nt \rfloor} = \frac{1}{\lfloor nt \rfloor - \lfloor ns \rfloor} \sum_{i=\lfloor ns \rfloor+1}^{\lfloor nt \rfloor} X_i \nu\{F(X_i)\}. \qquad (\mathrm{C.6})$$

In addition, let $U_1, \ldots, U_n$ be the unobservable sample obtained from $X_1, \ldots, X_n$ by the probability integral transformations $U_i = F(X_i)$, $i \in \{1, \ldots, n\}$, and recall

the definition of the empirical process $\mathbb{W}_{\alpha,n}$ in (B.2) and its weak limit $\mathbb{W}_\alpha$ given in Proposition B.1. The first step of the proof is to show that

$$(\mathbb{Y}_{\nu_1,n}, \ldots, \mathbb{Y}_{\nu_p,n}, \mathbb{W}_{\alpha,n}) \rightsquigarrow (\mathbb{Y}_{\nu_1}, \ldots, \mathbb{Y}_{\nu_p}, \mathbb{W}_\alpha) \qquad (C.7)$$

in $\{\ell^\infty(\Delta)\}^p \times \ell^\infty(\Delta \times [0,1])$.

The proof of the convergence in distribution of the finite-dimensional distributions of $(\mathbb{Y}_{\nu_1,n}, \ldots, \mathbb{Y}_{\nu_p,n}, \mathbb{W}_{\alpha,n})$ to those of $(\mathbb{Y}_{\nu_1}, \ldots, \mathbb{Y}_{\nu_p}, \mathbb{W}_\alpha)$ is omitted for the sake of brevity. To prove (C.7), it remains thus to show joint asymptotic tightness, which is implied by marginal asymptotic tightness. From Proposition B.1, we have that $\mathbb{W}_{\alpha,n} \rightsquigarrow \mathbb{W}_\alpha$ in $\ell^\infty(\Delta \times [0,1])$. Next, fix $\nu \in \{\nu_1, \ldots, \nu_p\}$ and notice that, since $\sigma_\nu^2 = \mathrm{Var}[X_1 \nu\{F(X_1)\}] < \infty$, we have, from the functional central limit theorem, that $\mathbb{Y}_{\nu,n}(0, \cdot) \rightsquigarrow \mathbb{Y}_\nu(0, \cdot)$ in $\ell^\infty([0,1])$, where $\mathbb{Y}_\nu(0, \cdot) = \sigma_\nu \mathbb{U}$ with $\mathbb{U}$ a standard Brownian motion. Since, for any $(s,t) \in \Delta$, $\mathbb{Y}_{\nu,n}(s,t) = \mathbb{Y}_{\nu,n}(0,t) - \mathbb{Y}_{\nu,n}(0,s)$, an immediate consequence of the continuous mapping theorem is that $\mathbb{Y}_{\nu,n} \rightsquigarrow \mathbb{Y}_\nu$ in $\ell^\infty(\Delta)$, where, for any $(s,t) \in \Delta$, $\mathbb{Y}_\nu(s,t) = \sigma_\nu\{\mathbb{U}(t) - \mathbb{U}(s)\}$. This completes the proof of (C.7).

Fix again $\nu \in \{\nu_1, \ldots, \nu_p\}$. Since $\sup_{x \in \mathbb{R}} |H_\alpha(x)| < \infty$, $\sup_{u \in [0,1]} |\nu'(u)| < \infty$ and F is continuous, the map $\psi_{\nu,\alpha} : \ell^\infty(\Delta \times [0,1]) \to \ell^\infty(\Delta)$ defined as

$$\psi_{\nu,\alpha}(f)(s,t) = \int_\mathbb{R} \nu'\{F(x)\} H_\alpha(x) f\{s, t, F(x)\} \mathrm{d}F(x), \qquad (s,t) \in \Delta, \qquad f \in \ell^\infty(\Delta \times [0,1]),$$
$$(C.8)$$

is continuous. The weak convergence in (C.7) and the continuous mapping theorem then imply that

$$\left(\mathbb{Y}_{\nu_1,n} + \psi_{\nu_1,\alpha}(\mathbb{W}_{\alpha,n}), \ldots, \mathbb{Y}_{\nu_p,n} + \psi_{\nu_p,\alpha}(\mathbb{W}_{\alpha,n}) \right) \rightsquigarrow \left(\mathbb{Y}_{\nu_1} + \psi_{\nu_1,\alpha}(\mathbb{W}_\alpha), \ldots, \mathbb{Y}_{\nu_p} + \psi_{\nu_p,\alpha}(\mathbb{W}_\alpha) \right)$$

in $\{\ell^\infty(\Delta)\}^p$. To complete the proof, it remains thus to show that, for any $\nu \in \{\nu_1, \ldots, \nu_p\}$,

$$\sup_{(s,t) \in \Delta} |\mathbb{X}_{\nu,n}(s,t) - \psi_{\nu,\alpha}(\mathbb{W}_{\alpha,n})(s,t)\}| = o_P(1), \qquad (C.9)$$

where $\mathbb{X}_{\nu,n}$ is defined in (C.5). Fix $\nu \in \{\nu_1, \ldots, \nu_p\}$ and consider the empirical process

$$\mathbb{X}'_{\nu,n}(s,t) = \int_\mathbb{R} \nu'\{F(x)\} H_\alpha(x) \mathbb{W}_{\alpha,n}\{s, t, F(x)\} \mathrm{d}F_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(x), \qquad (s,t) \in \Delta,$$
$$(C.10)$$

where $F_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}$ is defined in (2.2) with $\gamma = 0$. To prove (C.9), we shall first show that

$$\sup_{(s,t) \in \Delta} |\mathbb{X}_{\nu,n}(s,t) - \mathbb{X}'_{\nu,n}(s,t)| = o_P(1), \qquad (C.11)$$

and then that

$$\sup_{(s,t) \in \Delta} |\mathbb{X}'_{\nu,n}(s,t) - \psi_{\nu,\alpha}(\mathbb{W}_{\alpha,n})(s,t)| = o_P(1). \qquad (C.12)$$

$\square$

Proof of (C.11). On one hand, starting from (C.5), we have that, for any $(s, t) \in \Delta$,

$$\mathbb{X}_{\nu,n}(s, t) = \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} X_i \left[\nu\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(X_i)\} - \nu\{F(X_i)\} \right]$$

$$= \int_{\mathbb{R}} x \sqrt{n} \lambda_n(s, t) \left[\nu\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x)\} - \nu\{F(x)\} \right] \mathrm{d}F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x).$$

On the other hand, (C.10) can be rewritten as

$$\mathbb{X}'_{\nu,n}(s, t) = \int_{\mathbb{R}} x \nu'\{F(x)\} \sqrt{n} \lambda_n(s, t)\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x) - F(x)\} \mathrm{d}F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x), \quad (s, t) \in \Delta.$$

Then, the supremum on the left of (C.11) is smaller than

$$\sup_{(s,t) \in \Delta} \sqrt{n} \lambda_n(s, t) \int_{\mathbb{R}} |x| \left| \nu\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x)\} - \nu\{F(x)\} \right.$$

$$\left. - \nu'\{F(x)\}\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x) - F(x)\} \right| \mathrm{d}F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x). \quad (C.13)$$

By the mean value theorem, for any $x \in \mathbb{R}$, there exists $U^*_{n,s,t,x} \in (0, 1)$ such that

$$\nu\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x)\} - \nu\{F(x)\} = \nu'[F(x) + U^*_{n,s,t,x}\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x) - F(x)\}]$$

$$\times \{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x) - F(x)\}.$$

Hence, (C.13) is smaller than

$$\sup_{(s,t) \in \Delta} \int_{\mathbb{R}} |x| \left| \nu'[F(x) + U^*_{n,s,t,x}\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x) - F(x)\}] - \nu'\{F(x)\} \right|$$

$$\times |\sqrt{n} \lambda_n(s, t)\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x) - F(x)\}| \mathrm{d}F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x),$$

that is,

$$\sup_{(s,t) \in \Delta} \int_{\mathbb{R}} |H_\alpha(x)| \left| \nu'[F(x) + U^*_{n,s,t,x}\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x) - F(x)\}] - \nu'\{F(x)\} \right|$$

$$\times |\mathbb{W}_{\alpha,n}\{s, t, F(x)\}| \mathrm{d}F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x).$$

Therefore,

$$\sup_{(s,t) \in \Delta} |\mathbb{X}_{\nu,n}(s, t) - \mathbb{X}'_{\nu,n}(s, t)|$$

$$\leq \sup_{(s,t,x) \in \Delta \times \mathbb{R}} \left| \nu'[F(x) + U^*_{n,s,t,x}\{F_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(x) - F(x)\}] - \nu'\{F(x)\} \right|$$

$$\times \sup_{(s,t,x) \in \Delta \times \mathbb{R}} |\mathbb{W}_{\alpha,n}\{s, t, F(x)\}| \times \sup_{x \in \mathbb{R}} |H_\alpha(x)|. \quad (C.14)$$

Fix $\varepsilon, \eta > 0$. From the previous inequality, since $\sup_{u \in [0,1]} |\nu'(u)| < \infty$, $\sup_{x \in \mathbb{R}} |H_\alpha(x)| < \infty$, and using the fact that $\mathbb{W}_{\alpha,n}(s, t, \cdot) = 0$ when $s = t$ and

the asymptotic uniform equicontinuity in probability of $\mathbb{W}_{\alpha,n}$ (a consequence of Proposition B.1), there exists $\delta \in (0, 1)$ such that, for all sufficiently large n,

$$
\mathrm{P}\left(\sup_{\substack{(s,t)\in\Delta \\ t-s<\delta}} |\mathbb{X}_{v,n}(s, t) - \mathbb{X}'_{v,n}(s, t)| > \varepsilon \right)
$$

$$
\leq \mathrm{P}\left(2 \sup_{u\in[0,1]} |v'(u)| \times \sup_{x\in\mathbb{R}} |H_\alpha(x)| \times \sup_{\substack{(s,t,u)\in\Delta\times[0,1] \\ t-s<\delta}} |\mathbb{W}_{\alpha,n}(s, t, u)| > \varepsilon \right) < \eta/2.
$$

To prove (C.11), it remains therefore to show that, for all sufficiently large n,

$$
\mathrm{P}\left(\sup_{\substack{(s,t)\in\Delta \\ t-s\geq\delta}} |\mathbb{X}_{v,n}(s, t) - \mathbb{X}'_{v,n}(s, t)| > \varepsilon \right) < \eta/2.
$$

Let $\Delta^\delta = \{(s, t) \in \Delta : t - s \geq \delta\}$. To show the latter, it suffices to show that

$$
\sup_{(s,t)\in\Delta^\delta} |\mathbb{X}_{v,n}(s, t) - \mathbb{X}'_{v,n}(s, t)| = o_\mathrm{P}(1).
$$

The fact that $\mathbb{B}_n \rightsquigarrow \mathbb{B}$ in $\ell^\infty(\Delta \times [0, 1])$, where $\mathbb{B}_n$ is defined in (B.1), implies that

$$
\sup_{(s,t,u)\in\Delta^\delta\times[0,1]} |G_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(u) - u| = \sup_{(s,t,x)\in\Delta^\delta\times\mathbb{R}} |F_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(x) - F(x)| = o_\mathrm{P}(1),
$$

$$(\text{C.15})$$

since $G_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}\{F(x)\} = F_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(x)$ for all $(s, t, x) \in \Delta \times \mathbb{R}$. The previous display then implies that

$$
\sup_{(s,t,x)\in\Delta^\delta\times\mathbb{R}} \left| v'[F(x) + U^*_{n,s,t,x}\{F_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(x) - F(x)\}] - v'\{F(x)\} \right| = o_\mathrm{P}(1)
$$

since v' is (uniformly) continuous on $[0, 1]$. The latter combined with (C.14) in which the suprema are restricted to $(s, t) \in \Delta^\delta$ and the fact that $\sup_{(s,t,x)\in\Delta\times\mathbb{R}} |\mathbb{W}_{\alpha,n}\{s, t, F(x)\}| = O_\mathrm{P}(1)$ implies finally (C.11). $\qquad\square$

Proof of (C.12). We have

$$
\sup_{(s,t)\in\Delta} |\mathbb{X}'_{v,n}(s, t) - \psi_{v,\alpha}(\mathbb{W}_{\alpha,n})(s, t) \}| \leq \sup_{(s,t)\in\Delta} |\mathbb{X}'_{v,n}(s, t)| + \sup_{(s,t)\in\Delta} |\psi_{v,\alpha}(\mathbb{W}_{\alpha,n})(s, t)|
$$

$$
\leq 2 \sup_{u\in[0,1]} |v'(u)| \times \sup_{x\in\mathbb{R}} |H_\alpha(x)| \times \sup_{(s,t,u)\in\Delta\times[0,1]} |\mathbb{W}_{\alpha,n}(s, t, u)|.
$$

Fix $\varepsilon, \eta > 0$. Proceeding as for the proof of (C.12), from the previous inequality, the fact that $\mathbb{W}_{\alpha,n}(s, t, \cdot) = 0$ when $s = t$ and the asymptotic uniform equicontinuity in probability of $\mathbb{W}_{\alpha,n}$, there exists $\delta \in (0, 1)$ such that, for all sufficiently large n,

$$
\mathrm{P}\left(\sup_{\substack{(s,t)\in\Delta \\ t-s<\delta}} |\mathbb{X}'_{v,n}(s, t) - \psi_{v,\alpha}(\mathbb{W}_{\alpha,n})(s, t)\}| > \varepsilon \right) < \eta/2,
$$

and thus, to prove (C.12), it remains to show that

$$\sup_{(s,t)\in\Delta^\delta} |\mathbb{X}'_{v,n}(s,t) - \psi_{v,\alpha}(\mathbb{W}_{\alpha,n})(s,t)| = o_{\mathrm{P}}(1), \tag{C.16}$$

where $\Delta^\delta = \{(s,t) \in \Delta : t - s \geq \delta\}$. Proving the previous display is equivalent to showing that $\mathbb{X}'_{v,n} - \psi_{v,\alpha}(\mathbb{W}_{\alpha,n})$ converges weakly to $(s,t) \mapsto 0$ in $\ell^\infty(\Delta^\delta)$. Fix $(s,t) \in \Delta^\delta$ and let $G(u) = u$ for all $u \in [0,1]$. From (C.15), $(\mathbb{W}_{\alpha,n}(s,t,\cdot), G_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}) \rightsquigarrow (\mathbb{W}_\alpha(s,t,\cdot), G)$ in $\{\ell^\infty([0,1])\}^2$, which implies that

$$\left(x \mapsto v'\{F(x)\}H_\alpha(x)\mathbb{W}_{\alpha,n}\{s,t,F(x)\}, F_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}\right)$$
$$\rightsquigarrow \left(x \mapsto v'\{F(x)\}H_\alpha(x)\mathbb{W}_\alpha\{s,t,F(x)\}, F\right)$$

in $\{\ell^\infty(\overline{\mathbb{R}})\}^2$, where $\overline{\mathbb{R}} = [-\infty,\infty]$. The assumptions of the proposition ensure that $x \mapsto v'\{F(x)\}H_\alpha(x)$ is a continuous function on $\overline{\mathbb{R}}$ and therefore that the process

$$x \mapsto v'\{F(x)\}H_\alpha(x)\mathbb{W}_\alpha\{s,t,F(x)\} \in \ell^\infty(\overline{\mathbb{R}})$$

has continuous trajectories. Combining the latter fact with Lemma 3 of Holmes et al. (2013) and the continuity of the map $\psi_{v,\alpha}$ defined in (C.8), we obtain, from the continuous mapping theorem, that $\mathbb{X}'_{v,n}(s,t) - \psi_{v,\alpha}(\mathbb{W}_{\alpha,n})(s,t) \rightsquigarrow 0$. The latter weak convergence being equivalent to convergence in probability to zero, we have that the finite-dimensional distributions of $\mathbb{X}'_{v,n} - \psi_{v,\alpha}(\mathbb{W}_{\alpha,n})$ converge weakly to those of $(s,t) \mapsto 0$ in $\ell^\infty(\Delta^\delta)$.

It remains to show asymptotic tightness. To do so, we shall prove the $\|\cdot\|_1$-asymptotic uniform equicontinuity in probability of $\mathbb{X}'_{v,n} - \psi_{v,\alpha}(\mathbb{W}_{\alpha,n})$ in $\ell^\infty(\Delta^\delta)$. By the triangle inequality, it suffices to show the latter for $\mathbb{X}'_{v,n}$ only as $\psi_{v,\alpha}(\mathbb{W}_{\alpha,n})$ converges weakly to $\psi_{v,\alpha}(\mathbb{W}_\alpha)$ in $\ell^\infty(\Delta^\delta)$, the limit having continuous trajectories. By Problem 2.1.5 in Van der Vaart and Wellner (2000), it remains thus to prove that, for any positive sequence $a_n \downarrow 0$,

$$\sup_{\substack{(s,t),(s',t')\in\Delta^\delta \\ |s-s'|+|t-t'|\leq a_n}} |\mathbb{X}'_{v,n}(s,t) - \mathbb{X}'_{v,n}(s',t')| = o_{\mathrm{P}}(1). \tag{C.17}$$

The supremum on the left of the previous display is smaller than $A_n + B_n$, where

$$A_n = \sup_{\substack{(s,t),(s',t')\in\Delta^\delta \\ |s-s'|+|t-t'|\leq a_n}} \left| \int_{\mathbb{R}} v'\{F(x)\}H_\alpha(x)\mathbb{W}_{\alpha,n}\{s,t,F(x)\}\mathrm{d}F_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(x) \right.$$
$$\left. - \int_{\mathbb{R}} v'\{F(x)\}H_\alpha(x)\mathbb{W}_{\alpha,n}\{s',t',F(x)\}\mathrm{d}F_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(x) \right|$$

and

$$B_n = \sup_{\substack{(s,t),(s',t')\in\Delta^\delta \\ |s-s'|+|t-t'|\leq a_n}} \left| \int_{\mathbb{R}} v'\{F(x)\}H_\alpha(x)\mathbb{W}_{\alpha,n}\{s',t',F(x)\}\mathrm{d}F_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(x) \right.$$
$$\left. - \int_{\mathbb{R}} v'\{F(x)\}H_\alpha(x)\mathbb{W}_{\alpha,n}\{s',t',F(x)\}\mathrm{d}F_{\lfloor ns' \rfloor+1:\lfloor nt' \rfloor}(x) \right|.$$

Concerning A_n, we have

$$A_n \leq \sup_{u \in [0,1]} |v'(u)| \times \sup_{x \in \mathbb{R}} |H_\alpha(x)| \times \sup_{\substack{(s,t),(s',t') \in \Delta^\delta, u \in [0,1] \\ |s-s'|+|t-t'| \leq a_n}} |\mathbb{W}_{\alpha,n}(s,t,u) - \mathbb{W}_{\alpha,n}(s',t',u)| = o_P(1).$$

The term B_n is smaller than

$$\sup_{\substack{(s,t),(s',t') \in \Delta^\delta \\ |s-s'|+|t-t'| \leq a_n}} \left| \frac{1}{\lfloor nt \rfloor - \lfloor ns \rfloor} \left\{ \sum_{i=\lfloor ns \rfloor+1}^{\lfloor nt \rfloor} v'\{F(X_i)\}H_\alpha(X_i)\mathbb{W}_{\alpha,n}\{s',t',F(X_i)\} \right. \right.$$

$$\left. \left. - \sum_{i=\lfloor ns' \rfloor+1}^{\lfloor nt' \rfloor} v'\{F(X_i)\}H_\alpha(X_i)\mathbb{W}_{\alpha,n}\{s',t',F(X_i)\} \right\} \right| +$$

$$\sup_{\substack{(s,t),(s',t') \in \Delta^\delta \\ |s-s'|+|t-t'| \leq a_n}} \left| \left(\frac{1}{\lfloor nt \rfloor - \lfloor ns \rfloor} - \frac{1}{\lfloor nt' \rfloor - \lfloor ns' \rfloor} \right) \sum_{i=\lfloor ns' \rfloor+1}^{\lfloor nt' \rfloor} v'\{F(X_i)\}H_\alpha(X_i)\mathbb{W}_{\alpha,n}\{s',t',F(X_i)\} \right|,$$

which is smaller than

$$2 \times \sup_{\substack{(s,t),(s',t') \in \Delta^\delta \\ |s-s'|+|t-t'| \leq a_n}} \frac{|\lfloor nt \rfloor - \lfloor nt' \rfloor| + |\lfloor ns \rfloor - \lfloor ns' \rfloor|}{\lfloor nt \rfloor - \lfloor ns \rfloor} \times \sup_{u \in [0,1]} |v'(u)|$$

$$\times \sup_{x \in \mathbb{R}} |H_\alpha(x)| \times \sup_{(s,t,u) \in \Delta \times [0,1]} |\mathbb{W}_{\alpha,n}(s,t,u)| = o_P(1).$$

Hence, (C.17) and thus (C.16) hold, which implies (C.12).

Finally, (C.3) is a consequence of the fact that, for any $v \in \{v_1, \ldots, v_p\}$,

$$\mathbb{M}_v(s,t) = \mathbb{Y}_v(s,t) + \int_{\mathbb{R}} v'\{F(x)\}H_\alpha(x)\mathbb{W}_\alpha\{s,t,F(x)\}\mathrm{d}F(x)$$

and the analogue properties for $\mathbb{Y}_v$ and $\mathbb{W}_\alpha$, while (C.4) follows from standard calculations. $\square$

Appendix D: Proof of Proposition 2.3

Proposition 2.3 is a corollary of a slightly more general result that we shall first prove using Proposition C.2 shown in Appendix C. To be consistent with Appendix C, we keep using the richer notation introduced therein. Let $v_1, \ldots, v_p$ be p functions on $[0,1]$, let g be a function from $\mathbb{R}^p$ to $\mathbb{R}$ defined on $\mathcal{D}$ containing $\beta_v = (\beta_{v_1}, \ldots, \beta_{v_p})$, and consider the generic statistic

$$S_{v,g,n} = \max_{1 \leq k \leq n-1} \frac{k(n-k)}{n^{3/2}} \mathbf{1}(\hat{\beta}_{v,1:k} \in \mathcal{D}, \hat{\beta}_{v,k+1:n} \in \mathcal{D}) \left| g(\hat{\beta}_{v,1:k}) - g(\hat{\beta}_{v,k+1:n}) \right|,$$

where $\hat{\beta}_{v,1:k}$ and $\hat{\beta}_{v,k+1:n}$ are defined analogously to (C.1). When $p = 3$, $v_1(x) = 1$, $v_2(x) = x$, $v_3(x) = x^2$, $x \in [0, 1]$, $\mathcal{D} = \mathcal{D}_\xi$ in (2.5) and $g \in \{g_\mu, g_\sigma, g_\xi\}$, $S_{v,g,n}$ is $S_{g,n}$ in (2.6).

Proposition D.1 (Limiting null distribution of $S_{v,g,n}$) *If Condition C.1 holds,* $\mathrm{E}(|X_1|) < \infty$, $\mathcal{D}$ *is an open convex subset of* $\mathbb{R}^p$ *containing* β_v *and* g *is continuously differentiable on* $\mathcal{D}$, $S_{v,g,n}$ *converges in distribution to* $\sigma_{v,g} \sup_{s \in [0,1]} |\mathbb{U}(s) - s\mathbb{U}(1)|$, *where* $\mathbb{U}$ *is a standard Brownian motion on* $[0, 1]$ *and*

$$\sigma_{v,g}^2 = \sum_{i,j=1}^{p} \partial_i g(\beta_v) \partial_j g(\beta_v) \mathrm{Cov}(Y_{v_i}, Y_{v_j}),$$

with Y_{v_i} *defined in* (2.9).

Proof For any $s \in [0, 1]$, let

$$\mathbb{D}_{v,g,n}(s) = \sqrt{n}\lambda_n(0, s)\lambda_n(s, 1)\mathbf{1}(\hat{\beta}_{v,1:\lfloor ns \rfloor} \in \mathcal{D}, \hat{\beta}_{v,\lfloor ns \rfloor+1:n} \in \mathcal{D})\{g(\hat{\beta}_{v,1:\lfloor ns \rfloor}) - g(\hat{\beta}_{v,\lfloor ns \rfloor+1:n})\},$$

and notice that

$$S_{v,g,n} = \sup_{s \in [0,1]} |\mathbb{D}_{v,g,n}(s)|. \tag{D.1}$$

Furthermore, let

$$\mathbb{M}_{v,g,n}(s, t) = \sqrt{n}\lambda_n(s, t)\{g(\hat{\beta}_{v,\lfloor ns \rfloor+1:\lfloor nt \rfloor}) - g(\beta_v)\}, \qquad (s, t) \in \Delta.$$

Under H_0 in (1.2) implied by Condition C.1, it is easy to verify that

$$\mathbb{D}_{v,g,n}(s) = \lambda_n(s, 1)\mathbf{1}(\hat{\beta}_{v,1:\lfloor ns \rfloor} \in \mathcal{D}, \hat{\beta}_{v,\lfloor ns \rfloor+1:n} \in \mathcal{D})\mathbb{M}_{v,g,n}(0, s)$$
$$- \lambda_n(0, s)\mathbf{1}(\hat{\beta}_{v,1:\lfloor ns \rfloor} \in \mathcal{D}, \hat{\beta}_{v,\lfloor ns \rfloor+1:n} \in \mathcal{D})\mathbb{M}_{v,g,n}(s, 1), \qquad s \in [0, 1]. \quad \text{(D.2)}$$

To show the desired result, we shall first prove that

$$A_n = \sup_{s \in [0,1]} \mathbf{1}(\hat{\beta}_{v,1:\lfloor ns \rfloor} \in \mathcal{D}, \hat{\beta}_{v,\lfloor ns \rfloor+1:n} \in \mathcal{D}) \left| \mathbb{M}_{v,g,n}(0, s) - \sum_{i=1}^{p} \partial_i g(\beta_v)\mathbb{M}_{v_i,n}(0, s) \right| = o_\mathrm{P}(1) \tag{D.3}$$

and that

$$B_n = \sup_{s \in [0,1]} \mathbf{1}(\hat{\beta}_{v,1:\lfloor ns \rfloor} \in \mathcal{D}, \hat{\beta}_{v,\lfloor ns \rfloor+1:n} \in \mathcal{D}) \left| \mathbb{M}_{v,g,n}(s, 1) - \sum_{i=1}^{p} \partial_i g(\beta_v)\mathbb{M}_{v_i,n}(s, 1) \right| = o_\mathrm{P}(1), \tag{D.4}$$

where $\mathbb{M}_{v_1,n}, \ldots, \mathbb{M}_{v_p,n}$ are defined in (C.2).

Let us start with (D.3). Using the mean value theorem when the indicators are equal to one, A_n is smaller than $\sum_{i=1}^{p} A_{i,n}$, where

$$A_{i,n} = \sup_{s \in [0,1]} \mathbf{1}(\hat{\beta}_{v,1:\lfloor ns \rfloor} \in \mathcal{D}) |\partial_i g(\hat{\beta}^*_{v,1:\lfloor ns \rfloor}) - \partial_i g(\beta_v)| |\mathbb{M}_{v_i,n}(0,s)|, \tag{D.5}$$

and where $\hat{\beta}^*_{v,1:\lfloor ns \rfloor}$ lies on the segment between β_v and $\hat{\beta}_{v,1:\lfloor ns \rfloor}$. To show (D.3), it suffices therefore to prove that $A_{i,n} = o_{\mathrm{P}}(1)$ for all $i \in \{1, \ldots, p\}$.

Fix $v \in \{v_1, \ldots, v_p\}$. Since $\beta_v < \infty$, from the strong law of large number, we have that $\tilde{\beta}_{v,1:n} \xrightarrow{\text{a.s.}} \beta_v$, where $\tilde{\beta}_{v,1:n}$ is defined analogously to (C.6). Thus,

$$\left| \tilde{\beta}_{v,1:n} - \hat{\beta}_{v,1:n} \right| \leq \sup_{x \in \mathbb{R}} |v\{F_{1:n}(x)\} - v\{F(x)\}| \times \frac{1}{n} \sum_{i=1}^{n} |X_i| \xrightarrow{\text{a.s.}} 0$$

since $\mathrm{E}(|X_1|) < \infty$, and the supremum on the right converges to 0 almost surely as a consequence of the Glivenko-Cantelli lemma and the continuous mapping theorem. It follows that $\hat{\beta}_{v,1:n} \xrightarrow{\text{a.s.}} \beta_v$ for all $v \in \{v_1, \ldots, v_p\}$, and therefore that $\hat{\beta}_{v,1:n} \xrightarrow{\text{a.s.}} \beta_v$.

Fix $i \in \{1, \ldots, p\}$. The continuous mapping theorem next implies that $\partial_i g(\hat{\beta}_{v,1:n}) \xrightarrow{\text{a.s.}} \partial_i g(\beta_v)$, which in turn implies that $\sup_{k \geq 1} \mathbf{1}(\hat{\beta}_{v,1:k} \in \mathcal{D}) |\partial_i g(\hat{\beta}_{v,1:k}) - \partial_i g(\beta_v)| < \infty$ almost surely.

Let $\eta > 0$. Then, there exists $M > 0$ such that

$$\mathrm{P} \left\{ \sup_{k \geq 1} \mathbf{1}(\hat{\beta}_{v,1:k} \in \mathcal{D}) \left| \partial_i g(\hat{\beta}_{v,1:k}) - \partial_i g(\beta_v) \right| > M \right\} \leq \eta/4. \tag{D.6}$$

Furthermore, let $\varepsilon > 0$ and consider $\varepsilon/M > 0$. Using the asymptotic uniform equicontinuity in probability of $\mathbb{M}_{v_i,n}$ (a consequence of Proposition C.2), there exists $\delta \in (0, 1)$ such that, for all sufficiently large n,

$$\mathrm{P} \left\{ \sup_{s \in [0,\delta)} |\mathbb{M}_{v_i,n}(0,s)| > \varepsilon/M \right\} \leq \eta/4. \tag{D.7}$$

Then,

$$\mathrm{P}(A_{i,n} > \varepsilon) \leq \mathrm{P} \left\{ \sup_{s \in [0,\delta)} \mathbf{1}(\hat{\beta}_{v,1:\lfloor ns \rfloor} \in \mathcal{D}) |\partial_i g(\hat{\beta}^*_{v,1:\lfloor ns \rfloor}) - \partial_i g(\beta_v)| |\mathbb{M}_{v_i,n}(0,s)| > \varepsilon \right\}$$

$$+ \mathrm{P} \left\{ \sup_{s \in [\delta,1]} \mathbf{1}(\hat{\beta}_{v,1:\lfloor ns \rfloor} \in \mathcal{D}) |\partial_i g(\hat{\beta}^*_{v,1:\lfloor ns \rfloor}) - \partial_i g(\beta_v)| |\mathbb{M}_{v_i,n}(0,s)|\} > \varepsilon \right\}. \tag{D.8}$$

The first term on the right is smaller, for all sufficiently large n, than

$$\mathrm{P} \left\{ \sup_{k \geq 1} \mathbf{1}(\hat{\beta}_{v,1:k} \in \mathcal{D}) \left| \partial_i g(\hat{\beta}_{v,1:k}) - \partial_i g(\beta_v) \right| > M \right\} + \mathrm{P} \left\{ \sup_{s \in [0,\delta)} |\mathbb{M}_{v_i,n}(0,s)| > \varepsilon/M \right\} \leq \eta/2,$$

where the last inequality follows by (D.6) and (D.7). The second term on the right of (D.8) is smaller, for all sufficiently large n, than $\eta/2$. Indeed,

$$\sup_{s\in[\delta,1]} \mathbf{1}(\hat{\beta}_{\nu,1:\lfloor ns\rfloor} \in \mathcal{D})|\partial_i g(\hat{\beta}^*_{\nu,1:\lfloor ns\rfloor}) - \partial_i g(\beta_\nu)||\mathbb{M}_{\nu_i,n}(0,s)|$$

$$\leq \sup_{s\in[\delta,1]} |\partial_i g(\hat{\beta}^*_{\nu,1:\lfloor ns\rfloor}) - \partial_i g(\beta_\nu)| \times \sup_{s\in[0,1]} |\mathbb{M}_{\nu_i,n}(0,s)| = o_P(1)$$

since the second supremum is $O_P(1)$ from the weak convergence of $\mathbb{M}_{\nu_i,n}$, while the first supremum is $o_P(1)$ as a consequence of the fact that, for any $i \in \{1,\dots,p\}$,

$$\sup_{s\in[\delta,1]} |\hat{\beta}_{\nu_i,1:\lfloor ns\rfloor} - \beta_{\nu_i}| = n^{-1/2} \sup_{s\in[\delta,1]} \lambda_n(0,s)^{-1} \times \sup_{s\in[0,1]} |\mathbb{M}_{\nu_i,n}(0,s)| = o_P(1)$$

and the continuous mapping theorem. Hence, $A_{i,n}$ defined in (D.5) converges in probability to zero for all $i \in \{1,\dots,p\}$, which implies that (D.3) holds. The fact that (D.4) holds follows immediately since B_n in (D.4) (written as a maximum) is nothing else than A_n in (D.3) computed from the sequence $X_n,\dots,X_1$. Therefore, for any $\varepsilon > 0$, $P(B_n > \varepsilon) = P(A_n > \varepsilon) \to 0$ by (D.3).

Next, let

$$S'_{\nu,g,n} = \sup_{s\in[0,1]} \mathbf{1}(\hat{\beta}_{\nu,1:\lfloor ns\rfloor} \in \mathcal{D}, \hat{\beta}_{\nu,\lfloor ns\rfloor+1:n} \in \mathcal{D})|\mathbb{D}'_{\nu,g,n}(s)|,$$

where

$$\mathbb{D}'_{\nu,g,n}(s) = \sum_{i=1}^p \partial_i g(\beta_\nu)\{\lambda_n(s,1)\mathbb{M}_{\nu_i,n}(0,s) - \lambda_n(0,s)\mathbb{M}_{\nu_i,n}(s,1)\}, \qquad s \in [0,1]. \tag{D.9}$$

Then, starting from (D.1) and (D.2), and using the reverse triangle inequality, we obtain that

$$\left| S_{\nu,g,n} - S'_{\nu,g,n} \right| \leq A_n + B_n = o_P(1), \tag{D.10}$$

where A_n and B_n are defined in (D.3) and (D.4).

We shall now verify that

$$S'_{\nu,g,n} - S'''_{\nu,g,n} = o_P(1), \tag{D.11}$$

where

$$S'''_{\nu,g,n} = \sup_{s\in[0,1]} |\mathbb{D}'_{\nu,g,n}(s)|. \tag{D.12}$$

To do so, we use the decomposition $S'_{\nu,g,n} - S'''_{\nu,g,n} = S'_{\nu,g,n} - S''_{\nu,g,n} + S''_{\nu,g,n} - S'''_{\nu,g,n}$, where

$$S''_{\nu,g,n} = \sup_{s\in[0,1]} \mathbf{1}(\hat{\beta}_{\nu,\lfloor ns\rfloor+1:n} \in \mathcal{D})|\mathbb{D}'_{\nu,g,n}(s)|.$$

Since $\hat{\beta}_{\nu,1:k} \xrightarrow{\text{a.s.}} \beta_\nu$ as $k \to \infty$, we have that, for almost every ω, there exists k_ω such that $k \geq k_\omega$ implies that $\hat{\beta}_{\nu,1:k}(\omega) \in \mathcal{D}$. Hence, for almost every ω,

$$S'_{\nu,g,n}(\omega) - S''_{\nu,g,n}(\omega) = \sup_{s\in[k_\omega/n,1]} \mathbf{1}\{\hat{\beta}_{\nu,\lfloor ns\rfloor+1:n}(\omega) \in \mathcal{D}\}|\mathbb{D}'_{\nu,g,n}(s,\omega)|$$

$$- \sup_{s\in[0,1]} \mathbf{1}\{\hat{\beta}_{\nu,\lfloor ns\rfloor+1:n}(\omega) \in \mathcal{D}\}|\mathbb{D}'_{\nu,g,n}(s,\omega)| \to 0,$$

which implies that $S'_{\nu,g,n} - S''_{\nu,g,n} = o_P(1)$. Similarly, we obtain that

$$\sup_{s \in [0,1]} \mathbf{1}(\hat{\beta}_{\nu,1:\lfloor ns \rfloor} \in \mathcal{D})|\mathbb{D}'_{\nu,g,n}(s)| - S'''_{\nu,g,n} = o_P(1).$$

Then, using the fact that $S''_{\nu,g,n} - S'''_{\nu,g,n}$ is $\sup_{s \in [0,1]} \mathbf{1}(\hat{\beta}_{\nu,1:\lfloor ns \rfloor} \in \mathcal{D})|\mathbb{D}'_{\nu,g,n}(s)| - S'''_{\nu,g,n}$ computed from the sequence $X_n, \ldots, X_1$, we obtain that, for every $\varepsilon > 0$,

$$P(|S''_{\nu,g,n} - S'''_{\nu,g,n}| > \varepsilon) = P\left\{ \left| \sup_{s \in [0,1]} \mathbf{1}(\hat{\beta}_{\nu,1:\lfloor ns \rfloor} \in \mathcal{D})|\mathbb{D}'_{\nu,g,n}(s)| - S'''_{\nu,g,n} \right| > \varepsilon \right\} \to 0.$$

Therefore, (D.11) holds. Next, from Proposition C.2 and the continuous mapping theorem, we have that $\mathbb{D}'_{\nu,g,n} \rightsquigarrow \mathbb{D}'_{\nu,g}$ in $\ell^\infty([0,1])$, where $\mathbb{D}'_{\nu,g,n}$ is defined in (D.9),

$$\mathbb{D}'_{\nu,g}(s) = \sum_{i=1}^{p} \partial_i g(\beta_\nu)\{(1-s)\mathbb{M}_{\nu_i}(0,s) - s\mathbb{M}_{\nu_i}(s,1)\} = \sum_{i=1}^{p} \partial_i g(\beta_\nu)\{\mathbb{M}_{\nu_i}(0,s) - s\mathbb{M}_{\nu_i}(0,1)\},$$

for $s \in [0,1]$, and $\mathbb{M}_{\nu_1}, \ldots, \mathbb{M}_{\nu_p}$ are defined in Proposition C.2. The desired convergence in distribution of $S_{\nu,g,n}$ finally follows from (D.10), (D.11), (D.12) and standard calculations. $\qquad\qquad\square$

Proof of Proposition 2.3 The convergence in distribution of $S_{g,n}$ is an immediate consequence of Proposition D.1. It remains to prove that $S_{g,n} - T_{g,n} = o_P(1)$. The latter is a consequence of the fact that, for any $i \in \{1, 2, 3\}$,

$$I_n = \sup_{s \in [0,1]} |\mathbb{M}_{\nu_i,n}(0,s) - \sqrt{n}\lambda_n(0,s)(\hat{b}_{i,1:\lfloor ns \rfloor} - \beta_{\nu_i})| = o_P(1) \qquad (D.13)$$

and

$$J_n = \sup_{s \in [0,1]} |\mathbb{M}_{\nu_i,n}(s,1) - \sqrt{n}\lambda_n(s,1)(\hat{b}_{i,\lfloor ns \rfloor+1:n} - \beta_{\nu_i})| = o_P(1), \qquad (D.14)$$

where $\hat{b}_{i,1:\lfloor ns \rfloor}$ and $\hat{b}_{i,\lfloor ns \rfloor+1:n}$ are defined analogously to (2.4). The above statements are clearly true for $i = 1$ as $\hat{b}_{1,k:l} = \hat{\beta}_{\nu_1,k:l}$ for all integers $1 \le k, l \le n$. For $i = 2$, it can be verified that the supremum on the left of (D.13) restricted to $s \in [2/n, 1]$ can be rewritten as

$$\sup_{s \in [2/n,1]} \frac{1}{\sqrt{n}} \left| \sum_{i=1}^{\lfloor ns \rfloor} X_i F_{1:\lfloor ns \rfloor}(X_i) - \sum_{i=1}^{\lfloor ns \rfloor} X_i \frac{\lfloor ns \rfloor F_{1:\lfloor ns \rfloor}(X_i) - 1}{\lfloor ns \rfloor - 1} \right|$$

$$\le \frac{2}{\sqrt{n}} \sup_{s \in [2/n,1]} \frac{1}{\lfloor ns \rfloor - 1} \sum_{i=1}^{\lfloor ns \rfloor} |X_i| \le \frac{4}{\sqrt{n}} \sup_{k \ge 1} \frac{1}{k} \sum_{i=1}^{k} |X_i| \xrightarrow{\text{a.s.}} 0$$

since $E(|X_1|) < \infty$. The case $i = 3$ is similar. To show (D.14), we use the fact that J_n is I_n computed from the sequence $X_n, \ldots, X_1$ and conclude from (D.13).

Starting from (D.13) and (D.14), it is possible to adapt the proof of Proposition D.1 to show that $S_{g,n} - T_{g,n} = o_P(1)$. We omit the details for the sake of brevity. $\quad\square$

References

Bernard, E., Naveau, P., Vrac, M., Mestre, O.: Clustering of maxima: Spatial dependencies among heavy rainfall in France. J. Climat. **26**(2), 7929–7937 (2013). http://cse.ipsl.fr/outils-logiciels/90-pakage-r-clustermax-clustering-pour-valeurs-extremes

Bücher, A., Kojadinovic, I.: Dependent multiplier bootstraps for non-degenerate u-statistics under mixing conditions with applications. J. Stat. Plann. Infer. **170**, 83–105 (2016)

Csörgő, M., Horváth, L.: Limit theorems in change-point analysis. Wiley Series in Probability and Statistics. Wiley, Chichester, UK (1997)

de Haan, L., Ferreira, A.: Extreme value theory: An introduction. Springer (2006)

Diebolt, J., Guillou, A., Naveau, P., Ribereau, P.: Improving probability-weighted moment methods for the generalized extreme-value distribution. REVSTAT **6**(1), 33–50 (2008)

Ferreira, A., de Haan, L.: On the block maxima method in extreme value theory: PWM estimators. Ann. Stat. **43**(1), 276–298 (2015)

Genest, C., Segers, J.: Rank-based inference for bivariate extreme-value copulas. The Annals of Statistics **37**, 2990–3022 (2009)

Gilleland, E., Katz, R.W.: New software to analyze how extremes change over time. Eos **92**(2), 13–14 (2011)

Greenwood, J.A., Landwehr, J.M., Wallis, J.R.: Probability-weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. Water Resour. Res. **15**, 1049–1054 (1979)

Guillou, A., Naveau, P., Diebolt, J., Ribereau, P.: Return level bounds for discrete and continuous random variables. Test **18**, 584–604 (2009)

Gumbel, E.J.: Statistics of extremes. Columbia University Press, New York (1958)

Heffernan, J.E., Stephenson, A.G.: ismev: An Introduction to Statistical Modeling of Extreme Values. http://CRAN.R-project.org/package=ismev. R package version 1.40 (2014)

Holmes, M., Kojadinovic, I., Quessy, J.-F.: Nonparametric tests for change-point detection à la Gombay and Horváth. J. Multivar. Anal. **115**, 16–32 (2013)

Hosking, J.R.M., Wallis, J.R.: Parameter and quantile estimation for the generalized Pareto distribution. Technometrics **29**, 339–349 (1987)

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

Jaruskova, D., Rencová, M.: Analysis of annual maximal and minimal temperatures for some European cities by change point methods. Environmetrics **19**, 221–233 (2008)

Kojadinovic, I.: npcp: Some Nonparametric Tests for Change-Point Detection in Possibly Multivariate Observations. http://CRAN.R-project.org/package=npcp. R package version 0.1-5 (2015)

Kosorok, M.R.: Introduction to empirical processes and semiparametric inference. Springer, New York (2008)

Landwehr, J.M., Matalas, N.C., Wallis, J.R.: Probability-weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. Water Resour. Res. **15**, 1055–1064 (1979)

Leadbetter, M.R., Lindgren, G., Rootzén, H.: Extremes and related properties of random sequences and processes. Springer Series in Statistics, Springer-Verlag, New York (1983)

Phillips, P.C.B.: Time series regression with unit roots. Econometrica **55**, 277–301 (1987)

R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2016). http://www.R-project.org

Stephenson, A.G.: evd: Extreme Value Distributions. R News **2**(2) (2002). http://CRAN.R-project.org/doc/Rnews/

Van der Vaart, A.W., Wellner, J.A. Weak convergence and empirical processes, 2nd edn. Springer, New York (2000)