

# Measuring and comparing risks of different types

Maximilian Aigner, Valérie Chavez-Demoulin, Armelle Guillou

## ▶ To cite this version:

Maximilian Aigner, Valérie Chavez-Demoulin, Armelle Guillou. Measuring and comparing risks of different types. Insurance: Mathematics and Economics, 2022, 102, pp.1-21. 10.1016/j.insmatheco.2021.11.001. hal-03109579

# HAL Id: hal-03109579 https://hal.science/hal-03109579

Submitted on 13 Jan2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Measuring and comparing risks of different types

Maximilian Aigner<sup>(1)</sup>, Valérie Chavez-Demoulin<sup>(1)</sup>, Armelle Guillou<sup>(2)</sup>

<sup>(1)</sup> University of Lausanne, Faculty of Business and Economics, Switzerland

<sup>(2)</sup> Institut Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67084 Strasbourg cedex, France

#### Abstract

Being able to compare risk measures in practice is crucial in many applications such as in finance, insurance or environmental science. The difficulty is that the variables of interest are not always of the same nature, nor of the same type or scale. Thus the usual risk measures are often misleading and to solve this issue we propose to use the Expected Proportional Shortfall (EPS) which is scale invariant and thus, which does not depend on the unit of measurement. To estimate the EPS, an estimator of the tail index  $\gamma$  is required. The main asymptotic properties of our EPS estimator are provided under very general assumptions in case of d-variate  $\beta$ -mixing processes with Pareto-type marginals. Then, we propose a test statistic based on the EPS estimators to compare different risks, whatever their nature/type/scale are. Since the performances of the test statistic are poor when a biased estimate of  $\gamma$  is used, we propose to perform our EPS estimation with an asymptotically unbiased estimator for  $\gamma$ . The efficiency of our test statistic is illustrated in a simulation experiment and validated on an environmental dataset.

Keywords: Asymptotically unbiased estimator;  $\beta$ -mixing; Convergence in distribution; Expected Proportional Shortfall.

#### 1 Introduction and motivation

In many fields where assessing risk matters, it is of main importance to compare several risk measures but often these measures are calculated from variables of different types or nature or have different unit scales. An insurance company, for instance, may be interested in comparing its claims linked to fire and its claims related to car accidents. In finance, one may want to compare market risks of assets of different currencies. In this paper, we consider an environmental application, where the question of interest is: what is the most risky climate extremes, floods or heatwaves, in a certain area? Answering this question is crucial for policy-makers and insurers to prepare for a future of climate change. Although humans can prevent the effects of floods by building dams, extreme rainfalls still strike Europe, destroying constructions, lands and, sometimes, killing people. Likewise, severe heatwaves can lead to fatalities. The European heatwave of 2003, for instance, resulted in at least 30 000 deaths in total, with more than 14 000 recorded in France. These two risks linked to heavy rainfall and heatwaves originate in phenomena of different natures and measured on different scales. To help in addressing the question above, namely, which is the higher risk, we use daily data of precipitation (hourly sum over day, in mm), of temperatures (daily minimum, in degree Celsius) and of humidity at 2m above ground (daily maximum, in %) observed in Lausanne, Switzerland, during three years (2016–2019).

In this context of comparing risks linked to different types of data, the use of traditional risk measures such as the Value-at-Risk (VaR), commonly used in finance, or its equivalent used in environment (return level), can be misleading. To solve this issue, the Expected Proportional Shortfall (EPS) has been introduced by Belzunce *et al.* (2012), and is defined for a non-negative random variable X with finite mean and continuous distribution function  $F_X$ , as

$$EPS_p = \mathbb{E}\left[\left(\frac{X - VaR_p}{VaR_p}\right)_+\right],\tag{1}$$

for all  $p \in \{p \in (0,1) : \operatorname{VaR}_p > 0\}$ , where  $\operatorname{VaR}_p = \inf\{x \in \mathbb{R} : F_X(x) \ge 1 - p\}$  and  $a_+ = \max(a, 0)$ . The EPS measures the thickness of the upper tail to a fixed VaR, proportional to the VaR, and thus it can be useful to measure risks of different nature. In particular, it has the nice property of being unaltered when all the values are multiplied by a positive scalar value.

In this paper, we consider d-variate  $(X^{(1)}, \ldots, X^{(d)})^T$  time series where each marginal distribution function is of Pareto-type with tail index  $\gamma_j < 1$ ,  $j = 1, \ldots, d$ , respectively. We further assume that the vector is coming from a  $\beta$ -mixing process as we typically encounter in environmental applications, often dealing with auto-correlated and cross-correlated components.

In a recent paper, Hoga (2018) constructs tests for equality of "tail risk" in the multivariate context of  $\beta$ -mixing processes. The test is applied to financial data with an estimate of the test statistic based on the tail index estimator proposed by Hill (1975). As is well-known in the literature, this estimator may suffer from bias in some circumstances, which affects the performance of the test. To illustrate this point, we simulate 1 000 independent bivariate samples of size 1 000 (corresponding to the size of observations in our application) where each marginal distribution function  $F_j$ , j = 1, 2, follows a  $\text{Burr}(\lambda_j, \tau_j)$  distribution, namely:

$$F_j(x) = 1 - (1 + x^{\tau_j})^{-\lambda_j}, \qquad (2)$$

where  $\gamma_j = \frac{1}{\tau_j \lambda_j}$ , j = 1, 2. In our simulations below, we set  $\gamma_1 = \gamma_2 = 0.2$ , mimicking the estimated values obtained in our application (Section 5). Our objectif is to test the null hypothesis:  $\mathcal{H}_0: \gamma_1 = \gamma_2$ . To this aim, we use the test statistic proposed by Hoga (2018) with  $\gamma_j, j = 1, 2$ , estimated by the Hill estimator (as done in the latter paper), but also with the bias-corrected estimator proposed by Chavez-Demoulin and Guillou (2018). The choice of the tuning parameters involved in these estimators is the one suggested in each of the two abovementioned papers. Since we want to point out that the bias in the tail index estimation has an impact on the performance of the test, an important parameter in this illustration is the second order parameter, denoted subsequently by  $\rho_j$ , and which is equal to  $\rho_j = -\frac{1}{\lambda_j}, j = 1, 2$ , for a Burr-type distribution as defined in (2). More precisely, the closer  $|\rho_j|$  is to 0, the more biased is the Hill estimator.

Table 1 shows the rejection frequency of the null hypothesis  $\mathcal{H}_0$ :  $\gamma_1 = \gamma_2$ . We choose,  $\rho_1 = \rho_2 = -1$ , corresponding to a situation where the Hill estimator does not suffer from bias. The columns of the table correspond to the test statistic proposed by Hoga (2018) in case where the tail index is estimated by the Hill estimator (left, Hill-estimate) and the Chavez-Demoulin and

Guillou (2018) estimator (right, CG-estimate), for two levels 5% and 10% corresponding to the rows of the table. Clearly, both methods perform well with rejection frequency close to the level. This was expected since the second order parameters  $\rho_j$ , j = 1, 2, are far from 0, and thus the test statistic is not too much sensitive on the tail index estimation method used.

	Hill-estimate	CG-estimate
5%	0.055	0.047
10%	0.12	0.092

Table 1: Burr distribution. Rejection frequency of the null hypothesis  $\mathcal{H}_0$ :  $\gamma_1 = \gamma_2$  in case  $\gamma_1 = \gamma_2 = 0.2$  and  $\rho_1 = \rho_2 = -1$ , based on 1 000 samples of size 1 000.

Similarly, Table 2 shows also the rejection frequency of the null hypothesis, but this time, when  $\rho_1 = -0.5$  and  $\rho_2 \in \{-0.7, -0.6, -0.5, -0.4, -0.3\}$ . As  $|\rho_2|$  becomes closer to 0, the rejection frequency based on the test statistic computed with the Hill-estimate increases drastically, leading to an important type I error. On the contrary, with the CG-estimate, the rejection frequency stays close to the level.

$\rho_2$	-0.7	-0.6	-0.5	-0.4	-0.3
5%					
Hill-estimate	0.036	0.060	0.091	0.146	0.274
CG-estimate	0.046	0.047	0.052	0.052	0.049
10%					
Hill-estimate	0.101	0.123	0.160	0.236	0.421
CG-estimate	0.088	0.100	0.099	0.096	0.110

Table 2: Burr distribution. Rejection frequency of the null hypothesis  $\mathcal{H}_0$ :  $\gamma_1 = \gamma_2$  in case  $\gamma_1 = \gamma_2 = 0.2$  and  $\rho_1 = -0.5$ , based on 1 000 samples of size 1 000.

Since the bias in estimating the tail index parameter clearly affects the performance of the test statistic, we only use throughout the paper the bias-corrected estimator proposed by Chavez-Demoulin and Guillou (2018). First, we extend the theoretical results of the latter paper to the

multivariate context of *d*-variate  $\beta$ -mixing processes. Then, since a test based on equality of tail indices is difficult to interpret for practitioners, except if they are familiar with extreme value theory, we focus in this paper on a test statistic based on the EPS which, in addition to be easily interpretable, allows us to compare risks of different nature, and as such it provides an answer to our initial environmental question.

The remainder of the paper is organized as follows. In Section 2, an estimator of the Expected Proportional Shortfall is proposed and its asymptotic properties are established. Then, in Section 3, a test statistic based on this risk measure has been constructed in order to compare risks of different nature or different unit scales. The performance of our test statistic in terms of size and power is illustrated in Section 4 through a small simulation study. Finally, Section 5 is devoted to the analysis of the real data presented in the introduction in order to answer our environmental question. All the proofs are postponed to the appendix.

#### 2 Estimation of the Expected Proportional Shortfall

Let  $\{\aleph_i = (X_i^{(1)}, \cdots, X_i^{(d)})^T\}_{i=1,\cdots,n}$  be a d-variate time series, where the continuous distribution function of  $X_i^{(j)}, j \in \{1, \cdots, d\}$ , are denoted by  $F_j$  and is assumed to be of Pareto-type with index  $\gamma_j > 0$ , that is

$$1 - F_j(x) = x^{-1/\gamma_j} \ell_j(x),$$

where  $\ell_j(.)$  is a slowly varying function at infinity. Define also the associated tail quantile function by  $U_j(x) = \inf\{y : F_j(y) \ge 1 - 1/x\}$ . Our estimator of the Expected Proportional Shortfall is based on an intermediate sequence k, i.e., an integer sequence such that  $k = k_n \to \infty$ with  $k/n \to 0$  as  $n \to \infty$ . To establish its convergence towards some Gaussian process, we need to assume, as usual in extreme value theory, the following second order condition (SOC):

**Assumption** (SOC). For  $j \in \{1, \dots, d\}$ , there exist  $\rho_j < 0$  and a positive or negative function  $A_j(\cdot)$  with  $\lim_{t\to\infty} A_j(t) = 0$ , such that for all y > 0,

$$\lim_{t \to \infty} \frac{\frac{U_j(ty)}{U_j(t)} - y^{\gamma_j}}{A_j(t)} = y^{\gamma_j} \frac{y^{\rho_j} - 1}{\rho_j}, \qquad \text{for } y > 0.$$

The function  $A_j(\cdot)$  in Assumption (SOC) is regularly varying with index  $\rho_j < 0$ , and the bias of the tail estimators, e.g., extreme value index or EPS estimators, can be explained by the rate of convergence of this function  $A_j(\cdot)$  towards 0: the closer  $\rho_j$  is to 0, the larger the bias.

More precisely, our time series  $\{\aleph_i\}_{i=1,\dots,n}$  are assumed to be strictly stationary sequences of  $\mathbb{R}^d$ -valued random vectors, which are  $\beta$ -mixing, i.e., such that

$$\beta(\ell) := \sup_{m \in \mathbb{N}} \mathbb{E} \left[ \sup_{A \in \mathcal{F}_{m+\ell+1}^{\infty}} |\mathbb{P}(A|\mathcal{F}_{1}^{m}) - \mathbb{P}(A)| \right] \longrightarrow 0, \quad \text{as } \ell \to \infty$$

where  $\mathcal{F}_{m+\ell+1}^{\infty}$  and  $\mathcal{F}_{1}^{m}$  are the  $\sigma$ -algebra generated by  $(\aleph_{m+\ell+1}, \aleph_{m+\ell+2}, \cdots)$  and  $(\aleph_{1}, \cdots, \aleph_{m})$ , respectively.

All our results require some additional assumptions which can be viewed as multivariate analogue of conditions required in extreme value theory for dependent data, as already used in, e.g., Drees (2000, 2003), and defined as follows:

Assumption (C<sub>1</sub>). {ℵ<sub>i</sub>}<sub>i∈ℕ</sub> is a strictly stationary β−mixing process with continuous marginals and mixing coefficients β(·) such that

$$\lim_{n \to \infty} \left\{ \frac{n}{r_n} \beta(\ell_n) + \frac{r_n}{\sqrt{k}} \log^2(k) \right\} = 0,$$

for sequences  $\{\ell_n\}_{n\in\mathbb{N}}\subset\mathbb{N}, \{r_n\}_{n\in\mathbb{N}}\subset\mathbb{N}$  tending to infinity with  $\ell_n=o(r_n)$  and  $r_n=o(n)$ ;

Assumption (C<sub>2</sub>). There exists a function r<sub>i,j</sub>(x, y), (i, j) ∈ {1, · · · , d}<sup>2</sup>, such that for some δ > 0, we have for all (x, y) ∈ (0, 1 + δ]<sup>2</sup>:

$$\lim_{n \to \infty} \frac{n}{r_n k} Cov\left(\sum_{\ell=1}^{r_n} \mathbb{1}_{\{X_{\ell}^{(i)} > U_i(\frac{n}{kx})\}}, \sum_{m=1}^{r_n} \mathbb{1}_{\{X_m^{(j)} > U_j(\frac{n}{ky})\}}\right) = r_{i,j}(x,y);$$

• Assumption (C<sub>3</sub>). For some constants C > 0,  $\delta > 0$ , we have for  $i \in \{1, \dots, d\}$ :

$$\frac{n}{r_n k} \mathbb{E}\left[\sum_{\ell=1}^{r_n} \mathbb{1}_{\{U_i(\frac{n}{ky}) < X_{\ell}^{(i)} \le U_i(\frac{n}{kx})\}}\right]^4 \le C(y-x), \, \forall \, 0 < x < y \le 1+\delta, n \in \mathbb{N}.$$

Examples of time series satisfying Assumptions  $(C_1) - (C_3)$  can be found, for instance, in Drees (2000, 2003), and include, among others, linear and ARCH models.

Our aim is to estimate the Expected Proportional Shortfall defined in (1). To this goal, observe that the EPS based on  $X^{(j)}, j \in \{1, \dots, d\}$ , can be rewritten, for  $\gamma_j < 1$ , as follows

$$\begin{aligned} \text{EPS}_{p,j} &= \int_{\text{VaR}_{p,j}}^{\infty} \left( \frac{x}{\text{VaR}_{p,j}} - 1 \right) dF_j(x) \\ &= p \int_1^{\infty} \frac{1 - F_j(z \text{VaR}_{p,j})}{1 - F_j(\text{VaR}_{p,j})} dz \\ &= p \left\{ \int_1^{\infty} z^{-1/\gamma_j} dz + A_j \left( \frac{1}{p} \right) \int_1^{\infty} z^{-1/\gamma_j} \frac{z^{\rho_j/\gamma_j} - 1}{\gamma_j \rho_j} dz \\ &+ A_j \left( \frac{1}{p} \right) \int_1^{\infty} \left[ \frac{\frac{1 - F_j(z \text{VaR}_{p,j})}{1 - F_j(\text{VaR}_{p,j})} - z^{-1/\gamma_j}}{A_j \left( \frac{1}{p} \right)} - z^{-1/\gamma_j} \frac{z^{\rho_j/\gamma_j} - 1}{\gamma_j \rho_j} \right] dz \right\} \\ &= p \left\{ \frac{\gamma_j}{1 - \gamma_j} + O \left( A_j \left( \frac{1}{p} \right) \right) \right\}, \end{aligned}$$
(3)

since, according to de Haan and Ferreira (2006, p. 161), under our Assumption (SOC), we have the following inequality: for any  $\varepsilon, \tilde{\delta} > 0$ , there exists  $s_0 = s_0(\varepsilon, \tilde{\delta}) > 1$  such that for all  $s, sz \ge s_0$  and  $j \in \{1, \dots, d\}$ :

$$\frac{\frac{1-F_j(zs)}{1-F_j(s)} - z^{-1/\gamma_j}}{A_j(1/(1-F_j(s)))} - z^{-1/\gamma_j} \frac{z^{\rho_j/\gamma_j} - 1}{\gamma_j \rho_j} \left| \le \varepsilon z^{-\frac{1}{\gamma_j} + \frac{\rho_j}{\gamma_j}} \max\left(z^{\widetilde{\delta}}, z^{-\widetilde{\delta}}\right).$$

$$\tag{4}$$

Thus, to estimate the Expected Proportional Shortfall, we need to estimate the extreme value index  $\gamma_j$ .

Chavez-Demoulin and Guillou (2018) introduced a class of estimators for the extreme value index in the case of a univariate  $\beta$ -mixing time series, say X. This class can be viewed as statistical tail functionals,  $T(Q_n)$ , where  $Q_n$  is the tail quantile function defined as  $Q_n(u) :=$  $X_{n-\lfloor ku \rfloor,n}, 0 < u < n/k$ , and  $T(\cdot)$  is a suitable functional, see also Goegebeur and Guillou (2013). We propose to extend this class of estimators in the case of d-variate  $\beta$ -mixing time series, as follows:

$$\widehat{\gamma}_{j,k}^{K}(t) := \int_{0}^{1} \log \frac{Q_{n,t}^{(j)}(z)}{Q_{n,t}^{(j)}(1)} d(zK(z)),$$
(5)

where

$$Q_{n,t}^{(j)}(z) := X_{\lfloor nt \rfloor - \lfloor \lfloor kt \rfloor z \rfloor, \lfloor nt \rfloor}^{(j)}, \text{ for } 0 < z < n/k, \ t \in [0,1], \text{ and } j \in \{1, \cdots, d\},$$
(6)

and where K satisfies the following assumption:

Assumption (K). Let K be a function such that  $\int_0^1 K(t) dt = 1$ . Suppose that K is continuously differentiable on (0,1) and that there exist M > 0 and  $\tau \in [0,1/2)$  such that  $|K(t)| \leq M t^{-\tau}$ .

These conditions on the function K(.) are not restrictive, in particular they are satisfied by the usual power kernel function  $K(u) = (1 + \nu)u^{\nu}, \nu \ge 0$ , and the log-weight function  $K(u) = (-\log u)^{\nu}/\Gamma(1 + \nu), \nu \ge 0$ . Note also that the usual tail quantile function  $Q_n$  has been slightly generalized in (6) since, later in Section 3, we use the self-normalization approach in order to construct the test statistic, and thus we need to evaluate this function for different values of t. Clearly, the estimator proposed in Chavez-Demoulin and Guillou (2018) is a particular case of (5) corresponding to the specific value t = 1, whereas the well-known Hill estimator (Hill, 1975) corresponds to the case K(z) = 1 and t = 1.

Before stating our first theorem on the convergence in distribution of our kernel-type estimator for the index, we first define, for any  $t_0 \in (0, 1)$ , the space  $D[t_0, 1]$  of càdlàg functions on  $[t_0, 1]$ equipped with the Skorohod metric, and by  $D^d[t_0, 1] = D[t_0, 1] \times \cdots \times D[t_0, 1]$  the product space endowed with the product metric.

**Theorem 1.** Let  $\{\aleph_i\}_{i=1,\dots,n}$  be a time series satisfying Assumptions  $(C_1) - (C_3)$  together with the second order condition (SOC) with  $\sqrt{k}A_j(\frac{n}{k}) \rightarrow \lambda_j \in \mathbb{R}$  for  $j \in \{1,\dots,d\}$ , as  $n \rightarrow \infty$ . Assume that the distributions  $F_j$  are continuous, for all  $j \in \{1,\dots,d\}$ , and that the kernel Ksatisfies Assumption (K). Then we have

$$t\sqrt{k} \begin{pmatrix} \widehat{\gamma}_{1,k}^{K}(t) - \gamma_{1} - A_{1}\left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{1}} - 1}{\rho_{1}} d(sK(s)) \\ \vdots \\ \widehat{\gamma}_{d,k}^{K}(t) - \gamma_{d} - A_{d}\left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{d}} - 1}{\rho_{d}} d(sK(s)) \end{pmatrix} \stackrel{d}{\longrightarrow} \Sigma_{d}^{1/2} \mathcal{B}_{d}(t) \text{ in } D^{d}[t_{0}, 1],$$

where  $\mathcal{B}_d(\cdot)$  denotes a standard d-dimensional Brownian motion and

$$\Sigma_{d} := \left(\gamma_{i}\gamma_{j}\int_{0}^{1}\int_{0}^{1}\left[\frac{\sigma_{i,j}(x,y)}{xy} - \frac{\sigma_{i,j}(x,1)}{x} - \frac{\sigma_{i,j}(1,y)}{y} + \sigma_{i,j}(1,1)\right]d(xK(x))d(yK(y))\right)_{(i,j)\in\{1,\cdots,d\}^{2}}$$
  
with  $\sigma_{i,j}(x,y) := \frac{1}{2}\left[r_{i,j}(x,y) + r_{i,j}(y,x)\right], (i,j)\in\{1,\cdots,d\}^{2}.$ 

This theorem can be viewed as a generalization of Theorem 1 in Chavez-Demoulin and Guillou (2018) to the multivariate context. Recall that our aim is to construct asymptotically unbiased

estimators. To reach this goal, we have to use a mixture between two kernels, of the form  $\Delta K_1(\cdot) + (1 - \Delta)K_2(\cdot)$ , where  $\Delta \in \mathbb{R}$  is selected in such a way that the bias of our kernel-type estimator computed with this mixture is 0. This yields to a class of asymptotically unbiased estimators. Although the question of finding the asymptotically unbiased estimator with minimal variance is open in the case of a  $\beta$ -mixing time series, in the independent and identically distributed (iid) framework it has been solved by Goegebeur and Guillou (2013), who found the following "optimal" kernel

$$K_{\rho}(z) := \left(\frac{1-\rho}{\rho}\right)^2 - \frac{(1-\rho)(1-2\rho)}{\rho^2} z^{-\rho}, \qquad z \in (0,1).$$

In the sequel, we propose to use this advocated kernel for our  $\beta$ -mixing time series, as already done in Chavez-Demoulin and Guillou (2018) in the case of a univariate  $\beta$ -mixing time series. From (3), a natural estimator of the Expected Proportional Shortfall in case  $p \to 0$  is given by

$$\widehat{\operatorname{EPS}}_{p,j}^{\rho_j}(t) = p \frac{\widehat{\gamma}_{j,k}^{K_{\rho_j}}(t)}{1 - \widehat{\gamma}_{j,k}^{K_{\rho_j}}(t)}, \quad \text{for } j \in \{1, \cdots, d\}.$$

$$(7)$$

We are now able to show the convergence in distribution of our EPS estimator defined in (7).

**Corollary 1.** Let  $\{\aleph_i\}_{i=1,\dots,n}$  be a time series satisfying Assumptions  $(C_1) - (C_3)$  together with the second order condition (SOC) with  $\sqrt{k}A_j(\frac{n}{k}) \rightarrow \lambda_j \in \mathbb{R}$  for  $j \in \{1,\dots,d\}$ , as  $n \rightarrow \infty$ . Assume that the distributions  $F_j$  are continuous with  $\gamma_j < 1$ , for all  $j \in \{1,\dots,d\}$ , and that  $p = p_n$  is such that  $k/(np) \rightarrow \infty$ . Then we have

$$t\sqrt{k} \begin{pmatrix} \frac{\bar{\mathrm{EPS}}_{p,1}^{\rho_1}(t)}{\bar{\mathrm{EPS}}_{p,1}} - 1\\ \vdots\\ \frac{\bar{\mathrm{EPS}}_{p,d}^{\rho_d}(t)}{\bar{\mathrm{EPS}}_{p,d}} - 1 \end{pmatrix} \stackrel{d}{\longrightarrow} \overline{\Sigma}_d^{1/2} \mathcal{B}_d(t) \text{ in } D^d[t_0, 1],$$

where

$$\overline{\Sigma}_{d} := \left(\frac{1}{(1-\gamma_{i})(1-\gamma_{j})} \int_{0}^{1} \int_{0}^{1} \left[\frac{\sigma_{i,j}(x,y)}{xy} - \frac{\sigma_{i,j}(x,1)}{x} - \frac{\sigma_{i,j}(1,y)}{y} + \sigma_{i,j}(1,1)\right] d(xK_{\rho_{i}}(x))d(yK_{\rho_{j}}(y))\right)_{(i,j)\in\{1,\cdots,d\}^{2}}$$
  
and  $\sigma_{i,j}(x,y), (i,j)\in\{1,\cdots,d\}^{2}$ , defined as in Theorem 1.

Our Expected Proportional Shortfall estimator is thus asymptotically unbiased in the sense that the expectation of the limit is zero whatever  $\lambda_j, j \in \{1, \dots, d\}$ , is.

Unfortunately the second order parameters  $\rho_j, j \in \{1, ..., d\}$ , are unknown in practice and thus we can either replace them in (7) by canonical values or estimators. First, we consider the case where  $\rho_j$  is replaced by some canonical value  $\tilde{\rho}_j, j \in \{1, ..., d\}$ . In that case, the convergence in distribution of our EPS estimators given in Corollary 1 is modified as follows.

**Corollary 2.** Let  $\{\aleph_i\}_{i=1,\dots,n}$  be a time series satisfying Assumptions  $(C_1) - (C_3)$  together with the second order condition (SOC) with  $\sqrt{k}A_j(\frac{n}{k}) \rightarrow \lambda_j \in \mathbb{R}$  for  $j \in \{1,\dots,d\}$ , as  $n \rightarrow \infty$ . Assume that the distributions  $F_j$  are continuous with  $\gamma_j < 1$ , for all  $j \in \{1,\dots,d\}$ , and that  $p = p_n$  is such that  $k/(np) \rightarrow \infty$ . Then we have

$$t\sqrt{k} \begin{pmatrix} \frac{\widehat{\mathrm{EPS}}_{p,1}^{\rho_1}(t)}{\mathrm{EPS}_{p,1}} - 1 - A_1\left(\frac{n}{k}\right) \frac{(1-\widetilde{\rho}_1)(\widetilde{\rho}_1 - \rho_1)}{\widetilde{\rho}_1(1-\rho_1)(1-\rho_1 - \widetilde{\rho}_1)(1-\gamma_1)\gamma_1} \\ \vdots \\ \frac{\widehat{\mathrm{EPS}}_{p,d}^{\widetilde{\rho}_d}(t)}{\mathrm{EPS}_{p,d}} - 1 - A_d\left(\frac{n}{k}\right) \frac{(1-\widetilde{\rho}_d)(\widetilde{\rho}_d - \rho_d)}{\widetilde{\rho}_d(1-\rho_d)(1-\rho_d - \widetilde{\rho}_d)(1-\gamma_d)\gamma_d} \end{pmatrix} \stackrel{d}{\longrightarrow} \underline{\Sigma}_d^{1/2} \mathcal{B}_d(t) \text{ in } D^d[t_0, 1],$$

where

$$\underline{\Sigma}_{d} := \left( \frac{1}{(1-\gamma_{i})(1-\gamma_{j})} \int_{0}^{1} \int_{0}^{1} \left[ \frac{\sigma_{i,j}(x,y)}{xy} - \frac{\sigma_{i,j}(x,1)}{x} - \frac{\sigma_{i,j}(1,y)}{y} + \sigma_{i,j}(1,1) \right] d(xK_{\widetilde{\rho}_{i}}(x)) d(yK_{\widetilde{\rho}_{j}}(y)) \right)_{(i,j) \in \{1, \cdots, d\}^{2}}$$
and  $\sigma_{i,j}(x,y), (i,j) \in \{1, \cdots, d\}^{2}$ , defined as in Theorem 1.

One clearly loses the bias correction, except if  $\tilde{\rho}_j$  corresponds to the true value of  $\rho_j$ ,  $j \in \{1, \dots, d\}$ . However, the Expected Proportional Shortfall estimators are not very sensitive to such a mis-specification and thus our estimator  $\widehat{\operatorname{EPS}}_{p,j}^{\tilde{\rho}_j}(t), j \in \{1, \dots, d\}$ , can still outperform the estimators that are not corrected for bias. If we want nevertheless to keep the property of being asymptotically unbiased, we need to estimate  $\rho_j$  in a consistent way. This leads to the following general result.

**Corollary 3.** Let  $\{\aleph_i\}_{i=1,\dots,n}$  be a time series satisfying Assumptions  $(C_1) - (C_3)$  together with the second order condition (SOC) with  $\sqrt{k}A_j(\frac{n}{k}) \rightarrow \lambda_j \in \mathbb{R}$  for  $j \in \{1,\dots,d\}$ , as  $n \rightarrow \infty$ . Assume that the distributions  $F_j$  are continuous with  $\gamma_j < 1$ , for all  $j \in \{1,\dots,d\}$ , and that  $p = p_n$  is such that  $k/(np) \to \infty$ . Let  $\hat{\rho}_{j,k_{\rho}} := \hat{\rho}_j$  be an external estimator for  $\rho_j$ ,  $j \in \{1, \dots, d\}$ , consistent in probability, which depends on an intermediate sequence  $k_{\rho}$ . Then we have

$$t\sqrt{k} \begin{pmatrix} \frac{\widehat{\operatorname{EPS}}_{p,1}^{\widehat{\rho}_{1}}(t)}{\operatorname{EPS}_{p,1}} - 1\\ \vdots\\ \frac{\widehat{\operatorname{EPS}}_{p,d}^{\widehat{\rho}_{d}}(t)}{\operatorname{EPS}_{p,d}} - 1 \end{pmatrix} \stackrel{d}{\longrightarrow} \overline{\Sigma}_{d}^{1/2} \mathcal{B}_{d}(t) \text{ in } D^{d}[t_{0}, 1],$$

where  $\overline{\Sigma}_d$  is defined as in Corollary 1.

A possible choice for  $\hat{\rho}_{j,k_{\rho}}, j \in \{1, \dots, d\}$ , is that proposed by Gomes *et al.* (2002) and defined as

$$\widehat{\rho}_{j,k_{\rho}} := \frac{-4 + 6 S_{j,k_{\rho}}^{(2)} + \sqrt{3 S_{j,k_{\rho}}^{(2)} - 2}}{4 S_{j,k_{\rho}}^{(2)} - 3} \quad \text{provided } S_{j,k_{\rho}}^{(2)} \in \left(\frac{2}{3}, \frac{3}{4}\right),$$

where

$$S_{j,k_{\rho}}^{(2)} := \frac{3}{4} \frac{\left[ M_{j,k_{\rho}}^{(4)} - 24 \left( M_{j,k_{\rho}}^{(1)} \right)^4 \right] \left[ M_{j,k_{\rho}}^{(2)} - 2 \left( M_{j,k_{\rho}}^{(1)} \right)^2 \right]}{\left[ M_{j,k_{\rho}}^{(3)} - 6 \left( M_{j,k_{\rho}}^{(1)} \right)^3 \right]^2}$$

with

$$M_{j,k_{\rho}}^{(\alpha)} := \frac{1}{k_{\rho}} \sum_{\ell=1}^{k_{\rho}} \left( \log X_{n-\ell+1,n}^{(j)} - \log X_{n-k_{\rho},n}^{(j)} \right)^{\alpha}, \alpha \in \mathbb{N}.$$

According to Theorem 2.1 in Gomes *et al.* (2002), this estimator  $\hat{\rho}_{j,k_{\rho}}$  is consistent in probability as soon as the intermediate sequence  $k_{\rho}$  satisfies  $\sqrt{k_{\rho}}A_j(n/k_{\rho}) \to \infty$  and our second order condition (SOC) hold.

### 3 Test statistic based on the Expected Proportional Shortfall

We have now all the ingredients to move to our final goal of the paper, i.e., the construction of our test statistic, based on the Expected Proportional Shortfall, in order to compare risks of different nature or of different unit scales. Our null hypothesis is formulated as follows

$$\mathcal{H}_{0}^{\text{EPS}}: \lim_{p \to 0} \frac{\text{EPS}_{p,j}}{\text{EPS}_{p,d}} = 1, \text{ or equivalently } \lim_{p \to 0} \log \frac{\text{EPS}_{p,j}}{\text{EPS}_{p,d}} = 0, \qquad \forall j = 1, \cdots, d-1, \forall j =$$

and allows us to compare the EPS of  $\beta$ -mixing series as p goes to zero.

To construct our test statistic we propose to use the self-normalization approach, originally introduced by Kiefer *et al.* (2000), Lobato (2001) and Shao (2010). See also Shao (2015) for a recent review. The idea behind this approach is to use a ratio-type statistic in order to cancel the unknown asymptotic variance  $\overline{\Sigma}_d$  appearing in both the numerator and denominator. Thus a pivotal limiting distribution results, whose critical values may be computed through simulations. The main advantages of this approach is the simplicity of implementation together with the fact that it does not depend on any tuning parameters.

To this aim, define the two  $(d-1) \times 1$  matrices:

$$\widehat{\mathrm{LREPS}}_{p}(t) := \begin{pmatrix} \log \frac{\widehat{\mathrm{EPS}}_{p,1}^{\widehat{\rho}_{1}}(t)}{\widehat{\mathrm{EPS}}_{p,d}^{\widehat{\rho}_{d}}(t)} \\ \vdots \\ \log \frac{\widehat{\mathrm{EPS}}_{p,d-1}^{\widehat{\rho}_{d-1}}(t)}{\widehat{\mathrm{EPS}}_{p,d}^{\widehat{\rho}_{d}}(t)} \end{pmatrix} \quad \text{and} \quad \overline{\mathrm{LREPS}}_{p}(t) := \begin{pmatrix} \frac{1}{d} \sum_{j=1}^{d} \log \frac{\widehat{\mathrm{EPS}}_{p,j}^{\widehat{\rho}_{j}}(t)}{\widehat{\mathrm{EPS}}_{p,d}^{\widehat{\rho}_{d}}(t)} \\ \vdots \\ \frac{1}{d} \sum_{j=1}^{d} \log \frac{\widehat{\mathrm{EPS}}_{p,j}^{\widehat{\rho}_{j}}(t)}{\widehat{\mathrm{EPS}}_{p,d}^{\widehat{\rho}_{d}}(t)} \end{pmatrix}.$$

Then, our test statistic for  $\mathcal{H}_0^{\mathrm{EPS}}$  is given by

$$T_{p}^{\text{EPS}} := \left[\widehat{\text{LREPS}}_{p}(1) - \overline{\text{LREPS}}_{p}(1)\right]^{T} \times \left[\int_{t_{0}}^{1} t^{2} \left\{ \left[\widehat{\text{LREPS}}_{p}(t) - \overline{\text{LREPS}}_{p}(t)\right] - \left[\widehat{\text{LREPS}}_{p}(1) - \overline{\text{LREPS}}_{p}(1)\right] \right\} \times \left\{ \left[\widehat{\text{LREPS}}_{p}(t) - \overline{\text{LREPS}}_{p}(t)\right] - \left[\widehat{\text{LREPS}}_{p}(1) - \overline{\text{LREPS}}_{p}(1)\right] \right\}^{T} dt \right]^{-1} \times \left[\widehat{\text{LREPS}}_{p}(1) - \overline{\text{LREPS}}_{p}(1)\right], \qquad (8)$$

and its behaviour, under the null hypothesis and its alternative, is summarized in the following theorem.

**Theorem 2.** Under the same conditions as in Corollary 3, if  $\overline{\Sigma}_d$  is a positive-definite matrix, then we have:

• under  $\mathcal{H}_0^{\text{EPS}}$ :

$$T_p^{\text{EPS}} \xrightarrow{d} \mathcal{B}_{d-1}^T(1) \left[ \int_{t_0}^1 \left\{ \mathcal{B}_{d-1}(t) - t\mathcal{B}_{d-1}(1) \right\} \left\{ \mathcal{B}_{d-1}(t) - t\mathcal{B}_{d-1}(1) \right\}^T dt \right]^{-1} \mathcal{B}_{d-1}(1),$$

where  $\mathcal{B}_{d-1}(\cdot)$  denotes a standard (d-1)-dimensional Brownian motion; • under  $\mathcal{H}_1^{\text{EPS}}$ :

$$\frac{T_p^{\text{EPS}}}{k} = O_{\mathbb{P}}(1)$$

#### 4 Simulation experiment

In this section, we perform a simulation study in order to illustrate the performance in terms of size and power of our test statistic defined in (8). We limit ourselves to the case d = 2, although similar results can be obtained in higher dimensions. Note that our EPS estimator (7) is defined in case p tends to 0, but since our test statistic involves ratios between two EPS estimators, it no longer depends on p. In practice this means that the conclusion of the test is stable over levels p (small enough) or equivalently over time horizons if the interpretation of the EPS is in terms of the return period. This seems realistic in the environmental context: If, in a given area and over a certain (long-term) horizon, flood is more risky than heatwave, this is likely to be true for higher time horizon, under same underlying conditions of climate or human intervention. We choose throughout this section p = 1/n corresponding to the boundary of observations.

Now, we need to select  $t_0$ . This parameter cannot be too large or too small. Indeed, a small value of  $t_0$  means that the EPS estimator (7) is based on samples that are too small for t close to  $t_0$ , whereas a too large value, implies that the integral-term in our test statistic (8) is based on too few differences. An extensive simulation study (not reported in the paper) shows that the value  $t_0 = 0.3$  provides a reasonable trade-off and performs well in all the models considered and whatever the sample size is. Then, we have to compute the critical values of the limiting distribution in Theorem 2 for  $t_0 = 0.3$ . This is done by a Monte Carlo method using 100 000 paths of standard Brownian motions, each path coming from 100 000 standard normal random variables. Table 3 gives these critical values.

Next, the intermediate sequence k used in our test statistic needs to be selected. To this aim,

Level	5%	10%
Critical value	67	41

Table 3: Critical values of the limiting distribution in Theorem 2 for  $t_0 = 0.3$  and d = 2, for two levels.

the following algorithm is used. For a given series j, j = 1, ..., d, since the Chavez-Demoulin and Guillou (2018) estimator for  $\gamma$  is asymptotically unbiased, we expect to see a plateau when we look at this estimate as a function of the intermediate sequence. Thus, the latter is selected as the smallest value from which the median of the estimates of  $\gamma$  over all the replications exhibit a stable part. We denote this value by  $k_j, j = 1, ..., d$ . Finally, the retained value of k used in our test statistic corresponds to  $k = \min_{1 \le j \le d} k_j$ . We use here the minimum since a too large value of this parameter might induce a bias in the estimation procedure.

Concerning the choice of the sequence  $k_{\rho}$  required in the estimation of the second order parameter  $\rho$ , as in Chavez-Demoulin and Guillou (2018) we use

$$k_{\rho} := \sup\left\{k : k \le \min\left(m-1, \frac{2m}{\log\log m}\right) \text{ and } \widehat{\rho}_k \text{ exists}\right\},$$

with m the number of positive observations in the sample.

In this simulation study, we use two data-generating  $\beta$ -mixing processes. One corresponds to the iid case of two Burr-type random variables as defined in (2) and the second one, to a VMA process, which is a more realistic process encountered in real environmental context. More precisely, it is a  $\beta$ -mixing process that allows both auto-correlation and cross-correlation, and thus it can naturally appear in environmental data, auto-correlation reflecting persistence of weather, whereas cross-correlation reflects the dynamics of weather. Our VMA(1) process considered is of the following form

$$\begin{cases} X_i^{(1)} = W_i^{(1)} + W_{i-1}^{(1)} + \nu W_{i-1}^{(2)} \\ X_i^{(2)} = W_i^{(2)} + W_{i-1}^{(2)}, \end{cases}$$
(9)

 $i = 1, \ldots, n$ , where  $(W^{(1)}, W^{(2)})^T$ , is a bivariate random vector with margins  $W^{(j)} \sim \text{Burr}(\lambda_j, \tau_j)$ ,

j = 1, 2 and with dependence structure coming from a Gumbel copula with parameter  $\theta = 2$ . The choice of a Gumbel copula is to exhibit upper tail dependence. We take  $\nu = -0.45$  to mimic the (negative) cross correlation we observe between temperature and humidity in our application.

Table 4, respectively Table 5, shows the rejection frequency of the null hypothesis  $\mathcal{H}_0^{\text{EPS}}$  in the iid case of two Burr-type random variables, respectively our VMA(1) process defined in (9), with  $\gamma_1 = 0.2$ , mimicking the estimated values obtained in our application and with  $\rho_1 = -1$  and  $\rho_2 = -0.25$ . The index  $\gamma_2$  varies from 0.2 to 0.9. We run 1 000 samples of size 500, 1 000 and 2 000.

From these two tables, one can draw the following conclusions:

- If  $\gamma_2 = \gamma_1$ , the frequencies of rejection, which correspond to the type I error, are rather close to the level of the test, especially in Table 4 when the sample size is sufficiently large;
- If  $\gamma_2$  is different from  $\gamma_1$ , but without being too far away from it (e.g.,  $\gamma_1 = 0.2$  and  $\gamma_2 = 0.25$ ), the frequencies of rejection are already high, especially in Table 5. It underlines the quality of the test which is able to detect deviations, even small, with the null hypothesis;
- When  $\gamma_2$  deviates significantly from  $\gamma_1$ , the frequencies of rejection drastically increase for both levels and all sample sizes, illustrating the high power of the test.

#### 5 Application to an environmental dataset

The past few decades have seen extreme climate events with catastrophic impacts on human society and the economy. In addition to global warming, many climate projections suggest an increase in climate extremes in Europe (IPCC, 2012). With the warming of the oceans, periods of excessive rainfall will appear considerably more often. The European Environment Agency (www.eea.europa.eu) supports this projection, endorsing, in November 2019, the statement that "Heavy precipitation events are likely to become more frequent in most parts of Europe" and

$\gamma_2$	0.2	0.25	0.5	0.7	0.9
n = 500					
5%	0.034	0.055	0.478	0.839	0.986
10%	0.056	0.116	0.705	0.904	0.991
n = 1000					
5%	0.040	0.151	0.757	0.951	1.000
10%	0.087	0.270	0.872	0.971	1.000
n = 2000					
5%	0.050	0.220	0.891	0.862	0.996
10%	0.098	0.341	0.954	0.941	0.998

Table 4: Rejection frequency of the null hypothesis  $\mathcal{H}_0^{\text{EPS}}$  in the iid case of two Burr-type random variables based on 1 000 samples, for different values of  $\gamma_2$ , levels and sample sizes. The parameters are  $\gamma_1 = 0.2$ ,  $\rho_1 = -1$  and  $\rho_2 = -0.25$ .

$\gamma_2$	0.2	0.25	0.5	0.7	0.9	
n = 5	n = 500					
5%	0.040	0.102	0.602	0.678	0.933	
10%	0.064	0.177	0.763	0.808	0.964	
n = 1	n = 1000					
5%	0.066	0.242	0.874	0.862	0.981	
10%	0.122	0.371	0.949	0.942	0.994	
n = 2000						
5%	0.065	0.342	0.943	0.818	0.954	
10%	0.119	0.492	0.978	0.927	0.980	

Table 5: Rejection frequency of the null hypothesis  $\mathcal{H}_0^{\text{EPS}}$  in case of the VMA(1) process defined in (9) based on 1 000 samples, for different values of  $\gamma_2$ , levels and sample sizes. The parameters are  $\gamma_1 = 0.2$ ,  $\rho_1 = -1$  and  $\rho_2 = -0.25$ .

the possible consequences include erratic climate situations, leading to crop failures, increases in food prices, and economic and social instabilities (Vogel *et al.*, 2019). In Europe, another source of societal damages and fatalities is the extremes of temperature causing severe winter conditions or heatwaves. A large majority (91%) of deaths due to extreme temperatures are the result of heatwaves. Between 2000 and 2019, Europe accounts for 88% of all deaths, due to the heatwave in 2003. Although human intervention can, to a certain extent, return significant damage due to flooding, it is impossible to divert a heatwave or an extreme precipitation. These points are crucial in managing societal risk and in making public safety decisions. It is therefore important to ask the question: in a specific region of Europe and during the same season of the year, what is the highest risk between a heatwave or an heavy precipitation? These two events have alternatively stricken Lausanne several times in late spring season.

In this study, our aim is to compare the different risks of observing extreme precipitation, extreme humidity and extreme daily maximum temperature in Lausanne. The three being measured on different scales, it does not make sense to compare their VaR whereas the EPS is precisely an adequate risk measure in this context. Comparing the tail index  $\gamma$  of the three time series would be an alternative but unfortunately it is not easily interpretable for practitioners or people who are not familiar with extreme value theory. On the other hand, the EPS can claim to be a universal measure for risks of different natures.

Our dataset concern 1003 observations measured daily in Lausanne, from July 9, 2016 to April 12, 2019. The data consist of humidity at 2m above ground, in %, daily maximum (Figure 1); precipitation, in mm, hourly sum over day (Figure 2); and temperature, in degree Celsius, daily maximum (Figure 3). ACF plots (not shown here) of the series show some auto-correlation, so we apply ARMA models to the series individually and keep the residuals series to work with. According to the CCF plots of Figure 4, the residuals look cross-correlated, especially when the residual series of precipitation is involved, which means that a VMA process can be adjusted to our data. Figure 5 represents the estimators for  $\gamma$ , with, on the left panel, the Chavez-Demoulin and Guillou (2018) estimators and, on the right panel, the Hill estimators, for temperature (solid line), humidity (dotted line) and precipitation (dashed line). In the three cases, the estimates of  $\gamma$  based on the Chavez-Demoulin and Guillou (2018) method look more stable as a function



Figure 1: Humidity at 2m above ground in Lausanne (in %, daily maximum) from July 9, 2016 to April 12, 2019.



Figure 2: Precipitation in Lausanne (in mm, hourly sum over day) from July 9, 2016 to April 12, 2019.

of k than the Hill estimators. This is explained by the fact that the Chavez-Demoulin and Guillou (2018) estimators are asymptotically unbiased and, as explained in the introductory



Figure 3: Temperature in Lausanne (in degree Celsius, daily maximum) from July 9, 2016 to April 12, 2019.

section, for this reason, they are the reference in the estimation of the EPS. The retained value of k is 50, corresponding to the smallest value of the three intermediate sequences after which a stability in the left panel of Figure 5 appears. Choosing a level p = 1/n, for humidity, we get  $\widehat{\text{EPS}}_{\text{H}} = 2.79 \times 10^{-4}$ , for precipitation we have  $\widehat{\text{EPS}}_{\text{P}} = 5.38 \times 10^{-4}$ , and for temperature,  $\widehat{\text{EPS}}_{\text{T}} = 7.78 \times 10^{-5}$ . As expected, and in accordance with climate science literature, the risk of extreme rainfall in this region is higher than the risk of high humidity or high daily temperature. Lausanne has been struck by huge spring precipitation in June 2018 and 2019. On June 11th, 2018, 41.0 mm fell within 10 minutes, making that period of three days (three crossed points in Figure 2) an extreme climate event. Although the risk of humidity is lower than the risk of high precipitation, it is linked to precipitation by physical law. Table 6 provides the values of the test statistic based on the comparison of the three series together and then of all combinations of pairs. According to Table 3, the critical value corresponding to a level of 5% is 67 in dimension 2 and a similar computation yields to 168 in dimension 3. Since we reject the null hypothesis when using the three series, it is important to perform pairwise tests. We notice that we reject the null hypothesis based on the EPS of the pair (Precipitation, Temperature)



Figure 4: Cross Correlation plot (CCF) between Precipitation-Humidity, Precipitation-Temperature and Humidity-Temperature.

whereas we do not reject the null hypothesis when we use the pairs (Humidity, Precipitation) and (Humidity, Temperature). As shown in the left panel of Figure 5, the estimator for  $\gamma$ based on the precipitation variable is larger than the ones for humidity and temperature. This means that we can expect more rainfall extremes than high temperature or humidity extremes in Lausanne. The fact that we get significantly much higher risk (EPS) of extreme precipitations than temperature is in accordance with the increase of extreme rainfalls recently observed in Switzerland and planned by the climate extremes projections. Although there is an increase of heatwaves observed in Switzerland, note that the daily maximum temperature does not really define an heatwave but just a punctual peak, the definition of an heatwave including also the heat duration. For the pair (Humidity, Precipitation), the non-rejection of the null hypothesis may be due to the physical link between these two variables, whereas for the pair (Humidity,



Figure 5: The Chavez-Demoulin and Guillou estimator (left panel) and the Hill estimator (right panel) for temperature (solid line), humidity (dotted line) and precipitation (dashed line).

$\mathcal{H}_0^{EPS}$	Test statistic	Decision
$\mathrm{EPS}_{\mathrm{H}} = \mathrm{EPS}_{\mathrm{P}} = \mathrm{EPS}_{\mathrm{T}}$	190.88	Rejection
$\mathrm{EPS}_{\mathrm{H}} = \mathrm{EPS}_{\mathrm{P}}$	3.77	Non - rejection
$\mathrm{EPS}_{\mathrm{H}} = \mathrm{EPS}_{\mathrm{T}}$	6.37	Non - rejection
$\mathrm{EPS}_{\mathrm{P}} = \mathrm{EPS}_{\mathrm{T}}$	73.17	Rejection

Temperature), it shows an equivalent and limited risk for each variable, in Lausanne.

Table 6: Test statistic based on the three series (Humidity, Precipitation, Temperature) and the pairs (Humidity, Precipitation), (Humidity, Temperature), (Precipitation, Temperature).

## 6 Appendix: Proofs of the results

Before proving our Theorem 1, we first remark that our kernel-type estimator of the index,  $\widehat{\gamma}_{j,k}^{K}(t)$ , can be rewritten as

$$\begin{split} \widehat{\gamma}_{j,k}^{K}(t) &= \int_{0}^{1} \log \frac{Q_{n,t}^{(j)}(z)}{Q_{n,t}^{(j)}(1)} d(zK(z)) \\ &= \sum_{i=1}^{\lfloor kt \rfloor} \log \frac{X_{\lfloor it \rfloor - i+1, \lfloor nt \rfloor}^{i}}{X_{\lfloor nt \rfloor - i+1, \lfloor nt \rfloor}^{(j)}} \int_{\frac{l-1}{\lfloor kt \rfloor}}^{\frac{i}{l+1}} d(zK(z)) \\ &= \sum_{i=1}^{\lfloor kt \rfloor} \log \frac{X_{\lfloor it \rfloor - i+1, \lfloor nt \rfloor}^{(j)}}{X_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)}} \left\{ \frac{i}{\lfloor kt \rfloor} K\left(\frac{i}{\lfloor kt \rfloor}\right) - \frac{i-1}{\lfloor kt \rfloor} K\left(\frac{i-1}{\lfloor kt \rfloor}\right) \right\} \\ &= \sum_{i=1}^{\lfloor kt \rfloor} \log \frac{X_{\lfloor it \rfloor - i+1, \lfloor nt \rfloor}^{(j)}}{X_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)}} \left\{ \frac{i}{\lfloor kt \rfloor} K\left(\frac{i}{\lfloor kt \rfloor}\right) - \frac{i-1}{\lfloor kt \rfloor} K\left(\frac{i-1}{\lfloor kt \rfloor}\right) \right\} \\ &= \sum_{i=1}^{\lfloor kt \rfloor} \log X_{\lfloor nt \rfloor - i+1, \lfloor nt \rfloor}^{(j)} \left\{ \frac{i}{\lfloor kt \rfloor} K\left(\frac{i}{\lfloor kt \rfloor}\right) - \frac{i-1}{\lfloor kt \rfloor} K\left(\frac{i-1}{\lfloor kt \rfloor}\right) \right\} \\ &- K(1) \log X_{\lfloor nt \rfloor - i+1, \lfloor nt \rfloor}^{(j)} \left\{ \frac{i}{\lfloor kt \rfloor} K\left(\frac{i-1}{\lfloor kt \rfloor}\right) \log \frac{X_{\lfloor nt \rfloor - i+1, \lfloor nt \rfloor}^{(j)}}{X_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)}} \right\} \\ &= \sum_{i=0}^{\lfloor kt \rfloor - 1} \frac{\lfloor kt \rfloor - i}{\lfloor kt \rfloor} K\left(\frac{\lfloor kt \rfloor - i}{\lfloor kt \rfloor}\right) \log \frac{X_{\lfloor nt \rfloor - i+1, \lfloor nt \rfloor}^{(j)}}{X_{\lfloor nt \rfloor - \lfloor kt \rfloor + i, \lfloor nt \rfloor}^{(j)}} \right\} \\ &= \sum_{i=0}^{\lfloor kt \rfloor - 1} \int_{\substack{x_{\perp}^{(j)} - \lfloor kt \rfloor + i+1, \lfloor nt \rfloor}{X_{\lfloor nt \rfloor}^{(j)} - \lfloor kt \rfloor + i, \lfloor nt \rfloor}} \left\{ \int_{0}^{\lfloor kt \rfloor - i} \frac{J_{\lfloor kt \rfloor + i, \lfloor nt \rfloor}}{X_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)}} \right\} \right\}$$

where

$$\begin{split} F_{n,j}(t,y) &:= \frac{1}{\lfloor kt \rfloor} \sum_{i=1}^{\lfloor nt \rfloor} \mathbbm{1}_{\{X_i^{(j)} > yX_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)}\}} \\ &= \begin{cases} \frac{\lfloor kt \rfloor - i}{\lfloor kt \rfloor} & \text{if } y \in \left[\frac{X_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)}}{X_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)}}, \frac{X_{\lfloor nt \rfloor - \lfloor kt \rfloor + i + 1, \lfloor nt \rfloor}^{(j)}}{X_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)}}\right) \\ &0 & \text{if } y \ge \frac{X_{\lfloor nt \rfloor, \lfloor nt \rfloor}^{(j)}}{X_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)}}. \end{split}$$

Thus to study the asymptotic properties of  $\widehat{\gamma}_{j,k}^{K}(t)$ , we need first to look at the behaviour of  $F_{n,j}(t,\cdot)$ . This preliminary result is given and proved in Section 6.1 below.

#### 6.1 Preliminary result

**Theorem A.** Let  $\{\aleph_i\}_{i=1,\dots,n}$  be a time series satisfying Assumptions  $(C_1) - (C_3)$  together with the second order condition (SOC) with  $\sqrt{k}A_j(\frac{n}{k}) \rightarrow \lambda_j \in \mathbb{R}$  for  $j \in \{1,\dots,d\}$ , as  $n \rightarrow \infty$ . Assume that the distributions  $F_j$  are continuous, for all  $j \in \{1,\dots,d\}$ . Then, for any  $t_0 \in (0,1)$ and  $\nu \in [0,1/2)$ , under a Skorohod construction, we have, as  $n \rightarrow \infty$ 

$$\sup_{t \in [t_0,1], z \ge 1} \left| \begin{pmatrix} z^{\frac{\nu}{\gamma_1}} \left\{ t\sqrt{k} \left[ F_{n,1}(t,z) - z^{-\frac{1}{\gamma_1}} - A_1\left(\frac{n}{k}\right) z^{-\frac{1}{\gamma_1}} \frac{z^{\frac{\rho_1}{\gamma_1}} - 1}{\gamma_1 \rho_1} \right] - \left[ W_1\left(t,z^{-\frac{1}{\gamma_1}}\right) - z^{-\frac{1}{\gamma_1}} W_1(t,1) \right] \right\} \\ \vdots \\ z^{\frac{\nu}{\gamma_d}} \left\{ t\sqrt{k} \left[ F_{n,d}(t,z) - z^{-\frac{1}{\gamma_d}} - A_d\left(\frac{n}{k}\right) z^{-\frac{1}{\gamma_d}} \frac{z^{\frac{\rho_d}{\gamma_d}} - 1}{\gamma_d \rho_d} \right] - \left[ W_d\left(t,z^{-\frac{1}{\gamma_d}}\right) - z^{-\frac{1}{\gamma_d}} W_d(t,1) \right] \right\} \end{pmatrix} \right| \stackrel{a.s.}{\longrightarrow} 0,$$

where  $\mathbf{W}(.,.) = (W_1(\cdot,\cdot), \cdots, W_d(\cdot,\cdot))^T$  is a *d*-variate continuous centered Gaussian process with covariance function given by

$$Cov(\mathbf{W}(t_1, y_1), \mathbf{W}(t_2, y_2)) = \left(\frac{\min(t_1, t_2)}{2} \left[r_{i,j}(y_1, y_2) + r_{i,j}(y_2, y_1)\right]\right)_{(i,j) \in \{1, \cdots, d\}^2}.$$

**Proof of Theorem A.** We follow the lines of the proof of Theorem 1 in Hoga (2018), in particular, we consider random variables  $U_i^{(j)}$ ,  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, d\}$ , uniformly distributed on [0, 1], such that

• Assumption  $(U_1)$ .  $\left\{ \mathbf{U}_i = (U_i^{(1)}, \cdots, U_i^{(d)})^T \right\}_{i \in \mathbb{N}}$  is a strictly stationary  $\beta$ -mixing process with uniform[0, 1] marginals and mixing coefficients  $\beta(\cdot)$  such that

$$\lim_{n \to \infty} \left\{ \frac{n}{r_n} \beta(\ell_n) + \frac{r_n}{\sqrt{k}} \log^2(k) \right\} = 0$$

for sequences  $\{\ell_n\}_{n\in\mathbb{N}}\subset\mathbb{N}$ ,  $\{r_n\}_{n\in\mathbb{N}}\subset\mathbb{N}$  tending to infinity with  $\ell_n=o(r_n)$  and  $r_n=o(n)$ ;

Assumption (U<sub>2</sub>). There exists a function r<sup>U</sup><sub>i,j</sub>(u, v), (i, j) ∈ {1, · · · , d}<sup>2</sup>, such that for some δ > 0, we have for all (u, v) ∈ (0, 1 + δ]<sup>2</sup>:

$$\lim_{n \to \infty} \frac{n}{r_n k} Cov \left( \sum_{\ell=1}^{r_n} 1\!\!1_{\{U_\ell^{(i)} > 1 - \frac{k}{n}u\}}, \sum_{m=1}^{r_n} 1\!\!1_{\{U_m^{(j)} > 1 - \frac{k}{n}v\}} \right) = r_{i,j}^U(u,v);$$

• Assumption (U<sub>3</sub>). For some constants C > 0,  $\delta > 0$ , we have for  $i \in \{1, \dots, d\}$ :

$$\frac{n}{r_n k} \mathbb{E}\left[\sum_{\ell=1}^{r_n} 1\!\!1_{\{1-\frac{k}{n}v < U_\ell^{(i)} \le 1-\frac{k}{n}u\}}\right]^4 \le C(v-u), \, \forall \, 0 < u < v \le 1+\delta, n \in \mathbb{N}.$$

Then, as  $n \to \infty$ , we have, for any  $\nu \in [0, 1/2)$ 

$$y^{-\nu}\sqrt{k} \begin{pmatrix} \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} \left[ \mathbb{1}_{\{U_{\ell}^{(1)} > 1 - \frac{k}{n}y\}} - \frac{k}{n}y \right] \\ \vdots \\ \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} \left[ \mathbb{1}_{\{U_{\ell}^{(d)} > 1 - \frac{k}{n}y\}} - \frac{k}{n}y \right] \end{pmatrix} \stackrel{d}{\longrightarrow} y^{-\nu} \begin{pmatrix} W_{1}(t,y) \\ \vdots \\ W_{d}(t,y) \end{pmatrix} \text{ in } D^{d}([t_{0},1] \times (0,1+\delta]).$$
(11)

Note that Assumptions  $(C_1) - (C_3)$  for  $X_{\ell}^{(j)}$ ,  $j \in \{1, \dots, d\}$ , imply Assumptions  $(U_1) - (U_3)$  for  $U_{\ell}^{(j)} = F_j(X_{\ell}^{(j)})$ . Also

$$X_{\ell}^{(j)} > U_j\left(\frac{n}{ky}\right) \Leftrightarrow U_{\ell}^{(j)} > 1 - \frac{k}{n}y$$

which implies by (11) and since  $nt - 1 < \lfloor nt \rfloor \le nt$ , that, as  $n \to \infty$ ,

$$y^{-\nu}\sqrt{k} \begin{pmatrix} \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} 1\!\!1_{\{X_{\ell}^{(1)} > U_{1}(\frac{n}{ky})\}} - ty \\ \vdots \\ \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} 1\!\!1_{\{X_{\ell}^{(d)} > U_{d}(\frac{n}{ky})\}} - ty \end{pmatrix} \stackrel{d}{\longrightarrow} y^{-\nu} \begin{pmatrix} W_{1}(t,y) \\ \vdots \\ W_{d}(t,y) \end{pmatrix} \text{ in } D^{d}([t_{0},1] \times (0,1+\delta]).$$

By the Skorohod's representation theorem, this convergence holds a.s. on a suitable probability space, i.e.,

$$\sup_{t \in [t_0,1], y \in (0,1+\delta]} \left| y^{-\nu} \sqrt{k} \begin{pmatrix} \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} 1\!\!1_{\{X_{\ell}^{(1)} > U_1(\frac{n}{ky})\}} - ty \\ \vdots \\ \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} 1\!\!1_{\{X_{\ell}^{(d)} > U_d(\frac{n}{ky})\}} - ty \end{pmatrix} - y^{-\nu} \begin{pmatrix} W_1(t,y) \\ \vdots \\ W_d(t,y) \end{pmatrix} \right| \xrightarrow{a.s.} 0, \quad (12)$$

the limit being continuous, the convergence is thus uniform.

Now, we need to show that  $U_j(\frac{n}{ky})$  can be replaced by  $U_j(\frac{n}{k})y^{-\gamma_j}$ ,  $j \in \{1, \dots, d\}$ . To this aim, we use the inequality (4) by setting  $s = U_j(\frac{n}{k})$ ,  $z = y^{-\gamma_j}$ , from which we deduce that, as  $n \to \infty$ ,

$$\sup_{y \in (0,1+\delta]} \left| \frac{\frac{n}{k} [1 - F_j(y^{-\gamma_j} U_j(\frac{n}{k}))] - y - A_j(\frac{n}{k}) y \frac{y^{-\rho_j} - 1}{\gamma_j \rho_j}}{y A_j(\frac{n}{k})} \right| = o(1).$$
(13)

Set  $y_{n,j} := \frac{n}{k} [1 - F_j(y^{-\gamma_j}U_j(\frac{n}{k}))]$ . In the *j*-th row of the matrix in (12),  $j = 1, \dots, d$ , inserting  $y_{n,j}$  for y, implies that

$$\sup_{t \in [t_0,1], y \in (0,1+\delta]} \left| \begin{pmatrix} y_{n,1}^{-\nu} \left\{ \sqrt{k} \left[ \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} 1\!\!1_{\{X_{\ell}^{(1)} > U_1(\frac{n}{ky_{n,1}})\}} - ty_{n,1} \right] - W_1(t, y_{n,1}) \right\} \\ \vdots \\ y_{n,d}^{-\nu} \left\{ \sqrt{k} \left[ \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} 1\!\!1_{\{X_{\ell}^{(d)} > U_d(\frac{n}{ky_{n,d}})\}} - ty_{n,d} \right] - W_d(t, y_{n,d}) \right\} \end{pmatrix} \right| \xrightarrow{a.s.}{0},$$

i.e.,

$$\sup_{t \in [t_0,1], y \in (0,1+\delta]} \left| \begin{pmatrix} y_{n,1}^{-\nu} \left\{ \sqrt{k} \left[ \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} 1\!\!1_{\{X_{\ell}^{(1)} > y^{-\gamma_1} U_1(\frac{n}{k})\}} - ty_{n,1} \right] - W_1(t, y_{n,1}) \right\} \\ \vdots \\ y_{n,d}^{-\nu} \left\{ \sqrt{k} \left[ \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} 1\!\!1_{\{X_{\ell}^{(d)} > y^{-\gamma_d} U_d(\frac{n}{k})\}} - ty_{n,d} \right] - W_d(t, y_{n,d}) \right\} \end{pmatrix} \right| \stackrel{a.s.}{\longrightarrow} 0.$$

Using (13), the fact that  $\sqrt{k}A_j(\frac{n}{k}) \to \lambda_j \in \mathbb{R}$  and a uniform continuity argument for  $W_j(t, \cdot)$ ,  $j \in \{1, \cdots, d\}$ , we have

$$\sup_{t \in [t_0,1], y \in (0,1+\delta]} \left| \begin{pmatrix} y^{-\nu} \left\{ \sqrt{k} \left[ \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} \mathbb{1}_{\{X_{\ell}^{(1)} > y^{-\gamma_1} U_1(\frac{n}{k})\}} - ty - A_1(\frac{n}{k}) t \, y \, \frac{y^{-\rho_1} - 1}{\gamma_1 \rho_1} \right] - W_1(t,y) \right\} \\ \vdots \\ y^{-\nu} \left\{ \sqrt{k} \left[ \frac{1}{k} \sum_{\ell=1}^{\lfloor nt \rfloor} \mathbb{1}_{\{X_{\ell}^{(d)} > y^{-\gamma_d} U_d(\frac{n}{k})\}} - ty - A_d\left(\frac{n}{k}\right) t \, y \, \frac{y^{-\rho_d} - 1}{\gamma_d \rho_d} \right] - W_d(t,y) \right\} \end{pmatrix} \right| \xrightarrow{a.s.} 0,$$

or by changing y into  $z^{-\frac{1}{\gamma_j}}$  (j-th component)

$$\sup_{\substack{1,z \ge 1-\widetilde{\delta} \\ |z|,z \ge 1-\widetilde{\delta} \\ |$$

 $t \in [t_0,$ 

for some  $\tilde{\delta} > 0$ . The next step consists to replace  $U_j(\frac{n}{k}), j \in \{1, \dots, d\}$ , by suitable order statistics. This can be done since from (14) and the generalized Vervaat Lemma of Einmahl etal. (2010, Lemma 5), we have for all  $j \in \{1, \dots, d\}$ , as  $n \to \infty$ ,

$$\sup_{t\in[t_0,1]} \left| \sqrt{k} \left\{ \left( \frac{X_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)}}{U_j(\frac{n}{k})} \right)^{-\frac{1}{\gamma_j}} - 1 \right\} + \frac{W_j(t,1)}{t} \right| \xrightarrow{a.s.} 0.$$

Finally, replacing in (14) z by  $z_n = z X_{\lfloor nt \rfloor - \lfloor kt \rfloor, \lfloor nt \rfloor}^{(j)} / U_j(\frac{n}{k})$  (j-th component) and since  $\frac{\lfloor kt \rfloor}{k} =$  $t + O\left(\frac{1}{k}\right)$  uniformly in t, Theorem A follows. Note that the covariance structure of the Gaussian process  $\mathbf{W}(\cdot, \cdot)$  comes from Lemma 1 in Hoga (2018).

We have now all the ingredients in order to prove our main results.

#### 6.2 Proofs of the main results

**Proof of Theorem 1.** To prove this theorem, we make use of the Cramér-Wold device (see, e.g., Severini, 2005, p. 337). More precisely, we show that

$$\beta^{T} t \sqrt{k} \begin{pmatrix} \widehat{\gamma}_{1,k}^{K}(t) - \gamma_{1} - A_{1}\left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{1}} - 1}{\rho_{1}} d(sK(s)) \\ \vdots \\ \widehat{\gamma}_{d,k}^{K}(t) - \gamma_{d} - A_{d}\left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{d}} - 1}{\rho_{d}} d(sK(s)) \end{pmatrix} \stackrel{d}{\longrightarrow} \beta^{T} \begin{pmatrix} \gamma_{1} \int_{0}^{1} \left[s^{-1}W_{1}(t,s) - W_{1}(t,1)\right] d(sK(s)) \\ \vdots \\ \gamma_{d} \int_{0}^{1} \left[s^{-1}W_{d}(t,s) - W_{d}(t,1)\right] d(sK(s)) \end{pmatrix}$$

for all  $\beta \in \mathbb{R}^d$ .

To this aim, first remark that, using (10), and since, under Assumption (K), we have

$$\gamma_j = \int_1^\infty \left\{ \int_0^{y^{-1/\gamma_j}} d(zK(z)) \right\} \frac{dy}{y},$$

direct computations lead to

$$\begin{split} \beta^{T} t \sqrt{k} \begin{pmatrix} \widehat{\gamma}_{1,k}^{K}(t) - \gamma_{1} - A_{1}\left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{1}} - 1}{\rho_{1}} d(sK(s)) \\ \vdots \\ \widehat{\gamma}_{d,k}^{K}(t) - \gamma_{d} - A_{d}\left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{d}} - 1}{\rho_{d}} d(sK(s)) \end{pmatrix} \\ &= \sum_{\ell=1}^{d} \beta_{\ell} t \sqrt{k} \left[ \widehat{\gamma}_{\ell,k}^{K}(t) - \gamma_{\ell} - A_{\ell}\left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{\ell}} - 1}{\rho_{\ell}} d(sK(s)) \right] \\ &= \sum_{\ell=1}^{d} \beta_{\ell} t \sqrt{k} \left[ \gamma_{\ell} \int_{0}^{1} \left\{ \int_{0}^{F_{n,\ell}(t,z^{-\gamma_{\ell}})} d(sK(s)) - \int_{0}^{z} d(sK(s)) \right\} \frac{dz}{z} \\ &-A_{\ell}\left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{\ell}} - 1}{\rho_{\ell}} d(sK(s)) \right]. \end{split}$$

Then, applying the mean value theorem, we get for some  $\xi_z \in (0, 1)$ 

$$\int_0^{F_{n,\ell}(t,z^{-\gamma_\ell})} d(sK(s)) = \int_0^z d(sK(s)) + \left\{ F_{n,\ell}(t,z^{-\gamma_\ell}) - z \right\} \frac{d(sK(s))}{ds} |_{s=z+\xi_z \left[ F_{n,\ell}(t,z^{-\gamma_\ell}) - z \right]}$$

This implies that, uniformly in  $t \in [t_0, 1]$ :

$$\begin{split} \beta^{T} t \sqrt{k} \begin{pmatrix} \widehat{\gamma}_{1,k}^{K}(t) - \gamma_{1} - A_{1}\left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{1}} - 1}{\rho_{1}} d(sK(s)) \\ \vdots \\ \widehat{\gamma}_{d,k}^{K}(t) - \gamma_{d} - A_{d}\left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{d}} - 1}{\rho_{d}} d(sK(s)) \end{pmatrix} \\ &= \sum_{\ell=1}^{d} \beta_{\ell} t \sqrt{k} \left[ \gamma_{\ell} \int_{0}^{1} \left\{ F_{n,\ell}(t, z^{-\gamma_{\ell}}) - z \right\} \frac{d(sK(s))}{ds} |_{s=z+\xi_{z}} [F_{n,\ell}(t, z^{-\gamma_{\ell}}) - z] \frac{dz}{z} \\ &- A_{\ell} \left(\frac{n}{k}\right) \int_{0}^{1} \frac{s^{-\rho_{\ell}} - 1}{\rho_{\ell}} d(sK(s)) \right] \\ \xrightarrow{d} \sum_{\ell=1}^{d} \beta_{\ell} \gamma_{\ell} \int_{0}^{1} \left[ s^{-1} W_{\ell}(t, s) - W_{\ell}(t, 1) \right] d(sK(s)) \end{split}$$

by our Theorem A. This achieves the proof of Theorem 1.

**Proof of Corollary 1.** Combining (3) with (7), together with the fact that  $|A_j(.)|$  is regularly varying at infinity with index  $\rho_j < 0$ , we deduce that

$$t\sqrt{k}\left(\frac{\widehat{\mathrm{EPS}}_{p,j}^{\rho_j}(t)}{\mathrm{EPS}_{p,j}} - 1\right) = t\sqrt{k}\left(\frac{\widehat{\gamma}_{j,k}^{K_{\rho_j}}(t)}{1 - \widehat{\gamma}_{j,k}^{K_{\rho_j}}(t)} \frac{1 - \gamma_j}{\gamma_j + O(A_j(\frac{1}{p}))} - 1\right)$$

$$= \frac{1}{\gamma_j \left(1 - \widehat{\gamma}_{j,k}^{K_{\rho_j}}(t)\right)} t\sqrt{k} \left(\widehat{\gamma}_{j,k}^{K_{\rho_j}}(t) - \gamma_j\right) + O_{\mathbb{P}}\left(\sqrt{k}A_j\left(\frac{1}{p}\right)\right)$$

$$= \frac{1}{\gamma_j (1 - \gamma_j)} t\sqrt{k} \left(\widehat{\gamma}_{j,k}^{K_{\rho_j}}(t) - \gamma_j\right) (1 + o_{\mathbb{P}}(1)) + o_{\mathbb{P}}\left(\sqrt{k}A_j\left(\frac{n}{k}\right)\right)$$

$$= \frac{1}{\gamma_j (1 - \gamma_j)} t\sqrt{k} \left(\widehat{\gamma}_{j,k}^{K_{\rho_j}}(t) - \gamma_j\right) + o_{\mathbb{P}}(1),$$

uniformly in  $t \in [t_0, 1]$ .

Now, using Theorem 1 and the fact that for all  $j \in \{1, \dots, d\}, \int_0^1 \frac{s^{-\rho_j} - 1}{\rho_j} d(sK_{\rho_j}(s)) = 0$ , Corollary 1 is established.

Proof of Corollary 2. Following the lines of proof of our Corollary 1,

$$t\sqrt{k}\left(\frac{\widehat{\operatorname{EPS}}_{p,j}^{\widetilde{\rho}_{j}}(t)}{\operatorname{EPS}_{p,j}} - 1 - A_{j}\left(\frac{n}{k}\right)\frac{(1 - \widetilde{\rho}_{j})(\widetilde{\rho}_{j} - \rho_{j})}{\widetilde{\rho}_{j}(1 - \rho_{j})(1 - \rho_{j} - \widetilde{\rho}_{j})(1 - \gamma_{j})\gamma_{j}}\right)$$
$$= \frac{1}{\gamma_{j}(1 - \gamma_{j})}t\sqrt{k}\left(\widehat{\gamma}_{j,k}^{K_{\widetilde{\rho}_{j}}}(t) - \gamma_{j} - A_{j}\left(\frac{n}{k}\right)\frac{(1 - \widetilde{\rho}_{j})(\widetilde{\rho}_{j} - \rho_{j})}{\widetilde{\rho}_{j}(1 - \rho_{j})(1 - \rho_{j} - \widetilde{\rho}_{j})}\right) + o_{\mathbb{P}}(1),$$

uniformly in  $t \in [t_0, 1]$ .

Now, combining Theorem 1 with the fact that for all  $j \in \{1, \cdots, d\}$ 

$$\int_0^1 \frac{s^{-\rho_j} - 1}{\rho_j} d(sK_{\widetilde{\rho}_j}(s)) = \frac{(1 - \widetilde{\rho}_j)(\widetilde{\rho}_j - \rho_j)}{\widetilde{\rho}_j(1 - \rho_j)(1 - \rho_j - \widetilde{\rho}_j)}$$

Corollary 2 follows.

Proof of Corollary 3. Again, following the lines of proof of Corollary 1, we have

$$t\sqrt{k}\left(\frac{\widehat{\operatorname{EPS}}_{p,j}^{\widehat{\rho}_{j}}(t)}{\operatorname{EPS}_{p,j}}-1\right) = t\sqrt{k}\left(\frac{\widehat{\gamma}_{j,k}^{K_{\widehat{\rho}_{j}}}(t)}{1-\widehat{\gamma}_{j,k}^{K_{\widehat{\rho}_{j}}}(t)}\frac{1-\gamma_{j}}{\gamma_{j}+O(A_{j}(\frac{1}{p}))}-1\right)$$

$$= \frac{1}{\gamma_{j}\left(1-\widehat{\gamma}_{j,k}^{K_{\widehat{\rho}_{j}}}(t)\right)}t\sqrt{k}\left(\widehat{\gamma}_{j,k}^{K_{\widehat{\rho}_{j}}}(t)-\gamma_{j}\right)+O_{\mathbb{P}}\left(\sqrt{k}A_{j}\left(\frac{1}{p}\right)\right)$$

$$= \frac{1}{\gamma_{j}(1-\gamma_{j})}t\sqrt{k}\left(\widehat{\gamma}_{j,k}^{K_{\widehat{\rho}_{j}}}(t)-\gamma_{j}\right)+o_{\mathbb{P}}(1),$$

uniformly in  $t \in [t_0, 1]$ .

Now, we use the decomposition

$$t\sqrt{k} \begin{pmatrix} \widehat{\gamma}_{1,k}^{K_{\widehat{\rho}_{1}}}(t) - \gamma_{1} \\ \vdots \\ \widehat{\gamma}_{d,k}^{K_{\widehat{\rho}_{d}}}(t) - \gamma_{d} \end{pmatrix} = t\sqrt{k} \begin{pmatrix} \widehat{\gamma}_{1,k}^{K_{\rho_{1}}}(t) - \gamma_{1} \\ \vdots \\ \widehat{\gamma}_{d,k}^{K_{\rho_{d}}}(t) - \gamma_{d} \end{pmatrix} + t\sqrt{k} \begin{pmatrix} \widehat{\gamma}_{1,k}^{K_{\widehat{\rho}_{1}}}(t) - \widehat{\gamma}_{1,k}^{K_{\rho_{1}}}(t) \\ \vdots \\ \widehat{\gamma}_{d,k}^{K_{\widehat{\rho}_{d}}}(t) - \widehat{\gamma}_{d,k}^{K_{\rho_{d}}}(t) \end{pmatrix} =: T_{1} + T_{2}.$$

The convergence in distribution of  $T_1$  follows from Theorem 1. Concerning  $T_2$ , we have to show that each component is uniformly negligible as  $n \to \infty$ . To this aim, remark that, using (10) combining with the mean value theorem as in the proof of Theorem 1, we have for some  $\xi_z \in (0, 1)$ 

$$\begin{split} t\sqrt{k}\left(\widehat{\gamma}_{j,k}^{K_{\widehat{\rho}_{j}}}(t)-\widehat{\gamma}_{j,k}^{K_{\rho_{j}}}(t)\right) &= t\sqrt{k}\int_{1}^{\infty} \left\{\int_{0}^{F_{n,j}(t,y)} [d(sK_{\widehat{\rho}_{j}}(s))-d(sK_{\rho_{j}}(s))]\right\} \frac{dy}{y} \\ &= \gamma_{j}t\sqrt{k}\int_{0}^{1} \left\{\int_{0}^{F_{n,j}(t,z^{-\gamma_{j}})} [d(sK_{\widehat{\rho}_{j}}(s))-d(sK_{\rho_{j}}(s))]\right\} \frac{dz}{z} \\ &= \gamma_{j}t\sqrt{k}\int_{0}^{1} \left\{\int_{0}^{z} [d(sK_{\widehat{\rho}_{j}}(s))-d(sK_{\rho_{j}}(s))]\right\} \frac{dz}{z} \end{split}$$
(15)  
 
$$&+ \gamma_{j}t\sqrt{k}\int_{0}^{1} \left[F_{n,j}(t,z^{-\gamma_{j}})-z\right] \frac{d(s[K_{\widehat{\rho}_{j}}(s)-K_{\rho_{j}}(s)])}{ds}|_{s=z+\xi_{z}[F_{n,j}(t,z^{-\gamma_{j}})-z]} \frac{dz}{z} \\ &= \gamma_{j}t\sqrt{k}\int_{0}^{1} \left[F_{n,j}(t,z^{-\gamma_{j}})-z\right] \frac{d(s[K_{\widehat{\rho}_{j}}(s)-K_{\rho_{j}}(s)])}{ds}|_{s=z+\xi_{z}[F_{n,j}(t,z^{-\gamma_{j}})-z]} \frac{dz}{z} \end{split}$$

since the term in (15) is equal to 0. Now, using again the mean value theorem, we have

$$\frac{d(s[K_{\widehat{\rho}_j}(s) - K_{\rho_j}(s)])}{ds} = (\widehat{\rho}_j - \rho_j) \left\{ \frac{-2(1 - \widetilde{\rho}_j)}{\widetilde{\rho}_j^3} + \frac{2(1 - \widetilde{\rho}_j)(1 - \widetilde{\rho}_j - \widetilde{\rho}_j^2)}{\widetilde{\rho}_j^3} s^{-\widetilde{\rho}_j} + \frac{(1 - \widetilde{\rho}_j)^2(1 - 2\widetilde{\rho}_j)}{\widetilde{\rho}_j^2} s^{-\widetilde{\rho}_j} \ln s \right\}$$

with  $\tilde{\rho}_j$  an intermediate value between  $\hat{\rho}_j$  and  $\rho_j$ . Since  $\hat{\rho}_j$  is a consistent estimator for  $\rho_j$ , Theorem A achieves the proof of Corollary 3.

**Proof of Theorem 2.** First remark that

$$t\sqrt{k}\left[\widehat{\mathrm{LREPS}}_{p}(t) - \overline{\mathrm{LREPS}}_{p}(t)\right] =: A \times t\sqrt{k} \begin{pmatrix} \log \frac{\widehat{\mathrm{EPS}}_{p,1}^{\rho_{1}}(t)}{\mathrm{EPS}_{p,1}} \\ \vdots \\ \log \frac{\widehat{\mathrm{EPS}}_{p,d}^{\rho_{d}}(t)}{\mathrm{EPS}_{p,d}} \end{pmatrix} + t\sqrt{k} \begin{pmatrix} \log \frac{\mathrm{EPS}_{p,1}}{\mathrm{EPS}_{p,d}} - \frac{1}{d} \sum_{\ell=1}^{d} \log \frac{\mathrm{EPS}_{p,\ell}}{\mathrm{EPS}_{p,d}} \\ \vdots \\ \log \frac{\mathrm{EPS}_{p,d}(t)}{\mathrm{EPS}_{p,d}} \end{pmatrix}$$

where A is the  $(d-1) \times d$  matrix defined as

$$A := \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} - \frac{1}{d} \begin{pmatrix} 1 & \dots & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \dots & \dots & 1 \end{pmatrix}.$$

Then, we can remark that under  $\mathcal{H}_0^{\text{EPS}}$ , according to (3), we have all the indices equal, i.e.,  $\gamma_1 = \cdots = \gamma_d$ . This implies that, under  $\mathcal{H}_0^{\text{EPS}}$ ,

$$\begin{split} \sqrt{k} \log \frac{\mathrm{EPS}_{p,\ell}}{\mathrm{EPS}_{p,d}} &= \sqrt{k} \log \left\{ \frac{\gamma_{\ell}}{1 - \gamma_{\ell}} \frac{1 - \gamma_{d}}{\gamma_{d}} \left( 1 + O\left(A_{\ell}\left(\frac{1}{p}\right)\right) + O\left(A_{d}\left(\frac{1}{p}\right)\right) \right) \right\} \\ &= O\left(\sqrt{k}A_{\ell}\left(\frac{1}{p}\right)\right) + O\left(\sqrt{k}A_{d}\left(\frac{1}{p}\right)\right) \\ &= O\left(\sqrt{k}A_{\ell}\left(\frac{n}{k}\right)\right) + O\left(\sqrt{k}A_{d}\left(\frac{n}{k}\right)\right) \\ &= O(1), \end{split}$$

since  $|A_{\ell}(.)|$  and  $|A_{d}(.)|$  are regularly varying with negative parameters and  $k/(np) \to \infty$ . Moreover, we have

$$\operatorname{Cov}\left(A\overline{\Sigma}_{d}^{1/2}\mathcal{B}_{d}(t), A\overline{\Sigma}_{d}^{1/2}\mathcal{B}_{d}(s)\right) = \min(s, t)A\overline{\Sigma}_{d}A^{T},$$

where  $A\overline{\Sigma}_d A^T$  is a  $(d-1) \times (d-1)$  Hermitian positive-definite matrix, which thus can be decomposed into  $LL^T$ , where L is a lower triangular matrix with positive diagonal entries. Thus

$$t\sqrt{k}\left[\widehat{\mathrm{LREPS}}_p(t) - \overline{\mathrm{LREPS}}_p(t)\right] \xrightarrow{d} L\mathcal{B}_{d-1}(t) \text{ in } D^{d-1}[t_0, 1].$$

The first part of Theorem 2 then follows from an application of the continuous mapping theorem. The second part of Theorem 2 is a direct consequence of our Corollary 3 combining with the fact that, under  $\mathcal{H}_1^{\text{EPS}}$ ,

$$\lim_{p \to 0} \log \frac{\text{EPS}_{p,\ell}}{\text{EPS}_{p,d}} = \log \left( \frac{\gamma_{\ell}}{1 - \gamma_{\ell}} \frac{1 - \gamma_{d}}{\gamma_{d}} \right) \neq 0,$$

for all  $\ell = 1, \cdots, d - 1$ .

### Acknowledgement

The research of Armelle Guillou was supported by the French National Research Agency under the grant ANR-19-CE40-0013-01/ExtremReg project and an International Emerging Action (IEA-00179). The authors would like to thank Yannick Hoga for letting the R-code of its paper published in 2018 available.

### References

Belzunce, F., Pinar, J.F., Ruiz, J.F. and Sordo, M.A. (2012). Comparison of risks based on the expected proportional shortfall. *Insurance: Mathematics and Economics*, **51**, 292–302.

Chavez-Demoulin, V. and Guillou, A. (2018). Extreme quantile estimation for  $\beta$ -mixing time series and applications. *Insurance: Mathematics and Economics*, 83, 59–74.

Drees H. (2000). Weighted approximations of tail processes for  $\beta$ -mixing random variables. The Annals of Applied Probability, **10**, 1274–1301.

Drees H. (2003). Extreme quantile estimation for dependent data, with applications to finance. Bernoulli, 9, 617–657.

Einmahl, J.H.J., Gantner, M. and Sawitzki, G. (2010). Asymptotics of the shorth plot. *Journal* of Statistical Planning and Inference, **140**, 3003–3012.

Goegebeur, Y. and Guillou, A. (2013). Asymptotically unbiased estimation of the coefficient of tail dependence. *Scandinavian Journal of Statistics*, **40**, 174–189.

Gomes, M.I., de Haan, L. and Peng, L. (2002). Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes*, **5**, 387–414.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory. An Introduction*. Springer Series in Operations Research and Financial Engineering, Springer, New York.

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

Hoga, Y. (2017). Change point tests for the tail index of  $\beta$ -mixing random variables. *Econometric Theory*, **33**, 915–954.

Hoga, Y. (2018). Detecting tail risk differences in multivariate time series. *Journal of Time Series Analysis*, **39**, 665–689.

IPCC (2012). Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. *Special Report of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press), Hov et al., Extreme Weather Events in Europe: Preparing for Climate Change Adaptation.

Kiefer, N.M., Vogelsang, T.J. and Bunzel, H. (2000). Simple robust testing of regression hypotheses. *Econometrica*, **68**, 695–714.

Lobato, I.N. (2001). Testing that a dependent process is uncorrelated. *Journal of the American Statistical Association*, **96**, 1066–1076.

Severini, T.A. (2005). *Elements of distribution theory*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge.

Shao, X. (2010). A self-normalized approach to confidence interval construction in time series. Journal of the Royal Statistical Society: Series B (Statistical Methodology), **72**, 343–366.

Shao, X. (2015). Self-normalization for time series: a review of recent developments. *Journal of the American Statistical Association*, **110**, 1797-1817.

Vogel, E., Donat, M. G., Alexander, L. V., Meinshausen, M., Ray, D. K., Karoly, D., Meinshausen, N. and Frieler, K. (2019). The effects of climate extremes on global agricultural yields. *Environmental Research Letters*, **14**, 054010.