Extremes 3:4, 311–329, 2000

Estimation of a Bivariate Extreme Value Distribution

PHILIPPE CAPÉRAÀ
Département de mathématiques et de statistique, Université Laval, Québec, Canada G1K 7P4

ANNE-LAURE FOUGÈRES
Département de Génie Mathématique et Modélisation, Institut National des Sciences Appliquées, 135 Avenue de Rangueil, 31077 Toulouse Cedex 04, France

[Received December 7, 1999; Revised February 6, 2001; Accepted February 7, 2001]

Abstract. Several threshold methods have been proposed for the purpose of estimating a bivariate extreme value distribution from a sample of data whose distribution is only in its domain of attraction. An integrated view of these methods is presented which leads to the introduction of a new asymptotically consistent estimator of the dependence function characterizing the extreme dependence structure. Through Monte Carlo simulations, the new estimator is also shown to do as well as its competitors and to outperform them in cases of weak dependence. To the authors' knowledge, this is the first time that the small-sample behavior of nonparametric bivariate threshold methods has ever been investigated.

Key words. Archimax copulas, bivariate threshold methods, dependence functions, extreme value distributions, nonparametric estimation

AMS 1991 Subject Classification. Primary—62H05 60G70 60G05.

1. Introduction

System failures, resistance of materials, and hydrological problems such as river floods are but a few examples of areas of applications where extreme value modeling is appropriate and useful. Although they are not yet as standard as univariate techniques, several approaches have been developed recently to deal with multivariate (and particularly with bivariate) extreme value data. Most probabilistic results on which these methods are based can be found in Galambos (1987) and Resnick (1987).

Given a bivariate random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$, much of extreme value theory is concerned with the limiting behavior of a suitable normalization of the componentwise maxima $(M_{1,n}, M_{2,n})$, where $M_{1,n} = \max(X_1, \ldots, X_n)$, $M_{2,n} = \max(Y_1, \ldots, Y_n)$. More precisely, it is assumed that there exists a non-degenerate bivariate distribution function L such that, as $n \to \infty$,

$$P\{(M_{1,n} - b_{1,n})/a_{1,n} \le x, (M_{2,n} - b_{2,n})/a_{2,n} \le y\} \to L(x, y),$$

for sequences $a_{j,n} > 0$, $b_{j,n} \in \mathbb{R}$, $j = 1, 2$. These limiting distributions coincide with the max-stable distributions (cf. e.g. Resnick, 1987). To analyze separately the behavior of the marginals and the dependence structure of the distribution, it is convenient to write $L(x, y) = C\{F(x), G(y)\}$ in terms of univariate extreme value margins F and G and a ''copula'' C (cf. e.g. Sklar, 1959) defined for all $0 \le u, v \le 1$ by

$$C(u, v) = P\{F(X) \le u, G(Y) \le v\} = \exp[\log(uv)A^\star\{\log(u)/\log(uv)\}], \tag{1}$$

where $A^\star$ is a convex function on $[0, 1]$ such that $\max(t, 1-t) \le A^\star(t) \le 1$ for all $0 \le t \le 1$ (Pickands, 1981). Given this representation, bivariate extreme value theory for componentwise maxima can focus on the so-called univariate dependence function $A^\star$.

As pointed out by Smith et al. (1990), relatively little is known about the statistical properties of the estimation methods that have been developed in this context. These techniques can be classified in two sets, according to the nature of the data available. Situations where samples of componentwise maxima are at hand were originally studied by Pickands (1981). More recent contributions include those of Tawn (1988), Tiago de Oliveira (1989), Smith et al. (1990), Deheuvels (1991) and Capéraà et al. (1997). In these papers, the representation (1) is used to estimate the bivariate extreme value dependence structure directly.

In many situations, however, samples of maxima are not available. In such circumstances, one option briefly mentioned by Pickands (1975) would consist of forming groups of data and applying the above methods to group maxima. Unfortunately, simulations reported by Jacques (1998) show that this approach often yields poor results. The problem has been treated in a different way by Coles and Tawn (1991), Joe et al. (1992), de Haan and Resnick (1993), Einmahl et al. (1993, 1997, 1998), as well as de Haan and de Ronde (1998). These authors exploit the asymptotic Poisson process representation of the point process associated with the original bivariate random sample (e.g. de Haan, 1985, or Resnick, 1987, Chapter 5). The main steps of this approach will be described in the following section, as well as the different methods of estimation of the dependence structure to which it gives rise. It appears that there are practically no results concerning the finite-sample behavior of these so-called nonparametric bivariate threshold methods, as already mentioned by Capéraà et al. (2000). Note that some results can be found for small sample studies of parametric methods, see for example Coles' contribution to Davison and Smith's paper (1990).

The purpose of this paper is two-fold. To begin with, a new estimator of the dependence function will be proposed, for situations where data arise from a distribution belonging to the domain of attraction of a bivariate extreme value distribution. Next, the proposed estimator and alternative threshold methods will be compared, via Monte Carlo simulations. Suitable transformations based on the ranks of the original data are used to circumvent the fact that the marginals are not known. The estimation of the extreme value distribution can therefore be reduced to that of the dependence function $A^\star$, and the comparison is made in terms of different estimators of $A^\star$. In addition to being consistent, the new estimator is also shown to do as well as its competitors and to outperform them in terms of L_1 error in cases of weak dependence.

The main results concerning the point process representation of bivariate extremes are recalled in Section 2. Various threshold methods of estimation based on this representation are reviewed in Section 3, as well as estimators of the dependence function $A^\star$ and their properties. The new estimator of $A^\star$ is then defined and studied in Section 4. The following section describes the bivariate distributions used in the Monte Carlo experiment. After discussing practical issues concerning the implementation of the estimators in Section 6, the results of the comparisons are reported in Section 7, where an illustration of the extremal independence case is also presented, as the limit of the bivariate Normal distribution. Concluding remarks and further developments are discussed in Section 8. Mathematical developments are relegated to appendices.

2. Point process representation

The results sketched here are due to de Haan and Resnick (1977) and can also be found in Resnick (1987, Chapter 5). Let us consider a random sample $(X_1, Y_1), \dots, (X_n, Y_n)$ of a bivariate distribution, and without loss of generality, transform these data to $(Z_{1i}, Z_{2i}), i = 1, \dots, n$ with unit Fréchet margins $P(Z_{ji} \leq z) = e^{-1/z}$ for all $z > 0$ and for $j = 1, 2$ (see Section 6 for more practical details). Observations (Z_{1i}, Z_{2i}) are assumed to belong to the domain of attraction of an extreme value distribution L with unit Fréchet margins and dependence function $A^\star$. In view of (1), L is then of the form5

$$L(z_1, z_2) = \exp[-(1/z_1 + 1/z_2)A^\star\{z_2/(z_1 + z_2)\}]. \tag{2}$$

As is well known, such an assumption on the (Z_{1i}, Z_{2i})'s is equivalent to the statement that

$$nP\{1/n(Z_{1i}, Z_{2i}) \in \cdot\} \xrightarrow{v} \nu(\cdot), \tag{3}$$

in terms of a measure ν on $\mathbb{R}_+^2 \backslash \{(0,0)\}$ satisfying $t\nu(tB) = \nu(B)$, for all Borel sets of $\mathbb{R}_+^2 \backslash \{(0,0)\}$ and $t > 0$. Another formulation of (3) is that the point process on $\mathbb{R}_+^2$ associated with $\left\{\frac{1}{n}(Z_{11}, Z_{21}), \dots, \frac{1}{n}(Z_{1n}, Z_{2n})\right\}$ converges weakly to a non homogeneous Poisson process on $\mathbb{R}_+^2 \backslash \{(0,0)\}$ with intensity measure ν. Applying the polar transformation $T : \mathbf{z} = (z_1, z_2) \mapsto (r, \mathbf{w}) = (\|\mathbf{z}\|, \mathbf{z}/\|\mathbf{z}\|)$, this homogeneity property allows us to express ν as a product measure

$$\nu \circ T^{-1}(\mathrm{d}r \times \mathrm{d}\mathbf{w}) = 1/r^2 \mathrm{d}r\, \mathrm{d}S(\mathbf{w}),$$

in terms of a finite measure S on $\mathcal{N} = \left\{\omega \in \mathbb{R}_+^2 : \|\omega\| = 1\right\}$ such that $= \int_{\mathcal{N}} \omega_i\, \mathrm{d}S(\mathbf{w}) = 1$ for $i = 1, 2$, where $\mathbf{w} = (\omega_1, \omega_2)$. Here, $\|\cdot\|$ is an arbitrary norm on $\mathbb{R}^2$, which differs according to the authors: for example, Joe et al. (1992) use $\|\mathbf{z}\| = |z_1 + z_2|$, whereas Einmahl et al. (1993) take $\|\mathbf{z}\| = (z_1^2 + z_2^2)^{1/2}$, and Einmahl et al. (1997) consider $\|\mathbf{z}\| = z_1 \vee z_2$, as we will see in the following section.

An alternative representation of L in terms of S, due to Pickands (1981), is then (cf. e.g. Proposition 5.11 of Resnick, 1987)

$$L(z_1, z_2) = \exp\left\{-\int_{\mathcal{N}} \max(w_1/z_1, w_2/z_2) \mathrm{d}S(\mathbf{w})\right\}. \tag{4}$$

Putting equations (2) and (4) together, the dependence function $A^\star$ may thus be expressed in the form

$$A^\star(t) = \int_{\mathcal{N}} \max\{tw_1, (1-t)w_2\} \mathrm{d}S(\mathbf{w}) \tag{5}$$

for all $0 \le t \le 1$. Because of this relation, estimation of $A^\star$ can be deduced from estimation of S. In most methods, the latter is estimated via

$$\lim_{t\to\infty} tP(W \in \Delta, R > t) = S(\Delta), \tag{6}$$

an equivalent form of (3) in which (R, W) denotes $T(Z_1, Z_2)$ and Δ is any Borel set on $\mathcal{N}$. In view of relation (6), it is clear why the estimation techniques are referred to threshold methods.

Several different estimators then emerge, depending on the choice of norm. If $\|\mathbf{z}\| = |z_1 + z_2|$ for example, then (5) reduces to

$$A^\star(t) = \int_0^1 \max\{t\omega, (1-t)(1-\omega)\} \mathrm{d}S(\omega), \tag{7}$$

where S and T are respectively renamed as $S(\omega) = \nu \circ T^{-1}((0,\omega) \times (1,\infty))$ for all $0 \le \omega \le 1$, and $T(z_1, z_2) = (r, w) = (z_1 + z_2, z_1/(z_1 + z_2))$ for all $z_1, z_2 > 0$. In this case, the conditions on S are simply

$$\int_0^1 \omega \,\mathrm{d}S(\omega) = \int_0^1 (1-\omega)\mathrm{d}S(\omega) = 1. \tag{8}$$

Finally, equation (6) can be rewritten as

$$\lim_{t\to\infty} tP\{(W, R/t) \in (0,\omega] \times (1,\infty)\} = S(\omega). \tag{9}$$

Roughly speaking, an estimator of S therefore arises from the use of the empirical measure of the $(W_i, R_i/t)$s, namely $1/n \sum_{i=1}^n \mathbb{1}\{(W_i, R_i/t) \in \cdot\}$, provided that t is a suitable function of n, chosen in such a way as to ensure convergence. More precisely, this is achieved for all t of the form n/k_n, in terms of a sequence $\{k_n, n \in \mathbb{N}\}$ of integers such that $k_n \to \infty$ and $n/k_n \to \infty$ (Resnick, 1986).

3. Existing estimators and their properties

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a random sample with unknown marginals, and $(Z_{11}, Z_{21}), \ldots, (Z_{1n}, Z_{2n})$ be the corresponding transformed data as specified in Section 2 (a suitable transformation is given in Section 6). Suppose that this sample belongs to the domain of attraction of an extreme value distribution L with unit Fréchet margins and dependence function $A^\star$.

Three possible estimators of $A^\star$ based on threshold methods will be described in this section. The first two are derived from the work of Einmahl et al. (1993, 1997, 1998). In these papers, attention is given to a polar transformation T that slightly differs from that given in Section 2, namely $T : \mathbf{z} = (z_1, z_2) \mapsto (||\mathbf{z}||, \arctan(z_2/z_1))$. In this case, the measure ν can be expressed as

$$\nu \circ T^{-1}(\mathrm{d}r \times \mathrm{d}\theta) = 1/r^2 \mathrm{d}r\, \mathrm{d}\Phi(\theta),$$

where Φ is a finite measure on $[0, \pi/2]$ called the spectral measure that satisfies specific constraints (cf. Einmahl et al., 1997). Einmahl et al. (1993) opt for the Euclidean norm $||\mathbf{z}|| = (z_1^2 + z_2^2)^{1/2}$ and make use of equation (6), which they rewrite, for all $\theta \in [0, \pi/2]$, as

$$\lim_{t \to \infty} tP(\Theta \leq \theta, R > t) = \Phi(\theta),$$

where $(R, \Theta) = T(Z_1, Z_2) = ((Z_1^2 + Z_2^2)^{1/2}, \arctan(Z_2/Z_1))$. The empirical distribution function of the variables Θ_i, whose corresponding R_i are taken over a suitable threshold, provides an estimator of $\Phi(\cdot)$. Einmahl et al. (1993) make assumptions which ensure that $\Phi(\pi/2) = 1$. In order to relax these assumptions, an estimation of $\Phi(\pi/2)$ is proposed in this paper (see Section 6), via the property that $\Phi(\pi/2) = \lim_{r \to \infty} r P(R > r)$ (cf. e.g. Proposition 5.17 of Resnick, 1987).

To deduce an estimator of $A^\star$ from that of Φ, introduce $b_t = \{1 + t^2/(1-t)^2\}^{-1/2}$ and $c_t = \arccos b_t$ for all $0 \leq t \leq 1$. A simple calculation shows (see Appendix 1) that

$$A^\star(t) = 1 - t \int_{c_t}^{\pi/2} \Phi(\theta) \sin\theta \,\mathrm{d}\theta + (1-t) \int_0^{c_t} \Phi(\theta) \cos\theta \,\mathrm{d}\theta. \tag{10}$$

An estimate $\widetilde{A}_n^{(1)}$ of $A^\star$ is then deduced from this expression when $\Phi(\cdot)/\Phi(\pi/2)$ is replaced by the empirical distribution function

$$\mathbb{1}\varphi_n(\theta) = 1/k_n \sum_{i=1}^{n} \mathbb{1}\{\Theta_i \leq \theta, R_i \geq n/k_n\}, \tag{11}$$

where $\{k_n, n \in \mathbb{N}\}$ is a sequence of integers such that $k_n \to \infty$. Einmahl et al. (1993) prove the strong consistency and asymptotic normality of φ_n under the additional assumptions that $k_n/n \to 0$ and $k_n/\log n \to \infty$. Although their results assume equal margins, the general situation has been considered by Huang (1992, Chapter 2).

The second estimator considered is derived in an analogous way, but using the norm $||\mathbf{z}|| = z_1 \vee z_2$ as Einmahl et al. (1997) do. Making use of the sup-norm yields an expression of the form $A^\star(t) = A_1(t)\mathbb{1}\{t \leq 1/2\} + A_2(t)\mathbb{1}\{t > 1/2\}$, where

$$A_1(t) = 1 - t + (1-t)\int_0^{d_t} \Phi(\theta)/\cos^2\theta \, \mathrm{d}\theta,$$

$$A_2(t) = t + (1-t)\Phi(\pi/2) - t\int_{d_t}^{\pi/2} \Phi(\theta)/\sin^2\theta \, \mathrm{d}\theta,$$

and $d_t = \arctan\{t/(1-t)\}$, for $t \in (0,1)$. The estimate $\widetilde{A}_n^{(2)}$ is then obtained by replacing $\Phi(\cdot)/\Phi(\pi/2)$ by φ_n, where R_i is now equal to $Z_{1i} \vee Z_{2i}$ instead of $(Z_{1i}^2 + Z_{2i}^2)^{1/2}$. Strong consistency as well as asymptotic normality are obtained in the same way and under analogous conditions as above. Note that $\widetilde{A}_n^{(\ell)}(0) = \widetilde{A}_n^{(\ell)}(1) = 1$, for $\ell = 1, 2$.

The third approach considered is due to Joe et al. (1992). These authors make use of the fact that (3) implies

$$nP[1/n \max\{Z_{1i}, Z_{2i}(1-t)/t\} \in \cdot\,] \xrightarrow{v} \eta_t(\cdot), \tag{12}$$

for all $0 < t < 1$, in terms of a measure η_t on $\mathbb{R}_+^\star$ such that $\eta_t\{(u,\infty)\} = \mu\{1, t/(1-t)\}/u$ for all $u > 0$, where $\mu(z_1, z_2) = \nu\{([0, z_1] \times [0, z_2])^c\}$, c denoting complementation. As Joe et al. consider the norm $||\mathbf{z}|| = |z_1 + z_2|$, the dependence function $A^\star$ is related to μ by $A^\star(t) = t\,\mu\{1, t/(1-t)\}$, for all $t \in (0,1)$. An estimator $\widetilde{A}_n^{(3)}$ of $A^\star$ can thus be immediately deduced from an estimator of $\eta_t\{(1,\infty)\}$. This latter is obtained using (12), which implies that

$$\lim_{s \to \infty} sP(T_i(t) > s) = \eta_t\{(1,\infty)\},$$

where $T_i(t) = \max\{Z_{1i}, Z_{2i}(1-t)/t\}$. As a result,

$$\hat{\eta}_{t,n}\{(1,\infty)\} = 1/k_n \sum_{i=1}^{n} \mathbb{1}\{T_i(t) \geq n/k_n\}$$

is a convergent estimator of $\eta_t\{(1,\infty)\}$ provided that k_n is such that $k_n \to \infty$ and $k_n/n \to 0$. Although Joe et al. (1992) did not describe the properties of this estimator, its convergence and asymptotic normality can be obtained using similar techniques to those of Einmahl et al. (1993, 1997).

4. A new estimator

Combining (7) and (9), it would seem natural, as suggested by de Haan (1985), to estimate $A^\star(t)$ by

$$\widetilde{A}_n(t) = \frac{1}{k_n}\sum_{i=1}^{n} \max\{tW_i, (1-t)(1-W_i)\}\mathbb{1}\{R_i \geq n/k_n\},$$

where $\{k_n, n\in\mathbb{N}\}$ is a sequence of integers such that $k_n \to \infty$ and $k_n/n \to 0$. If this estimator is to be used to derive a bivariate extreme value distribution, however, it would be essential that the estimator of $A^\star$ satisfy the properties of a dependence function. In order to achieve this, we propose the following convenient modification of $\widetilde{A}_n$, viz.

$$\widetilde{A}_n^{(4)}(t) = \max\Big\{\Psi(t), \widetilde{A}_n(t) + (2t-1)(1-\Lambda_n)\Big\}, \tag{13}$$

where $\Lambda_n = 1/k_n \sum_{i=1}^{n} W_i \mathbb{1}\{R_i \geq n/k_n\}$ and $\Psi(t) = \max(t, 1-t)$. The term $(2t-1)(1-\Lambda_n)$ ensures that $\widetilde{A}_n^{(4)}(0) = \widetilde{A}_n^{(4)}(1) = 1$. The basic properties of $\widetilde{A}_n^{(4)}$ are summarized in the following proposition, whose proof is given in Appendix 2.

Proposition 3.1: *Let $\{k_n, n\in\mathbb{N}\}$ be a sequence of integers such that $k_n \to \infty$ and $k_n/n \to 0$. The estimator $\widetilde{A}_n^{(4)}$ defined by (13) is a consistent convex estimator of $A^\star$ such that* $\max(t, 1-t) \leq \widetilde{A}_n^{(4)}(t) \leq 1$.

Note that the strong consistency can also be proved under the additional assumption that $k_n/\log n \to \infty$, as it is the case in de Haan and Resnick (1993).

5. Distributions used in the Monte Carlo experiment

To compare the estimators presented in Sections 3–4 via a Monte Carlo experiment, a selection of bivariate extreme value distributions is required, as well as a choice of laws belonging to each domain of attraction selected. Capéraà et al. (2000) provide a way of constructing distributions with a given maximum attractor, and introduce for this purpose a family of bivariate distributions with uniform margins, called Archimax copulas, which may be expressed in the form

$$C_{\phi,A}(x,y) = \phi^{-1}\left[\{\phi(x)+\phi(y)\}A\left\{\frac{\phi(x)}{\phi(x)+\phi(y)}\right\}\right]$$

for all $0 \leq x, y \leq 1$ in terms of

i. a convex function $A : [0,1] \to [1/2,1]$ such that $\max(t, 1-t) \le A(t) \le 1$ for all $0 \le t \le 1$;
ii. a convex, decreasing function $\phi : (0,1] \to [0,\infty)$ verifying $\phi(1) = 0$, with the convention that $\phi(0) \equiv \lim_{t \to 0^+} \phi(t)$ and $\phi^{-1}(s) = 0$ when $s \ge \phi(0)$.

As shown by Capéraà et al. (2000), who determine the maximum attractor of $C_{\phi,A}$ under suitable regularity conditions, this system of copulas encompasses various extreme value behaviors, whereas most classical parametric families lead to independence.

Different dependence functions $A^\star$ and different pairs of generators A and ϕ were selected for the present study. In particular, four choices of $A^\star$ were used, namely $A_1^\star(t) = 3(t^2 - t)/10 + 1$, $A_2^\star(t) = 2(t^2 - t)/3 + 1$,

$$A_3^\star(t) = \begin{cases} 15t^2/8 - 3t/4 + 1 & \text{if } t \le 1/5, \\ (15t^2 - 6t + 119)/128 & \text{if } t > 1/5, \end{cases}$$

and

$$A_4^\star(t) = \begin{cases} 3t^2/2 - t + 1 & \text{if } t \le 1/3, \\ (3t^2 - 2t + 7)/8 & \text{if } t > 1/3. \end{cases}$$

The first two, which are special cases of Tawn's (1988) mixed model, yield symmetric dependence structures between the extremes with two different degrees of dependence, as measured by Kendall's tau, say. The other two generators are such that $\tau_{A_3^\star} = \tau_{A_1^\star} \approx 0.1$ and $\tau_{A_4^\star} = \tau_{A_2^\star} \approx 0.26$ but are asymmetric.

For each choice of dependence function $A_i^\star$, it was decided to take $A_i = A_i^\star$ and four generators ϕ were selected. Two of them were from Clayton's (1978) family, $\phi_{1,\alpha}(t) = (t^{-\alpha} - 1)/\alpha$, $\alpha > 0$, and the others were from Frank's (1979) system, viz.

$$\phi_{2,\alpha}(t) = -\log\left\{\frac{1 - \exp(-t\alpha)}{1 - \exp(-\alpha)}\right\}, \quad \alpha \in \mathbb{R}.$$

With these choices, $C_{\phi,A}$ belongs to the domain of attraction of the extreme value distribution with dependence function $A = A^\star$. Values of α were chosen to ensure predetermined initial degrees of dependence, namely $\tau_{\phi,A_1} = \tau_{\phi,A_3} = 0.3$ or 0.5 and $\tau_{\phi,A_2} = \tau_{\phi,A_4} = 0.5$ or 0.7. The various cases considered are summarized in Table 1.

6. Implementation of the estimators

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a random sample from a bivariate distribution which is assumed to be in the domain of attraction of an extreme value distribution with dependence function $A^\star$. While the marginal distributions are not assumed to be known, a convenient transformation of the data can be used to ensure that, asymptotically, they have

Table 1. Generators (A, ϕ) used for simulations of Archimax distributions with predetermined Kendall's tau.

	Values of $\tau_{\phi,A}$		
	0.3	0.5	0.7
$\tau_{A^\star} = 0.1$	$(A_1, \phi_{1,4/7})$ $(A_1, \phi_{2,2.09})$ $(A_3, \phi_{1,4/7})$ $(A_3, \phi_{2,2.09})$	$(A_1, \phi_{1,1.6})$ $(A_1, \phi_{2,4.8})$ $(A_3, \phi_{1,1.6})$ $(A_3, \phi_{2,4.8})$	
$\tau_{A^\star} = 0.26$		$(A_2, \phi_{1,0.972})$ $(A_2, \phi_{2,3.23})$ $(A_4, \phi_{1,0.972})$ $(A_4, \phi_{2,3.23})$	$(A_2, \phi_{1,2.953})$ $(A_2, \phi_{2,7.83})$ $(A_4, \phi_{1,2.953})$ $(A_4, \phi_{2,7.83})$

unit Fréchet margins. For example, Joe et al. (1992) suggest the transformation $Z_{ji} = 1/\log\{n/(R_{ji} - 1/2)\}$, $j = 1, 2$, where R_{1i} (resp. R_{2i}) is the rank of X_i (resp. Y_i) among $X_1, \ldots, X_n$ (resp. $Y_1, \ldots, Y_n$). From now on, the transformed random sample $(Z_{11}, Z_{21}), \ldots, (Z_{1n}, Z_{2n})$ will be used, so that the estimation problem is concentrated on the dependence structure.

The estimators of $A^\star$ presented in Sections 3–4 are exclusively based on observations whereof a given function, say $\delta_\ell(Z_{1i}, Z_{2i})$, $i = 1, \ldots, n$, is greater than a fixed threshold. Choosing such a threshold a priori seems delicate in practice, as it could happen that no datum is retained. Another method will be used here, which is both more tractable and asymptotically equivalent. It consists in keeping a given proportion of the observations that correspond to the greatest $\delta_\ell(Z_{1i}, Z_{2i})$. Estimates $\widetilde{A}_n^{(\ell)}$, $\ell = 1, 2$ of $A^\star$ are then rewritten using (11), with condition (a) $R_i \geq n/k_n$ replaced by (b) $R_i \geq R_{[k_{\ell,n}]}$, where $R_{[j]}$ is the $(n - j + 1)$th order statistic of the R_is. These two conditions are shown in Appendix 3 to be asymptotically equivalent whenever $k_n \to \infty$ and $n/k_n \to \infty$.

In the sequel, $\Phi(\pi/2)$ is also estimated by $1/(k_{\ell,n} - i_{\ell,n} + 1) \sum_{i=i_{\ell,n}}^{k_{\ell,n}} iR_{[i]}/n$, $\ell = 1, 2$, where $i_{\ell,n} \in \mathbb{N}$ is a parameter introduced to exclude the largest R_is and to increase the stability of this estimation.

Joe et al. (1992) observe that the number of $T_i(t)$s greater or equal to z is equal to j if $z = T_{[j]}(t)$, for $j \in \mathbb{N}$, so that $\mu(1, t/(1-t))$ can be estimated by $jT_{[j]}(t)/n$. Note at this point that the expectation of $jT_{[j]}(t)/n$ is approximately constant, provided that j is not too small. This arises (cf. Proposition 5.17 of Resnick, 1987) from the equivalent form of (3) in terms of the distribution function H of the (Z_{1i}, Z_{2i})s, viz.

$$\lim_{T \to \infty} \{1 - H(Tz_1, Tz_2)\}/\{1 - H(T, T)\} = \nu(z_1, z_2)/\nu(1, 1).$$

As mentioned by Joe et al. the conditional distribution of the $T_i(t)$s that are bigger than a large enough T has density T/y^2 for $y \geq T$, and hence

$$\mathbb{E}\{jT_{[j]}(t)/n\} = j/(j-1)T. \tag{14}$$

Therefore, large js have to be eliminated, so that the $T_{[j]}(t)$s are large enough to meet the asymptotic requirement; furthermore, small js also have to be excluded to ensure that (14) is approximately constant in j. This leads Joe et al. to estimate $\mu(1, t/(1-t))$ by

$$1/(k_{3,n} - i_{3,n} + 1) \sum_{j=i_{3,n}}^{k_{3,n}} jT_{[j]}(t)/n,$$

in terms of two parameters $i_{3,n}$ and $k_{3,n}$.

Note that the estimators $\widetilde{A}_n^{(\ell)}, \ell = 1, \ldots, 3$, do not satisfy all the conditions of a dependence function [i.e. (i) A convex, and (ii) $\Psi(t) = \max(t, 1-t) \le A(t) \le 1$ for all $0 \le t \le 1$]. However, it is simple enough to modify them as follows, in order to meet condition (ii):

$$\widehat{A}_n^{(\ell)}(t) = \min\Big[\max\Big\{\Psi(t), \widetilde{A}_n^{(\ell)}(t)\Big\}, 1\Big],\ \ \ell = 1, \ldots, 3.$$

This modification is most useful when $A^\star \equiv 1$, as illustrated in Figure 1, where it can be seen that the estimator $\widetilde{A}_n^{(3)}$ of Joe et al. (1992) exceeds 1 on part of the domain.

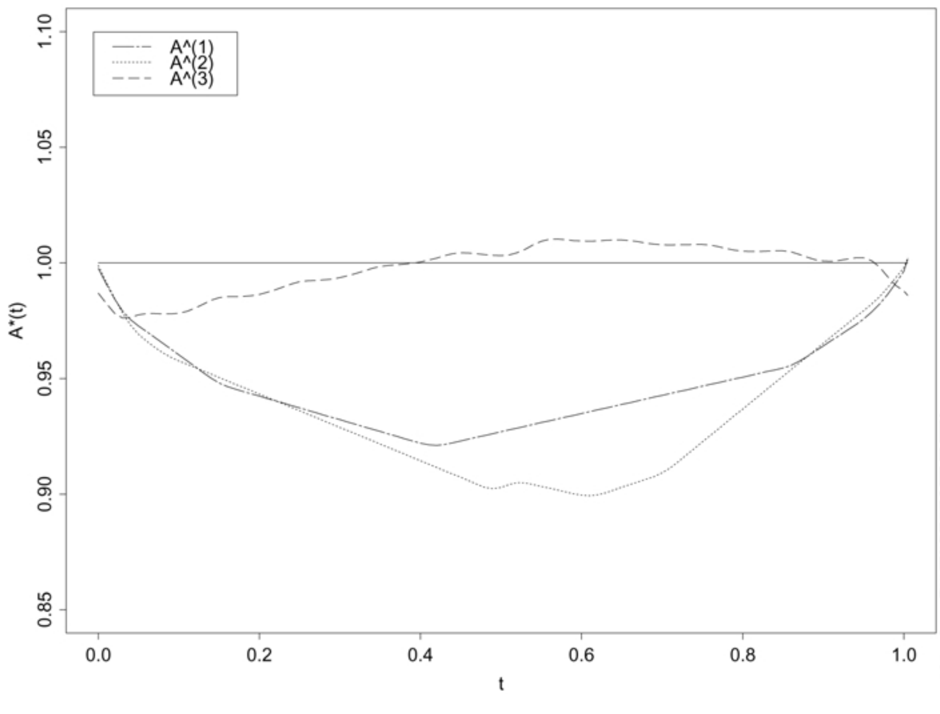

Figure 1. Dependence function estimates $(\widetilde{A}_n^{(\ell)}, \ell = 1, \ldots, 3)$ based on 500 observations of a Normal copula with correlation coefficient $\rho = 0.25$. The solid line is the asymptotic dependence function $A^\star \equiv 1$.

To achieve condition (i), one could also consider, as suggested by Tiago de Oliveira (1989), taking the convex hull of the $\widehat{A}_n^{(\ell)}$s, but as this complicates matters, this avenue is not pursued here. While Figure 1 shows that this additional modification would be necessary for the estimators $\widehat{A}_n^{(2)}$ and $\widehat{A}_n^{(3)}$ of Einmahl et al. (1997) and Joe et al. (1992), it seems likely that the other estimator of Einmahl et al. (1993), viz. $\widehat{A}_n^{(1)}$, is convex, though a formal proof remains to be found.

In small samples, simulations indicate that $\widetilde{A}_n^{(4)}$ defined by (13) has a tendency to overestimate the dependence. A slight modification of this new estimator is therefore proposed, which does not affect its asymptotic behavior. For $1 \leq i \leq k_n$, let $\widetilde{W}_i$ denote the W_is that correspond to the R_is greater than $R_{[k_n]}$, and denote by $\widetilde{W}_{(i)}$ the associated order statistics. The proposed modification consists in allocating lighter weights to central $\widetilde{W}_i$s by setting

$$\widehat{A}_n^{(4)}(t) = \max\left[\Psi(t), 2/K_n \sum_{i=1}^{k_{4,n}} p_i \max\{t\widetilde{W}_{(i)}, (1-t)(1-\widetilde{W}_{(i)})\} + (2t-1)(1-2\widetilde{\Lambda}_n)\right],$$

in terms of weights $p_i = \mathbb{1}_{\{i \leq i_{4,n}\} \cup \{i \geq k_{4,n} - i_{4,n} + 1\}} + 1/2\,\mathbb{1}_{\{i_{4,n}+1 \leq i \leq k_{4,n} - i_{4,n}\}}$, with $K_n = \sum_{i=1}^{k_{4,n}} p_i$ and $\widetilde{\Lambda}_n = 1/K_n \sum_{i=1}^{k_{4,n}} p_i \widetilde{W}_{(i)}$.

In the end, all the $\widehat{A}_n^{(\ell)}$'s depend on two parameters $i_{\ell,n}$ and $k_{\ell,n}$. While optimal selection procedures for these parameters are yet to be found, choices must be made for any comparative study. In order to circumvent this problem, an initial comparison of the $\widehat{A}_n^{(\ell)}$s was made, using several choices of $(i_{\ell,n}, k_{\ell,n})$. For each of them, 100 samples of size $n = 500$, 1000, 2500 were generated from Archimax copulas presented in Section 5, and boxplots of the L_1 errors were compared. In the light of this study, ''empirically optimal'' choices of $(i_{\ell,n}, k_{\ell,n})$ were made for each estimator. It turned out that the estimators $\widehat{A}_n^{(1)}$ and $\widehat{A}_n^{(4)}$ yielded different ''optimal choices'' according to the strength of dependence of the attractor, which is unknown in practice. In the end, however, a single pair had to be used for the simulations. The values selected are given in Table 2. Observe that the number of observations used is really small for $\widehat{A}_n^{(2)}$ and much larger for $\widehat{A}_n^{(3)}$. In the sequel, $\widehat{A}_n^{(\ell)}$ refers to the estimator with parameters $i_{\ell,n}$ and $k_{\ell,n}$ specified in Table 2. Figure 2 provides an illustration of the form taken by these four estimators, based on 500 observations of an Archimax copula with generators A_4 and $\phi_{2,7.83}$.

Table 2. Values of the ''empirically optimal'' parameters $(i_{\ell,n}, k_{\ell,n})$ of the estimators $\widehat{A}_n^{(\ell)}$, $\ell = 1, \ldots, 4$, in terms of sample size n.

	Estimators			
	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(4)}$
$n = 500$	(5,15)	(10,15)	(10,40)	(4,16)
$n = 1000$	(5,15)	(15,20)	(15,60)	(4,16)
$n = 2500$	(5,30)	(15,30)	(15,120)	(7,30)

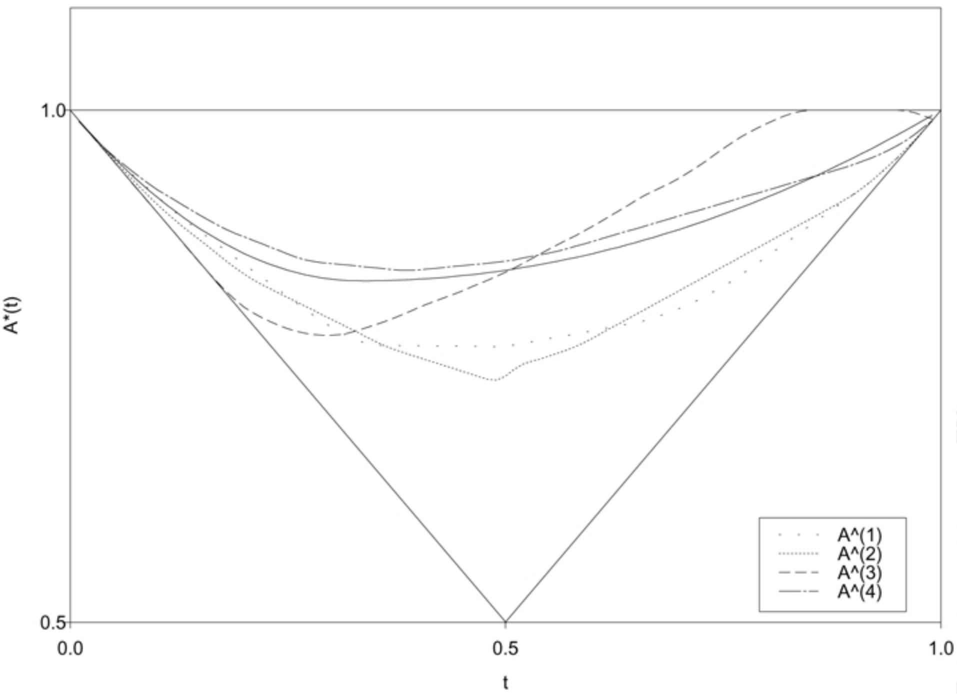

Figure 2. Dependence function estimates $(\widehat{A}_n^{(\ell)}, \ell = 1, \ldots, 4)$ based on 500 observations of an Archimax copula with generators A_4 and $\phi_{2,7.83}$. The solid line is the asymptotic dependence function $A_4^\star$.

7. Comparative study

Data were generated according to a factorial design for three sample sizes $n = 500$, 1000 and 2500, and two strengths of dependence of the attractor $\tau_{A^\star} = 0.1$ and 0.26. Three factors were selected and considered as fixed, namely

i. the choice of estimator $(\widehat{A}_n^{(\ell)}, \ell = 1, \ldots, 4)$.
ii. the presence or absence of symmetry about 1/2 in the function $A^\star$, as represented by generators A_1 or A_3 (resp. A_2 or A_4) in case of weak (resp. strong) dependence, as defined in Section 5.
iii. the type of Archimax distribution, as embodied by generators ϕ specified in Table 1 (four levels in each case).

For each treatment, 100 independent pseudo-random samples of size n were generated, and corresponding L_1 errors of the estimators $\widehat{A}_n^{(\ell)}$ were computed. The empirical mean and standard deviation of the samples of L_1 errors for $n = 500$ are given in Table 3 in order to provide some quantitative informations for comparative purposes.

Table 3. Mean and standard deviation of samples of size 100 of the L_1 errors of the estimators $\widehat{A}_n^{(\ell)}$, $\ell = 1, \ldots, 4$, obtained from samples of size $n = 500$. Means (resp. standard deviations) are multiplied by factor 10^2 for convenience. Standard deviations are specified between parentheses.

	Estimators			
	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(4)}$
$A_1, \phi_{1,4/7}$	2.719 (1.887)	2.963 (1.793)	2.782 (1.520)	1.744 (1.298)
$A_1, \phi_{1,1.6}$	3.315 (1.920)	3.977 (2.134)	2.913(1.466)	1.993 (1.342)
$A_1, \phi_{2,2.09}$	3.344 (1.714)	3.524 (2.397)	3.048 (1.677)	1.866(1.243)
$A_1, \phi_{2,4.8}$	4.628(1.844)	4.482 (2.277)	3.401 (1.922)	2.943 (1.753)
$A_3, \phi_{1,4/7}$	2.607 (1.642)	3.277 (1.870)	2.569 (1.463)	1.881 (1.276)
$A_3, \phi_{1,1.6}$	3.082 (1.819)	3.785 (2.112)	2.967 (1.876)	2.072 (1.315)
$A_3, \phi_{2,2.09}$	3.222 (2.029)	4.024 (2.153)	2.527 (1.331)	2.150 (1.396)
$A_3, \phi_{2,4.8}$	4.945 (2.475)	4.677 (2.299)	3.098 (1.783)	2.882 (1.490)
$A_2, \phi_{1,0.972}$	2.167 (1.350)	2.816 (1.395)	3.281 (1.864)	2.485 (1.358)
$A_2, \phi_{1,2.953}$	2.937 (1.677)	3.445 (1.570)	3.383 (1.715)	2.494 (1.762)
$A_2, \phi_{2,3.23}$	2.496 (1.590)	3.129 (2.051)	3.348 (1.737)	2.030 (1.103)
$A_2, \phi_{2,7.83}$	2.462 (1.383)	3.461 (1.799)	3.394 (2.062)	2.311 (1.330)
$A_4, \phi_{1,0.972}$	2.610 (1.437)	3.498 (1.921)	2.999 (1.613)	2.256 (1.343)
$A_4, \phi_{1,2.953}$	2.707 (1.367)	3.283 (1.759)	3.156 (1.810)	2.299(1.295)
$A_4, \phi_{2,3.23}$	2.759 (1.690)	3.313 (1.768)	2.815 (1.634)	2.301(1.347)
$A_4, \phi_{2,7.83}$	2.444 (1.654)	3.486 (1.744)	3.250 (1.903)	2.448 (1.351)

The experiment yielded six balanced factorial designs with 100 random samples in each cell. The dependent variable used was the logarithm of the L_1 distance between $A^\star$ and its estimation, so that the standard hypotheses of the analysis of variance model be approximately met. The analysis shows that for $n = 500$ and $\tau_{A^\star} = 0.1$, the interaction between the first and second factors is significant (P-value $= 0.003$), as well as the interaction between the first and the third (P-value < 0.0001). When $n = 1000$ and $\tau_{A^\star} = 0.26$, only the latter interaction is significant (P-value < 0.0001). In the two other cases, no interaction turned out to be significant. Tables 4 and 5 summarize the results of paired comparisons between the four estimators by level of the two other factors. This is done for $n = 500$ and 1000, and for the two strengths of dependence of the attractor, though this is not strictly necessary when the interactions are not significant. Essentially identical results were obtained for $n = 2500$. The estimators are ordered from left to right in terms of increasing L_1 error. Non-significant differences at the 0.01 level are underlined. A star ($\star$) at the beginning of a line indicates that the standard hypotheses of the analysis of variance model (e.g. homoscedasticity) were not verified.

First examine the case of weak extremal dependence $\tau_{A^\star} = 0.1$. The results of Table 4 show that $\widehat{A}_{500}^{(4)}$ dominates the other estimators except for generator $\Phi_{2,4.8}$, where $\widehat{A}_{500}^{(4)}$ and $\widehat{A}_{500}^{(3)}$ are not significantly different. This generator also yields different results than the others for estimators $\widehat{A}_n^{(\ell)}$, $\ell = 1, \ldots, 3$. The choice of copula thus seems to have a significant influence in this case. In contrast, symmetry or lack thereof does not appear to play a role, even though for the asymmetric A_3, the estimator $\widehat{A}_{500}^{(2)}$ is not as good as the

Table 4. Paired comparisons of the $\log L_1$ error of four estimators $\widehat{A}_n^{(\ell)}$, $\ell = 1, \ldots, 4$, of the dependence function $A^\star$ of the maximum attractor of a bivariate Archimax distribution, with sample size $n = 500$.

$\tau_{A^\star} = 0.1$

	$A_1, \phi_{1,4/7}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$
$\star$	$A_1, \phi_{1,1.6}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$
	$A_1, \phi_{2,2.09}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(1)}$
$\star$	$A_1, \phi_{2,4.8}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(1)}$
	$A_3, \phi_{1,4/7}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$
	$A_3, \phi_{1,1.6}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$
	$A_3, \phi_{2,2.09}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$
	$A_3, \phi_{2,4.8}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(1)}$

$\tau_{A^\star} = 0.26$

$A_2, \phi_{1,0.972}$:	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(3)}$
$A_2, \phi_{1,2.953}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$
$A_2, \phi_{2,3.23}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(3)}$
$A_2, \phi_{2,7.83}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$
$A_4, \phi_{1,0.972}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$
$A_4, \phi_{1,2.953}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$
$A_4, \phi_{2,3.23}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$
$A_4, \phi_{2,7.83}$:	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$

others in three of the four cases considered. As the interactions are non-significant when $n = 1000$, the results of the analysis of variance show directly that $\widehat{A}_{1000}^{(4)}$ dominates $\widehat{A}_{1000}^{(\ell)}$, $\ell = 1, \ldots, 3$. In this case, $\widehat{A}_{1000}^{(1)}$ and $\widehat{A}_{1000}^{(3)}$ are actually equivalent and dominate $\widehat{A}_{1000}^{(2)}$. Symmetry and copula type don't seem to make a difference.

When extremal dependence is strong $(\tau_{A^\star} = 0.26)$, the analysis of variance results for the case $n = 500$ show in the absence of significant interactions that $\widehat{A}_{500}^{(4)}$ and $\widehat{A}_{500}^{(1)}$ are equivalent; they jointly dominate the other two, which are also undistinguishable. In this case, neither symmetry nor choice of copula seem to matter. When $n = 1000$, similar conclusions can be derived from Table 5, although the difference between the two pairs of equivalent estimators is not so great.

In summary, it stems from these simulations that (1) symmetry or absence thereof matters very little, if at all, in comparing the estimators, (2) the nature of the underlying copula has only a slight influence, mainly for $n = 500$, and (3) the new estimator

Table 5. Paired comparisons of the $\log L_1$ error of four estimators $\widehat{A}_n^{(\ell)}$, $\ell = 1, \ldots, 4$, of the dependence function $A^\star$ of the maximum attractor of a bivariate Archimax distribution, with sample size $n = 1000$.

$\tau_{A^\star} = 0.1$					$\tau_{A^\star} = 0.26$				
$A_1, \phi_{1,4/7}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$	$A_2, \phi_{1,0.972}$:	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(3)}$
$A_1, \phi_{1,1.6}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$	$A_2, \phi_{1,2.953}$:	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$
$A_1, \phi_{2,2.09}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$	$A_2, \phi_{2,3.23}$:	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$
$A_1, \phi_{2,4.8}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$	$A_2, \phi_{2,7.83}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(3)}$
$A_3, \phi_{1,4/7}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$	$A_4, \phi_{1,0.972}$:	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(2)}$
$A_3, \phi_{1,1.6}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$	$A_4, \phi_{1,2.953}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$
$A_3, \phi_{2,2.09}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$	$A_4, \phi_{2,3.23}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(3)}$
$A_3, \phi_{2,4.8}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$	$A_4, \phi_{2,7.83}$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$

dominates the other three when extremal dependence is weak $(\tau_{A^\star} = 0.1)$, whereas when $\tau_{A^\star} = 0.26$, it is preferable to $\widehat{A}_n^{(2)}$ and $\widehat{A}_n^{(3)}$ but not significantly different from $\widehat{A}_n^{(1)}$.

In order to illustrate the case of extremal independence, we proceed next to a comparative analysis using the Normal copula with different values of the correlation coefficient, namely $\rho = 0.25$, 0.50 and 0.75. Table 6 gives for these cases quantitative informations as empirical mean and standard deviation of the samples of L_1 errors for $n = 500$. Unfortunately, it was not possible to carry out an analysis of variance of the results with ρ and the estimators as fixed factors, because of heteroscedasticity (P-value < 0.0001). Three separate analyses were thus conducted using the Welch option of the

Table 6. Mean and standard deviation of samples of size 100 of the L_1 errors of the estimators $\widehat{A}_n^{(\ell)}$, $\ell = 1, \ldots, 4$, obtained from samples of size $n = 500$ of a Normal copula with correlation coefficient ρ. Means (resp. standard deviations) are multiplied by factor 10^2 so that results are on the same scale as in Table 3. Standard deviations are specified between parentheses.

	Estimators			
	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(4)}$
$\rho = 0.25$	4.915 (1.698)	5.180 (1.940)	2.361 (1.940)	3.628 (1.037)
$\rho = 0.50$	8.465 (2.331)	8.368 (2.273)	6.327 (2.680)	6.515 (1.353)
$\rho = 0.75$	13.751 (2.041)	13.493 (2.620)	12.625 (2.948)	11.615 (2.155)

Table 7. Paired comparisons of the $\log L_1$ error of four estimators $\widehat{A}_n^{(\ell)}$, $\ell = 1, \ldots, 4$, of the dependence function $A^\star \equiv 1$ of the maximum attractor of a Normal copula with correlation coefficient ρ, with sample size $n = 500$ and 1000.

	$n = 500$					$n = 1000$			
$\rho = 0.25$:	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$	$\rho = 0.25$:	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$
$\rho = 0.50$:	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(1)}$	$\rho = 0.50$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$
$\rho = 0.75$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(2)}$	$\widehat{A}_n^{(1)}$	$\rho = 0.75$:	$\widehat{A}_n^{(4)}$	$\widehat{A}_n^{(3)}$	$\widehat{A}_n^{(1)}$	$\widehat{A}_n^{(2)}$

SAS ANOVA procedure. The results are given in Table 7 for $n = 500$ and 1000. In this case, the estimator $\widehat{A}_n^{(3)}$ of Joe et al. (1992) turns out to be best for $\rho = 0.25$ and 0.50.

8. Conclusion

A new estimator of the dependence function $A^\star$ of a bivariate extreme value distribution has been proposed in this paper. Comparisons with three competitors show that this estimator has the smallest L_1-error in the case of weak extremal dependence. When dependence is strong, it is equivalent to the estimator $\widehat{A}_n^{(1)}$ of Einmahl et al. (1993), but remains preferable to the other two. In the case of extremal independence obtained from Normal data, the estimator of Joe et al. (1992) provides the best result, except when the correlation in the original data is strong. As pointed out both by Ledford and Tawn (1996) and de Haan and Sinha (1999), however, extreme value models are not adapted to the study of bivariate phenomena whose components are asymptotically independent.

While this paper emphasized the estimation of the dependence function $A^\star$, practical applications of extreme value theory often require the estimation of the probability of extreme events, which also depend on estimates of the first and second derivatives of $A^\star$. This raises additional issues that must be the object of future investigations.

Appendix 1

Derivation of equation (10)

Starting from (5) and using Euclidean norm, $A^\star$ can be rewritten as in (7) in the form

$$A^\star(t) = \int_0^1 \max\left\{t\omega, (1-t)(1-\omega^2)^{1/2}\right\} \mathrm{d}S(\omega),$$

with

$$T(z_1, z_2) = (r, \omega) = \left((z_1^2 + z_2^2)^{1/2}, z_1/(z_1^2 + z_2^2)^{1/2}\right).$$

The transformation $(r, \omega) \mapsto (r, \theta = \arctan\ (z_1/z_2))$ then relates S and Φ via

$$S(\omega) = S(1) - \Phi(\arccos\ \omega),$$

so that

$$\begin{aligned} A^\star(t) &= 1 + t\int_0^{b_t} S(\omega)\mathrm{d}\omega - (1-t)\int_{b_t}^1 \omega(1-\omega^2)^{-1/2} S(\omega)\mathrm{d}\omega \\ &= 1 - t\int_{c_t}^{\pi/2} \Phi(\theta)\sin\theta\,\mathrm{d}\theta + (1-t)\int_0^{c_t} \Phi(\theta)\cos\theta\,\mathrm{d}\theta. \end{aligned}$$

Appendix 2

Proof of Proposition 3.1

Convexity of $\widetilde{A}_n^{(4)}$ is clear, as $\widetilde{A}_n^{(4)}$ is a sum of linear functions of t, and conditions on bounds are immediate consequences of the definition (13). The consistency of $\widetilde{A}_n^{(4)}$ follows from rewriting (13) as

$$\widetilde{A}_n^{(4)}(t) = \max\left[\Psi(t), \int_0^1 \max\{t\omega, (1-t)(1-\omega)\}\mathrm{d}S_n(\omega) + (2t-1)(1-\Lambda_n)\right],$$

where

$$S_n(\omega) = 1/k_n \sum_{i=1}^n \mathbb{1}\{W_i \le \omega\}\mathbb{1}\{R_i \ge n/k_n\}.$$

Equation (9) ensures that S_n is a consistent estimator of S, provided that k_n is such that $k_n \to \infty$ and $k_n/n \to 0$. Finally, note that $\Lambda_n = \int_0^1 \omega\,\mathrm{d}S_n(\omega)$ converges to $\int_0^1 \omega\,\mathrm{d}S(\omega)$ which is equal to 1, thanks to condition (8). So the second summand of (13) tends in probability to 0, whereas the first one converges in probability to $A^\star(t)$ because of (7).

Appendix 3

Asymptotic equivalence of conditions (a) and (b)

If H (resp. H_n) is the (resp. empirical) survival function of the R_is, one must show that the number of R_is greater than n/k_n is of the same order in probability as k_n. Equivalently, one must prove that

$$n/k_n H_n(n/k_n) \xrightarrow{p} S(\mathcal{N}).$$

Note that $tH(t) \to S(\mathcal{N})$ when $t \to \infty$ because of (6). This ensures the quadratic convergence of $n/k_n(H_n - H)(n/k_n)$ to 0 whenever $n/h_n \to \infty$ and $h_n \to \infty$, whence (15).

Acknowledgments

The authors thank Christian Genest, Jonathan Tawn and a referee for helpful comments, as well as Julie Jacques for her participation in the Monte Carlo experiment. Research funds in partial support of this work were granted by the Natural Sciences and Engineering Research Council of Canada and by the Fonds pour la formation de chercheurs et l'aide à la recherche du Gouvernement du Québec.

References

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

Capéraà, P., Fougères, A.-L., and Genest, C., ''Bivariate distributions with given extreme value attractor,'' *J. Multivariate Anal.* 72, 30–49, (2000).

Clayton, D.G., ''A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence,'' *Biometrika* 65, 141–151, (1978).

Coles, S.G. and Tawn, J.A., ''Modeling multivariate extreme events,'' *J. R. Statist. Soc.* B 53, 377–392, (1991).

Davison, A.C. and Smith, R.L., ''Models for exceedances over high thresholds,'' *J. R. Statist. Soc.* B 52, 393–442, (1990).

Deheuvels, P., ''On the limiting behavior of the Pickands estimator for bivariate extreme value distributions,'' *Statist. Probab. Lett.* 12, 429–439, (1991).

Einmahl, J., de Haan, L., and Huang, X., ''Estimating a multidimensional extreme value distribution,'' *J. Multivariate Anal.* 47, 35–47, (1993).

Einmahl, J., de Haan, L., and Sinha, A.K., ''Estimating the spectral measure of an extreme value distribution,'' *Stoch. Proc. Appl.* 70, 143–171, (1997).

Einmahl, J., de Haan, L., and Piterbarg, V., ''Nonparametric estimation of the spectral measure of an extreme value distribution.'' Technical Report 98-01, 1998.

Frank, M.J., ''On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$,'' *Æquationes Math.* 19, 194–226, (1979).

Galambos, J., *The Asymptotic Theory of Extreme Order Statistics*, 2nd edition. Kreiger, Malabar, FL, 1987.

de Haan, L., ''Extremes in higher dimensions: The model and some statistics.'' In: *Proc. 45th Session Inter. Statist. Inst.*, Amsterdam, paper 26.3. International Statistical Institute, Amsterdam, 1985.

de Haan, L. and Resnick, S.I., ''Limit theory for multidimensional sample extremes,'' *Z. Wahr. verw. Geb.* 40, 317–337, (1977).

de Haan, L. and Resnick, S.I., ''Estimating the limit distribution of multivariate extremes,'' *Comm. Statist. Stoch. Models* 9, 275–309, (1993).

de Haan, L. and de Ronde, J., ''Sea and wind: multivariate extremes at work,'' *Extremes* 1, 7–45, (1998).

de Haan, L. and Sinha, A.K., ''Estimating the probability of a rare event,'' *Ann. Statist.* 27, 732–759, (1999).

Huang, X., *Statistics of bivariate extreme values*, Thesis, Erasmus University Rotterdam, Tinbergen Institute Research Series 22, 1992.

Jacques, J., *Méthodes d'estimation de la fonction de dépendance des copules de valeurs extrêmes*. Unpublished M.Sc. thesis number 16333, Université Laval, Québec, Canada, 1998.

Joe, H., Smith, R.L., and Weissman, I., ''Bivariate threshold models for extremes,'' *J. R. Statist. Soc.* B 54, 171–183, (1992).

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

Pickands, J., ''Statistical inference using extreme order statistics,'' *Ann. Statist.* 3, 119–131, (1975).

Pickands, J., ''Multivariate extreme value distributions,'' *Bull. Int. Statist. Inst.* 859–878, (1981).

Resnick, S.I., ''Point processes, regular variation and weak convergence,'' *Adv. Appl. Probab.* 18, 66–138, (1986).

Resnick, S.I. *Extreme Values, Regular Variation, and Point Processes*, Springer, New York, 1987.

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

Smith, R.L., Tawn, J.A., and Yuen, H.K., ''Statistics of multivariate extremes,'' *Inter. Statist. Rev.* 58, 47–58, (1990).

Tawn, J.A., ''Bivariate extreme value theory: Models and estimation,'' *Biometrika* 75, 397–415, (1988).

Tiago de Oliveira, J., ''Intrinsic estimation of the dependence structure for bivariate extremes,'' *Statist. Probab. Lett.* 8, 213–218, (1989).