Afrika Statistika
Vol. 12(1), 2017, pages 1171–1184.
DOI: http://dx.doi.org/10.16929/as/2017.1171.98

Afrika Statistika
ISSN 2316-090X

Optimal number of upper order statistics used in estimation for the coefficient of tail dependence

Samah Betteka[1,2] **and Brahim Brahimi**[1,*]

[1]Laboratory of Applied Mathematics, Mohamed Khider University, Biskra, Algeria
[2]University of Mentouri Brothers, Constantine, Algeria

Received : December 31, 2016; Accepted : March 5, 2017

Abstract. Beirlant *et al.* (2011) introduced a bias-reduced estimator for the coefficient of tail dependence and for bivariate tail probability in bivariate extreme value statistics. In this paper, we are interested in the problem of choice of the number of extreme order statistics of bivariate observations exceeding high thresholds, we want to optimize the estimators on the choice we expose different methods for determining this number. The efficiency of our methods is illustrated on a simulation study and by an application to real data.

Résumé. Beirlant *et al.* (2011) ont introduit un estimateur à bais-réduit pour le coefficient de dépendance de la queue et la probabilité de la queue dans les statistiques de valeur extrême bivariées. Dans cet article, nous nous intéressons au problème du choix du nombre de statistiques d'ordre extrême des observations bivariées dépassant les seuils élevés, nous voulons optimiser les estimateurs sur le choix que nous exposons par des différentes méthodes pour déterminer ce nombre. L'efficacité de nos méthodes est illustrée par une étude de simulation et par une application sur des données réelles.

Key words: Coefficient of tail dependence; Bias reduction; Extended Pareto distribution; Tail probability; Copula; Hill estimator; Moment estimator.
AMS 2010 Mathematics Subject Classification : 62G32; 34L20; 60G70.

---

1. Introduction and motivation

In the classical setting of bivariate extreme value theory, the procedures to estimate the probability of an extreme event are not applicable if the component-wise maxima of the

*Corresponding author Brahim Brahimi: brah.brahim@gmail.com
Samah Betteka : Samah_bateka@yahoo.fr

observations are asymptotically independent. To address this problem, Ledford and Tawn (1996) proposed a sub-model in which the tail dependence is characterized by an additional parameter named coefficient of tail dependence η and it satisfies $\eta \in (0,1]$. This coefficient generalized in Ledford and Tawn (1997) without establishing their asymptotic properties. It has turned out to be a useful tool for describing the tail behavior of a bivariate distribution (not necessarily max-stable, distribution function (df) F). In the case that the margins to have a standard (unit Fréchet), then η is defined by

$$P(Z_1 > z_1, Z_2 > z_2) = z_1^{-c_1} z_2^{-c_2} \mathcal{L}(z_1, z_2) \quad \text{with } c_1,\ c_2 > 0. \tag{1}$$

where $\eta = (c_1 + c_2)^{-1}$ and the function $\mathcal{L}$ is a bivariate slowly varying function, i.e.

$$\mathcal{L}(z_1, z_2) \sim g_1(z_1, z_2)\,(1 + g_2(z_1, z_2) z_1^{\rho_1} z_2^{\rho_2}) \quad \text{as } z_1, z_2 \to \infty \tag{2}$$

Then $\eta = 1$ in the case of asymptotic dependence, whereas $\eta < 1$ implies asymptotic independence. For some counter-examples of the Ledford and Tawn's model see Ledford and Tawn (1997).

Various methods to estimate the coefficient of tail dependence are proposed see Peng (1999), Draisma *et al.* (2004) and Beirlant and Vandewalle (2002). We start with Ledford and Tawn (1996), they proposed first to standardize the marginal to the unit Fréchet distribution, using either the empirical marginal distributions (that is, using the ranks of the components) or extreme value estimators for the marginal tails, and then to estimate η as the shape parameter of the minimum of the components, by a classical estimator for the extreme value index, e.g. the Hill estimator (Hill, 1975) or the moment estimator.

In Beirlant *et al.* (2011), an asymptotically unbiased estimator for η was proposed, based on fitting the extended Pareto distribution with the method of maximum likelihood to properly transformed random variables. Goegebeur and Guillou (2013) obtained asymptotic unbiasedness by taking a properly weighted sum of two biased estimators for η. Dutanga *et al.* (2014) introduced a robust and asymptotically unbiased estimator for η, his estimator is obtained by fitting a second order model to the data by means of the minimum density power divergence criterion. In this paper we use estimator for the coefficient of tail dependence proposed by Beirlant *et al.* (2011) based on second order model.

This paper is organized as follows. In Section 2 we present bivariate tail estimation and estimating the coefficient of tail dependence. In Section 3 we introduced some of the methods proposed for balancing between bias and variance in order to obtain an optimal number k_{opt} of order statistic. In Section 4 we carry out a simulation study to compare three different estimators of the coefficient of tail dependence with special emphasis on their mean square error, a diagnostic for selecting the number of data to be used in estimating is provided.

Draisma *et al.* (2004) interpreted an extension of Ledford and Tawn's condition as a bivariate second order (SO) regular variation condition, the latter is not too restrictive and commonly used in estimation problems involving the bivariate case.

Corollary 1. *Let (X, Y) be a random vector with joint df F and continuous marginal df's F_X and F_Y such that*

$$\lim_{t \downarrow 0} q_1(t)^{-1} \left(\frac{P(1 - F_X(X) < tx, 1 - F_Y(Y) < ty)}{q(t)} - c(x, y) \right) =: c_1(x, y) \tag{3}$$

exists for all $x \geq 0, y \geq 0$ with $x+y>0$, a positive function q and a function q_1 both tending to zero as $t \downarrow 0$, and c_1 a function neither constant nor a multiple of c. Moreover, we assume that the convergence is uniform on $\{(x,y) \in [0,\infty)^2 | x^2+y^2=1\}$, that c_1 is continuous and $c_1(x,x) = x^{1/\eta}(x-1)/\tau$.

Recall that $q(t) := P(1-F_X(X)<t,\ 1-F_Y(Y)<t)$ (in the paper of Ledford and Tawn, 1996, $q(t)$ is equals to $t^{1/\eta}$). It can be shown by (3) implies that q and $|q_1|$ are regularly varying at zero with index $1/\eta$ and $\tau \geq 0$, respectively. The function is homogeneous of order $1/\eta$, that is $c(tx,ty) = t^{1/\eta}c(x,y)$.

The parameter η has the same meaning as in Ledford and Tawn (1997) and condition (3) is similar to condition (2).

We assume that $l := \lim_{t\downarrow 0} q(t)/t$ exists, where $l>0$ implies asymptotic independence. Hence $\eta<1$ implies asymptotic independence.

Lemma 1. *Model (1) satisfies SO condition. Many commonly used joint df satisfy the model (1). Note that this model is in fact a condition on the copula function C. Indeed, one easily verifies that*

$$P(1-F_X(X)<x,\ 1-F_Y(Y)<y) = x+y-1+C(1-x,\ 1-y).$$

We give now some examples of distributions satisfying condition SO.

1.1. The Farlie Gumbel Morgenstern distribution

The Farlie Gumbel Morgenstern copula function is given by

$$C(u,\ v) = uv[1+\alpha(1-u)(1-v)], \qquad (u,v) \in [0,\ 1]^2,$$

where $\alpha \in [-1,1]$. Straightforward calculations lead to

$$P(1-F_X(X)<tx,\ 1-F_Y(Y)<ty) = t^2xy[1+\alpha-\alpha t(x+y)+\alpha t^2xy].$$

In the case where $\alpha \in [-1,1]$

$$\frac{P(1-F_X(X)<tx,\ 1-F_Y(Y)<ty)}{P(1-F_X(X)<t,\ 1-F_Y(Y)<t)} = xy\left[1-\frac{\alpha t}{1+\alpha}(x+y-2)+O(t^2)\right],$$

from which one easily verifies that SO condition is satisfied with $\eta=0.5$, $c(x,y) = xy$, $c_1(x,y) = xy(x+y-2)/2$, $q_1(t) = -2\alpha t/(1+\alpha)$, so $\tau=1$.

For the case $\alpha=-1$

$$\frac{P(1-F_X(X)<tx,\ 1-F_Y(Y)<ty)}{P(1-F_X(X)<t,\ 1-F_Y(Y)<t)} = xy\left[\frac{x+y}{2}+\frac{1}{4}(x+y-2xy)+O(t^2)\right],$$

and hence condition SO is satisfied with $\eta=1/3$, $c(x,y) = xy(x+y)/2$, $c_1(x,y) = xy(2xy-x-y)/2$, $q_1(t) = -t/2$, so $\tau=1$.

1.2. Bivariate normal distribution

The bivariate normal distribution with mean 0, variance 1 and correlation coefficient $\rho \in (-1, 1)$ satisfies the condition SO with $\eta = (1+\rho)/2$. We consider $\rho = -0.5$ for $\eta = 0.25$. We refer to Ledford and Tawn (1996) and Draisma *et al.* (2004) for further details.

2. Estimating the coefficient of tail dependence

Beirlant *et al.* (2009) introduced the extended Pareto distribution as approximate model for relative excesses over a threshold. We use the estimation of bivariate tail suggest by Beirlant *et al.* (2011), this estimation for the coefficient of tail dependence based on the second order condition named bias reduced estimator, summarized as following

$$P\left(Z_1 > z_1,\ Z_2 > z_2\right) = P(\min(Z_1,\ Z_2 w/(1-w)) > z_1) =: P(Y_w > z_1).$$

Or $w = \dfrac{z_1}{z_1 + z_2}$ is a tuning parameter, i.e. the transformed variable $\min(Z_1, Z_2)$ follows a Pareto-type model with index $1/\eta$,

$$P(Y_w > z_1) = z_1^{-1/\eta} C_w \left[1 + \frac{1}{\eta}\delta_w(z_1)\right],$$

where

$$C_w = \left(\frac{w}{1-w}\right)^{c_2} g_1^*(w),$$

and

$$\delta_w = \eta g_2^*(w) \left(\frac{w}{1-w}\right)^{-\rho_2} z_1^{\tau}(1 + o(1)),$$

with $tz = z_1$ and t is a suitably chosen threshold. We have

$$P(Z_1 > z_1,\ Z_2 > z_2) = \overline{G}_{\eta,\delta_w,\tau}\left(z_1/t\right) P(Y_w > t) + o(|\delta_w(t)| P(Y_w(t))).$$

The model $\overline{G}_{\eta,\delta_w,\tau}$ was introduced in Beirlant *et al.* (2009), and named by the extended Pareto distribution (EPD).

It is well known that choosing the threshold t is a difficult problem even if the estimators have an explicit form, the number of excesses over t must be sufficiently large to make inference feasible.

The choice of an adaptive data threshold $t = Y_{n-k,n}$ within the ordered sequence $Y_{1,n} \leq Y_{2,n} \leq ... \leq Y_{n,n}$ of the observed values of Y we can ensure both criteria by choosing $k \to \infty$ and $k/n \to 0$ as $n \to \infty$.

For full details on selecting the number of extreme order statistics, we refer to Cheng and Peng (2001), Neves and Fraga (2004), Danielsson *et al.* (2001), Dekkers et al (1993), Draisma *et al.* (1999) and Drees and Kaufmann (1998).

Definition 1 (EPD). The EPD with parameter vector (η, δ, τ) in the range $\tau < 0 < \eta$ and $\delta >$ max $(-1,\ 1/\tau)$ is defined by its df

$$\overline{G}_{\eta,\delta,\tau}(z) = \begin{cases} 1 - \{z(1 + -\delta z)^{\tau}\}^{-1/\eta} & if \ \ z > 1 \\ 0 & if \ \ z \leq 1. \end{cases}$$

2.1. Estimation of the EPD parameters

The estimation of bivariate tail distributions requires first the estimation of the EPD parameters, see Beirlant *et al.* (2009). Since

$$P(Y_w > tz \backslash Y_w > t) = \overline{G}_{\eta,\delta_w,\tau}(z) = o(|\delta_w(t)|).$$

The estimation of the EPD parameters is based on the relative excesses Y/t when $Y > t$. As discussed in Beirlant *et al.* (2009), the asymptotic distribution of the tail estimator will not depend on the asymptotic distribution of the estimator of τ. Therefore, the unknown second-order parameter τ will be replaced by a consistent estimator which will not affect the asymptotic distributions of the other estimators. In practice, the parameter τ is often replaced by -1.

2.1.1. Maximum Likelihood Estimation (MLE)

The estimators of η and δ will be found by maximizing an approximation to the EPD likelihood given the sample of k relative excesses $Y_{n-k+i,n}/Y_{n-k,n}$, $i = 1, ..., k$. The log-likelihood of the EPD is given by

$$\log L(z) = -\log \eta - (1/\eta + 1)\log z - (1/\eta + 1)\log(1 + \delta - \delta z^{\tau}) + \log(1 + \delta - \delta z^{\tau} - \delta\tau z^{\tau}).$$

define

$$\begin{aligned}\hat{\eta}_B &= \frac{\log L(z)}{\delta\eta} = H_{k,n} = \frac{1}{k}\sum_{j=1}^{k}\log(Y_{n-k+j,n}/Y_{n-k,n}),\\ \hat{\delta}_B &= \frac{\log\ L(z)}{\delta\delta} = E_{k,n}(s) = \frac{1}{k}\sum_{j=1}^{k}\log(Y_{n-k+j,n}/Y_{n-k,n})^{s}, \qquad s \leq 0.\end{aligned} \tag{4}$$

Note that $H_{k,n}$ is the Hill estimator (see Hill, 1975).

Another asymptotically equivalent estimator of η found in Beirlant *et al.* (2009) where the authors proposed simplified estimators for the model based on the functions EPD notes are linearized in δ as a result

$$\begin{aligned}\hat{\delta}_L = &\quad H_{k,n}(1-2\hat{\beta})(1-\hat{\beta})^3\hat{\beta}^{-4}\left(\frac{1}{k}\sum_{j=1}^{k}(Y_{n-k+j,n}/Y_{n-k,n})^{\hat{\beta}/H_{k,n}} - \tfrac{1}{1-\hat{\beta}}\right),\\ \hat{\eta}_L = &\quad H_{k,n} - \hat{\delta}\tfrac{\hat{\beta}}{1-\hat{\beta}}.\end{aligned} \tag{5}$$

Let $\hat{\beta}$ be a weakly consistent estimator sequence for $\beta = \eta\tau$. Beirlant *et al.* (2009) used the estimator introduced in Fraga et al. (2003).

2.1.2. Estimation of the rare probability

Estimating $P(Y_w > t)$ by the empirical proportion k/n, we then obtain the estimates

$$\hat{p} = \frac{k}{n}\overline{G}_{\hat{\eta},\hat{\delta},\hat{\tau}}(z_1/Y_{n-k,n}). \tag{6}$$

when we omit the second-order part of the model (δ_w and τ) we obtain a classical Weissman (1978) type estimator for the bivariate tail probability:

$$\hat{p}_W = \frac{k}{n}(z_1/Y_{n-k,n})^{-1/\hat{\eta}}. \tag{7}$$

3. Optimal sample fraction selection

In this Section, we start with the graphical method for selection of the optimal sample fraction k and we note that when k is small the variance is large, and the use of a large value of k introduces a large bias in the estimation. Next the optimum value k can be attained through the minimization of the mean squared error of the considered estimator. Recently, adaptive procedures for automated selection of the optimal sample fraction proposed to compute $\hat{k}_{opt}$ for k_{opt} in the sense $\hat{k}_{opt}/k_{opt} \xrightarrow{p} 1$ as $n \to \infty$, such as bootstrap methods Danielsson *et al.* (2001), Draisma *et al.* (1999) and sequential procedures Drees and Kaufmann (1998), we will apply the method of Reiss and Thomas and the method of Cheng and Peng. Some simulations are discussed in Section 4, in which we recall the criterion on $k = k_n$.

$$k \to \infty,\ 1 \leq k \leq n \text{ and } k/n \to 0 \text{ when } n \to \infty, \tag{8}$$

and we compare the last two methods.

3.1. A graphical approach

The graphical method should be applied prior to any numerical investigation, this method consists of using the plot

$$\{(k,\ \hat{\eta}\,(k)) : k = 1, \ldots, n\}.$$

The selection of k_{opt} is graphically illustrated in Figure 1.

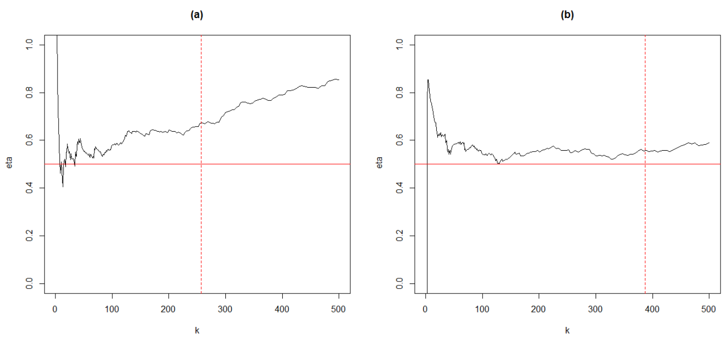

Fig. 1. Estimators of η versus k for Farlie-Gumbel-Morgenstern distribution, average over 500 simulations, (a) Hill estimator, (b) Moment estimator, horizontal line: true η, vertical line: k_{opt}.

3.2. Mean squared error

The MSE of coefficient of tail dependence estimator $\hat{\eta}$ is defined by

$$MSE(\hat{\eta}) := E_{\infty}(\hat{\eta} - \eta)^2,$$

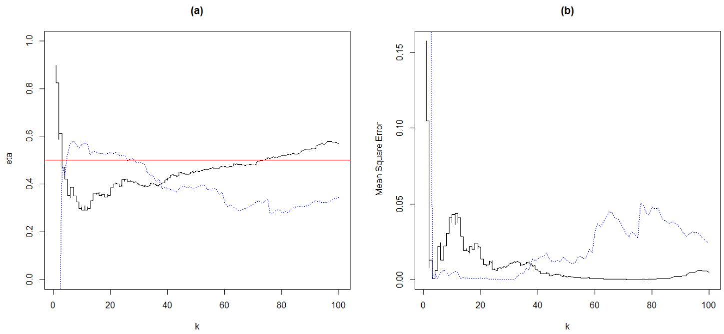

Fig. 2. Estimators of η versus k for Farlie-Gumbel-Morgenstern distribution, average over 500 simulations, (a) Hill estimator, (b) Moment estimator, horizontal line: true η, vertical line: k_{opt}.

where E_∞ denotes the expectation with respect to the limit distribution. The optimal choice k_{opt}, corresponds to the smallest MSE, i.e.

$$k_{opt} := \arg\min_k MSE(\hat{\eta}).$$

3.3. Approach of Reiss and Thomas

Neves and Fraga (2004) proposed an procedure that the adequate number k should be the value which minimizes a mean distance encapsulating a penalty term by minimizing

$$\frac{1}{k}\sum_{i\leq k} i^{\beta}\left|\hat{\eta}_n(i) - \text{median}(\hat{\eta}_1, ..., \hat{\eta}_n)\right|, \qquad 0 \leq \beta \leq 1/2. \tag{9}$$

They also suggested minimizing equation (9):

$$\hat{k}_{opt} = \arg\min_k \frac{1}{k-1}\sum_{i<k} i^{\beta}\left(\hat{\eta}_i - \hat{\eta}_k\right)^2, \qquad 0 \leq \beta \leq 1/2. \tag{10}$$

3.4. Approach of Cheng and Peng

The coverage accuracy for Hill estimator is evaluated see Cheng and Pan (1998) and the theoretical optimal choice of the sample fraction for the one-sided confidence interval is given in terms of minimizing the absolute coverage error by Cheng and Peng (2001) but it depends on some unknown quantities. Cheng and Peng (2001) proposed a plug-in estimator for the optimal sample fraction as a result

$$\hat{k}_{opt} := \begin{cases} \left(\dfrac{(1+2z_\alpha^2)}{3\hat{\delta}(1+2\hat{\rho})}\right)^{1/(1+\hat{\rho})} n^{\hat{\rho}/(1+\hat{\rho})} & \text{if } \hat{\delta} > 0, \\ \left(\dfrac{(1+2z_\alpha^2)}{-3\hat{\delta}}\right)^{1/(1-\rho)} n^{-\hat{\rho}/(1-\hat{\rho})} & \text{if } \hat{\delta} < 0, \end{cases}$$

or

$$\hat{\rho} := -\log\left(\left|\frac{M_n^{(2)}\left(n/\left(2\sqrt{\log n}\right)\right) - 2\left\{M_n^{(1)}\left(n/\left(2\sqrt{\log n}\right)\right)\right\}^2}{M_n^{(2)}\left(n/\sqrt{\log n}\right) - 2\left\{M_n^{(1)}\left(n/\sqrt{\log n}\right)\right\}}\right|\right)/\log 2,$$

and

$$\hat{\delta} := (1+\hat{\rho})\left(\log n\right)^{\hat{\rho}/2}\frac{M_n^{(2)}\left(n/\sqrt{\log n}\right) - 2\left\{M_n^{(1)}\left(n/\sqrt{\log n}\right)\right\}^2}{2\hat{\rho}\left\{M_n^{(1)}\left(n/\sqrt{\log n}\right)\right\}^2},$$

where $M_n^{(1)}$ and $M_n^{(2)}$ are Hill estimator and Moment estimator respectively.

4. Simulation results

In the simulation experiment we compare the optimal number of upper order statistics used in estimation for the coefficient of tail dependence (in the sense of mean square error) obtained with two methods listed in Section 3.

We examine the small sample behavior of the linear estimator given in (5), the Hill estimator and Moment estimator for different distributions, with Fréchet margins. For each of the distributions we generated 1000 samples of size $n = 500$ and selected the number of upper order statistics used in the estimation then we compute the estimators of η.

The parameter w was introduced to estimate probabilities in joint tail regions, but has little practical relevance for the estimation of η, therefore we fix it at 0.5.

First, we performed simulations from a Farlie-Gumbel-Morgenstern (FGM) distribution with Fréchet marginal ($\alpha = -1$), and from the Normal distribution with Fréchet marginal ($\rho = -0.5$).

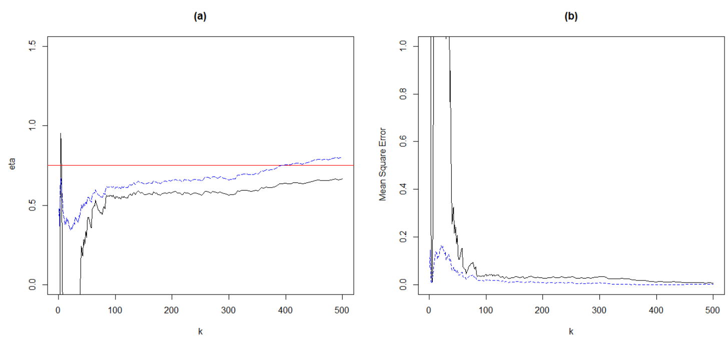

Fig. 3. Farlie Gumbel Morgenstern copula with $\alpha = -1$: (a) estimators for η, (b) MSE: (full line) Hill estimator, (dashed line) Moment estimator, based on 1000 simulations. The horizontal line corresponds to the true value of η.

Next we examine the sample behavior of the estimators for three different distributions, For each of the distributions we compute the MSE, the true η and k_{opt}. In Table 1, we present

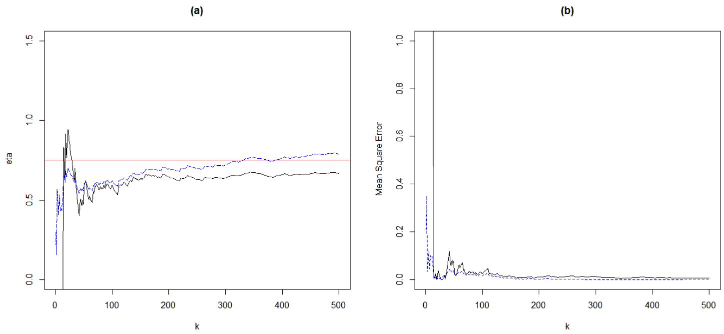

Fig. 4. Farlie Gumbel Morgenstern copula with $\alpha = 0.75$: (a) estimators for η, (b) MSE: (full line) linear estimator, (dashed line) Hill estimator, based on 1000 simulations. The horizontal line corresponds to the true value of η.

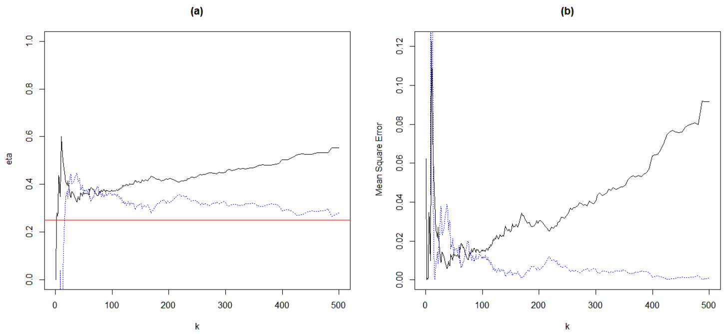

Fig. 5. Normal copula with $\rho = -0.5$: (a) estimators for η, (b) MSE: (full line) Hill estimator, (dashed line) Moment estimator, based on 1000 simulations. The horizontal line corresponds to the true value of η.

some simulation results to FGM ($\alpha = 0.6$), bivariate Normal ($\rho = -0.5$) and Frank ($\alpha = 1$) based on 1000 simulated samples of sizes $n = 300, 500$ and 1000.

In Table 2, we present some simulation results based on simplified procedure of Cheng and Peng method and Reiss and Thomas method to the bivariate FGM ($\alpha = -0.25$ and 0.75) and bivariate Normal ($\rho = -0.5$ and 0.5) distribution, both with Fréchet marginal, based on 1000 simulated samples of size $n = 500$. The comparison results are presented in Tables 2 and 3 illustrated in Figure 6.

We apply the method of Reiss and Thomas with the Moment estimator. For $\hat{\eta}_M < 0$ we take $\beta = 0.35$ in version (10). If $\hat{\eta}_M \geq 0$, select $\beta = 0.4$ in version (9). Table 3

By applying adaptive methods for selecting a good value of extreme order statistics (Cheng and Peng and Reiss and Thomas), we select the optimal sample fraction k_{opt} the optimal number of upper extremes used in the computation of the coefficient of tail dependence estimate, the results are summarized in Table 4 and from this Table, we can say that the

	$\hat{\eta}_H$			$\hat{\eta}_M$		
	η	K_{opt}	MSE	η	K_{opt}	MSE
FGM distribution						
$n = 300$	0.5008	179	6.2821×10^{-7}	0.5001	193	1.1189×10^{-8}
$n = 500$	0.4945	257	3.0081×10^{-7}	0.4993	387	4.2669×10^{-7}
$n = 1000$	0.5008	294	6.8282×10^{-7}	0.4996	957	1.3123×10^{-7}
Bivariate Normal distribution						
$n = 300$	0.2505	143	2.8621×10^{-6}	0.2497	193	5.2328×10^{-7}
$n = 500$	0.2486	229	1.9785×10^{-6}	0.2494	390	2.9245×10^{-7}
$n = 1000$	0.2521	336	4.7927×10^{-6}	0.2521	890	4.6871×10^{-7}
Frank distribution						
$n = 300$	0.4999	185	2.6563×10^{-7}	0.4998	289	2.6915×10^{-8}
$n = 500$	0.4976	321	5.6776×10^{-7}	0.5001	346	3.7154×10^{-8}
$n = 1000$	0.5080	501	6.4849×10^{-7}	0.4993	691	3.9494×10^{-8}

Table 1. Simulation results of the threshold selection procedure for Hill estimator and Moment estimator.

Distribution	MSE	$\hat{K}_{opt}$	$\hat{\eta}_H$
FGM $\alpha = 0.6$	6.8282×10^{-7}	51	0.4784
FGM $\alpha = -0.25$	2.1494×10^{-7}	38	0.4906
Bivariate normal $\rho = -0.5$	1.7963×10^{-7}	25	0.2442
Bivariate normal $\rho = 0.25$	4.7452×10^{-7}	60	0.7049

Table 2. Simulation results of the threshold selection procedure for Hill estimator with Chang and Peng method.

Distribution	MSE	$\hat{K}_{opt}$	$\hat{\eta}_M$
FGM ($\alpha = 0.6$)	1.1667×10^{-4}	201	0.5532
FGM ($\alpha = -0.25$)	1.1047×10^{-4}	260	0.2498
Bivariate normal ($\rho = -0.5$)	1.1446×10^{-4}	246	0.1846
Bivariate normal ($\rho = 0.5$)	1.1524×10^{-7}	335	0.7254

Table 3. Simulation results of the threshold selection procedure for Moment estimator with Reiss and Thomas method.

two methods applied to determine k give results very close to the true value of η, but the method of Cheng and Peng is faster than the other method. The method of Reiss and Thomas is long since it is based on the calculation of the median.

4.1. Real data

We consider here the estimation of η in a real case: Loss-ALAE data in the log scale (for details see Frees and Valdez, 1998). Each claim consists of an indemnity payment (the loss,

Method	# of extremes	% of extremes	Estimate/True η
Cheng and Peng	90	9%	0.5092/0.5
Reiss and Thomas	131	13.1%	0.4679/0.5

Table 4. Optimal numbers of upper order statistics used in the computation of Hill's estimate of FGM (0.75) distribution, based on 1000 observations.

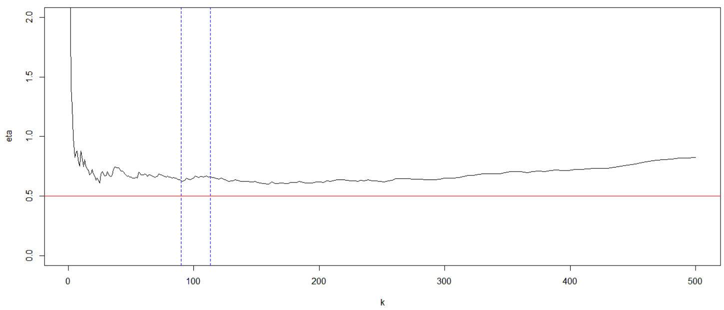

Fig. 6. Hill estimator of Farlie Gumbel Morgenstern copula with $\alpha = 0.75$, based on 1000 simulations. The horizontal line represents the true value of η whereas the vertical lines correspond to the optimal numbers of extremes of Cheng and Peng (solid) and Reiss and Thomas (dashed).

X) and an allocated loss adjustment expense (ALAE, Y). Examples of ALAE are the fees paid to outside attorneys, experts, and investigators used to defend claims. The data size is $n = 1500$.

First the bivariate data are transformed to Fréchet marginal (Z_1, Z_2) using the empirical distribution functions $\hat{F}_X(x)$, $\hat{F}_Y(x)$:

$$Z_1 = -1/\log \hat{F}_X(X), \quad Z_2 = -1/\log \hat{F}_Y(Y).$$

In Figures 6(a), 7(a) and 8(a), we give different estimators for η and we note that the meaning of k_{opt} is different from the one in the estimators of $\hat{\eta}_H$, $\hat{\eta}_M$ and $\hat{\eta}_L$ respectively.

We compare two estimators, Hill and Moment in ALAE-loss data, the results in Figure 10(a) and their MSE in Figure 10(b).

We illustrate the performance of k_{opt} in a little simulation study, the results summarized in Table 5, where we present also the percentage of extremes and MSE.

5. Conclusion

In this paper, we introduced the estimation of the coefficient of tail dependence in bivariate extreme value statistics. We have compared three different estimators of this coefficient. We focus on selection of upper order statistics used in estimation. We note that most methods are based on minimizing the asymptotic mean square error. We described two adaptive selection

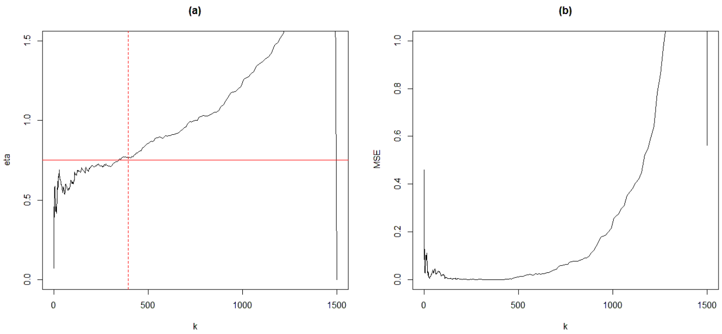

Fig. 7. Loss-ALAE data : (a) Hill estimator for η, (b) MSE: The horizontal reference line corresponds to the true value of η.

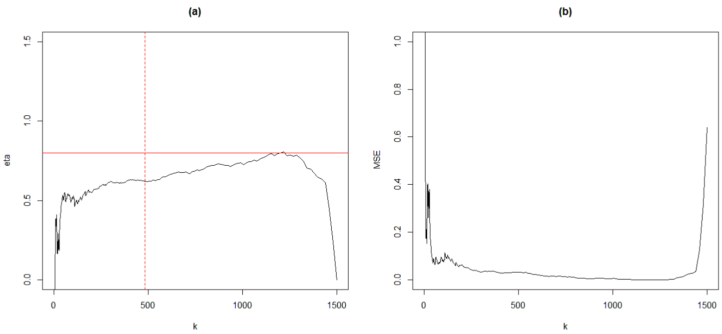

Fig. 8. Loss-ALAE data: (a) Moment estimator for η, (b) MSE: The horizontal reference line corresponds to the true value of η.

Estimator	# of extremes	% of extremes	Estimate/True η	MSE
$\hat{\eta}_H$	419	27.93%	0.8019/0.8	3.8823×10^{-6}
$\hat{\eta}_M$	484	32.26%	0.7990/0.8	9.0713×10^{-7}
$\hat{\eta}_L$	454	30.26%	0.8052/0.8	2.7151×10^{-5}

Table 5. Optimal numbers of upper order statistics used in the computation of estimates coefficient of tail dependence for loss-ALAE data.

methods for the Hill and Moment estimator and applied them to a practical example, the simulations indicated that in general the results are quite similar.

Acknowledgments

We are grateful to the reviewer for his pertinent comments which allowed us to improve our paper.

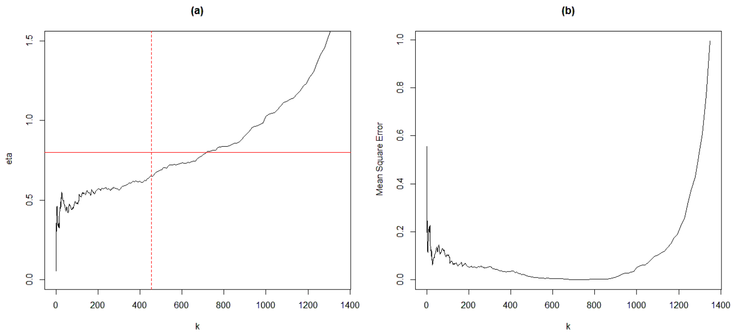

Fig. 9. Loss-ALAE data: (a) linear estimator for η, (b) MSE: The horizontal reference line corresponds to the true value of η.

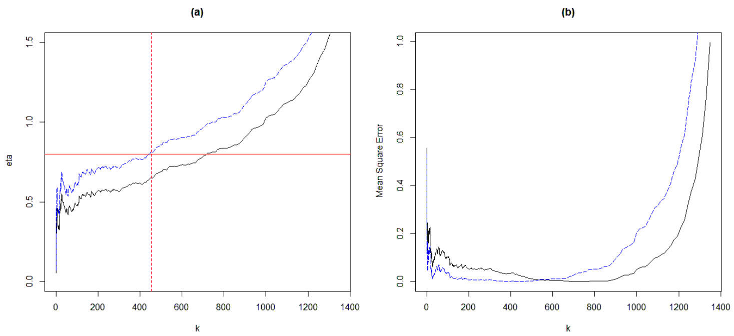

Fig. 10. Loss-ALAE data: (a) estimators for η, (b) MSE: (full line) linear estimator, (dashed line) Hill estimator, based on 1000 simulations. The horizontal line corresponds to the true value of η.

References

Beirlant, J., Joossens, E., Segers, J., 2009. Second-order refined peaks-over-threshold modelling for heavy-tailed distribution. *Journal of Statistical Planning and Inference.* **139**, 2800–2815.

Beirlant J., Dierckx G. and Guillou A., 2011. Reduced bias estimators in joint tail modelling. *Insurance: Mathematics and Economics.* **49**, 18–26.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. *Statistics of Extremes Theory and Applications.* Wiley.

Beirlant, J. and Vandewalle, B., 2002. Some comments on the estimation of a dependence index in bivariate extreme value in statistics. *Statist. Probab. Lett.* **60**, 265–278.

Cheng, S. and Pan, J., 1998. Asymptotic expansions of estimators for the tail index with applications. *Scand. J. Statist.* **25**, 717–728.

Cheng, S. and Peng, L., 2001. Confidence intervals for the tail index. *Bernoulli.* **7**, 751-760.

Danielsson, J., de Haan, L., Peng, L. et de Vries, C.G., 2001. Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation. *Journal of Multivariate Analysis.* **76**, 226-248.

Dekkers, A. L. M., Einmahl, J. H. J. & de Haan, L., 1989. A moment estimator for the index of an extreme-value distribution. *Ann. Statist.* **17**, 1833–1855.

Dekkers, A.L.M. et de Haan, L., 1993. Optimal choice of sample fraction in extreme-value estimation. *J. Multivariate Anal.* **47**, 173–195.

Draisma, G., de Haan, L., Peng, L. et Perreira, T.T., 1999. A Bootstrap-Based Method to Achieve Optimality in Estimating the Extreme-Value Index. *Extremes.* **2**, 367-404.

Draisma G., Drees H., Ferreira A. and de Haan L., 2004. Bivariate tail estimation: dependence in asymptotic independence, *Bernoulli.* **10**, 251-280.

Drees, H. et Kaufmann, E., 1998. Selection of the Optimal Sample Fraction in Univariate Extreme Value Estimation. *Stochastic Processes and their Applications.* **75**, 149-195.

Dutanga, C., Goegebeurb, Y., Armelle Guillou, A., 2014. Robust and bias-corrected estimation of the coefficient of tail dependence. *Insurance: Mathematics and Economics.* **57**, 46–57.

Fraga Alves, M.I., Gomes, M.I., de Haan, L., 2003. A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica.* **60**, 193–214.

Frees, E.W., Valdez, E.A., 1998. Understanding relationships using copulas. *North American Actuarial Journal.* **2**, 1–15.

Goegebeur, Y., Guillou, A., 2013. Asymptotically unbiased estimation of the coefficient of tail dependence. *Scand. J. Stat.* **40**, 174–189.

Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. *Ann.Statist.* **3**, 1163–1174.

Statistics for near independence in multivariate extreme values. *Biometrika.* **83**, 169-187.

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

Neves, C. et Fraga Alves, M.I., 2004. Reiss and Thomas' Automatic Selection of the Number of Extremes. *Computational Statistics and Data Analysis.* **47**, 689–704.

Peng, L., 1999. Estimation of the coefficient of tail dependence in bivariate extremes. *Statistics and Probability Letters.* **43**, 399–409.

Weissman, I., 1978. Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association.* **73**, 812–815.