REVSTAT – Statistical Journal
Volume 11, Number 1, March 2013, 1–16

NONPARAMETRIC ESTIMATION OF THE TAIL-DEPENDENCE COEFFICIENT

Author: Marta Ferreira
– Center of Mathematics of Minho University, Portugal
msferreira@math.uminho.pt

Abstract:

- A common measure of tail dependence is the so-called tail-dependence coefficient. We present a nonparametric estimator of the tail-dependence coefficient and prove its strong consistency and asymptotic normality in the case of known marginal distribution functions. The finite-sample behavior as well as robustness will be assessed through simulation. Although it has a good performance, it is sensitive to the extreme value dependence assumption. We shall see that a block maxima procedure might improve the estimation. This will be illustrated through simulation. An application to financial data shall be presented at the end.

Key-Words:

- *extreme value theory; stable tail dependence function; tail-dependence coefficient.*

AMS Subject Classification:

- 62G32.

1. INTRODUCTION

Modern risk management is highly interested in assessing the amount of tail dependence. Many minimum-variance portfolio models are based on correlation, but correlation itself is not enough to describe a tail dependence structure and often results in misleading interpretations (Embrechts *et al.*, [7]). Multivariate extreme value theory (EVT) is the natural tool to measure and model such extremal dependence. The importance of this issue has led to several developments and applications in literature, e.g., Sibuya ([25]), Tiago de Oliveira ([27]), Joe ([16]), Coles *et al.* ([5]), Embrechts *et al.* ([8]), Frahm *et al.* ([11]), Schmidt and Stadtmüller ([23]), Ferreira and Ferreira ([9]); see de Carvalho and Ramos ([6]) for a recent survey.

The *tail-dependence coefficient* (TDC) measures the probability of occurring extreme values for one random variable (r.v.) given that another assumes an extreme value too. More precisely, it is defined as

$$\lambda = \lim_{t\to\infty} P\Big(F_1(X_1) > 1-1/t \mid F_2(X_2) > 1-1/t\Big)\,, \tag{1.1}$$

where F_1 and F_2 are the distribution functions (d.f.'s) of r.v.'s X_1 and X_2, respectively. Observe that it can be formulated as

$$\lambda = \lim_{\alpha\to 0} P\Big(X_1 > \mathit{VaR}_{1-\alpha}(X_1) \mid X_2 > \mathit{VaR}_{1-\alpha}(X_2)\Big)\,,$$

where $\mathit{VaR}_{1-\alpha}(X_i)$ $(i=1,2)$ is the Value-at-Risk of X_i at probability level $1-\alpha$ given by the quantile function evaluated at $1-\alpha$, $F_i^{-1}(1-\alpha) = \inf\{x : F_i(x) \geq 1-\alpha\}$ (see e.g., Schmidt and Stadtmüller, [23]). The TDC can also be defined via the notion of copula, introduced by Sklar ([26]). A copula C is a cumulative distribution function whose margins are uniformly distributed on $[0,1]$. If C is the copula of (X_1,X_2) having joint d.f. F, i.e., $F(x_1,x_2) = C\big(F_1(x_1),F_2(x_2)\big)$, observe that

$$\begin{aligned} \lambda &= 2 - \lim_{t\to\infty} t\,P\Big(F_1(X_1) > 1-1/t \ \text{ or } \ F_2(X_2) > 1-1/t\Big) \\ &= 2 - \lim_{t\to\infty} t\Big\{1 - C\big(1-1/t,\, 1-1/t\big)\Big\}\,. \end{aligned} \tag{1.2}$$

The TDC was the first tail dependence concept appearing in literature in a Sibuya's paper, where it was shown that, no matter how high we choose the correlation of normal random pairs, if we go far enough into the tail, extreme events tend to occur independently in each margin (Sibuya, [25]). It characterizes the dependence in the tail of a random pair (X_1,X_2), in the sense that, $\lambda > 0$ corresponds to tail dependence whose degree is measured by the value of λ, whereas $\lambda = 0$ means tail independence. The well-known bivariate t-distribution presents tail dependence, whereas the above mentioned bivariate normal is an example of tail independent model.

The conventional multivariate extreme value theory has emphasized the asymptotically dependent class resulting in its wide use. However, if the series are truly asymptotically independent, i.e., $\lambda = 0$, an overestimation of extreme value dependence, and consequently of the risk, will take place (see, e.g., Poon *et al.*, [21]; for further details about asymptotically independent class and respective models and coefficients, see also Ledford and Tawn, [19, 20]). Therefore, it is important to conclude whether (X_1, X_2) is tail dependent or not. In practice, this is not an easy task and one must be careful by inferring tail dependence from a finite random sample. Tests for tail independence can be seen in, e.g., Zhang ([28]), Hüsler and Li ([15]) and references therein. Frahm *et al.* ([11]) presents illustrations of misidentifications of the dependence structure. The bad performance of several nonparametric TDC estimators under tail independence was also shown in this latter paper through simulation. We remark that the examples that were used only concern models whose dependence function is not of the extreme value type. Here we present a nonparametric estimator for the TDC derived from Ferreira and Ferreira ([10]) and thus under an extreme value dependence, which we denote $\widehat{\lambda}^{(\mathrm{FF})}$. Strong consistency and asymptotic normality are proved (this latter in the case of known marginal d.f.'s). The finite-sample behavior and robustness are analyzed through simulation. We also compare with other existing methods. The simulation studies reveal some sensitivity to an extreme value dependence assumption and a large bias problem in the particular case of tail independence. In practice this may be overcome by taking block maxima, but one must be careful with a bias-variance trade-off arising from the number of block maxima to be considered: the larger this number the smaller the variance but the larger the bias (Frahm *et al.*, [11]). The simulation studies present improvements in estimates in some cases and allow to conclude the best block length choice. We end with an application to financial data.

2. EVT AND TAIL DEPENDENCE

Let $\big\{\big(X_1^{(n)}, X_2^{(n)}\big)\big\}_{n\geq 1}$ be i.i.d. copies of 2-dimensional random vector, (X_1, X_2), with common d.f. $\mathbf{F}$, and let $M_j^{(n)} = \max_{1\leq i\leq n} X_j^{(i)}$, $j = 1, 2$, be the partial maxima for each marginal. If there exist sequences of constants $a_j^{(n)} > 0$, $b_j^{(n)} \in \mathbb{R}$, for $j = 1, 2$, and a distribution function G with non-degenerate margins, such that

$$
\begin{aligned}
(2.1)\qquad P\Big(M_1^{(n)} \leq a_1^{(n)} x_1 + b_1^{(n)},\ M_2^{(n)} \leq a_2^{(n)} x_2 + b_2^{(n)}\Big) &= \\
= \mathbf{F}^n\Big(a_1^{(n)} x_1 + b_1^{(n)},\ a_2^{(n)} x_2 + b_2^{(n)}\Big) &\underset{n\to\infty}{\longrightarrow} \mathbf{G}(x_1, x_2)\ ,
\end{aligned}
$$

for all continuity points of $\mathbf{G}(x_1, x_2)$, then it must be a bivariate extreme value distribution, given by

$$\mathbf{G}(x_1, x_2) \,=\, \exp\Big[-l\big\{-\log G_1(x_1),\, -\log G_2(x_2)\big\}\Big]\,, \tag{2.2}$$

for some bivariate function l, where G_j, $j = 1, 2$, is the marginal d.f. of $\mathbf{G}$. We also say that $\mathbf{F}$ belongs to the max-domain of attraction of $\mathbf{G}$, in short, $\mathbf{F} \in \mathcal{D}(\mathbf{G})$. The function l in (2.2) is called *stable tail dependence function*, sometimes denoted extreme value dependence. It can be verified that l is convex, is homogeneous of order 1, and that $\max(x_1, x_2) \le l(x_1, x_2) \le x_1 + x_2$ for all $(x_1, x_2) \in [0, \infty)^2$, where the upper bound is due to the positive dependence of extreme value models and corresponds to independence whilst the lower bound means complete dependence (see, e.g. Beirlant *et al.* [1], Section 8.2.2). These properties also hold in the d-variate case, with $d > 2$. The statement in (2.1) has a similar formulation for the respective copulas, say $C_{\mathbf{X}}$ and C:

$$C^n_{\mathbf{X}}(u_1^{1/n}, u_2^{1/n}) \underset{n\to\infty}{\longrightarrow} C(u_1, u_2)\;, \tag{2.3}$$

where

$$C(u_1, u_2) \,=\, \exp\Big\{-l\big(-\log u_1,\, -\log u_2\big)\Big\} \tag{2.4}$$

is called a bivariate extreme value copula. In the sequel it will be denoted BEV copula and we will also refer the extreme value dependence context as a BEV dependence. The defining feature of a BEV copula is the max-stability property, i.e., $C(u_1, u_2) = C(u_1^{1/m}, u_2^{1/m})^m$ for every integer $m \ge 1$, $\forall\,(u_1, u_2) \in [0, 1]^2$. The max-domain of attraction condition (2.1) implies (2.3) but the reciprocal is not true since it must also be imposed that each marginal belongs to some max-domain of attraction. Since we have

$$\begin{aligned}
&\lim_{t\to\infty} t\,P\Big(F_1(X_1) > 1 - 1/t,\; F_2(X_2) > 1 - 1/t\Big) = \\
&\qquad = \;2 \,-\, \lim_{t\to\infty} t\Big\{1 - C\big(1 - 1/t,\, 1 - 1/t\big)\Big\} \\
&\qquad = \;2 \,-\, \lim_{t\to\infty} \log C^t\big(1 - 1/t,\, 1 - 1/t\big) \\
&\qquad = \;2 \,-\, \lim_{t\to\infty} \log C\big((1 - 1/t)^t,\, (1 - 1/t)^t\big) \\
&\qquad = \;2 \,-\, l(1, 1)\;,
\end{aligned} \tag{2.5}$$

the TDC of a BEV copula can be obtained through the function l as

$$\lambda \,=\, 2 - l(1, 1)\;. \tag{2.6}$$

In the following we list some examples of stable tail dependence functions of BEV copulas and respective tail dependence:

- *Logistic:* $l(v_1, v_2) = (v_1^{1/r} + v_2^{1/r})^r$, with $v_j \ge 0$ and parameter $0 < r \le 1$; complete dependence is obtained in the limit as $r \to 0$ and independence when $r = 1$.

- *Asymmetric Logistic:* $l(v_1, v_2) = (1 - t_1)\, v_1 + (1 - t_2)\, v_2 + \big\{(t_1 v_1)^{1/r} + (t_2 v_2)^{1/r}\big\}^r$, with $v_j \geq 0$ and parameters $0 < r \leq 1$ and $0 \leq t_j \leq 1$, $j = 1, 2$; when $t_1 = t_2 = 1$ the asymmetric logistic model is equivalent to the logistic model; independence is obtained when either $r = 1$, $t_1 = 0$ or $t_2 = 0$. Complete dependence is obtained in the limit when $t_1 = t_2 = 1$ and r approaches zero.
- *Hüsler–Reiss:* $l(v_1, v_2) = v_1 \Phi\big(r^{-1} + \frac{1}{2}\, r \log(v_1/v_2)\big) + v_2\, \Phi\big(r^{-1} + \frac{1}{2}\, r \cdot \log(v_2/v_1)\big)$, with parameter $r > 0$ and where Φ is the standard normal d.f.; complete dependence is obtained as $r \to \infty$ and independence as $r \to 0$.

Non-BEV copulas cannot be obtained in the limit in (2.3), i.e., do not satisfy max-stability and cannot be expressed through formulation (2.4) based on the extreme value dependence function l with the given properties.

Examples of non-BEV copulas correspond, for instance, to the class of elliptical ones. The bivariate normal and the symmetric generalized hyperbolic distributions are tail independent models within this class. On the other hand, the bivariate t-distribution presents tail dependence with TDC,

$$\lambda \,=\, 2\, F_{t_{\nu+1}} \Big\{ -\sqrt{(\nu + 1)\,(1 - \rho)/(1 + \rho)} \Big\} \;,$$

where $\rho > -1$ and $F_{t_{\nu+1}}$ is the d.f. of the one dimensional $t_{\nu+1}$ distribution. See, e.g., Schmidt ([22]) and Frahm *et al.* ([11]).

Bivariate Archimedean copulas are another wide class that includes some tail independent non-BEV copulas such as Clayton, $C(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}$ with $\theta \geq 0$. Another special type which do not belong to either one of the three classes above is the tail independent Plackett-copula

$$C(u_1, u_2) \,=\, \frac{1 + (\theta - 1)\,(u_1 + u_2) - \Big[\big\{1 + (\theta - 1)\,(u_1 + u_2)\big\}^2 - 4\, u_1 u_2\, \theta(\theta - 1)\Big]^{1/2}}{2\,(\theta - 1)} \;,$$

with parameter $\theta \in \mathbb{R}^+ \backslash \{1\}$, and $C(u_1, u_2) = u_1 u_2$, if $\theta = 1$. For more details, see Joe ([16]).

3. ESTIMATION

The use of (semi)parametric estimators bears a model risk and may lead to wrong interpretations of the dependence structure. Nonparametric procedures avoid this type of misspecification but usually come along with a larger variance. Frahm *et al.* ([11]) confirms this assertion and shows that (semi)parametric estimators may have disastrous performances under wrong model assumptions.

So, in practice, if we are not sure about the type of model underlying data, nonparametric approach can be an alternative. Here we focus on nonparametric methods.

Huang ([14]), considered an estimator derived from the definition in (1.2) by plugging-in the respective empirical counterparts:

$$\widehat{\lambda}^{(\mathrm{H})} = 2 - \frac{1}{k_n} \sum_{i=1}^{n} \mathbf{1}_{\left\{\widehat{F}_1(X_1^{(i)})>1-\frac{k_n}{n} \text{ or } \widehat{F}_2(X_2^{(i)})>1-\frac{k_n}{n}\right\}}, \tag{3.1}$$

where $\widehat{F}_j$ is the empirical d.f. of F_j, $j = 1, 2$. Concerning estimation accuracy, some modifications of this latter may be used, like replacing the denominator n by $n + 1$, i.e., considering

$$\widehat{F}_j(u) = \frac{1}{n+1} \sum_{i=1}^{n} \mathbf{1}_{\left\{X_j^{(i)} \le u\right\}}$$

(Beirlant *et al.* [1], Section 9.4.1). A similar procedure was considered in Schimdt and Stadtmüler ([23]). For asymptotic properties, see the more recent results in Bücher and Dette ([2]). The consistency and asymptotic normality of the estimator $\widehat{\lambda}^{(\mathrm{H})}$ are derived with the asymptotics holding for an intermediate sequence $\{k_n\}$, $k_n \to \infty$ and $k_n/n \to 0$, as $n \to \infty$. The choice of $k \equiv k_n$ that allows for the 'best' bias–variance tradeoff is of major difficulty, since small values of k come along with a large variance whenever an increasing k results in a strong bias. A similar problem exists for univariate tail index estimations of heavy tailed distributions, for estimators of the stable tail dependence function l (Krajina, [18]) and other TDC estimators (e.g., Frahm *et al.* [11] and Schmidt and Stadtmüller [23]).

Under a BEV copula assumption, i.e., a copula with formulation (2.4), and given (2.6), estimators for the TDC can be obtained through the ones of the stable tail dependence function l. Within this context and motivated in Capéraà *et al.* ([4]), Frahm *et al.* ([11]) presented the estimator

$$2 - 2\exp\left[\frac{1}{n}\sum_{i=1}^{n}\log\left(\sqrt{\log\frac{1}{\widehat{F}_1(X_1^{(i)})}\,\log\frac{1}{\widehat{F}_2(X_2^{(i)})}}\Big/\log\frac{1}{\max\{\widehat{F}_1(X_1^{(i)}),\widehat{F}_2(X_2^{(i)})\}^2}\right)\right].$$

This rank-based estimator was shown to have the best performance among all nonparametric estimators considered in Frahm *et al.* ([11]). Optimally corrected versions can be seen in Genest and Segers ([12]) and alternative estimators are presented in Bücher *et al.* ([3]). In the sequel, we shall use a corrected version satisfying the boundary condition $l(1, 0) = l(0, 1) = 1$ considered in Genest and Segers ([12]), and here denoted $\widehat{\lambda}^{(\mathrm{CFG\text{-}C})}$.

Our approach is motivated by Ferreira and Ferreira ([10]) and has the same assumption of a BEV copula dependence structure. More precisely, it is based on the following representation of the stable tail dependence function:

$$l(x_1,x_2) \;=\; \frac{E\Big[\max\big\{F_1(X_1)^{1/x_1},\; F_2(X_2)^{1/x_2}\big\}\Big]}{1-E\Big[\max\big\{F_1(X_1)^{1/x_1},\; F_2(X_2)^{1/x_2}\big\}\Big]}\;, \tag{3.2}$$

where the expected values are estimated using sample means. Observe that the d.f. of $\max\big(F_1(X_1)^{1/x_1}, F_2(X_2)^{1/x_2}\big)$ is given by

$$\begin{aligned} P\Big(\max\big\{F_1(X_1)^{1/x_1},\, F_2(X_2)^{1/x_2}\big\} \le u\Big) \;&=\; C\big(u^{x_1}, u^{x_2}\big) \\ &=\; \exp\Big(-l\big(-\log u^{x_1}, -\log u^{x_2}\big)\Big) \\ &=\; \exp\Big(-\big(-\log u\big)\, l\big(x_1,x_2\big)\Big) \\ &=\; u^{l(x_1,x_2)}\;, \end{aligned} \tag{3.3}$$

where the penultimate step is due to the first order homogeneity property of function l. Hence

$$E\Big[\max\big\{F_1(X_1)^{1/x_1}, F_2(X_2)^{1/x_2}\big\}\Big] \;=\; \frac{l(x_1,x_2)}{1+l(x_1,x_2)}\;.$$

Therefore, based on (2.6) and (3.2), we propose the estimator

$$\widehat{\lambda}^{(\mathrm{FF})} \;=\; 3-\Big[1-\overline{\max\big\{\widehat{F}_1(X_1), \widehat{F}_2(X_2)\big\}}\,\Big]^{-1}, \tag{3.4}$$

where $\overline{\max\big\{\widehat{F}_1(X_1), \widehat{F}_2(X_2)\big\}}$ is the sample mean of $\max\big\{\widehat{F}_1(X_1), \widehat{F}_2(X_2)\big\}$, i.e.,

$$\overline{\max\big\{\widehat{F}_1(X_1), \widehat{F}_2(X_2)\big\}} \;=\; \frac{1}{n}\sum_{i=1}^{n}\max\big\{\widehat{F}_1(X_1^{(i)}), \widehat{F}_2(X_2^{(i)})\big\}\;.$$

Proposition 3.1. *The estimator $\widehat{\lambda}^{(FF)}$ in (3.4) is strongly consistent.*

Proof: Observe that

$$\begin{aligned} &\left|\frac{1}{n}\sum_{i=1}^{n}\max_{j\in\{1,2\}}\Big\{\widehat{F_j}\big(X_j^{(i)}\big)\Big\} \;-\; E\Big[\max_{j\in\{1,2\}}\Big\{F_j\big(X_j\big)\Big\}\Big]\right| \;\le \\ &\qquad \le\; \left|\frac{1}{n}\sum_{i=1}^{n}\max_{j\in\{1,2\}}\Big\{\widehat{F_j}\big(X_j^{(i)}\big)\Big\} \;-\; \frac{1}{n}\sum_{i=1}^{n}\max_{j\in\{1,2\}}\Big\{F_j\big(X_j^{(i)}\big)\Big\}\right| \\ &\qquad\quad +\; \left|\frac{1}{n}\sum_{i=1}^{n}\max_{j\in\{1,2\}}\Big\{F_j\big(X_j^{(i)}\big)\Big\} \;-\; E\Big[\max_{j\in\{1,2\}}\Big\{F_j\big(X_j\big)\Big\}\Big]\right|\;, \end{aligned} \tag{3.5}$$

where the second term converges *almost surely* to zero by the *Strong Law of Large Numbers* (by (3.3), $\max_{j\in\{1,2\}}\big\{F_j(X_j)\big\} \sim \mathrm{Beta}\big(l(1,1),1\big)$, $1 \le l(1,1) \le 2$, and all the moments exist).

The first term in (3.5) is upper bounded by

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j\in\{1,2\}}\left|\widehat{F}_j\big(X_j^{(i)}\big)-F_j\big(X_j^{(i)}\big)\right| ,$$

which converges *almost surely* to zero according to Gilat and Hill ([13]; Theorem 1.1). See also Ferreira and Ferreira ([10], Proposition 3.7). □

The asymptotic normality in case the marginal d.f.'s are known is derived from Ferreira and Ferreira ([10], Proposition 3.3) and the delta method. More precisely, denoting this version as $\widehat{\lambda}_*^{(\mathrm{FF})}$, we have

$$\sqrt{n}\,\big(\widehat{\lambda}_*^{(\mathrm{FF})}-\lambda\big)\;\to\;N(0,\sigma^2)\ ,$$

where

$$\sigma^2\;=\;\frac{l(1,1)\,\big(1+l(1,1)\big)^2}{2+l(1,1)}\ .$$

In the case of unknown marginals, we believe that the asymptotic normality of $\sqrt{n}\,\big(\widehat{\lambda}^{(\mathrm{FF})}-\lambda\big)$ may be derived from the weak convergence of the empirical copula process (Segers, [24]). This will be addressed in a future work.

Observe that estimators $\widehat{\lambda}^{(\mathrm{FF})}$ and $\widehat{\lambda}^{(\mathrm{CFG\text{-}C})}$ are obtained under the more restrictive assumption of an extreme value dependence but have a convergence rate of $\sqrt{n}$. On the other hand, estimator $\widehat{\lambda}^{(\mathrm{H})}$ has no restrictive assumptions but has to pay the price of a slower convergence rate $\sqrt{k_n}$, since only the largest $k_n=o(n)$ observations can be taken into account.

4. SIMULATION STUDY

In this section we analyze the finite-sample behavior of our estimator. We simulate 1000 independent random samples of sizes $n=50,100,500,1000$ from three BEV copulas with stable tail dependence functions: logistic, asymmetric logistic and Hüsler–Reiss. We consider the two types of dependence: tail dependence (Table 1) and tail independence (Table 2). The results obtained from the logistic and asymmetric logistic under tail independence are quite similar and thus we omit the latter case. In order to assess robustness we also analyze the case of non-BEV copulas, by considering, for tail dependence, a bivariate t-distribution with $\nu=1.5$ degrees of freedom and, for tail independence, a BSGH distribution (Table 3). In both cases we take a correlation parameter of $\rho=0.5$. Since the t-distribution is somewhat 'close' to being an extreme value copula (see Bücher, Dette and Volgushev [3], Section 2), we also consider a convex combination of a rotated Clayton copula (corresponding to negative dependence) and a t-distribution, more precisely, $C_\alpha(u_1,u_2)=\alpha\big(u_2-C_{\mathrm{Clayton}}(1-u_1,u_2)\big)+(1-\alpha)\,C_{t_\nu}(u_1,u_2)$. For comparison, we compute estimator $\widehat{\lambda}^{(\mathrm{CFG\text{-}C})}$ which works under the same assumptions (i.e, an extreme value dependence) and the more general estimator $\widehat{\lambda}^{(\mathrm{H})}$ which has no model restrictions (the required choice of k to balance the variance-bias problem is based on an heuristic procedure in Frahm *et al.* [11]). Absolute empirical bias and the root mean-squared error (rmse) for all implemented TDC estimations are in Tables 1, 2 and 3.

Table 1: Tail dependent ($\lambda > 0$) BEV copulas with stable tail dependence functions: Logistic and Asym. Logistic with $r = 0.4$ and Hüsler–Reiss with $r = 3$.

	$\widehat{\lambda}^{(\mathrm{FF})}$ bias (rmse)	$\widehat{\lambda}^{(\mathrm{CFG\text{-}C})}$ bias (rmse)	$\widehat{\lambda}^{(\mathrm{H})}$ bias (rmse)
$\lambda = 0.6805$		Logistic	
$(n = 50)$	0.0019 (0.0994)	0.0050 (0.0556)	0.0395 (0.1962)
$(n = 100)$	0.0052 (0.0711)	0.0044 (0.0395)	0.0389 (0.1412)
$(n = 500)$	0.0006 (0.0330)	0.0005 (0.0180)	0.0216 (0.0883)
$(n = 1000)$	0.0002 (0.0232)	0.0004 (0.0122)	0.0099 (0.1379)
$\lambda = 0.3402$		Asym. Logistic	
$(n = 50)$	0.0085 (0.1147)	0.0332 (0.1122)	0.0527 (0.1836)
$(n = 100)$	0.0053 (0.0824)	0.0203 (0.0754)	0.0635 (0.1363)
$(n = 500)$	0.0020 (0.0389)	0.0045 (0.0355)	0.0335 (0.0847)
$(n = 1000)$	0.0014 (0.0287)	0.0031 (0.0245)	0.0038 (0.1193)
$\lambda = 0.7389$		Hüsler–Reiss	
$(n = 50)$	0.0040 (0.0484)	0.0057 (0.0462)	0.0202 (0.1697)
$(n = 100)$	0.0003 (0.0331)	0.0020 (0.0323)	0.0075 (0.1094)
$(n = 500)$	0.0002 (0.0152)	0.0007 (0.0140)	0.0011 (0.0655)
$(n = 1000)$	0.0002 (0.0292)	0.0005 (0.0097)	0.0103 (0.0342)

Table 2: Tail independent ($\lambda = 0$) BEV copulas with stable tail dependence functions: Logistic with $r = 1$ and Hüsler–Reiss with $r = 0.03$.

	$\widehat{\lambda}^{(\mathrm{FF})}$ bias (rmse)	$\widehat{\lambda}^{(\mathrm{CFG\text{-}C})}$ bias (rmse)	$\widehat{\lambda}^{(\mathrm{H})}$ bias (rmse)
$\lambda = 0$		Logistic	
$(n = 50)$	0.0230 (0.1284)	0.0900 (0.1389)	0.1040 (0.1644)
$(n = 100)$	0.0062 (0.0956)	0.0467 (0.0952)	0.1004 (0.1348)
$(n = 500)$	0.0036 (0.0415)	0.0140 (0.0361)	0.0492 (0.0650)
$(n = 1000)$	0.0017 (0.0296)	0.0077 (0.0257)	0.0502 (0.0578)
$\lambda \approx 0$		Hüsler–Reiss	
$(n = 50)$	0.0254 (0.1370)	0.0875 (0.1353)	0.1002 (0.1660)
$(n = 100)$	0.0084 (0.0966)	0.0412 (0.0883)	0.0991 (0.1336)
$(n = 500)$	0.0009 (0.0415)	0.0100 (0.0361)	0.0492 (0.0653)
$(n = 1000)$	0.0003 (0.0299)	0.0061 (0.0265)	0.0081 (0.0298)

Table 3: Non-BEV tail dependent case: t_ν with $\nu = 1.5$ and $\rho = 0.5$ and a convex combination of a rotated Clayton and t_ν (RC&T), $C_{0.5}(u_1, u_2) = 0.5\left(u_2 - C_{\text{Clayton}}(1 - u_1, u_2)\right) + 0.5\, C_{t_\nu}(u_1, u_2)$; non-BEV tail independent case: BSGH distribution with $\rho = 0.5$.

	$\widehat{\lambda}^{(\text{FF})}$ bias (rmse)	$\widehat{\lambda}^{(\text{CFG-C})}$ bias (rmse)	$\widehat{\lambda}^{(\text{H})}$ bias (rmse)
$\lambda = 0.4406$	t-distribution		
$(n = 50)$	0.0099 (0.1043)	0.0318 (0.1022)	0.0084 (0.1970)
$(n = 100)$	0.0087 (0.0711)	0.0213 (0.0743)	0.0094 (0.1393)
$(n = 500)$	0.0124 (0.0339)	0.0130 (0.0348)	0.0044 (0.0884)
$(n = 1000)$	0.0122 (0.0267)	0.0123 (0.0266)	0.0120 (0.1403)
$\lambda = 0.3669$	RC&T		
$(n = 50)$	0.4396 (0.6562)	0.2832 (0.4736)	0.2990 (0.3064)
$(n = 100)$	0.4052 (0.6440)	0.2879 (0.4282)	0.1371 (0.2779)
$(n = 500)$	0.3800 (0.6411)	0.2793 (0.4681)	0.1350 (0.2772)
$(n = 1000)$	0.3791 (0.6342)	0.2650 (0.4571)	0.1314 (0.2743)
$\lambda = 0$	BSGH		
$(n = 50)$	0.4288 (0.4396)	0.4305 (0.4544)	0.3730 (0.4238)
$(n = 100)$	0.4287 (0.4346)	0.4239 (0.4294)	0.3704 (0.3926)
$(n = 500)$	0.4248 (0.4259)	0.4030 (0.4052)	0.3130 (0.3232)
$(n = 1000)$	0.4238 (0.4243)	0.4001 (0.4008)	0.2188 (0.2489)

Estimators $\widehat{\lambda}^{(\text{FF})}$ and $\widehat{\lambda}^{(\text{CFG-C})}$ behave well within BEV copulas (or ‘close’ of being BEV as t-distribution). Yet, they performed poorly on a non-BEV dependence context (see Table 3). Estimator $\widehat{\lambda}^{(\text{H})}$ tends to present a slight larger bias but performs better under non extreme value dependence. This is consistent with a slower rate of convergence and the fact that it holds in a general framework, as discussed in the previous section. All estimators also performed poorly on tail independent non-BEV copulas. Our results do not contradict however the ones in Frahm *et al.* ([11]), where the misbehavior of nonparametric estimation concerned tail independence within non-BEV copulas. By considering a block maxima procedure, i.e., divide n-length data into m blocks of size $b = \lfloor n/m \rfloor$ ($\lfloor x \rfloor$ denotes the largest integer not exceeding x) and take only the maximum observation within each block, we obtain a sample of maximum, which is more consistent with an extreme values model and thus a BEV copula. This methodology involves a bias–variance tradeoff arising from the number of block maxima (block length) to be considered: the larger (smaller) this number the smaller the variance but the larger the bias (Frahm *et al.*, [11]). It requires not too small sample sizes to also provide not too small maxima samples. A simulation study to find the value(s) of b that better accommodates this compromise will be implemented in the next section.

4.1. Block maxima procedure for non-BEV dependence

We consider 1000 independent random samples of sizes $n = 500, 1000, 1500, 2000, 5000$ generated from the tail independent and non-BEV copulas: bivariate normal (BN), BSGH and Plackett-copula (BPC). We estimate the TDC through a block maxima procedure for block lengths $b = 15, 30, 60, 90$. The absolute empirical bias and the rmse of all implemented TDC estimations are presented in Tables 4 and 5, for BN and BPC, respectively. The results obtained for the BSGH case (omitted here) were not good in all the three estimators and, in practice,

Table 4: Block maxima samples with given length b of BN model with $\rho = 0.5$ (the case $b = 1$ correspond to the whole sample).

BN	$\widehat{\lambda}^{(\text{FF})}$	$\widehat{\lambda}^{(\text{CFG-C})}$	$\widehat{\lambda}^{(\text{H})}$
$(n = 500)$	bias (rmse)	bias (rmse)	bias (rmse)
$(b = 1)$	0.4025 (0.4036)	0.3702 (0.3733)	0.3244 (0.3294)
$(b = 15)$	0.2319 (0.2578)	0.2520 (0.2966)	0.1986 (0.2594)
$(b = 30)$	0.2081 (0.2924)	0.2958 (0.3351)	0.2241 (0.3825)
$(b = 60)$	0.2798 (0.4187)	0.3703 (0.4486)	0.1900 (0.4594)
$(b = 90)$	0.2887 (0.4275)	0.3734 (0.4937)	0.3816 (0.7975)
$(n = 1000)$	bias (rmse)	bias (rmse)	bias (rmse)
$(b = 1)$	0.4023 (0.4029)	0.3587 (0.3692)	0.3238 (0.3262)
$(b = 15)$	0.2046 (0.2297)	0.2201 (0.2438)	0.2037 (0.2498)
$(b = 30)$	0.1941 (0.2428)	0.2251 (0.2720)	0.2185 (0.3012)
$(b = 60)$	0.1724 (0.2695)	0.2625 (0.3234)	0.2000 (0.3578)
$(b = 90)$	0.2888 (0.3582)	0.3692 (0.4259)	0.3625 (0.5874)
$(n = 1500)$	bias (rmse)	bias (rmse)	bias (rmse)
$(b = 1)$	0.4024 (0.4028)	0.3562 (0.3663)	0.3236 (0.3253)
$(b = 15)$	0.2011 (0.2180)	0.2114 (0.2242)	0.1848 (0.2165)
$(b = 30)$	0.1612 (0.2015)	0.2001 (0.2328)	0.1682 (0.2309)
$(b = 60)$	0.1546 (0.2311)	0.2310 (0.2760)	0.2064 (0.3200)
$(b = 90)$	0.1708 (0.2696)	0.2674 (0.3093)	0.1964 (0.3480)
$(n = 2000)$	bias (rmse)	bias (rmse)	bias (rmse)
$(b = 1)$	0.3230 (0.3243)	0.3559 (0.3661)	0.3230 (0.3243)
$(b = 15)$	0.2012 (0.2141)	0.2013 (0.2155)	0.2054 (0.2312)
$(b = 30)$	0.1601 (0.1912)	0.1628 (0.2116)	0.1810 (0.2293)
$(b = 60)$	0.1600 (0.2172)	0.1600 (0.2111)	0.2047 (0.2887)
$(b = 90)$	0.1829 (0.2535)	0.2029 (0.2986)	0.2269 (0.3489)
$(n = 5000)$	bias (rmse)	bias (rmse)	bias (rmse)
$(b = 1)$	0.4024 (0.4025)	0.3603 (0.3644)	0.3234 (0.3240)
$(b = 15)$	0.1936 (0.1988)	0.1801 (0.1884)	0.1973 (0.2068)
$(b = 30)$	0.1595 (0.1730)	0.1519 (0.1717)	0.1762 (0.1952)
$(b = 60)$	0.1348 (0.1648)	0.1550 (0.1846)	0.1701 (0.2093)
$(b = 90)$	0.1283 (0.1744)	0.1677 (0.1948)	0.1742 (0.2288)

Table 5: Block maxima samples with given length b of BPC (Plackett-copula) with $\theta = 2$ (the case $b = 1$ correspond to the whole sample).

BPC	$\widehat{\lambda}^{(\mathrm{FF})}$	$\widehat{\lambda}^{(\mathrm{CFG\text{-}C})}$	$\widehat{\lambda}^{(\mathrm{H})}$
$(n = 500)$	bias (rmse)	bias (rmse)	bias (rmse)
$(b = 1)$	0.2028 (0.2062)	0.1766 (0.1805)	0.1676 (0.1741)
$(b = 15)$	0.0894 (0.1779)	0.1592 (0.2026)	0.1555 (0.2462)
$(b = 30)$	0.0801 (0.2341)	0.1954 (0.2565)	0.1329 (0.2833)
$(b = 60)$	0.1981 (0.3407)	0.3397 (0.3913)	0.1244 (0.3718)
$(b = 90)$	0.2189 (0.3892)	0.2695 (0.4624)	0.3942 (0.4507)
$(n = 1000)$	bias (rmse)	bias (rmse)	bias (rmse)
$(b = 1)$	0.2012 (0.2030)	0.1708 (0.1733)	0.1684 (0.1720)
$(b = 15)$	0.0538 (0.1249)	0.0906 (0.1313)	0.1134 (0.1613)
$(b = 30)$	0.0701 (0.1668)	0.1457 (0.1880)	0.1579 (0.2459)
$(b = 60)$	0.0955 (0.2315)	0.2122 (0.2623)	0.1517 (0.3090)
$(b = 90)$	0.2262 (0.3168)	0.3295 (0.3714)	0.3000 (0.5177)
$(n = 1500)$	bias (rmse)	bias (rmse)	bias (rmse)
$(b = 1)$	0.2019 (0.2031)	0.1695 (0.1705)	0.1684 (0.1708)
$(b = 15)$	0.0547 (0.1081)	0.0798 (0.1089)	0.1053 (0.1408)
$(b = 30)$	0.0514 (0.1389)	0.1068 (0.1452)	0.1042 (0.1667)
$(b = 60)$	0.0545 (0.1943)	0.1504 (0.2077)	0.1480 (0.2647)
$(b = 90)$	0.1000 (0.2250)	0.2077 (0.2610)	0.1535 (0.3084)
$(n = 2000)$	bias (rmse)	bias (rmse)	bias (rmse)
$(b = 1)$	0.2012 (0.2021)	0.1684 (0.1692)	0.1695 (0.1712)
$(b = 15)$	0.0539 (0.0967)	0.0713 (0.0971)	0.1141 (0.1141)
$(b = 30)$	0.0384 (0.1175)	0.0887 (0.1252)	0.1135 (0.1613)
$(b = 60)$	0.0642 (0.1728)	0.1390 (0.1831)	0.1405 (0.2262)
$(b = 90)$	0.0953 (0.2072)	0.1819 (0.1861)	0.1678 (0.2942)
$(n = 5000)$	bias (rmse)	bias (rmse)	bias (rmse)
$(b = 1)$	0.2015 (0.2018)	0.1665 (0.1669)	0.1697 (0.1704)
$(b = 15)$	0.0430 (0.0671)	0.0404 (0.0641)	0.1108 (0.1215)
$(b = 30)$	0.0345 (0.0807)	0.0495 (0.0781)	0.1065 (0.1283)
$(b = 60)$	0.0366 (0.1079)	0.0729 (0.1065)	0.1210 (0.1629)
$(b = 90)$	0.0397 (0.1376)	0.0895 (0.1323)	0.0970 (0.1407)

may lead to wrongly infer tail dependence. If this is an adequate model for data, then (semi)parametric estimators considered in Frahm [11]) are a more sensible choice. We have also implemented a block maxima procedure for the non-BEV case of the convex combination copula considered in Table 3 with similar results of the BPC and thus omitted. Observe that block maxima procedure improves estimates in some cases, in particular for estimators $\widehat{\lambda}^{(\mathrm{FF})}$ and $\widehat{\lambda}^{(\mathrm{CFG\text{-}C})}$. The adequate choices for block-length b in sample sizes ranging from, approximately, 500 and 1000, are $b = 15, 30$, while for sample sizes between 1000 and 2000 we can choose $b = 30, 60$, and for larger sample sizes (ranging from 2000 to 5000) a block-length $b = 60, 90$ seems appropriate.

4.2. Application to financial data

We consider the negative log-returns of Dow Jones (USA) and FTSE100 (UK) indexes for the time period 1994–2004. The corresponding scatter plot and TDC estimate plot of $\widehat{\lambda}^{(\mathrm{H})}$ for various k (Figure 1) show the presence of tail dependence and the order of magnitude of the tail-dependence coefficient.

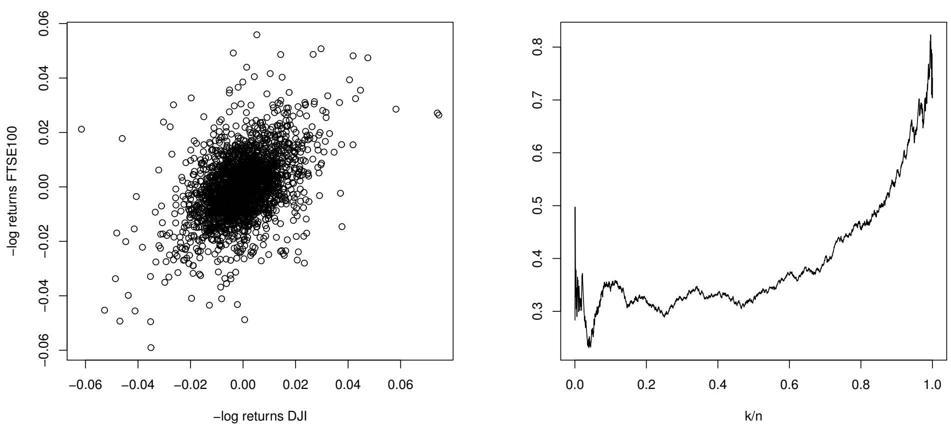

Figure 1: Scatter plot of Dow Jones versus FTSE100 negative log-returns ($n = 2529$ data points) and the corresponding TDC estimates $\widehat{\lambda}^H$ for various k/n.

Moreover, the typical variance-bias problem for various threshold values k can be observed, too. In particular, a small k induces a large variance, whereas an increasing k generates a strong bias of the TDC estimate. The threshold choosing procedure of k leads to a TDC estimate of $\widehat{\lambda}^{(\mathrm{H})} = 0.3397$ and from our estimator we derive $\widehat{\lambda}^{(\mathrm{FF})} = 0.3622$. In computing $\widehat{\lambda}^{(\text{CFG-C})}$ we obtain 0.354. The results from the three considered estimators are quite close, leading to a tail-dependence estimate that should be approximately 0.35.

5. DISCUSSION

One must be careful by inferring tail dependence/independence from a finite random sample and (semi)parametric and nonparametric procedures have pros and cons. Thus, the message is that there is no perfect strategy and the best way to protect against errors is the application of several methods to the same data set. A test of tail independence is advised (see, e.g., Zhang [28], Hüsler and Li [15] and references therein). The proposed estimator has revealed good performance even in the independent case. However the simulation results showed sensitivity to the assumption of an extreme value dependence structure and we recommend to test in advance for this hypothesis. See Kojadinovic, Yan and Segers ([17]) or Bücher, Dette and Volgushev ([3]) and references therein. A block maxima procedure may improve the estimates. A study focused on the asymptotic properties will be addressed in a future work.

ACKNOWLEDGMENTS

This research was financed by FEDER Funds through "Programa Operacional Factores de Competitividade — COMPETE" and by Portuguese Funds through FCT — "Fundação para a Ciência e a Tecnologia", within the Project Est-C/MAT/UI0013/2011.

We also acknowledge the valuable suggestions from the referees.

REFERENCES

[1] Beirlant, J.; Goegebeur, Y.; Segers, J. and Teugels, J. (2004). *Statistics Of Extremes: Theory and Application*, John Wiley, England.

[2] Bücher, A. and Dette, H. (2012). Multiplier bootstrap of tail copulas with applications, *Bernoulli*, to appear.

[3] Bücher, A.; Dette, H. and Volgushev, S. (2011). New estimators of the Pickands dependence function and a test for extreme-value dependence, *The Annals of Statistics*, **39**(4), 1963–2006.

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

[5] Coles, S.; Heffernan, J. and Tawn, J. (1999). Dependence measures for extreme value analysis, *Extremes*, **2**, 339–366.

[6] de Carvalho, M. and Ramos, A. (2012). Bivariate extreme statistics, II, *Revstat*, **10**, 83–107.

[7] Embrechts, P.; McNeil, A. and Straumann, D. (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, Cambridge, 176–223.

[8] Embrechts, P.; Lindskog, F. and McNeil, A. (2003). *Modelling dependence with copulas and applications to risk management.* In "Handbook of Heavy Tailed Distributions in Finance" (S. Rachev, Ed.), Elsevier, 329–384.

[9] Ferreira, H. and Ferreira, M. (2012). Tail dependence between order statistics, *Journal of Multivariate Analysis*, **105**(1), 176–192.

[10] Ferreira, H. and Ferreira, M. (2012). On extremal dependence of block vectors, *Kybernetika*, **48**(5), 988–1006.

[11] Frahm, G.; Junker, M. and Schmidt, R. (2005). Estimating the tail-dependence coefficient: properties and pitfalls, *Insurance: Mathematics & Economics*, **37**(1), 80–100.

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

[13] Gilat, D. and Hill, T. (1992). One-sided refinements of the strong law of large numbers and the Glivenko–Cantelli Theorem, *The Annals of Probability*, **20**, 1213–1221.

[14] Huang, X. (1992). *Statistics of Bivariate Extreme Values*, Ph. D. thesis, Tinbergen Institute Research Series 22, Erasmus University, Rotterdam.

[15] Hüsler, J. and Li, D. (2009). Testing asymptotic independence in bivariate extremes, *Journal of Statistical Planning and Inference*, **139**, 990–998.

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

[17] Kojadinovic, I.; Segers, J. and Yan, J. (2011). Large-sample tests of extreme-value dependence for multivariate copulas, *The Canadian Journal of Statistics*, **39**(4), 703–720.

[18] Krajina, A. (2010). *An M–Estimator of Multivariate Tail Dependence*, Tilburg University Press, Tilburg.

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

[20] Ledford, A. and Tawn, J.A. (1997). Modelling dependence within joint tail regions, *Journal of the Royal Statistical Society, Series B*, **59**, 475–499.

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

[22] Schmidt, R. (2002). Tail dependence for elliptically contoured distributions, *Mathematical Methods of Operations Research*, **55**, 301–327.

[23] Schmidt, R. and Stadtmüller, U. (2006). Nonparametric estimation of tail dependence, *Scandinavian Journal of Statistics*, **33**, 307–335.

[24] Segers, J. (2012). Asymptotics of empirical copula processes under nonrestrictive smoothness assumptions, *Bernoulli*, **18**, 764–782.

[25] Sibuya, M. (1960). Bivariate extreme statistics, *Annals of the Institute of Statistical Mathematics*, **11**, 195–210.

[26] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229–231.

[27] Tiago de Oliveira, J. (1962–1963). Structure theory of bivariate extremes: extensions, *Estudos de Matemática, Estatística e Econometria*, **7**, 165–195.

[28] Zhang, Z. (2008). Quotient correlation: a sample based alternative to Pearson's correlation, *The Annals of Statistics*, **36**(2), 1007–1030.