Research Article

Open Access

Alexandre Lekina, Fateh Chebana, and Taha B. M. J. Ouarda

On the tail dependence in bivariate hydrological frequency analysis

Abstract: In Bivariate Frequency Analysis (BFA) of hydrological events, the study and quantification of the dependence between several variables of interest is commonly carried out through Pearson's correlation (r), Kendall's tau (τ) or Spearman's rho (ρ). These measures provide an overall evaluation of the dependence. However, in BFA, the focus is on the extreme events which occur on the tail of the distribution. Therefore, these measures are not appropriate to quantify the dependence in the tail distribution. To quantify such a risk, in Extreme Value Analysis (EVA), a number of concepts and methods are available but are not appropriately employed in hydrological BFA. In the present paper, we study the tail dependence measures with their non-parametric estimations. In order to cover a wide range of possible cases, an application dealing with bivariate flood characteristics (peak flow, flood volume and event duration) is carried out on three gauging sites in Canada. Results show that r, τ and ρ are inadequate to quantify the extreme risk and to reflect the dependence characteristics in the tail. In addition, the upper tail dependence measure, commonly employed in hydrology, is shown not to be always appropriate especially when considered alone: it can lead to an overestimation or underestimation of the risk. Therefore, for an effective risk assessment, it is recommended to consider more than one tail dependence measure.

Keywords: Asymptotic; Extreme; Copula; Bivariate distribution; Non-parametric estimation

1 Introduction

Given economic, social and scientific issues related to floods, storms and droughts, no serious debate on these notions can be conducted without a reflection on the extreme nature of these events [*e.g.* 13, 47]. They require an accurate modelling and an appropriate analysis. In order to evaluate hydrological risk, some studies advocate univariate analysis based mainly on flood peaks [*e.g.* 6]. Nevertheless, hydrological processes are characterized by several variables. For instance, floods are mainly described with three variables (peak flow, flood volume and event duration) obtained from the hydrograph [*e.g.* 69, 74]. Thus, an effective risk assessment cannot be conducted by studying each variable separately since this does not take into account the dependence between variables and can lead to an overestimation or underestimation of the risk [*e.g.* 11, 18, 52, 59]. In such a situation, copulae are widely employed [*e.g.* 9, 28]. In hydrology, the quantification of the degree of dependence between the underlying variables with an indicator value in a scalar format is fundamental [63].

Alexandre Lekina: ITE - VEDECOM, Institut de la Transition Énergétique du Véhicule Décarboné et Communicant et de sa Mobilité. Fondation Partenariale de l'Université de Versailles Saint-Quentin-en-Yvelines, 77 rue des Chantiers, 78000- Versailles, France and Institut National de la Recherche Scientifique, INRS-ETE, 490 rue de la Couronne, Québec (Québec) G1K 9A9, Canada

Fateh Chebana: Institut National de la Recherche Scientifique, INRS-ETE, 490 rue de la Couronne, Québec (Québec) G1K 9A9, Canada

Taha B. M. J. Ouarda: Masdar Institute of Science and technology, PO Box 54224, Abu Dhabi, UAE and Institut National de la Recherche Scientifique, INRS-ETE, 490 rue de la Couronne, Québec (Québec) G1K 9A9, Canada

During the last years, the study of the dependence of hydrometeorological variables has gained increasing attention in hydrological risk assessment [see 11, and references therein]. In this framework, common measures such as Pearson's correlation r, Kendall's τ and Spearman's ρ have been largely employed by hydrologists. However, these indicators are not always appropriate for a proper understanding of dependencies in Bivariate Frequency Analysis (BFA) of extreme events [see *e.g.* 22, for a study in financial markets] since they cover the whole distribution without focusing on the tail of the distribution where extreme risks could occur. In particular, the coefficient r is based on the notions of linearity, normality and mean which are not appropriate when dealing with extreme events. The use of this indicator can lead to underestimation of the risk. Moreover, the Pearson coefficient may not even exist for heavy tailed distributions such as the Generalized Extreme Value or the Generalized Pareto. For instance, in the case of the Cauchy distribution, a theoretical value of Pearson's correlation does not exist. Embrechts et al. [21] showed that the Gaussian model is inadequate to quantify the extreme risks and indicated that the covariance gives incomplete information of joint extreme risks.

The non-parametric dependence measures, Spearman's ρ and Kendall's τ, do not assume linearity and are not based on normality. The Spearman's ρ can be seen as the Pearson's correlation coefficient between the ranked variables [*e.g.* 61] and measures the average departure from independence [see, *e.g.* 63, Section B.2.3]. The Kendall's τ is also based on the ranks of the observations [40] and measures the excess of concordance/discordance [see, *e.g.* 63, Section B.2.2]. These coefficients do not attribute sufficient weight to the extreme values. They are good overall indicators but are not appropriate when the focus is on the extremes and the distribution tail.

To study the dependence in the BFA of extreme events, a "local dependence measure" is required since the interest is in the distribution tails. In Extreme Value Analysis (EVA), a number of relevant concepts and methods are developed to locally study the dependence in a joint distribution [*e.g.* 26, 27]. These concepts are commonly used in actuarial sciences and finance [*e.g.* 1, 7, 22, 50]. For instance, the upper tail dependence parameter is introduced by Joe [37, p. 33]. However, to the best knowledge of the authors, there are no hydrological investigations of such methods for hydrological BFA except the upper and/or lower tail dependence parameter which is, for instance, briefly presented in Salvadori et al. [63], Genest and Favre [29], Poulin et al. [58], Serinaldi [68], Shiau et al. [70] and Lee et al. [46]. Nevertheless, this parameter is not always appropriate and should be combined with other complementary measures.

The aim of the present paper is to introduce and study different tail dependence measures for bivariate random variables (X, Y) in hydrological BFA. The paper is organized as follows. In Section 2, we present the recent and significant tail dependence measures in EVA. In Section 3, we focus on the special case of Bivariate Extreme Value (BEV) distributions due to their importance in EVA. Non-parametric estimators of the presented tail dependence measures are briefly developed in Section 4. Section 5 is devoted to the applications and Section 6 presents the conclusions.

2 Tail dependence measures for bivariate distributions

Let $(X_1, Y_2), \ldots, (X_n, Y_n)$ be independent random vectors in $\mathbb{R}^2$ with joint cumulative distribution function (CDF) $F(., .)$. We denote the marginal distributions of $F(., .)$ as $F_1(.)$ and $F_2(.)$ respectively for X and Y and by $C(., .)$ the copula function associated to $F(., .)$. A copula is a cumulative distribution function (CDF) whose margins are uniformly distributed on $[0, 1]$. The joint distribution function can be written in the form [71]:

$$F(x, y) = C\left(F_1(x), F_2(y)\right) \text{ for } x, y \in \mathbb{R}. \tag{1}$$

A copula function represents the dependence structure of a multivariate random vector. It contains complete information about the joint distribution apart from its margins. In this sense, a copula describes the association between X and Y in a form that is invariant with respect to strictly increasing marginal transformations [12]. The marginal distributions $F_1(.)$ and $F_2(.)$ are assumed to be continuous, which is the case for hydrological series. Therefore the copula $C(., .)$ is unique. The reader is referred to Nelsen [51] or Joe [37] for

further details on the theory of copulae and to Salvadori et al. [63] for practical illustrations. In the remainder of the Section, three tail dependence measures are presented. The first allows to group together the distributions into two classes whereas the second provides a complementary information to that provided by the first. The third links the first measure with the second and allows also to reinforce the findings given by the first two.

2.1 Tail dependence measure χ_{U}

The first concepts were discussed as far back as Geffroy [26, 27] and the following formal definition has been given by Joe [37, p. 33]:

$$\chi_{\text{U}} = \lim_{u \to 1} \mathbb{P}(F_1(X) > u | F_2(Y) > u). \tag{2}$$

The limit χ_{U} is called the *upper tail dependence parameter* (UTDP). It roughly corresponds to the probability that one margin exceeds a large quantile threshold u under the condition that the other margin exceeds u as well [24]. In other words, it is the probability that if one variable is extreme, then the other is also extreme.

The formulation in (2) is of interest for hydrological processes, since it is based on $F_1(X)$ and $F_2(Y)$ and not directly on X and Y and therefore, the variables describing the hydrological event do not need to have the same scale and to be of the same nature. The UTDP χ_{U} is defined as the limiting value of $\chi(u)$ as $u \to 1$ where [12]

$$\chi(u) = 2 - \frac{\log \mathbb{P}(F_1(X) < u, F_2(Y) < u)}{\log \mathbb{P}(F_1(X) < u)} \equiv 2 - \frac{\log C(u, u)}{\log u}, \quad 0 < u < 1. \tag{3}$$

Note that in EVA, the statistical study of the tail or the extreme risk is always established under asymptotic considerations. In the remainder of the paper, the term "asymptotic" refers to $u \to 1$. The function $\chi(u)$ can be interpreted as a quantile-dependent measure of dependence [12]. Its upper and lower bounds are given by:

$$2 - \frac{\log \left(\max(2u - 1, 0) \right)}{\log u} \le \chi(u) \le 1, \quad 0 < u < 1. \tag{4}$$

The left and right hand sides in (4) correspond respectively to perfect negative and perfect positive dependence [3, p.344]. The function $\chi(u)$ provides an insight to the dependence structure at lower quantile levels. The case $C(u, u) = u^2$ corresponds to exact independence $\chi(u) \equiv 0$. When $\chi_{\text{U}} \in (0, 1]$, then X and Y are said to be asymptotically dependent, whereas when $\chi_{\text{U}} = 0$, these variables are said to be asymptotically independent. In general, $\chi(u)$ is a non-trivial function of u and does not have explicit formula. As illustrated in Figure 1a, Coles et al. [12] showed that for a pair of Gaussian variables with correlation coefficient ρ, $\chi(u)$ increases with ρ, but as $u \to 1$ the effect of dependence diminishes and $\chi(u) \to 0$ for all $\rho < 1$. For an intermediate value of u, $\chi(u)$ is reasonably linear with distinctly different values for all ρ. For $\rho > 0$, $\chi(u)$ converges very slowly and ultimately abruptly. An important finding from this Figure is that the dependence in the center is clearly not maintained in the extremes. It is possible to pass from high dependence to independence. On the other hand, this means that it is possible to conclude erroneously that the extremes are asymptotically dependent simply because the extreme independence is not easily detectable due to inadequate sample size. This indicates that the bivariate extreme models are not adapted in the case of asymptotic independence, see, later, the remark after Eq. (13). Therefore, although these models clearly reflect the behaviour of extremes in the case of asymptotic dependence, in the case of asymptotic independence the result is very mixed.

In summary, in the extremes context, although χ_{U} is "better" than overall dependence measures r, ρ and τ, it is not always sufficient to quantify the dependence appropriately in all situations. It could fail to discriminate between the degrees of relative strength of dependence for asymptotically independent variables. Thus, it is important to overcome this limitation by introducing another characterization or a complementary dependence measure. Note that χ_{U} is the only measure employed in hydrological applications and it is only considered in few studies [18, 32, 46].
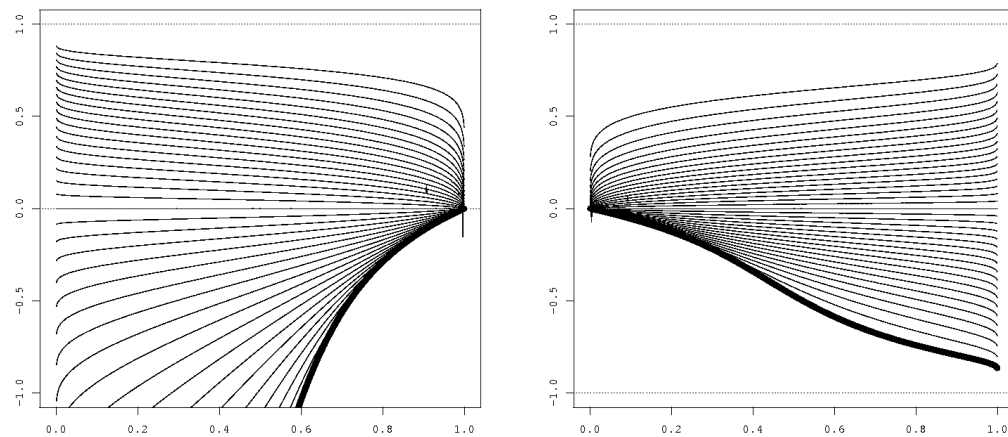
(a) Tail dependence function $\chi(u)$, $u \in (0, 1)$ **(b)** Tail dependence function $\bar{\chi}(u)$, $u \in (0, 1)$

Figure 1: Tail dependence functions for the Gaussian dependence model. The curves (bottom to top) correspond to correlation coefficient $\rho = -0.95, -0.90, -0.85, \ldots, 0.90, 0.95$. In dotted line, the upper and lower bounds on $\chi(.)$ and $\bar{\chi}(.)$.

Before introducing the complementary measure of $\chi(.)$ in the next section, it seems to be interesting to focus, briefly, on the auxiliary function $q_c(.)$ called the *tail concentration function* (TCF). It depends on the diagonal section of a copula and is defined as follows [72]

$$q_c(u) \quad = \quad \mathbb{P}(F_1(X) < u | F_2(Y) < u).\mathbb{1}_{\{(0,\,0.5]\}} + \mathbb{P}(F_1(X) > u | F_2(Y) > u).\mathbb{1}_{\{(0.5,\,1)\}} \tag{5}$$

$$= \quad \frac{C(u, u)}{u}\mathbb{1}_{\{(0,\,0.5]\}} + \frac{1 - 2u + C(u, u)}{1 - u}\mathbb{1}_{\{(0.5,\,1)\}}, \; 0 < u < 1. \tag{6}$$

The TCF can be seen as a tool to give a description of tail dependence at finite scale. In addition, it can be more suited to assess the risk of joint extremes than its limits given by χ_{U} and χ_{L} where $\chi_{\mathrm{L}} = \lim_{u \to 0} \mathbb{P}(F_1(X) < u | F_2(Y) < u)$. Thus, when the convergence speed of the TCF to 1 is slow, this implies that the dependence in the finite upper tail can be significantly stronger than in the limit [19]. The reader is referred to [19, see Fig. 3] for the discussion related to the practical effect of considering the TCF.

2.2 Tail dependence measure $\bar{\chi}_{\mathrm{U}}$

The function $\chi(.)$ given in (3) as a tail dependence measure is useful in the case where the variables are asymptotically dependent. It is not appropriate for discriminating asymptotic independence for which data exhibit positive or negative association, *i.e.* correlation, that only gradually disappears at more and more extreme levels. A complementary measure of $\chi(.)$, denoted $\bar{\chi}(.)$, has been introduced by Ledford and Tawn [43, 44] and developed by Coles et al. [12]. The function $\bar{\chi}(.)$ measures the strength of dependence within the class of asymptotically independent distributions. In a similar way to the function $\chi(u)$ given in (3), $\bar{\chi}(u)$ is defined as follows

$$\bar{\chi}(u) = \frac{2 \log \mathbb{P}(F_1(X) > u)}{\log \mathbb{P}(F_1(X) > u, F_2(Y) > u)} - 1 \equiv \frac{2 \log(1 - u)}{\log \bar{C}(u, u)} - 1, \; 0 < u < 1, \tag{7}$$

where $\bar{C}(u, v) = 1 - u - v + C(u, v)$. The function $\bar{\chi}(u)$ is also bounded from below and above as

$$\frac{2 \log(1 - u)}{\log(\max(1 - 2u, 0))} - 1 \leq \bar{\chi}(u) \leq 1, \; 0 < u < 1. \tag{8}$$

$\bar{\chi}(u)$ has the following properties [2, 3, 12]:

- If an exact independence occurs beyond u, then $\bar{\chi}(u) = 0$;
- If there is a perfect dependence beyond u, then $\bar{\chi}(u) = 1$;
- If $\bar{\chi}(u) \in (0, 1)$, then $\mathbb{P}\left(F_1(X) > u | F_2(Y) > u\right) > \mathbb{P}\left(F_2(Y) > u\right)$ and the extremes are positively associated; *i.e.* observations for which both $F_1(X) > u$ and $F_2(Y) > u$ for large threshold u are likely to occur more frequently than under exact independence between X and Y;
- If $\bar{\chi}(u) \in (-1, 0)$, then $\mathbb{P}\left(F_1(X) > u | F_2(Y) > u\right) < \mathbb{P}\left(F_2(Y) > u\right)$ and we say that the extremes are negatively associated, *i.e.* observations for which both $F_1(X) > u$ and $F_2(Y) > u$ for a large threshold u are likely to occur less frequently than under exact independence between X and Y;
- $|\bar{\chi}(u)|$ increases with the tail dependence.

To focus on extremal characteristics, by analogy to χ_U, one defines $\bar{\chi}_U$ as the limiting value of $\bar{\chi}(u)$ as $u \to 1$ for which $-1 \le \bar{\chi}_U \le 1$. This limit has the following properties:

- $\bar{\chi}_U = 1$ corresponds to the asymptotic dependence of extremes. The bivariate Gumbel-logistic extreme value distribution is an example where this case occurs;
- $\bar{\chi}_U < 1$ corresponds to the asymptotic independence of extremes and $\bar{\chi}_U$ provides a limiting measure that increases with relative dependence strength within this class;
- $\bar{\chi}_U$ allows to better characterize a possible asymptotic independence and it provides a complementary information to that provided by χ_U. For instance, as illustrated in Figure 1b, in the case of a Gaussian pair, we have $\bar{\chi}_U = \rho$ and $\bar{\chi}(u)$ is approximately linear for $0.5 < u < 1$. Therefore, one concludes an asymptotic independence, despite what might suggest a direct interpretation of $\chi(u)$ in Figure 1a [12].

In summary, the quantities χ_U and $\bar{\chi}_U$ allow to characterize the dependence of extremes as follows:

- $\chi_U \in [0, 1]$ with the set $(0, 1]$ corresponds to *asymptotic dependence*;
- $\bar{\chi}_U \in [-1, 1]$ with the set $[-1, 1)$ corresponds to *asymptotic independence*.

As a result, the pair $(\chi_U, \bar{\chi}_U)$ can be used as a summary of extreme dependence:

- If $(\chi_U > 0, \bar{\chi}_U = 1)$, the variables are *asymptotically dependent* and χ_U determines a measure of strength of dependence within the class of asymptotically dependent distributions;
- The case $(\chi_U = 0, \bar{\chi}_U < 1)$ corresponds to *asymptotic independence* between variables and $\bar{\chi}_U$ measures the strength of dependence within the class of asymptotically independent distributions.

2.3 Coefficient of tail dependence η

In this subsection, we assume that a joint distribution of (X, Y) has unit Fréchet margins, *i.e.*

$$F_1(x) = \exp(-1/x), \ x > 0 \text{ and } F_2(y) = \exp(-1/y), \ y > 0. \tag{9}$$

This restrictive assumption is without loss of generality since, if necessary, $F_1(.)$ and $F_2(.)$ can be transformed into unit Fréchet margins under suitable assumptions [see *e.g.* 43]. In order to analyse the asymptotic dependence structure between the Fréchet margins and to link χ_U with $\bar{\chi}_U$, Ledford and Tawn [43, 44] introduced the following model on the tail of the joint survival function of (X, Y):

$$\mathbb{P}(X > z, Y > z) \sim z^{-1/\eta} L(z), \ \text{ as } z \to \infty, \tag{10}$$

where L is a *univariate slowly varying function at infinity* [5, Theorem 1.5.12], *i.e.* ,

$$L(\lambda z)/L(z) \to 1 \text{ as } z \to \infty \ \text{ for all } \lambda > 0. \tag{11}$$

The rate of decay in (10) is primarily controlled by η. The coefficient η describes the type of limiting dependence between X and Y, and L is its relative strength given a particular value of η. By putting $T = \min(X, Y)$, it follows that $\mathbb{P}(X > z, Y > z) = \mathbb{P}(T > z) \sim z^{-1/\eta} L(z)$ and η is identified as the tail index of the variable T. Hence, the usual univariate techniques can be used to evaluate η [35, 55]. One can show that

$$\bar{\chi}_U = 2\eta - 1, \tag{12}$$

and the estimate of $\bar{\chi}_{\mathrm{U}}$ can be obtained from that of η which is more developed and studied since it is related to the tail index. As a consequence, we have [12, 34, 43, 45]:

- X and Y are asymptotically dependent if and only if $\eta = 1$ and $L(z) \to c \in (0, 1]$ as $z \to \infty$. In this situation, we have $(\chi_{\mathrm{U}} = c, \bar{\chi}_{\mathrm{U}} = 1)$. The constant c denotes the dependence degree where $c = 1$ corresponds to the perfect dependence in tail;
- The case $\eta \to 0$ and $L(z) = 1$ corresponds to *perfect negative dependence* (in tail);

In addition, within the class of asymptotically independent variables, *i.e.* $0 < \eta < 1$, three types of independence can be identified:

- The case $\eta = 1/2$ corresponds to near independence between the extremes of X and Y. These extremes are exactly independent when $c = 1$;
- If $1/2 < \eta < 1$ and $c > 0$, or $\eta = 1$ and $c = 0$, then the marginal variables are said to be *positively associated*;
- If $0 < \eta < 1/2$, then the marginal variables are said to be *negatively associated*.

To summarize, the degree of dependence between large values of Fréchet margins is determined by η, with increasing values of η corresponding to stonger association. For a given η, the relative dependence strength is characterized by the slowly varying function L [3, p.346]. For instance, for the Gaussian dependence model with correlation $\rho < 1$ illustrated in Figure 1, we have $\eta = (1 + \rho)/2$ and $L(z) = c_\rho (\log z)^{-\rho/(1+\rho)}$ where $c_\rho = (1+\rho)^{3/2}(1-\rho)^{-1/2}(4\pi)^{-\rho/(1+\rho)}$ [34]. In that case, positive association, negative association and exact independence arise respectively as $\rho > 0$, $\rho < 0$ and $\rho = 0$. The *perfect* positive and negative associations are reached as $\rho \to 1$ and $\rho \to -1$ respectively.

Figure 2 summarizes in a diagram the presented tail dependence measures by highlighting the concepts of the asymptotic independence/dependence. Figure 2 gives also additional information which is developed in the following Section. In Figure 2, the circle denotes the starting point, with several possible paths that can be followed. This Figure will be described later, at the end of Section 3.

3 Particular case of the BEV distributions

The BEV distributions are a particular case of bivariate distributions. They are characterized by some specific dependence functions which can be expressed through the previous tail dependence measures. In this Section, we briefly present the relevant measures of the tail dependence for these distributions since they play a prominent role in the studies of bivariate extreme events. In order to carry out a meaningful study about tail dependence in the BFA, we assume that $F(., .)$ belongs to the domain of attraction of a BEV distribution G, *i.e.* there exist standardizing sequences $a_n, \ c_n > 0$ and $b_n, \ d_n \in \mathbb{R}$ such that for all x and y [25, 60]

$$\lim_{n \to \infty} \mathbb{P}\left[\frac{\max(X_1, \ldots, X_n) - b_n}{a_n} \le x, \ \frac{\max(Y_1, \ldots, Y_n) - d_n}{c_n} \le y \right] = G(x, y). \tag{13}$$

It is shown in the literature that all BEV distributions are asymptotically dependent, otherwise, in the case of an asymptotic independence, the only possible situation is the exact independence [*e.g.* 2, 12]. For the latter, $\bar{\chi}_{\mathrm{U}} = 1$ and $\chi_{\mathrm{U}} > 0$, and in practice the dependency is the stronger as the UTDP χ_{U} is close to 1. Besides, the tail dependence function $\chi(.)$ is constant. Figure 3 illustrates the behaviour of the tail dependence functions $\chi(.)$ and $\bar{\chi}(.)$ for the bivariate Gumbel-logistic distribution with dependence parameter $0 < \theta \le 1$ which is a BEV distribution. Notice that, for the bivariate Gumbel-logistic distribution, the parameter θ measures the strength of the dependence and the limiting cases $\theta = 1$ and $\theta = 0$ correspond respectively to independence and perfect dependence. Figure 3a shows that $\chi(.)$ is positive, constant and close to 1 when θ is close to 0. In Figure 3b, for large values of θ, $\bar{\chi}(.)$ converges slowly to 1 as $u \to 1$.

An estimation of $\chi(.)$ significantly non-constant reflects an inadequacy of the BEV distribution to the data. This situation arises when (X, Y) are asymptotically independent and n, the block size maxima, is not large enough to meet the condition in (13) [12]. In hydrological BFA, since the peak flows are extracted as
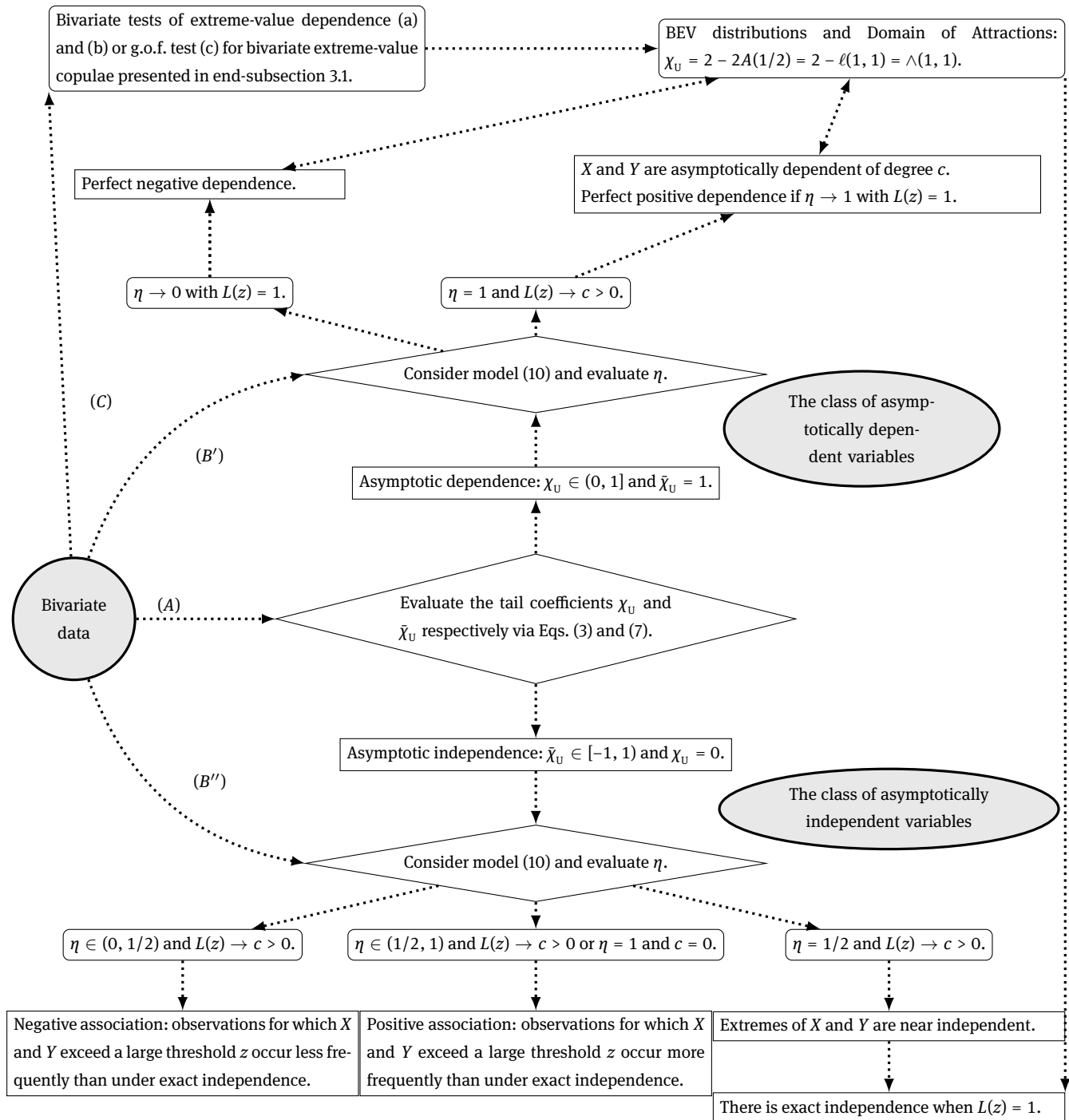
Figure 2: Diagram of the study of the tail dependence in bivariate frequency analysis. The circle denotes the starting point. We can follow the path (A), (B'), (B'') or (C), however we recommend to follow the path (A).

block maxima, hydrologists tend to jointly model flood characteristics with the component-wise maxima, *i.e.* a BEV distribution, without always checking first if $\bar{\chi}_U = 1$. The dependence structure of $G(.,.)$ in (13) is characterized by quantities given in the following subsections.
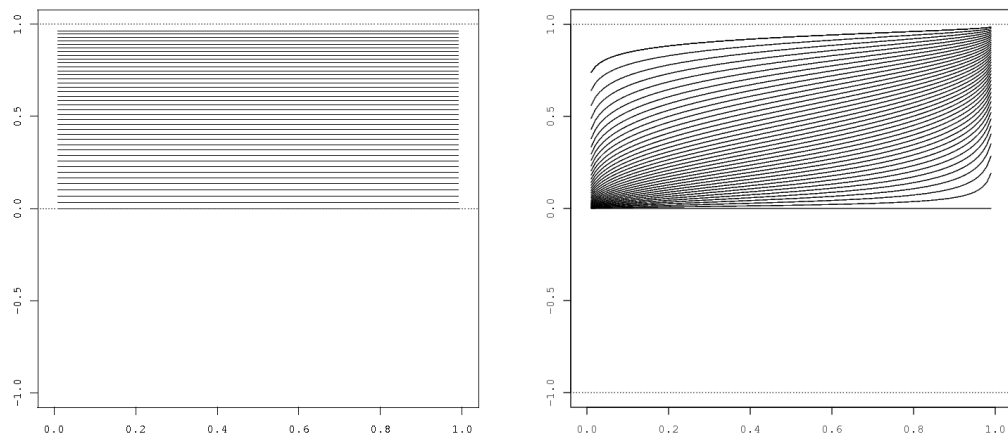
(a) Tail dependence function $\chi(u)$, $u \in (0, 1)$ **(b)** Tail dependence function $\bar{\chi}(u)$, $u \in (0, 1)$

Figure 3: Tail dependence functions for the Gumbel-logistic distribution. The curves (top to bottom) correspond to parameter of dependence $\theta = 0.025, 0.050, 0.075, \ldots, 0.950, 0.975$. In dotted line, the upper and lower bounds on $\chi(.)$ and $\bar{\chi}(.)$.

3.1 Pickands dependence function

The representation of dependence structure discovered by Pickands [56] turned out to be far more convenient than its predecessors [3, p. 270] such that:

$$G(x, y) = \exp\left[-\left(\frac{1}{x} + \frac{1}{y}\right) A\left(\frac{y}{x+y}\right)\right], \ x, y > 0, \tag{14}$$

where $A : [0, 1] \rightarrow [1/2, 1]$, known as *Pickands dependence function*, is a convex function such that $A(0) = A(1) = 1$. The independence case corresponds to $A(.) \equiv 1$ whereas, $A(w) = \max(w, 1 - w)$ leads to the perfect dependence. The copula of extreme value denoted by $\mathcal{C}(., .)$ is expressed as $\mathcal{C}(u, v) = \exp\left[\log(uv)A\left(\frac{\log(v)}{\log(uv)}\right)\right]$ and for $u = v = z$, it follows that

$$\mathcal{C}(z, z) = z^{2A(1/2)}. \tag{15}$$

Definition (15) implies the following relationship between χ_{U} and $A(.)$:

$$\chi_{\mathrm{U}} = 2 - 2A(1/2). \tag{16}$$

Thus, estimating χ_{U} is a particular case of estimating $A(.)$. The reader is referred to Salvadori and De Michele [62] for practical applications of Pickands dependence function in hydrology. The Pickands dependence function $A(.)$ and the extreme value copula $\mathcal{C}(., .)$ allow to check whether a sample comes from a BEV distribution $G(., .)$ or at least if $F(., .)$ belongs to the domain of attraction of a BEV distribution G. To this end, three statistical tests can be used: *(i)* the bivariate test of extreme-value dependence based on Kendall's process [33] or *(ii)* the one based on the Pickands dependence function [41] and then, if H_0 is accepted, *(iii)* the goodness-of-fit tests for bivariate extreme-value copulae [30] .

3.2 Stable tail dependence function

A bivariate CDF $F(., .)$ with continuous margins $F_1(.)$ and $F_2(.)$ is said to have a *stable dependence function* (STDF) $\ell(., .)$ if the following limit exists [36]:

$$\lim_{t \to 0} t^{-1}\mathbb{P}\left(1 - F_1(X) \le tx \ \text{or} \ 1 - F_2(Y) \le ty\right) = \ell(x, y), \ \text{ for } x, y \ge 0. \tag{17}$$

Referring to [17], $F(.,.)$ is equivalent to (13) if and only if *(i)* $F_1(.)$ and $F_2(.)$ are in the max-domains of attractions of extreme value distributions $G_1(.)$ and $G_2(.)$ respectively, and *(ii)* $F(.,.)$ has a STDF $\ell(.,.)$ defined by

$$\ell(x, y) = -\log G\left(\frac{x^{-\gamma_1} - 1}{\gamma_1}, \frac{y^{-\gamma_2} - 1}{\gamma_2}\right), \tag{18}$$

where $G(.)$ is a BEV distribution while γ_1 and γ_2 are real constants called the *marginal extreme value indices*. The STDF $\ell(.,.)$ can be seen as a starting point to construct non-parametric models or BEV distributions. For instance, one can cite the Gumbel-logistic model for which

$$\ell(x, y; \theta) = \left(x^{1/\theta} + y^{1/\theta}\right)^{\theta}, \quad x, y \geq 0 \text{ and } 0 < \theta \leq 1. \tag{19}$$

The STDF $\ell(.,.)$ and the Pickands dependence function $A(.)$ are related by $A(t) = \ell(1 - t, t)$ for $t \in [0, 1]$ [see 3, p.267]. It follows that χ_U and $\ell(.,.)$ are related as

$$\chi_U = 2 - \ell(1, 1) = 2 - 2\ell(1/2, 1/2) > 0. \tag{20}$$

Recall that for BEV distributions $\bar{\chi}_U = 1$.

3.3 Tail copula function

The tail copula is a function that describes the dependence structure in the tail of a joint CDF $F(.,.)$. Similar to (17), for all non-negative x and y, the quantity

$$\Lambda(x, y) = \lim_{t \to 0} t^{-1} \mathbb{P}\left(1 - F_1(X) \leq tx \text{ and } 1 - F_2(Y) \leq ty\right), \tag{21}$$

is called the *tail copula function* (TCF) of (X, Y), provided the limit exists. The relationship between $\Lambda(.,.)$ and $\ell(.,.)$ is given by

$$\Lambda(x, y) = x + y - \ell(x, y), \quad \text{for all } x, y \geq 0. \tag{22}$$

The quantity $\Lambda(1, 1)$ is the UTDP χ_U of (X, Y) [*e.g.* 20]. Schmidt and Stadtmuller [65] proposed $\Lambda(.,.)$ as a starting point to construct a multivariate distribution of extreme values. In addition, the TCF function is considered as an intuitive and straightforward generalization of the tail dependence function $\chi(.)$ via a function describing the dependence structure in the tail of a distribution [65]. To summarize, in these particular cases, χ_U is expressed explicitly with $\Lambda(.,.)$, $\ell(.,.)$ and $A(.)$ as follows:

$$\chi_U = \Lambda(1, 1) = 2 - \ell(1, 1) = 2 - 2A(1/2) > 0. \tag{23}$$

These relations are useful for estimating χ_U since the established properties of the functions $\Lambda(.,.)$, $\ell(.,.)$ and $A(.)$ are well developed.

Figure 2 summarizes in a diagram all presented tail dependence measures, χ_U, $\bar{\chi}_U$, η, $\Lambda(.,.)$, $\ell(.,.)$ and $A(.)$, by highlighting the concepts of the asymptotic independence and asymptotic dependence. In Figure 2, from the starting point circle, there are several possible paths $((A), (B'), (B'')$ and $(C))$ which can be followed. The choice of which path to take depends on the available information and the goal. However we recommend to follow the path (A). It can be seen as a procedure starting from data to obtain the tail behaviour via the presented measures. The path (A) describes as follows:

(i) From the bivariate data $\{(X_i, Y_i), i = 1, \ldots, n\}$ with joint distribution function $F(.,.)$, evaluate the tail coefficients χ_U and $\bar{\chi}_U$ respectively given by (3) and (7). If $(\chi_U > 0, \bar{\chi}_U = 1)$, we are within the class of asymptotically dependent distributions; otherwise if $(\chi_U = 0, \bar{\chi}_U < 1)$, we are within the class of asymptotically independent distributions.

(ii) Consider model (10) and evaluate η:

(1) Within the class of asymptotically independent distributions, depending on the values of η, four cases are possible: negative association, positive association, near independence or exact independence.

(2) Within the class of asymptotically dependent distributions, according to η, three types of dependence are possible: perfect negative dependence, perfect positive dependence and asymptotic dependence.

(iii) Except for the exact independence, all BEV distributions are asymptotically dependent.

4 Non-parametric estimation of tail dependence

Depending on the level of available information about the distribution of the data, there exist several approaches to estimate the tail dependence functions and coefficients. First, the bivariate distribution $F(.,.)$ could be either known [22] or belongs to a class of distributions [64, 65]. Second, the tail dependence can be estimated by using a specific copula [49, 53] or a class of copulae [39]. Finally, non-parametric estimation methods can be employed when no specific form is known or constrained on the copula or on the marginal distributions. In the present section, we focus on non-parametric methods. The tail dependence estimates are obtained from the empirical copula or based on the transformation of original data to Fréchet variables because $F(.,.)$ or $C(.,.)$ are generally unknown [57].

4.1 Estimators of tail dependence parameter χ_{U}

As shown implicitly in Section 2, the tail dependence parameter χ_{U} can be estimated by using the copula, the Pickands dependence function, the STDF or the TCF. In the following, one shows how an estimator of χ_{U} is obtained by using the latter functions.

4.1.1 Estimation via the empirical copula

An estimator of $\chi(.)$ is obtained via the empirical copula. [38] introduced the following estimator

$$\hat{\chi}^{\mathrm{SEC}}\left(\frac{n-k}{n}\right) = 2 - \left[1 - \hat{C}_n\left(\frac{n-k}{n}, \frac{n-k}{n}\right)\right] \bigg/ \left[1 - \frac{n-k}{n}\right], \; 0 < k < n, \tag{24}$$

where k denotes a threshold, that is a sample fraction, to be chosen and $\hat{C}_n(.,.)$ is the empirical copula defined by [14, 31]

$$\hat{C}_n(u, v) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{\{R_i^X/(n+1)\leq u, \; R_i^Y/(n+1)\leq v\}}, \; u, v \in [0, 1], \tag{25}$$

where $\mathbb{1}_{\{.\}}$ is the indicator function, while R_i^X and R_i^Y, respectively, stand for the ranks of X_i among $X_i, \ldots, X_n$ and Y_i among $Y_i, \ldots, Y_n$. Coles et al. [12] introduced, on the basis of (3), the following estimator of the tail dependence function

$$\hat{\chi}^{\mathrm{LOG}}\left(\frac{m-k}{m}\right) = 2 - \left[\log\left(\hat{C}_m\left(\frac{m-k}{m}, \frac{m-k}{m}\right)\right) \bigg/ \log\left(\frac{m-k}{m}\right)\right], \; 0 < k < m \leq n, \tag{26}$$

where $\hat{C}_m(.,.)$ is an empirical copula computed from m block maxima $X_{lj}^{\star}$ and $Y_{lj}^{\star}$, $j = 1, \ldots, m$, and where each block contains $l = n/m$ elements of the original data. The estimators $\hat{\chi}_{\mathrm{U}}^{\mathrm{SEC}}$ and $\hat{\chi}_{\mathrm{U}}^{\mathrm{LOG}}$ are deduced respectively from $\hat{\chi}^{\mathrm{SEC}}(.)$ and $\hat{\chi}^{\mathrm{LOG}}(.)$ by noting that $u = (n - k)/n$ is close to 1 when k is small.

The coefficient χ_{U} can also be estimated by the least-square method such that [15, 23]:

$$\hat{\chi}_{\mathrm{U}}^{\mathrm{FD}} = \hat{\chi}_{\mathrm{U},k}^{\mathrm{FD}} = \underset{\lambda \in [0,1]}{\arg\min} \sum_{i=1}^{k}\left(\hat{C}_n\left(\frac{n-i}{n}, \frac{n-i}{n}\right) - \left(\frac{n-i}{n}\right)^{2-\lambda}\right)^2, \; 0 < k < n, \tag{27}$$

where $\arg\min_{\lambda \in [0,1]} h(\lambda)$ gives an argument at which $h(.)$ is minimized over the domain $[0, 1]$. Dobric and Schmid [15] showed that $k \approx \sqrt{n}$ can be an appropriate choice to built the estimators $\hat{\chi}_U^{SEC}, \hat{\chi}_U^{LOG}$ and $\hat{\chi}_U^{FD}$. Frahm et al. [24] suggested to deduce $\hat{\chi}_U^{SEC}$ and $\hat{\chi}_U^{LOG}$ by choosing a threshold k based on the property of tail copula homogeneity as stated in Schmidt and Stadtmuller [65, Theorem 1]. This approach consists in identifying a plateau, which is induced by the homogeneity, on the graphs $(k, \hat{\chi}^\bullet(.))$. Nevertheless, the plateau-finding algorithm developed in Frahm et al. [24] requires a prior definition of some parameters.

4.1.2 Estimation via Pickands dependence function $A(.)$

As mentioned in subsection 3.1, estimating χ_U can be obtained by estimating $A(.)$ via (16). Since the margins $F_1(.)$ and $F_2(.)$ are rarely known in practice, a natural way to proceed is then to estimate them empirically by $\hat{F}_{1,n}(.)$ and $\hat{F}_{2,n}(.)$. This leads to estimating the copula $C(.,.)$ on the basis of the transformed observations $\{(\hat{F}_{1,n}(X_i), \hat{F}_{2,n}(Y_i)), \ i = 1, \ldots, n\}$. However, it is more convenient to consider scaled variables defined by Genest and Segers [31]:

$$\hat{U}_i = \hat{U}_{i,n} = \frac{1}{n+1} \sum_{j=1}^{n} \mathbb{1}_{\{X_j \leq X_i\}} \text{ and } \hat{V}_i = \hat{V}_{i,n} = \frac{1}{n+1} \sum_{j=1}^{n} \mathbb{1}_{\{Y_j \leq Y_i\}}, \tag{28}$$

The scaled pairs $\{(\hat{U}_i, \hat{V}_i), \ i = 1, \ldots, n\}$ are called the *pseudo-observations* from copula $C(.,.)$. They allow to avoid dealing with points at the boundary of the unit square. Genest and Segers [31] proposed the two following estimators of $A(.)$ which are the rank-based versions of the estimators given respectively by Pickands [56] and Capéraà et al. [8]:

$$\hat{A}_{n,r}^P(t) = 1 \left/ \left\{ \frac{1}{n} \sum_{i=1}^{n} \hat{\xi}_i(t) \right\} \right. \text{ and } \hat{A}_{n,r}^{CFG}(t) = \exp\left(-c_E - \frac{1}{n} \sum_{i=1}^{n} \log \hat{\xi}_i(t) \right), \tag{29}$$

where $c_E \approx 0.57721$ is the Euler's constant while, for $i \in \{1, \ldots, n\}$, the function $\hat{\xi}(.)$ is defined as $\hat{\xi}_i(0) = -\log \hat{U}_i, \hat{\xi}_i(1) = -\log \hat{V}_i$ and for all $t \in (0, 1)$ $\hat{\xi}_i(t) = \min\left\{ \frac{-\log \hat{U}_i}{1-t}, \frac{-\log \hat{V}_i}{t} \right\}$. The estimators in (29) lead to $\hat{\chi}_{U,r}^P = 2 - 2\hat{A}_{n,r}^P(1/2)$ and $\hat{\chi}_{U,r}^{CFG} = 2 - 2\hat{A}_{n,r}^{CFG}(1/2)$. Another estimator of the UTDP, motivated by Capéraà et al. [8] and studied by Frahm et al. [24], is given by

$$\hat{\chi}_U^F = 2 - 2 \exp\left(\frac{1}{n} \sum_{i=1}^{n} \log\left\{ \left(\log \frac{1}{F_1(X_i)} \log \frac{1}{F_2(Y_i)} \right)^{1/2} \left/ \log \frac{1}{\max(F_1(X_i).F_2(Y_i))^2} \right. \right\} \right). \tag{30}$$

The latter estimator relies on the hypothesis that the underlying empirical copula can be approximated by an extreme value copula. As the margins are unknown, in practice one can replace $F_1(X_i)$ and $F_2(Y_i)$ by the scaled variables $\hat{U}_i$ and $\hat{V}_i$. Note that, in some situations, the Pickands and the CFG estimators can be altered to meet the endpoint constraints $A(0) = A(1) = 0$. Therefore, for all $t \in [0, 1]$, Segers [67] suggested endpoint-correction versions of the Pickands and CFG estimators. Genest and Segers [31] showed that the endpoint correction to estimators (29) has no impact on their limiting distributions. In addition, they showed that that the CFG estimator is generally preferable to the Pickands one when the endpoint corrections are applied to both of them.

4.1.3 Estimation via stable tail dependence function $\ell(.,.)$

A natural non-parametric estimator of the STDF $\ell(.,.)$ based on $\{(X_i, Y_i), \ i = 1, \ldots, n\}$ is obtained by replacing in (17) t by k/n with $k \in \{1, \ldots, n\}$ and $\mathbb{P}, F_1$ and F_2 by their empirical counterparts [36]. Then, the empirical STDF is defined by:

$$\hat{\ell}_n^{\mathrm{H}}(x, y) = \frac{1}{k} \sum_{i=1}^{n} \mathbb{1}_{\{X_i \geq X_{n-kx+1,n} \text{ or } Y_i \geq Y_{n-ky+1,n}\}}, \ 0 \leq x, y \leq n/k, \tag{31}$$

where $X_{1,n} \leq \ldots \leq X_{n,n}$ and $Y_{1,n} \leq \ldots \leq Y_{n,n}$ are respectively the order statistics of $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$. The distribution of $\hat{\ell}_n^{\mathrm{H}}(., .)$ is independent of the continuous margins [17]. Einmahl et al. [20] proposed the following estimator that usually performs slightly better than $\hat{\ell}_n^{\mathrm{H}}(., .)$ for finite samples:

$$\hat{\ell}_n^{\mathrm{EKS}}(x, y) = \frac{1}{k} \sum_{i=1}^{n} \mathbb{1}_{\{R_i^X > n-kx+1/2 \text{ or } R_i^Y > n-ky+1/2\}}, \tag{32}$$

where R_i^X and R_i^Y as in (25). Definitions (31) and (32) lead to estimators of the UTDP given by $\hat{\chi}_{\mathrm{U}}^{\mathrm{H}} = 2 - 2\hat{\ell}_n^{\mathrm{H}}(1, 1)$ and $\hat{\chi}_{\mathrm{U}}^{\mathrm{EKS}} = 2 - 2\hat{\ell}_n^{\mathrm{EKS}}(1, 1)$ respectively.

4.2 Estimators of tail dependence parameter $\bar{\chi}_{\mathrm{U}}$

Estimator of $\bar{\chi}_{\mathrm{U}}$ can be obtained on the basis of the copula or the tail coefficient η. The function $\bar{\chi}(.)$ can be estimated by substituting in (7) the empirical estimate of survival copula function $\hat{\bar{C}}_n(., .)$ given by $\hat{\bar{C}}_n(u, v) = 1 - u - v + \hat{C}_n(u, v)$ for all $u, v \in [0, 1]$ where $\hat{C}_n(u, v)$ is given in (25). Thus, a non-parametric estimator of $\bar{\chi}(.)$ is given by [12]

$$\hat{\bar{\chi}}^{\mathrm{COLES}}\left(\frac{n-k}{n}\right) = \left[2 \log\left(1 - \frac{n-k}{n}\right) \middle/ \log \hat{\bar{C}}_n\left(\frac{n-k}{n}, \frac{n-k}{n}\right)\right] - 1, \ 0 < k < n. \tag{33}$$

On the other hand, according to (12) an estimation of η leads to an estimation of $\bar{\chi}_{\mathrm{U}}$. Since η is identified as the tail index of the univariate variable $T = \min(X, Y)$, one can estimate η with the estimator called Zipf [42, 66]

$$\eta_k^{\mathrm{Z}} = \frac{\frac{1}{k}\sum_{j=1}^{k} \log\frac{k+1}{j} \log T_{n-j+1,n} - \left(\frac{1}{k}\sum_{j=1}^{k}\log\frac{k+1}{j}\right)\left(\frac{1}{k}\sum_{j=1}^{k}\log T_{n-j+1,n}\right)}{\frac{1}{k}\sum_{j=1}^{k}\left(\log\frac{k+1}{j}\right)^2 - \left(\frac{1}{k}\sum_{j=1}^{k}\log\frac{k+1}{j}\right)^2}, \ 1 \leq k < n - 1. \tag{34}$$

where $T_{1,n} \leq \ldots \leq T_{n,n}$ denote the order statistics of the random variables T_i. It can also be estimated with the Hill [35] estimator given by:

$$\eta_k^{\mathrm{H}} = \frac{1}{k} \sum_{j=1}^{k} \log T_{n-j+1,n} - \log T_{n-k,n}, \ 1 \leq k < n - 1. \tag{35}$$

The two latter estimating procedures require the knowledge of the margins $F_1(.)$ and $F_2(.)$ since the model of Ledford and Tawn [43, 44] assumes that (X, Y) has unit Fréchet margins. However, when the margins are not identical or not Fréchet distributed, the original variables can be transformed to standard Fréchet margins defined by $X_{\mathrm{new}} = -1/\log \hat{F}_1(X_{\mathrm{original}})$ and $Y_{\mathrm{new}} = -1/\log \hat{F}_2(Y_{\mathrm{original}})$ [43]. However, these transformations could induce an uncertainty in the estimates of η [3, p.351]. Therefore, Peng [54] and Draisma et al. [16] respectively proposed non-parametric alternatives (36) and (37) based directly on the empirical distributions of the original observations given respectively by:

$$\hat{\eta}_k^{\mathrm{P}} = \log(2) \middle/ \log\left[S_n(2k)/S_n(k)\right], \tag{36}$$

$$\hat{\eta}_k^{\mathrm{D}} = \sum_{j=1}^{k} S_n(j) \middle/ \left(kS_n(k) - \sum_{j=1}^{k} S_n(j)\right), \tag{37}$$

where $S_n(k) = \sum_{i=1}^{n} \mathbb{1}_{\{X_i > X_{n-k+1,n}, \ Y_i > Y_{n-k+1,n}\}}$ with $k = 0, \ldots, n - 1$. For the choice of the threshold k, the reader is referred to Lekina et al. [48] and references therein.

5 Application to floods

In this Section, the presented estimators of the tail dependence are applied to a particular hydrological event, namely floods. Flood events are mainly described by three characteristics that are flood peak (Q), flood volume (V) and flood duration (D).

5.1 Data description

The data used in this case study consists in daily natural streamflow measurements from three stations in the province of Quebec (Canada). The reference numbers of the selected basins are 050301, 080101 and 050119 and the gauging stations are respectively denoted by ST_{050301}, ST_{080101} and ST_{050119}. Maximum annual flood events are described by their flood peaks, durations and volumes as extracted from the daily streamflow data.The three variables correspond to the same flood event each year. In particular, it corresponds to the spring flood event which is the important flood event in Quebec and is caused mainly by snow melting [4]. Gauging station ST_{050301} is located on the Batiscan River and the corresponding data are available from 1932 to 1990 with missing values in 1989-1983, 1979-1976 and 1972. Station ST_{080101} is located on the Harricana River 3.4 km downstream from the Route 111 bridge in Amos. Corresponding data are available from 1934 to 2002 with missing values in 1998 and 1999. The third gauging station ST_{050119} is located on the Matawin River 4.0 km downstream from the pont-route 131 in Saint-Michel-des-Saints. Data are available from 1932-2001 with missing values in 1940, 1941, 1943 and 1972. In Table 1, gauging station coordinates and record lengths are summarized.

Table 1: Station description.

Station	Localisation	Latitude	Longitude	Observation years
050301	Batiscan River	46°35′8″	−72°24′17″	1932 – 1990
080101	Harricana River	48°35′52″	−78°6′33″	1934 – 2002
050119	Matawin River	46°41′10″	−73°54′49″	1932 – 2001

The hydrological literature has highlighted issues concerning the correlation between the three characteristics of flood events. Due to space limitations, the focus will be made on the study of (Q, V) whereas brief results will be provided concerning (V, D) and (Q, D). The couple (Q, V) is generally highly correlated and represents the most studied in the literature [see *e.g.* 11, 69, 74, 75]. The tail dependence behaviours of the couples (Q, V), (D, V) and (V, D) are studied according to the proposed approach summarized in Figure 2. This step is performed before any modelling of the joint distributions. While in Chebana and Ouarda [10] the bivariate descriptives statistics based on the depth function are investigated, the present study focuses solely on the study of the tail dependance. Before presenting the results, it should be noted that the hydrological analysis is generally affected by the record length especially when we deal with extremes. Despite this limitation, asymptotic results are usually employed in the hydrological frequency analysis and in the multivariate setting in particular. This issue could have similar effect on the tail dependence measures. However, it is expected that in the future data will be more available and hence this issue would have less and less impact.

5.2 Tail dependence measures

Since the distributions of the series (Q, V), (Q, D) and (V, D) are unknown, the tail dependence function $\chi(.)$ and its complementary function $\bar{\chi}(.)$ are directly evaluated via the empirical copula. This allows to assess the strength of tail dependence.

The estimators $\hat{\chi}^{\mathrm{SEC}}(.)$ (with $m = n$), $\hat{\chi}^{\mathrm{LOG}}(.)$ and $\hat{\chi}_{\mathrm{U}}^{\mathrm{FD}}$, defined respectively in (24), (26) and (27), are evaluated. First, $\hat{\chi}_{\mathrm{U}}^{\mathrm{SEC}}$ and $\hat{\chi}_{\mathrm{U}}^{\mathrm{LOG}}$ are deduced respectively from the functions $\hat{\chi}^{\mathrm{SEC}}(.)$ and $\hat{\chi}^{\mathrm{LOG}}(.)$ by noting that u is close to 1 for small k. Second, the UTDP estimators $\hat{\chi}_{\mathrm{U}}^{\mathrm{SEC}}$, $\hat{\chi}_{\mathrm{U}}^{\mathrm{LOG}}$ and $\hat{\chi}_{\mathrm{U}}^{\mathrm{FD}}$ are obtained by fixing $k = \sqrt{n}$.

The coefficient $\bar{\chi}_{\mathrm{U}}$ is deduced from the function $\bar{\chi}(.)$ which is estimated via the empirical survival copula as well as via the coefficient of tail dependence η. Hence, one uses respectively the estimator $\hat{\bar{\chi}}^{\mathrm{COLES}}$ defined in (33) and the estimator $\hat{\bar{\chi}}_{\mathrm{U},k}^{\bullet} = 2\hat{\eta}_{k}^{\bullet} - 1$ where $\hat{\eta}_{k}^{\bullet}$ is defined in (34), (35) or (37). The symbol $\bullet$ denotes one of the indices Z, H or D. Since the margins are unknown, the estimators $\hat{\eta}_{k}^{Z}$ and $\hat{\eta}_{k}^{H}$, given in (34) and (35) respectively, are computed by first transforming the margins to standard Fréchet margins. The threshold k is chosen in the simultaneous stability range of the estimators $\hat{\bar{\chi}}_{\mathrm{U},k}^{H}, \hat{\bar{\chi}}_{\mathrm{U},k}^{Z}$ and $\hat{\bar{\chi}}_{\mathrm{U},k}^{D}$ and the corresponding estimator is denoted $\hat{\bar{\chi}}_{\mathrm{U}}^{\star}$. This technique is commonly used for the estimation of the tail index or extreme quantile in EVA [*e.g.* 48]. The overall estimated dependence coefficients of Pearson's, Kendall's and Spearman's are denoted respectively by $\hat{r}_{n}$, $\hat{\tau}_{n}$ and $\hat{\rho}_{n}$ are evaluated for comparison purposes.

Tail dependence for the series (Q, V)

In the remainder of the Section, the analysis is presented first according to gauging stations and then according to the measures of dependence. Figures 4 and 5 illustrate the different estimators of tail dependence functions $\chi(.)$ and $\bar{\chi}(.)$ respectively.

Generally, Figure 4 shows that the estimators of $\chi(.)$ are not too close to 1 when u is large enough. Therefore the tail dependence is not strong in the three considered stations. Since the degree of dependence is not strong, this suggests to restrict the analysis to the univariate case only. However, as it will be seen in the obtained results by an in-depth analysis in the remainder of this section, proceeding in this manner may be inappropriate or misleading.

For ST_{050301}, (Figure 4a), all estimators of $\chi(u)$ are considerably larger than 0 for $u < 0.9$. When u is close to 1, *i.e.* $u > 0.9$, the estimators $\hat{\chi}^{\mathrm{LOG}}(u)$, $\hat{\chi}^{\mathrm{SEC}}(u)$ and the estimated UTDP $\hat{\chi}_{\mathrm{U}}^{\mathrm{DF}}$ converge to 0 abruptly with respect to u or k and the difference between these estimators and $\hat{\chi}_{\mathrm{U}}^{\mathrm{DF}}$ is large. It would have been possible to conclude erroneously that the couple (Q, V) is asymptotically dependent in station ST_{050301}. Similar results are obtained by [12] for other data. In addition, for $u = 1 - k/n \approx 0.854$, the UTDP estimators are $\hat{\chi}_{\mathrm{U}}^{\mathrm{LOG}} = 0.309$, $\hat{\chi}_{\mathrm{U}}^{\mathrm{SEC}} = 0.395$ and $\hat{\chi}_{\mathrm{U}}^{\mathrm{DF}} = 0.170$. Figure 5a indicates that all estimators of $\bar{\chi}_{\mathrm{U}}$ are significantly different from 1. For instance, they are almost stable around the interval $[0.233, 0.458]$ for $u \in (0.40, 0.57)$ whereas $\hat{\bar{\chi}}_{\mathrm{U}}^{\mathrm{COLES}} = 0.399$ for $u = 1 - k/n \approx 0.854$. Indeed, according to the properties of the coefficient $\bar{\chi}_{\mathrm{U}}$ given in subsection 2.2 and summarized in Figure 2, the peak flow and the flood volume of ST_{050301} can not be described by BEV distributions since there is no asymptotic dependence. As indicated by the paths (A) or (B'') in Figure 2, we recall that in ST_{050301}, an analysis based only on the estimators of χ_{U} does not guarantee that the couple (Q, V) is asymptotically independent. Figure 5a indicates that: *(i)* $\hat{\bar{\chi}}_{\mathrm{U}}^{\star} \approx 0.456$ and $\hat{\bar{\chi}}_{\mathrm{U}}^{\mathrm{COLES}} = 0.399$, *(ii)* $\hat{\chi}^{\mathrm{DF}} \in [0, 0.170)$, $\hat{\chi}^{\mathrm{LOG}} \in [0, 0.309)$ and $\hat{\chi}^{\mathrm{SEC}} \in [0, 0.395)$ for $0.854 < u < 1$. This leads to conclude that the extremes are positively associated, *i.e.* in ST_{050301} the observations for which both $F_1(Q) > u$ and $F_2(V) > u$ for large thresholds u occur more frequently than under exact independence between Q and V.

For ST_{080101}, (Figure 4b), for $u \in [0.7, 1)$ all estimators of $\chi(.)$ are considerably larger than 0. More precisely, for $1 - k/n \approx 0.878 \leq u < 1$, we have $\hat{\chi}^{\mathrm{LOG}} \in (0, 0.624]$, $\hat{\chi}^{\mathrm{SEC}} \in (0, 0.656]$ and $\hat{\chi}^{\mathrm{DF}} \in (0, 0.522]$. Accordingly, with respect to path (A) in Figure 2, the couple (Q, V) seems to be asymptotically dependent. On the other hand, Figure 5b indicates that the estimators $\hat{\bar{\chi}}_{\mathrm{U},k}^{H}$ and $\hat{\bar{\chi}}_{\mathrm{U},k}^{Z}$ have a regular behavior and are almost stable for large thresholds. This indicates more accurate evaluation of $\bar{\chi}$. In other respects, the tail dependence estimators $\hat{\bar{\chi}}^{\mathrm{COLES}}$ and $\hat{\bar{\chi}}_{\mathrm{U},k}^{\bullet}$ are non-negative for $0 < u < 1$ and $1 < k < n$ respectively. More precisely, $\hat{\bar{\chi}}_{\mathrm{U}}^{\mathrm{COLES}} = 0.620$ for $u \approx 0.878$ whereas the estimators $\hat{\bar{\chi}}_{\mathrm{U},k}^{\bullet}$ are located in the range $[0.327, 1]$. In particular, $\hat{\bar{\chi}}_{\mathrm{U},k}^{D}$ fluctuates slightly around $u \approx 1$ and the estimators $\hat{\bar{\chi}}_{\mathrm{U},k}^{H}$ and $\hat{\bar{\chi}}_{\mathrm{U},k}^{Z}$ are almost stable and approximately equal to 1. This suggests that $\hat{\bar{\chi}}_{\mathrm{U}}^{\star} \approx 1$ and one can then conclude that the couple (Q, V) is asymptotically dependent, see path (B') in Figure 2. In addition $\hat{\chi}_{\mathrm{U}} \in [0.502, 0.656]$ denotes the strength of this dependence, *i.e.* this dependence is slightly high.
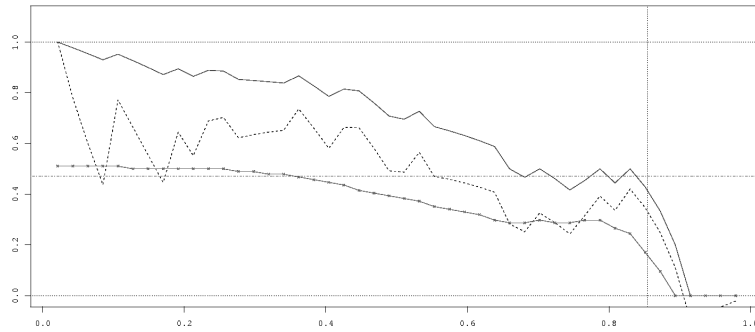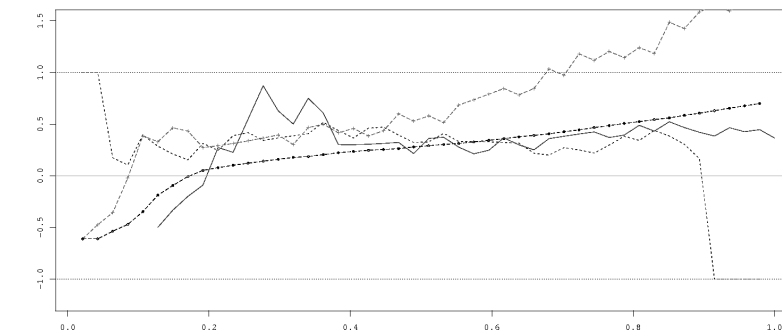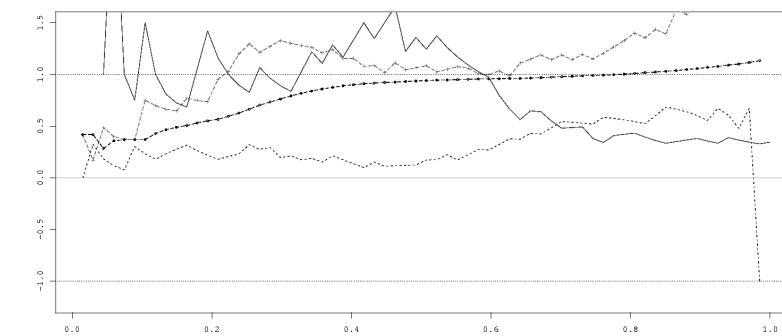
(a) Station 050301
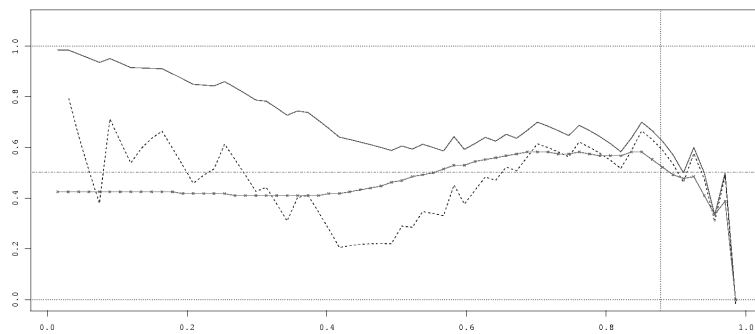
(b) Station 080101
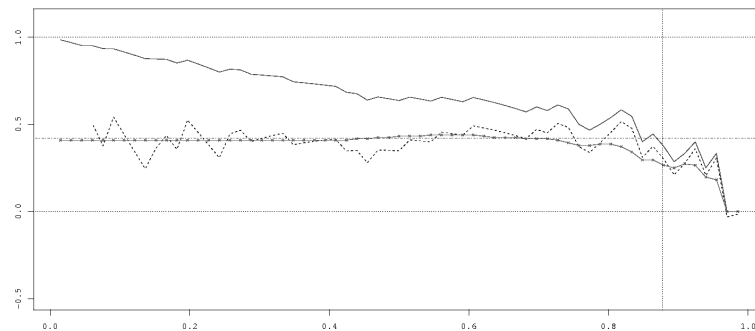
(c) Station 050119

Figure 4: Estimators of the tail dependence function $\chi(.)$ of the pair (Q, V). The vertical axis corresponds to $\hat{\chi}(u)$ with $\hat{\chi}^{\text{LOG}}(\text{- - -})$ and $\hat{\chi}^{\text{SEC}}(\text{—})$, $\hat{\chi}^{\text{DF}}_{\text{U},k}(\text{-×-×-})$ and $\hat{\chi}^{\text{F}}_{\text{U}}(\text{·····})$, while the horizontal axis corresponds to thresholds $u = 1/n, 2/n, 3/n, \ldots, 1 - 1/n$ and $k = n - 1, n - 2, \ldots, 1$. In horizontal dotted line, the upper and lower bounds on $\chi(.)$. In vertical dotted line, the chosen threshold.

For ST_{050119}, Figure 4c suggests that Q and V are asymptotically dependent since the estimators of $\chi(.)$ are inside $(0, 1)$ with $\hat{\chi}^{\text{SEC}}_{\text{U}} = 0.400$, $\hat{\chi}^{\text{LOG}}_{\text{U}} = 0.330$, $\hat{\chi}^{\text{DF}}_{\text{U}} = 0.265$ and $\hat{\chi}^{\text{F}}_{\text{U}} = 0.407$ for $u \approx 0.877$. Nevertheless, Figure 5c invalidates clearly these finding since $\hat{\tilde{\chi}}^{\bullet}_{\text{U},k}$ are located in the range $[0.345, 1]$. On the other hand, from $\hat{\tilde{\chi}}^{\text{COLES}}_{\text{U}}$ and $\hat{\tilde{\chi}}^{\text{D}}_{\text{U},k}$, one deduces that $\hat{\tilde{\chi}}^{\star}_{\text{U}} = 0.345 \neq 1$. In fact, according Figure 2, for the couple (Q, V), Figures 5c

(a) Station 050301

(b) Station 080101
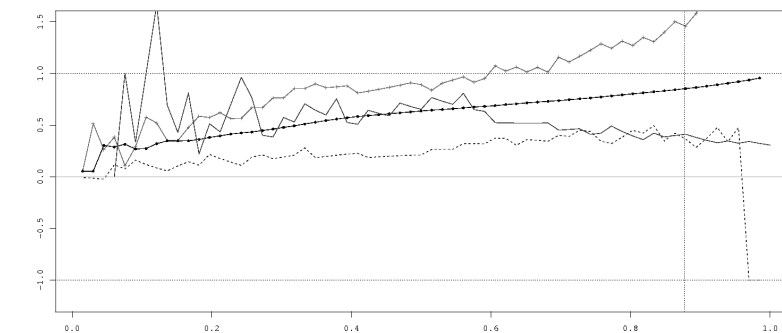
(c) Station 050119

Figure 5: Estimators of the tail dependence function $\bar{\chi}(.)$ of the pair (Q, V). The vertical axis corresponds to $\hat{\bar{\chi}}(u)$ with $\hat{\bar{\chi}}^{\text{COLES}}$(- - -), $\hat{\bar{\chi}}^{D}_{U,k}$(—), $\hat{\bar{\chi}}^{H}_{U,k}$(-+-+-) and $\hat{\bar{\chi}}^{Z}_{U,k}$(-○-○-), while the horizontal axis corresponds to threshold $u = k/n$ with $k = 1, \ldots, n - 1$. In horizontal dotted line, the upper and lower bounds on $\bar{\chi}(.)$.

and 4c do not allow to conclude since the tail dependence estimators do not satisfy ($\hat{\bar{\chi}}_{U} \in [-1, 1), \hat{\chi}_{U} = 0$) or ($\hat{\bar{\chi}}_{U} = 1, \hat{\chi}_{U} \in (0, 1]$).

The estimated overall dependence coefficients, the estimated tail dependence parameters and the estimated tail dependence functions for $u \in (1 - k/n, 1)$ are summarized in Table 2. Table 2 shows that, in all three stations, the overall coefficients lead to conclude that there are significant correlations which are not very high.

Table 2: Couple (Q, V). Estimated tail dependence measures $\hat{\chi}^{\text{SEC}}, \hat{\chi}^{\text{LOG}}, \hat{\chi}^{\text{DF}}_{\text{U}}, \hat{\bar{\chi}}^{\text{COLES}}$ for $u \in [1 - \sqrt{n}/n, 1)$ and $\hat{\chi}^{\text{F}}_{\text{U}}, \hat{\chi}^{\star}_{\text{U}}$ and $\hat{\eta}$. Estimated overall correlation coefficients $\hat{r}_n$, $\hat{\tau}_n$ and $\hat{\rho}_n$.

Station	$\hat{\chi}^{\text{SEC}}$	$\hat{\chi}^{\text{LOG}}$	$\hat{\chi}^{\text{DF}}_{\text{U}}$	$\hat{\bar{\chi}}^{\text{COLES}}$	$\hat{\chi}^{\text{F}}_{\text{U}}$	$\hat{\chi}^{\star}_{\text{U}}$	$\hat{\eta}$	$\hat{r}_n$	$\hat{\tau}_n$	$\hat{\rho}_n$
050301	[0, 0.395]	[0, 0.309]	[0, 0.170]	[−1, 0.399]	0.471	0.30	0.65	0.455	0.418	0.547
080101	(0, 0.656]	(0, 0.624]	(0, 0.522]	(−1, 0.620]	0.503	1.00	1.00	0.660	0.403	0.562
050119	[0, 0.400]	[0, 0.330]	[0, 0.265]	(0, 0.354]	0.407	0.345	0.673	0.521	0.360	0.511

Nevertheless, as concluded previously on the basis of tail dependence measures, in ST_{050301} the extremes are asymptotically independent. Thus, one observes that an analysis solely based on the overall dependence coefficients does not give enough information to reflect the nature of the relationship between extremes of the couple (Q, V) in ST_{050301}, ST_{080101} and ST_{050119}.

As previously mentioned (see Section 3), all BEV distributions are asymptotically dependent. Since it was concluded that the couple (Q, V) in ST_{080101} is asymptotically dependent, one of the BEV distributions could be a candidate for the sample of (Q, V), for instance, an extreme value copula or the Gumbel-Hougaard family. The bivariate tests presented in subsection 3.1 are used to check this. Results are provided in Table 3. The Gumbel-Hougaard copula family is commonly used for hydrological FA [*e.g.* 32]. Notice that for a given degree of dependence, the most popular extreme value copulae are strikingly similar [*e.g.* 30]. The tests require estimating the Pickands dependence function $A(.)$.

Table 3: Couple (Q, V). Estimators based on the Pickands dependence function $\hat{\chi}^{\text{P}}$ and $\hat{\chi}^{\text{CFG}}$. The p-values of the bivariate tests of extreme-value dependence based on (i) Kendall's process (pv.jac, pv.fsa, pv.asy), (ii) the Pickands dependence function (pv.ky), (iii) of the p-values goodness-of-fit test for extreme-value copulae (pv.mpl, pv.iτ, pv.iρ).

Station	$\hat{\chi}^{\text{P}}$	$\hat{\chi}^{\text{CFG}}$	pv.jac	pv.fsa	pv.asy	pv.ky	pv.mpl	pv.iτ	pv.iρ
050301	[0.570, 0.610]	[0.471, 0.516]	0.010	0.004	0.001	0.011	0.006	0.039	0.127
080101	[0.468, 0.503]	[0.503, 0.538]	0.863	0.850	0.834	0.372	0.277	0.309	0.356
050119	[0.387, 0.426]	[0.419, 0.456]	0.526	0.498	0.467	0.294	0.898	0.826	0.811

For the Ghoudi et al. [33] bivariate test, the approximative p-values obtained by jackknife, the finite sample plug-in and the asymptotic plug-in are noted pv.jac, pv.fsa and pv.asy respectively. For the Kojadinovic and Yan [41] bivariate test, the p-value is denoted by pv.ky. For the Genest et al. [30] goodness-of-fit tests, pv.mpl, pv.itau and pv.irho are the approximative p-values obtained respectively by using parametric bootstrap combined with the maximum pseudo-likelihood, the method of the inversion of Kendall's tau and the method of the inversion of Spearman's rho. In Figure 6, we present for $t \in [0, 1]$ the rank-based estimators $\hat{A}^{\text{P}}_{n,r}(t)$ and $\hat{A}^{\text{CFG}}_{n,r}(t)$ defined in (29), and the corresponding corrected endpoint estimators noted $\hat{A}^{\text{P}}_{n,c}(t)$ and $\hat{A}^{\text{CFG}}_{n,c}(t)$ [67]. The estimated UTDP $\hat{\chi}^{\text{P}}_{\text{U}}$ and $\hat{\chi}^{\text{CFG}}_{\text{U}}$ which are related, via (16), to the dependence function $A(.)$ and the obtained p-values of all bivariate statistical tests of extreme value dependence are summarized in Table 3. Notice that an analysis based on Pickands dependence function or the UTDP estimators lead to the same findings.

The analysis on tail dependence function $\chi(.)$ and its complementary function $\bar{\chi}(.)$ allow to conclude that the couple (Q, V) is asymptotically independent for ST_{050301}. To consolidate this finding, the function $A(.)$ is estimated and the p-values of the bivariate tests used previously are computed. Figure 6a suggests that in ST_{050301} the couple (Q, V) is asymptotically dependent since via (16) we have $\hat{\chi}^{\text{P}}_{\text{U}} \in [0.570, 0.610]$ and $\hat{\chi}^{\text{CFG}}_{\text{U}} \in [0.471, 0.516]$ which are lower than 1. Nevertheless, as shown in the previous analysis based on tail dependence function $\chi(.)$ and its complementary function $\bar{\chi}(.)$, this represents only a graphical indication. In fact, the bivariate statistical tests in Table 3 confirm that we can not model the couple (Q, V) by a BEV
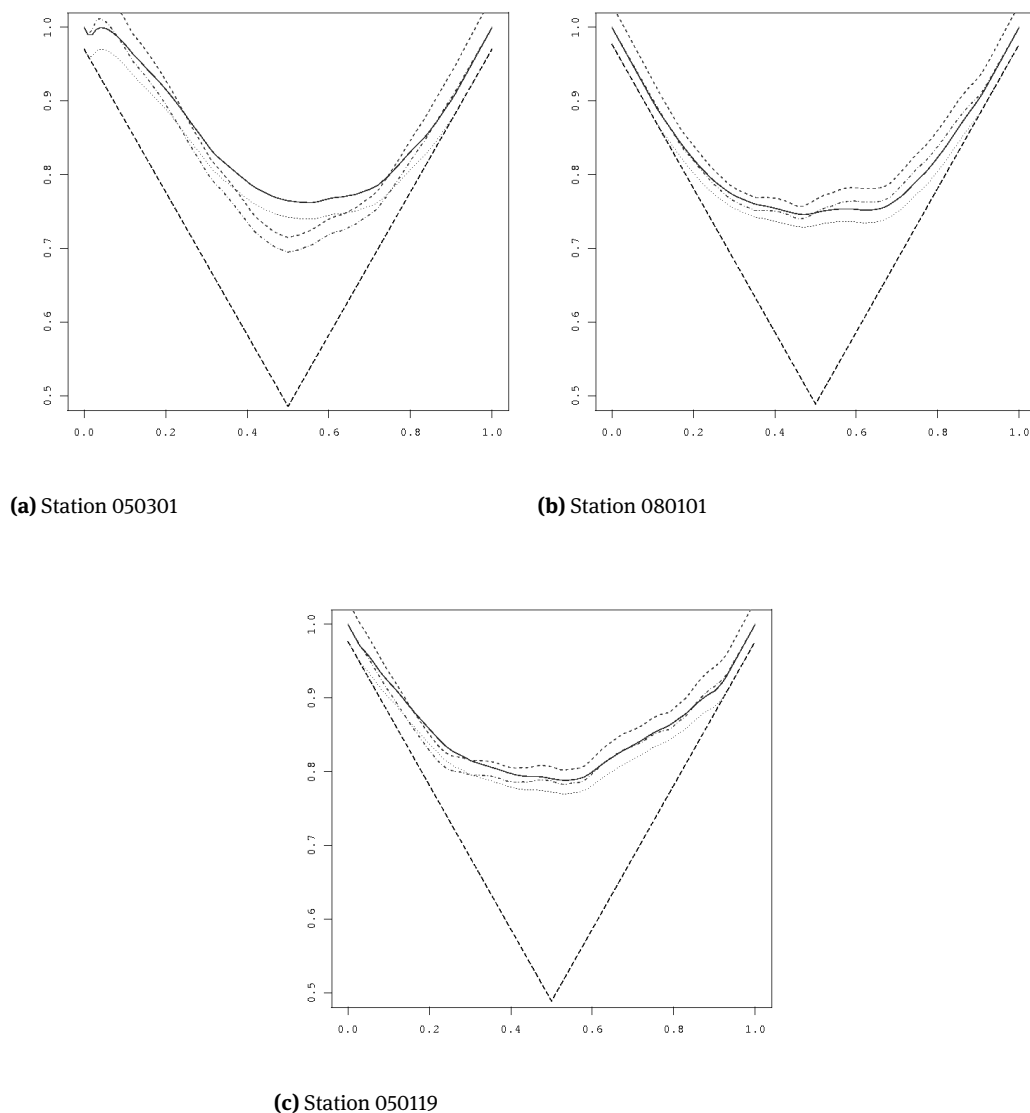
(a) Station 050301

(b) Station 080101

(c) Station 050119

Figure 6: Estimators of the Pickands dependence function $A(t)$ for $t \in [0, 1]$: corrected endpoint estimators $\hat{A}_{n,c}^{\mathrm{P}}(t)$ (·····), $\hat{A}_{n,c}^{\mathrm{CFG}}(t)$ (—) and rank-based estimators $\hat{A}_{n,r}^{\mathrm{P}}(t)$ (- - - -) and $\hat{A}_{n,r}^{\mathrm{CFG}}(t)$ (...).

distribution and especially by the Gumbel-Hougaard family copula since pv.jac, pv.fsa, pv.asy, pv.ky, pv.mpl and pv.itau are lower than 0.05.

Figure 6b indicates that in ST_{080101}, Q and V are asymptotically dependent since via the relationship (16) we have $\hat{\chi}_{\mathrm{U}}^{\mathrm{P}} \in [0.468, 0.503]$ and $\hat{\chi}_{\mathrm{U}}^{\mathrm{CFG}} \in [0.503, 0.538]$ which are lower than 1. The p-values of all bivariate tests of extreme value dependence summarized in Table 3 are higher than 0.05 which confirms a good fit with the BEV distributions. Then, for ST_{080101}, the dependence of (Q, V) can be modelled with the Gumbel-Hougaard family copula. In addition, from $\hat{\chi}_{\mathrm{U}}^{\mathrm{P}}$ and $\hat{\chi}_{\mathrm{U}}^{\mathrm{CFG}}$, one deduces that the degree of dependence between Q and V is within the interval $[0.468, 0.503]$. Notice that even though this degree of dependence is slightly lower than the previous values, *i.e.* $\hat{\chi}_{\mathrm{U}} \in [0.502, 0.656]$, where no assumption on the model was made, the same conclusion is obtained: *i.e.* there is asymptotic dependence.

In Figure 6c, the indication graph suggests an asymptotic dependence between Q and V in ST_{050119} since $\hat{\chi}_{\mathrm{U}}^{\mathrm{P}} \in [0.387, 0.426]$ and $\hat{\chi}_{\mathrm{U}}^{\mathrm{CFG}} \in [0.419, 0.456]$ are lower than 1. Moreover the bivariate statistical tests

in Table 3 confirm this graphical indication since the obtained p-values are higher than 0.05. However, this finding is not compatible with this result $\hat{\bar{\chi}}_U^\star = 0.345 \neq 1$ which means that there is asymptotic independence. This could be explained on the basis of construction of the tests used. Indeed, the tests used are based only on the function $A(.)$ and not on the tail-dependence measure $\bar{\chi}(.)$. In addition, the p-value is a measure of the evidence against the null hypothesis: the smaller the p-value, the stronger the evidence against the null hypothesis. A large p-value is not strong evidence in favour of null hypothesis. A large p-value can occur for two reasons: *(i)* null hypothesis is true or *(ii)* null hypothesis is false but the test has low power. The p-value is not the probability that the null hypothesis is true [see 73, p.157].

Table 4: Estimated tail dependence parameters $\hat{\chi}_U^{SEC}, \hat{\chi}_U^{LOG}, \hat{\chi}_U^{DF}$ for $u = 1 - \sqrt{n}/n$, $\hat{\chi}^{COLES}$ for $u \in [1 - \sqrt{n}/n, 1)$ and $\hat{\bar{\chi}}_U^\star$. Estimated overall correlation coefficients $\hat{r}_n$, $\hat{\tau}_n$ and $\hat{\rho}_n$.

Station	Couple	$\hat{\chi}_U^{SEC}$	$\hat{\chi}_U^{LOG}$	$\hat{\chi}_U^{DF}$	$\hat{\bar{\chi}}_U^\star$	$\hat{\bar{\chi}}^{COLES}$	$\hat{r}_n$	$\hat{\tau}_n$	$\hat{\rho}_n$
050301	(Q, D)	−0.041	−0.150	0	0	−1	−0.256	−0.112	−0.164
	(V, D)	0.249	0.130	0.255	0.5	$(-1, 0.250)$	0.507	0.301	0.460
080101	(Q, D)	−0.045	−0.094	0	−0.13	−1	−0.389	−0.239	−0.334
	(V, D)	0.045	−0.094	0	0	−1	0.364	0.335	0.468
050119	(Q, D)	0.092	0	0	0.050	$[-1, 0)$	0.361	0.259	0.373
	(V, D)	0.400	0.330	0.182	0.220	$[-1, 0.355]$	−0.450	−0.302	−0.445

Tail dependence for the series (Q, D) and (V, D)

In Figures 7 and 8, we respectively present the estimators of tail dependence functions $\chi(.)$ and $\bar{\chi}(.)$ for the pairs (Q, D) and (V, D) for each station. The estimated tail dependence parameters deduced from Figures 7 and 8 and the estimated overall correlation coefficients are summarized in Table 4.

Figures 7a, b, Figures 8a, b and Table 4 show that Q and D are asymptotically independent in both ST_{050301} and ST_{080101}. More precisely, in ST_{050301} we observe near independence since $\hat{\chi} \approx 0$ and $\hat{\bar{\chi}}_U^\star \approx 0$ whereas $\hat{\bar{\chi}}^{COLES} \approx 0$ for $0 < u < 0.854$ and $\hat{\bar{\chi}}^{COLES} = -1$ for $0.854 \leq u < 1$.

Figures 7c and 8c indicate that Q and D are asymptotically independent in ST_{050119} since $\hat{\chi} \approx 0$. Based on the estimators $\hat{\bar{\chi}}_{U,k}^D$ and $\hat{\bar{\chi}}^{COLES}$, one can deduce that the association between Q and D is negative. Otherwise, one can deduce a positive association when the analysis is only based on the estimators $\hat{\bar{\chi}}_{U,k}^H$ and $\hat{\bar{\chi}}_{U,k}^Z$.

Nevertheless following Figure 2, an analysis of Figures 7d and 8d does not allow to conclude for an asymptotic independence between V and D in ST_{050301} since the estimators of $\chi(.)$ are clearly inside $(0, 1)$ (see also Table 4).

In addition, the estimators of $\bar{\chi}(.)$ are non-negative with $\hat{\bar{\chi}}_U^\star \approx 0.5$ and $\hat{\bar{\chi}}^{COLES}$ is in the range $(0.220, 0.250)$ for $0.854 \leq u < 1$. Figures 7e, 8e and Table 4 suggest that V and D are asymptotically dependent in ST_{080101}. Figures 7f and 8f do not allow to conclude for an asymptotic dependence between V and D in ST_{050119} since the estimators of $\chi(.)$ converge to values around 0.340 for $u \leq 1 - k/n \approx 0.877$ whereas, the estimators of $\bar{\chi}(.)$ are clearly unstable and significantly different from 1 with $\hat{\bar{\chi}}^{COLES} \approx 0.355$ and $\hat{\bar{\chi}}_{U,k}^D \approx 0.220$ for $u \approx 0.877$. When $u > 0.877$, the estimators of $\chi(.)$ converge abruptly to 0.

In addition, as the UTDP estimators are in the range $(0.182, 0.400)$ for $u \approx 0.877$ (see Table 4), one might conclude at best for an asymptotic independence with positive association since the estimators of $\bar{\chi}(.)$ are non-negative for $u \in (0.2, 0.95)$.

In Table 4, the obtained estimated overall coefficients $(\hat{r}_n, \hat{\tau}_n, \hat{\rho}_n)$ lead simply to conclude that a significant overall correlation exists in all cases. This shows once again that the hydrological analyses based on overall coefficients are inadequate to quantify the extreme risks that occur at the tail of the distribution. More-
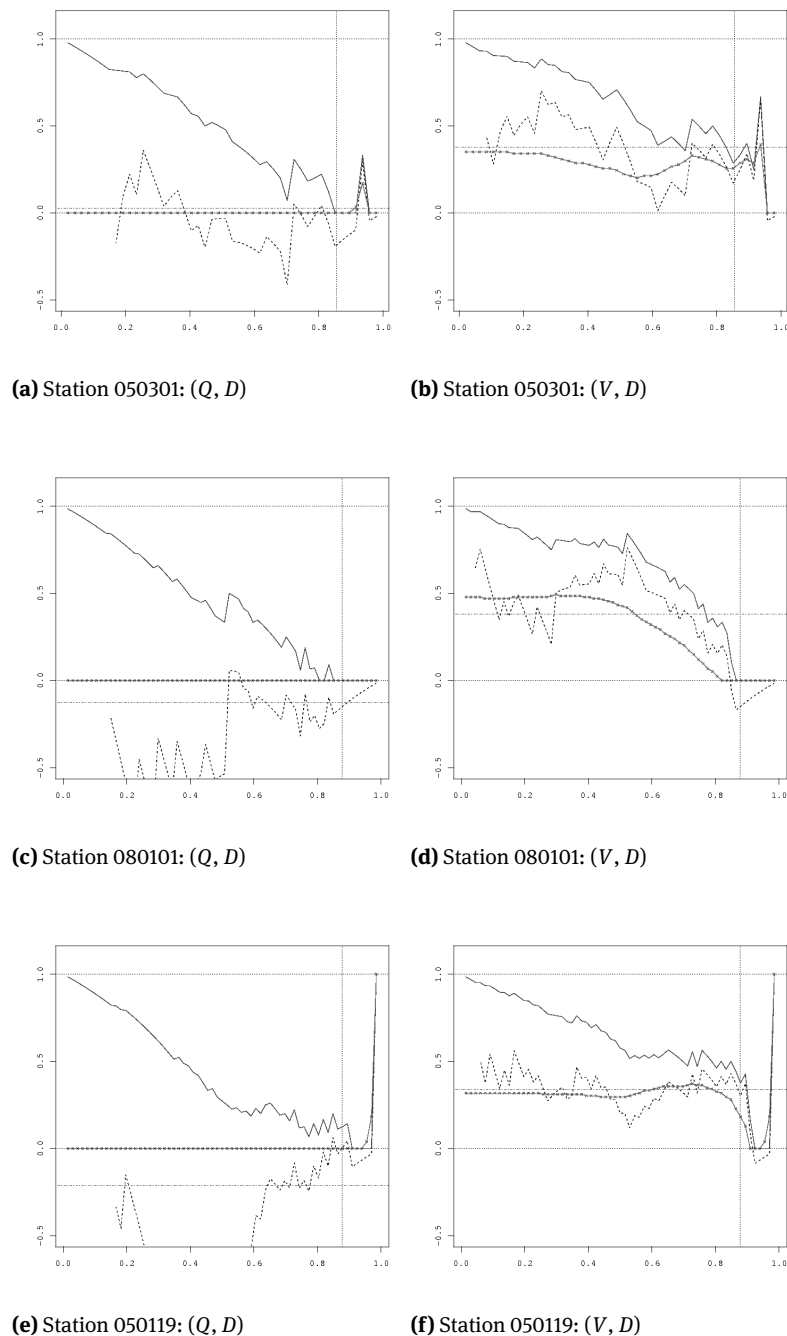
(a) Station 050301: (Q, D) **(b)** Station 050301: (V, D)

(c) Station 080101: (Q, D) **(d)** Station 080101: (V, D)

(e) Station 050119: (Q, D) **(f)** Station 050119: (V, D)

Figure 7: Estimators of the tail dependence function $\chi(.)$ of the pairs (Q, D) and (V, D). The vertical axis corresponds to $\hat{\chi}(u)$ with $\hat{\chi}^{\mathrm{LOG}}$(- - - -) and $\hat{\chi}^{\mathrm{SEC}}$(—), $\hat{\chi}^{\mathrm{DF}}_{\mathrm{U},k}$(-×-×-) and $\hat{\chi}^{\mathrm{F}}_{\mathrm{U}}$(·····), while the horizontal axis corresponds to thresholds $u = 1/n, 2/n, 3/n, \ldots, 1 - 1/n$ and $k = n - 1, n - 2, \ldots, 1$. In horizontal dotted line, the upper and lower bounds on $\chi(.)$. In vertical dotted line, the chosen threshold.

over, this case study shows that the measure $\chi(.)$ alone is not always sufficient to exhibit the relationship between the extremes. Notice that the overall coefficients for the couple (Q, D) in ST_{050119} are negative. This could indicate that the couple (Q, D) is in fact negatively associated which is not in accordance with the last finding based on tail dependence measures where the sign of the association was not clear.
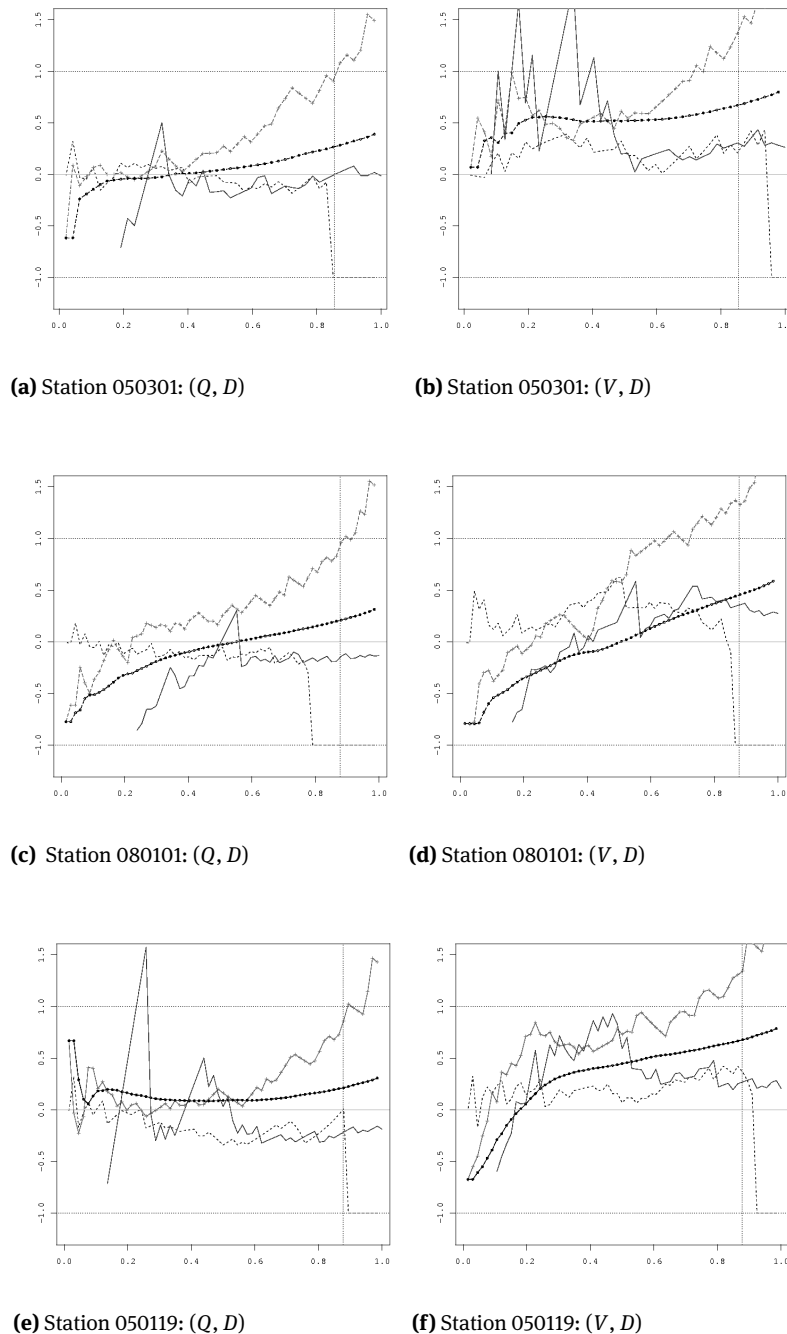
(a) Station 050301: (Q, D) **(b)** Station 050301: (V, D)

(c) Station 080101: (Q, D) **(d)** Station 080101: (V, D)

(e) Station 050119: (Q, D) **(f)** Station 050119: (V, D)

Figure 8: Estimators of the tail dependence function $\bar{\chi}(.)$ of the pairs (Q, D) and (V, D). The vertical axis corresponds to $\hat{\bar{\chi}}(u)$ with $\hat{\bar{\chi}}^{COLES}$(- - -), $\hat{\bar{\chi}}^{D}_{U,k}$(—), $\hat{\bar{\chi}}^{H}_{U,k}$(-+-+-) and $\hat{\bar{\chi}}^{Z}_{U,k}$(-○-○-), while the horizontal axis corresponds to threshold $u = k/n$ with $k = 1, \ldots, n-1$. In horizontal dotted line, the upper and lower bounds on $\bar{\chi}(.)$.

The summarized results in Table 5 show that the couple (Q, V) is asymptotically independent and positively associated in the Bastican River (ST_{050301}). In ST_{050301}, similar to the Gaussian dependence model studied by Coles et al. [12], one can not conclude using only χ_U. On the other hand, the overall coefficient values indicate that (Q,V) are relatively highly correlated. However, the conclusion on the tail is not the same where it is shown that the couple (Q,V) for this station is asymptotically independent. In the Harri-

Table 5: Summary of all the results. The case ⋆ means that either ($\bar{\chi}_U \in [-1, 1), \chi_U = 0$) or ($\bar{\chi}_U = 1, \chi_U \in (0, 1]$) is not fulfilled.

Station	Couple	Overall correlation (sign)	Tail dependence
050301	(Q, V)	relatively high (+)	asymptotic independence & positive association
	(Q, D)	very low (-)	near independence
	(V, D)	moderate (+)	no conclusion ⋆
080101	(Q, V)	high (+)	asymptotic dependence
	(Q, D)	low (-)	asymptotic independence & negative association
	(V, D)	low (+)	asymptotic dependence
050119	(Q, V)	relatively high (+)	no conclusion ⋆
	(Q, D)	moderate (-)	asymptotic independence & negative association
	(V, D)	low (+)	asymptotic independence & positive association

cana River (ST_{080101}), (Q, V) are asymptotically dependent and can be fitted with a BEV distribution such as the Gumbel-Hougaard copula family. In the Mattawin River (ST_{050119}) one can not conclude on the tail dependence whereas the overall coefficients indicate that (Q, V) are relatively highly correlated. In the Batiscan River, the couple (Q, D) is near independent. In the Harricana River, the couple (D, V) is asymptotically independent whereas in the Mattawin River it is asymptotically dependent with negative association. The couple (V, D) is shown to be asymptotically dependent in the Harricana River. In the Bastican River, it is difficult to conclude since the estimated UTDP's are approximately in the range $[0.130, 0.255]$ whereas, its complementary coefficients are approximately in $[0.220, 0.500]$. In the Mattawin River, it seems that V and D are asymptotically independent with positive association.

6 Conclusions

In the present paper, a number of dependence measures, which are more adapted to the treatment of bivariate extreme events in the hydrological BFA framework, are introduced. These measures focus on the tail of bivariate distributions. The statistical properties of these measures are reviewed and associated non-parametric estimations are provided.

The overall coefficients Kendall's τ, Spearman's ρ and Pearson's r do not give enough information to reflect the nature of the relationship between extremes. They do not allow to study the concomitant occurrence of extreme values since they do not attribute, for instance, sufficient weight to the extreme values. They are more adapted to the center body of the distribution. Furthermore, the tail dependence measure χ_U, employed solely in some hydrological studies, does not allow to conclude in all cases. For instance, it is not appropriate for discriminating variables that are asymptotically independent. Hence, it is recommended to consider the complementary measure $\bar{\chi}_U$. This one quantifies the strength of dependence within the class of asymptotically independent variables in bivariate extremes.

An application of the presented tail dependence measures is carried out on three gauging stations located in the province of Quebec (Canada). The application deals with flood peak (Q), volume (V) and duration (D). According to the proposed approach, it can be concluded that the couple (Q, D) is near independent in the Batiscan River whereas in the Harricana River, it is asymptotically independent.

List of symbols or abbreviations

$A(.)$ Pickands dependence function

$C(.,.)$ Copula function associated to $F(.,.)$

$F(.,.)$ Cumulative distribution function

$F_1(.)$ and $F_2(.)$ Marginal distributions of $F(.,.)$

$G(.,.)$ Bivariate extreme value distribution

$L(.)$ Univariate slowly varying function at infinity

$\Lambda(.,.)$ Tail copula function (TCF)

$\bar{C}(.,.)$ Survival copula function associated to $C(.,.)$, *i.e.* $\bar{C}(u,v) = 1 - u - v + C(u,v) \;\; \forall\, u,v \in [0,1]$

$\bar{\chi}(.)$ Tail dependence function :: complementary measure of $\chi(.)$

$\bar{\chi}_{\text{U}}$ Complementary tail dependence parameter of χ_{U}

$\chi(.)$ Tail dependence function :: quantile-dependent measure of dependence

χ_{U} Upper tail dependence parameter (UTDP)

$\ell(.,.)$ Stable tail dependence function (STDF)

η Coefficient of tail dependence

$\mathcal{C}(.,.)$ Copula of extreme value

BEV   Bivariate Extreme Value

BFA   Bivariate Frequency Analysis

EVA   Extreme Value Analysis

References

[1] Ané, T. and C. Kharoubi (2003). Dependence structure and risk measure. *J. Business* 76, 411–438.

[2] Bacro, J. N. (2005). Dépendance des extrêmes. Tech. rep., Institut de Mathématiques et de Modélisation de Montpellier.

[3] Beirlant, J., Y. Goegebeur, J. Teugels, and J. Segers (2004). *Statistics of Extremes: Theory and Applications*. John Wiley and Sons, Chichester.

[4] Ben Aissia, M.-A., F. Chebana, T. B. M. J. Ouarda, L. Roy, G. Desrochers, I. Chartier, and E. Robichaud (2012). Multivariate analysis of flood characteristics in a climate change context of the watershed of the Baskatong reservoir, Province of Québec, Canada. *Hydrol. Process.* 26(1), 130–142.

[5] Bingham, N. H., C. M. Goldie, and J. L. Teugels (1987). *Regular Variation*. Cambridge University Press.

[6] Bobée, B. and F. Ashkar (1991). *The Gamma Family and Derived Distributions Applied in Hydrology*. Water Resources Publications, Colorado.

[7] Bouchouicha, R. (2010). Dépendance entre risques extrêmes: application aux hedge funds. Tech. Rep. 1013, Groupe d'Analyse et de Théorie Economique (GATE), Centre national de la recherche scientifique (CNRS), Université Lyon 2, Ecole Normale Supérieure.

[8] Capéraà, P., A.-L. Fougères, and C. Genest (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* 84(3), 567–577.

[9] Chebana, F. (2013). Multivariate analysis of hydrological variables. In *Encyclopedia of Environmetrics*, Second Edition, A. H. El-Shaarawi and W. W. Piegorsch eds., 1676–1681. John Wiley and Sons, Chichester.

[10] Chebana, F. and T. Ouarda (2011). Depth-based multivariate descriptive statistics with hydrological applications. *J. Geophys. Res.* 116(D10), 1–19.

[11] Chebana, F. and T. Ouarda (2011). Multivariate quantiles in hydrological frequency analysis. *Environmetrics* 22(1), 63–78.

[12] Coles, S., J. Heffernan, and J. Tawn (1999). Dependence measures for extremes value analyses. *Extremes* 2(4), 339–365.

[13] Cont, R. (2009). La statistique face aux événements rares. *Pour la science* 385, 116–123.

[14] Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Acad. Roy. Belg. Bull. Cl. Sci.* 5(65), 274–292.

[15] Dobric, J. and F. Schmid (2005). Nonparametric estimation of the lower tail dependence λL in bivariate copulas. *J. Appl. Statist.* 32, 387–407.

[16] Draisma, G., H. Drees, A. Ferreira, and L. de Haan (2004). Bivariate tail estimation: dependence in asymptotic independence. *Bernoulli* 10(2), 251–280.

[17] Drees, H. and X. Huang (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *J. Multivariate Anal.* 64(1), 25–47.

[18] Dupuis, D. J. (2007). Using copulas in hydrology: benefits, cautions, and issues. *J. Hydrol. Eng.* 12(4), 381–393.

[19] Durante, F., J. Fernández-Sánchez, and R. Pappadà (2015). Copulas, diagonals, and tail dependence. *Fuzzy Set. Syst.* 264, 22–41.

[20] Einmahl, J. H. J., A. Krajina, and J. Segers (2008). A method of moments estimator of tail dependence. *Bernoulli* 14, 1003–1026.

[21] Embrechts, P., A. McNeil, and D. Straumann (1999). Correlation: Pitfalls and alternatives. *Risk Magazine* May, 69–71.

[22] Embrechts, P., A. Mcneil, and D. Straumann (2002). Correlation and dependence in risk management: Properties and pitfalls. In *Risk Management: Value at Risk and Beyond*, M.A.H. Dempster ed., Cambridge University Press, 176-223.

[23] Fischer, M. and M. Dörflinger (2006). A note on a non-parametric tail dependence estimator. Discussion Papers 76/2006, Friedrich-Alexander-University Erlangen-Nuremberg.

[24] Frahm, G., M. Junker, and R. Schmidt (2005). Estimating the tail-dependence coefficient: properties and pitfalls. *Insurance Math. Econ.* 37(1), 80–100.

[25] Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*. Wiley, New York.

[26] Geffroy, J., 1958: Contribution à la théorie des valeurs extrêmes. *Public. Inst. Stat. Paris 7* 7, 37–121.

[27] Geffroy, J. (1959). Contribution à la théorie des valeurs extrêmes. *Public. Inst. Stat. Paris 7* 8, 123–184.

[28] Genest, C. and F. Chebana (2015). Copula modeling in hydrological frequency analysis. To appear in *Chow's Handbook of Applied Hydrology, Second Edition*, McGraw-Hill.

[29] Genest, C. and A.-C. Favre (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *J. Hydrol. Eng.* 12(4), 347–368.

[30] Genest, C., I. Kojadinovic, J. Nešlehová, and J. Yan (2010). A goodness-of-fit test for bivariate extreme-value copulas. *Bernoulli* 17(1), 253–275.

[31] Genest, C. and J. Segers (2009). Rank-based inference for bivariate extreme-value copulas. *Ann. Stat.* 37, 2990–3022.

[32] Ghosh, S. (2010). Modelling bivariate rainfall distribution and generating bivariate correlated rainfall data in neighbouring meteorological subdivisions using copula. *Hydrol. Process.* 24(24), 3558–3567.

[33] Ghoudi, K., A. Khoudraji, and L.-P. Rivest (1998). Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *Canad. J.Stat.* 26(1), 187–197.

[34] Heffernan, J. (2000). A directory of coefficients of tail dependence. *Extremes* 3(3), 279–290.

[35] Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Stat.* 3(5), 1163–1174.

[36] Huang, X. (1992). *Statistics of Bivariate Extreme Values*. Ph.D. thesis, Erasmus University Rotterdam.

[37] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall, London.

[38] Joe, H., R. L. Smith, and I. Weissman (1992). Bivariate threshold methods for extremes. *J. R. Stat. Soc. Ser. B* 54(1), 171–183.

[39] Juri, A. and M. Wuthrich (2002). Copula convergence theorems for tail events. *Insurance Math. Econ.* 30(3), 405–420.

[40] Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika* 30(1/2), 81–93.

[41] Kojadinovic, I. and J. Yan (2010). Nonparametric rank-based tests of bivariate extreme-value dependence. *J. Multivariate Anal.* 101, 2234–2249.

[42] Kratz, M. and S. Resnick (1996). The QQ-estimator and heavy tails. *Stochastic Models* 12(4), 699–724.

[43] Ledford, A. W. and J. A. Tawn (1996). Statistics for near independence in multivariate extreme values. *Biometrika* 83(1), 169–187.

[44] Ledford, A. W. and J. A. Tawn (1997). Modelling dependence within joint tail regions. *J. R. Stat. Soc. Ser. B* 59(2), 475–499.

[45] Ledford, A. W. and J. A. Tawn (1998). Concomitants of extremes. *Adv. Appl. Probab.* 30(1), 197–215.

[46] Lee, T., R. Modarres, and T. Ouarda (2013). Data based analysis of bivariate copula tail dependence for drought duration and severity. *Hydrol. Process.* 27(10), 1454–1463.

[47] Lekina, A., (2010). *Estimation Non-paramétrique des Quantiles Extrêmes Conditionnels*. Ph.D. thesis, Université de Grenoble.

[48] Lekina, A., F. Chebana, and T. Ouarda (2014). Weighted estimate of extreme quantile: an application to the estimation of high flood return periods. *Stoch. Env. Res. Risk. A.* 28(2), 147–165.

[49] Longin, F. M. and B. Solnik (2001). Extreme correlations of international equity markets during extremely volatile periods. *J. Finance* 56, 649–676.

[50] Malevergne, Y. and D. Sornette (2004). Investigating extreme dependences. *Review Financ. Studies*, in press.

[51] Nelsen, R. B. (2006). *An Introduction to Copulas*, Second edition, Springer-Verlag New York.

[52] Ouarda, T. B. M. J., M. Haché, P. Bruneau, and B. Bobée (2000). Regional flood peak and volume estimation in northern canadian basin. *J. Cold Reg. Eng.* 14(4), 176–191.

[53] Patton, A. (2006). Estimation of multivariate models for time series of possibly different lengths. *J. Appl. Econom.* 21(2), 147–173.

[54] Peng, L. (1999). Estimation of the coefficient of tail dependence in bivariate extremes. *Stat. Probabil. Lett.* 43(4), 399–409.

[55] Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Stat.* 3(1), 119–131.

[56] Pickands, J. (1981). Multivariate extreme value distributions. *Bull. Inst. Internat. Statist.* 49(2), 859–878.

[57] Poon, S.-H., M. Rockinger, and J. Tawn (2004). Extreme value dependence in financial markets: diagnostics, models, and financial implications. *Review Financ. Studies* 17(2), 581–610.

[58] Poulin, A., D. Huard, A.-C. Favre, and S. Pugin (2007). Importance of tail dependence in bivariate frequency analysis. *J. Hydrol. Eng.* 12(4), 394–403.

[59] Raynal-Villasenor, J. and J. Salas (1987). Multivariate extreme value distributions in hydrological analyses. In *Water for the Future: Hydrology in Perspective*, Rodda, J. C. and N.C. Matalas eds., 111–119.

[60] Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Process.* Springer Verlag, New-York.

[61] Rohatgi, V. (1976). *An Introduction to Probability Theory and Mathematical Statistics*. John Wiley & Sons, New York.

[62] Salvadori, G. and C. De Michele (2011). Estimating strategies for multiparameter multivariate extreme value copulas. *Hydrol. Earth Syst. Sc.* 15(1), 141–150.

[63] Salvadori, G., C. De Michele, N. T. Kottegoda, and R. Rosso (2007). *Extremes in Nature: an Approach Using Copulas*. Springer, Dordrecht.

[64] Schmidt, R. (2002). Tail dependence for elliptically contoured distributions. *Math. Methods Oper. Res.* 55, 301–327.

[65] Schmidt, R. and U. Stadtmuller (2006). Non-parametric estimation of tail dependence. *Scand. J. Statist.* 33, 307–335.

[66] Schultze, J. and J. Steinebach (1996). On least squares estimates of an exponential tail coefficient. *Stat. Decisions* 14(3), 353–372.

[67] Segers, J. (2007). Non-parametric inference for bivariate extreme-value copulas. In *Topics in Extreme Values*, M. Ahsanullah and S. N. U. A. Kirmani eds., Nova Science Publishers Inc, New York, 181–203.

[68] Serinaldi, F. (2008). Analysis of inter-gauge dependence by Kendall's τ_K, upper tail dependence coefficient, and 2-copulas with application to rainfall fields. *Stoch. Env. Res. Risk. A.* 22(6), 671–688.

[69] Shiau, J. T. (2003). Return period of bivariate distributed extreme hydrological events. *Stoch. Env. Res. Risk. A.* 17, 42–57.

[70] Shiau, J.-T., R. Modarres, and S. Nadarajah (2012). Assessing multi-site drought connections in Iran using empirical copula. *Environ. Model. Assess.* 17(5), 469–482.

[71] Sklar, A. (1959). Fonstions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* 8, 229–231.

[72] Venter, G. G., 2002: Tails of copulas. *Proc. Casualty Actuarial Soc.* 89, 68–113.

[73] Wasserman, L. (2003). *All of Statistics: a Concise Course in Statistical Inference*. Springer, New York.

[74] Yue, S., T. Ouarda, B. Bobée, P. Legendre, and P. Bruneau (1999). The Gumbel mixed model for flood frequency analysis. *J. Hydrol.* 226(1-2), 88–100.

[75] Zhang, L. and V. P. Singh (2006). Bivariate flood frequency analysis using the copula method. *J. Hydrol. Eng.* 11(2), 150–164.