Tomáš Bacigál; Vladimír Jágr; Radko Mesiar

Non-exchangeable random variables, Archimax copulas and their fitting to real data

Kybernetika, Vol. 47 (2011), No. 4, 519--531

Persistent URL: `http://dml.cz/dmlcz/141656`

KYBERNETIKA — VOLUME 47 (2011), NUMBER 4, PAGES 519–531

NON-EXCHANGEABLE RANDOM VARIABLES, ARCHIMAX COPULAS AND THEIR FITTING TO REAL DATA

Tomáš Bacigál, Vladimír Jágr and Radko Mesiar

The aim of this paper is to open a new way of modelling non-exchangeable random variables with a class of Archimax copulas. We investigate a connection between powers of generators and dependence functions, and propose some construction methods for dependence functions. Application to different hydrological data is given.

Keywords: Archimax copula, dependence function, generator

Classification: 93E12, 62A10

1. INTRODUCTION

In recent years copulas turned out to be a promising tool in multivariate modelling, mostly with applications in actuarial sciences and hydrology.

In short, copula is a function which allows modelling dependence structure between stochastic variables. The main advantage is that the copula approach can split the problem of constructing multivariate probability distributions into a part containing the marginal one-dimensional distribution functions and a part containing the dependence structure. These two parts can be studied and estimated separately and then rejoined to form a joint distribution function.

Restricting ourselves to bivariate case, copula is a function $C\colon [0,1]^2 \to [0,1]$ which satisfies the boundary conditions, $C(t,0) = C(0,t) = 0$ and $C(t,1) = C(1,t) = t$ for $t \in [0,1]$ (uniform margins), and the 2-increasing property, $C(u_2,v_2)-C(u_2,v_1)-C(u_1,v_2)+C(u_1,v_1) \geq 0$ for all $u_1 \leq u_2, v_1 \leq v_2 \in [0,1]$. Copula is symmetric if $C(u,v) = C(v,u)$ for all $(u,v) \in [0,1]^2$ and is asymmetric otherwise. By $[a,b]$ we mean a closed interval with endpoints a and b, while $]a,b[$ will denote an open interval.

There are several approaches how to model exchangeable random variables. Most of them refer to Archimedean copulas [15], i.e., copulas $C_\varphi\colon [0,1]^2 \to [0,1]$ expressible in the form

$$C_\varphi(u,v) = \varphi^{(-1)}\left(\varphi(u) + \varphi(v)\right), \tag{1}$$

where $\varphi\colon [0,1] \to [0,\infty]$ is a continuous strictly decreasing convex function satisfying

$\varphi(1) = 0$, and its pseudo-inverse $\varphi^{(-1)}\colon [0,\infty] \to [0,1]$ is given by

$$\varphi^{(-1)}(x) = \varphi^{-1}\left(\min(\varphi(0), x)\right). \tag{2}$$

Among several approaches allowing to fit copulas to real data we recall [5] and references therein.

The aim of this paper is to open a new way of modelling non-exchangeable random variables which are related to asymmetric copulas. In the next section we recall Archimax copulas, a special class of copulas which are non-symmetric, in general. After some new theoretical results about the structure of Archimax copulas, in Section 3 we propose new construction methods for one part of this class of copulas. Section 4 gives a short overview of estimation methods used in the application to modelling hydrological data in Section 5.

2. ARCHIMAX COPULAS

Since there are much more cases in the nature when we feel the causality among stochastic processes flows in certain direction rather than the cases when we observe random variables equally affected by common underlying process, we find symmetry of the most used copulas quite restrictive. Among few classes of asymmetric copulas, convenient enough to model non-exchangeable random variables, we focus on the class of Archimax copulas [2] built up from a convex continuous decreasing function $\varphi\colon [0,1] \to [0,\infty]$, $\varphi(1) = 0$, called generator and a convex function $A\colon [0,1] \to [0,1]$, $\max(t, 1-t) \leq A(t) \leq 1$ for all $t \in [0,1]$, called dependence function. Then the corresponding Archimax copula is given by

$$C_{\varphi,A}(u,v) = \varphi^{(-1)}\left[(\varphi(u)+\varphi(v))\, A\left(\frac{\varphi(u)}{\varphi(u)+\varphi(v)}\right)\right] \quad \text{for all } u,v \in [0,1] \tag{3}$$

(with conventions $0/0 = \infty/\infty = 0$, where $\varphi^{(-1)}$ is given by (2)). Observe that Archimax copulas contains as special subclasses all Archimedean copulas (then $A \equiv A^* = 1$) and all extreme value copulas [16], in short EV copulas (then $\varphi(t) = -\log(t)$). For the weakest dependence function $A = A_*$,

$$A_*(t) = \max(t, 1-t),$$

we have $C_{\varphi,A_*} = M$, the strongest copula of co-monotone dependence, independently of the generator φ.

Moreover, it is easy to check that an Archimax copula $C_{\varphi,A}$ is symmetric if and only if $A(t) = A(1-t)$ for all $t \in [0,1]$ (i. e., A is symmetric wrt. axis $x = 1/2$). Recent results on measuring asymmetry can be found in [4].

Suppose that φ is a generator of a copula C_φ. Then also φ^λ, $\lambda > 1$, is a generator of a copula C_{φ^λ}. As an example recall the Gumbel family of copulas $\left(C^G_{(\lambda)}\right)_{\lambda\in[1,\infty[}$, generated by generators $\varphi^G_{(\lambda)} : [0,1] \to [0,\infty]$, $\varphi^G_{(\lambda)}(x) = (-\log x)^\lambda$, which bears from the product copula Π generated by $\varphi^G_{(1)}$.

Proposition 2.1. Let $\varphi : [0,1] \to [0,\infty]$ be a generator of a copula C_φ. For any dependence function A, and any $\lambda \geq 1$, the Archimax copula $C_{\varphi^\lambda,A}$ is also an Archimax copula based on generator φ, i. e., $C_{\varphi^\lambda,A} = C_{\varphi,B_{(A,\lambda)}}$, where $B_{(A,\lambda)}$: $[0,1] \to [0,1]$ is a dependence function given by

$$B_{(A,\lambda)} = A_{(\lambda)}(t)\left[A\left(\left(\frac{t}{A_{(\lambda)}(t)}\right)^\lambda\right)\right]^{1/\lambda}, \tag{4}$$

with $A_{(\lambda)}:\ [0,1] \to [0,1]$, $A_{(\lambda)}(t) = \left(t^\lambda + (1-t)^\lambda\right)^{1/\lambda}$.

Proof. Formula (4) is a matter of processing of the equality $C_{\varphi^\lambda,A} = C_{\varphi,B_{(A,\lambda)}}$. Indeed,

$$\begin{aligned} C_{\varphi^\lambda,A}(u,v) &= \varphi^{(-1)}\left(\left[(\varphi^\lambda(u)+\varphi^\lambda(v))A\left(\frac{\phi^\lambda(u)}{\phi^\lambda(u)+\phi^\lambda(v)}\right)\right]^{1/\lambda}\right) \\ &= \varphi^{(-1)}\left((\varphi(u)+\varphi(v))\left[\frac{\phi^\lambda(u)+\phi^\lambda(v)}{(\phi(u)+\phi(v))^\lambda}A\left(\frac{\phi^\lambda(u)}{\phi^\lambda(u)+\phi^\lambda(v)}\right)\right]^{1/\lambda}\right), \end{aligned}$$

while

$$\begin{aligned} &C_{\varphi^\lambda,B(A,\lambda)}(u,v) \\ &\quad = \varphi^{(-1)}\left((\varphi(u)+\varphi(v))\left[\left(\frac{\phi(u)}{\phi(u)+\phi(v)}\right)^\lambda + \left(\frac{\phi(v)}{\phi(u)+\phi(v)}\right)^\lambda\right]^{1/\lambda}\right. \\ &\qquad\qquad \left. A\left(\frac{\left(\frac{\phi(u)}{\phi(u)+\phi(v)}\right)^\lambda}{\left(\frac{\phi(u)}{\phi(u)+\phi(v)}\right)^\lambda + \left(\frac{\phi(v)}{\phi(u)+\phi(v)}\right)^\lambda}\right)\right) \\ &\quad = \varphi^{(-1)}\left((\varphi(u)+\varphi(v))\left[\frac{\phi^\lambda(u)+\phi^\lambda(v)}{(\phi(u)+\phi(v))^\lambda}A\left(\frac{\phi^\lambda(u)}{\phi^\lambda(u)+\phi^\lambda(v)}\right)\right]^{1/\lambda}\right). \end{aligned}$$

To see that $B_{(A,\lambda)}$ is indeed a dependence function, note that based on Gumbel family, we have also $C_{\varphi^G_{(\lambda)},A} = C_{\varphi^G_{(1)},B_{(A,\lambda)}}$. Due to [2], $C_{\varphi^G_{(\lambda)},A}$ is a copula. Moreover, for any power $p \in]0,\infty[$,

$$\begin{aligned} &C_{\varphi^G_{(\lambda)},A}(u^p,v^p) \\ &= \exp\left(-\left[\left((-\log u^p)^\lambda + (-\log v^p)^\lambda\right)A\left(\frac{(-\log u^p)^\lambda}{(-\log u^p)^\lambda + (-\log v^p)^\lambda}\right)\right]^{1/\lambda}\right) \\ &= \exp\left(-p\left[\left((-\log u)^\lambda + (-\log v)^\lambda\right)A\left(\frac{(-\log u)^\lambda}{(-\log u)^\lambda + (-\log v)^\lambda}\right)\right]^{1/\lambda}\right) \\ &= \left(C_{\varphi^G_{(\lambda)},A}(u,v)\right)^p, \end{aligned}$$

i. e., $C_{\varphi^G_{(\lambda)},A}$ is an EV-copula [15, 16]. Due to representation of EV-copulas, there is a dependence function $B: [0,1] \to [0,1]$ such that

$$C_{\varphi^G_{(\lambda)},A} = C_{\varphi^G_{(1)},B}$$

and evidently $B = B_{A,\lambda}$. □

Dependence functions $A_{(\lambda)}$, $\lambda \in [0,1]$, are called Gumbel dependence functions due to the fact that $C^G_{(\lambda)} = C_{\varphi^G_{(1)},A_{(\lambda)}}$. Observe that the Archimedean copula C_{φ^λ} is just an Archimax copula based on φ and $A_{(\lambda)}$, $C_{\varphi^\lambda} = C_{\varphi,A_{(\lambda)}}$, independently of the generator φ. Proposition 2.1 has an important impact for the structure of Archimax copulas. For any generator $\varphi: [0,1] \to [0,\infty]$, classes $\mathcal{A}_{\varphi^\lambda}$ of Archimax copulas based on generators φ^λ, $\lambda \in [1,\infty[$, are nested, and $\mathcal{A}_{\varphi^\lambda} \subsetneqq \mathcal{A}_{\varphi^\mu}$ whenever $1 \le \mu < \lambda \le \infty$, where $\mathcal{A}_{\varphi^\infty} = \bigcap_{\lambda=1}^{\infty} \mathcal{A}_{\varphi^\lambda} = \{M\}$. Therefore it is important to know the basic form η of each generator φ, $\varphi = \eta^\lambda$ with $\lambda \ge 1$, where $\eta: [0,1] \to [0,\infty]$ is a generator such that for any $\lambda \in]0,1[$, η^λ is no more convex. Such generators η will be called basic generators and they correspond to Archimedean copulas C_η such that for any $p > 1$, the corresponding L_p-norm $\|C_\eta\|_p > 1$ (for more details we recommend [3, 14]).

Proposition 2.2. Let $\varphi: [0,1] \to [0,\infty]$ be a generator. Let

$$\alpha = \inf\Big\{\frac{\varphi(x)\varphi''(x)}{(\varphi'(x))^2} \,\Big|\, x \in]0,1[\text{and } \varphi'(x), \varphi''(x) \text{ exist}\Big\}.$$

Then

$$\eta = \varphi^{1/p}, \quad \text{where } p = \frac{1}{1-\alpha},$$

is a basic generator.

Proof. The convexity of a generator $\varphi(\eta)$ is equivalent to the non-negativity of the derivatives $\varphi''(\eta'')$ in all points where they exist. Let $\varphi = \eta^p$, $p \ge 1$, where η is a basic generator. Then $\eta = \varphi^{1/p}$, $\eta' = \frac{1}{p}\varphi^{1/p-1}\varphi'$ and

$$\begin{aligned}\eta'' &= \frac{1}{p}\left(\frac{1}{p}-1\right)\varphi^{\frac{1}{p}-2}(\varphi')^2 + \frac{1}{p}\varphi^{\frac{1}{p}-1}\varphi'' \\ &= \frac{1}{p}\varphi^{\frac{1}{p}-2}\left(\left(\frac{1}{p}-1\right)(\varphi')^2 + \varphi'\varphi''\right) \ge 0\end{aligned}$$

if and only if $\alpha \le \varphi\varphi''/(\varphi')^2$, where $\alpha = 1 - 1/p$, where the last inequality should be satisfied in each point from $]0,1[$ where φ' and φ'' exist. The result follows. □

Based on Propositions 2.1 and 2.2, we propose to fit Archimax copulas based on basic generators η only. Thus before choosing the appropriate candidates for fitting of a generator, one should check their basic forms. The next lemma gives a sufficient condition for a generator η to be basic.

Lemma 2.3. Let $\eta:\ [0,1]\to[0,\infty]$ be a generator and let $\eta'(1^-)\neq 0$. Then η is a basic generator.

P r o o f. Due to continuity of η and $\eta(1)=0$, if $\eta'(1^-)\neq 0$ then $\alpha=\inf\left\{\frac{\eta(x)\eta''(x)}{(\eta'(x))^2}\mid x\in]0,1[\text{ and } \eta'(x),\eta''(x) \text{ exist}\right\}=0$ and thus $p=1$. □

Example 2.4.

(i) For each Gumbel generator $\varphi^G_{(\lambda)}$, the corresponding basic generator is $\eta=\varphi^G_{(1)}$ (the generator of the product copula).

(ii) The weakest copula $C^{(p)}$ which has minimal L_p-norm, $\|C^{(p)}\|_p=1$, $p\in[1,\infty[$, is an Archimedean copula generated by a generator $\varphi^Y_{(p)}:\ [0,1]\to[0,\infty]$, $\varphi^Y_{(p)}(x)=(1-x)^p$ (Y stands for Yager family, see [18], more details on L_p-norms and copulas can be found in [3]). Again, for any $p\in[1,\infty[$, the corresponding basic generator $\eta=\varphi^Y_{(1)}$ is unique (generator of the lower Frechet-Hoeffding bound W).

(iii) Based on Lemma 2.3 one can quickly check that the families of Clayton, Frank, Ali–Mikhail–Haq (see [10, 15]), are generated by basic generators only.

(iv) Taking generators from some two-parameter families given in [10], one may easily verify that

— BB1 generator $\varphi(t)=(t^{-a}-1)^b$ with $a>0,b\geq 1$ gains its basic form only for $b=1$, while BB3 with $\varphi(t)=e^{b(-\log t)^a}-1$ and $a\geq 1,b>0$ only for $a=1$. Then, both would result in strict Clayton copula;

— BB2 generator $\varphi(t)=e^{b(t^{-a}-1)}-1$ with $a>0,b>0$ is basic for any a,b;

— BB6 generator $\varphi(t)=\left[-\log(1-(1-t)^a)\right]^b$ with $a\geq 1,b>0$ reduces to basic form if $b=1/a$;

— each BB7 generator $\varphi(t)=\left[\left(1-(1-t)^a\right)^{-b}-1\right]^{1/a}$ with $a\geq 1,b>0$ is a basic generator.

3. SOME CONSTRUCTION METHODS FOR DEPENDENCE FUNCTIONS

Based on some known dependence functions, it is desirable to be able to construct new dependence functions to increase the fitting potential of our Archimax copulas buffer. Recall that for dependence functions $A_1,\dots,A_n$ also their convex sum $A=\sum_{i=1}^n\lambda_iA_i$ (with $\sum_{i=1}^n\lambda_i=1$) is a dependence function. Inspired by the bivariate construction [11] and based on the recent results [13], consider dependence functions $A_1,\dots,A_n$. Then the corresponding EV copulas $C_{A_1},\dots,C_{A_n}:\ [0,1]^2\to[0,1]$ are given by

$$C_{A_i}(u,v)=\exp\left((\log u+\log v)A_i\left(\frac{\log u}{\log u+\log v}\right)\right). \tag{5}$$

Take arbitrary two probability vectors $(a_1,\dots,a_n),(b_1,\dots,b_n)\in[0,1]^n$, $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 1$. Then due to [13] the function $C\colon [0,1]^2\to[0,1]$ given by

$$C(u,v)=\prod_{i=1}^n C_i\left(u^{a_i},v^{b_i}\right) \tag{6}$$

is also a copula. Note that EV copulas are characterized by the power stability $C\left(u^\lambda,v^\lambda\right)=(C(u,v))^\lambda$ for any $\lambda\in]0,\infty[$, $u,v\in[0,1]$. It is then easy to see that C given by (6) is also an EV copula, and thus there is a dependence function A so that $C=C_A$. For processing purposes, denote $t=\frac{\log u}{\log u+\log v}$. Then $\log v=\frac{1-t}{t}\log u$ and $\log u+\log v=\frac{\log u}{t}$. Moreover,

$$C(u,v)=\exp\left((\log u+\log v)A\left(\frac{\log u}{\log u+\log v}\right)\right)=\exp\left(\frac{\log u}{t}A(t)\right). \tag{7}$$

On the other hand, due to (6),

$$\begin{aligned} C(u,v) &= \prod_{i=1}^n \exp\left(\left(a_i\log u+b_i\frac{1-t}{t}\log u\right)A_i\left(\frac{a_i\log u}{a_i\log u+b_i\frac{1-t}{t}\log u}\right)\right)\\ &= \exp\left(\frac{\log u}{t}\sum_{i=1}^n\left(ta_i+(1-t)b_i\right)A_i\left(\frac{ta_i}{ta_i+(1-t)b_i}\right)\right). \end{aligned} \tag{8}$$

Comparing (7) and (8), we see that

$$A(t)=\sum_{i=1}^n\left(ta_i+(1-t)b_i\right)A_i\left(\frac{ta_i}{ta_i+(1-t)b_i}\right). \tag{9}$$

What was just shown is the following construction method.

Proposition 3.1. Let $A_1,\dots,A_n$ be dependence functions. Then for any probability vectors $(a_1,\dots,a_n),(b_1,\dots,b_n)\in[0,1]^n$, also the function $A\colon[0,1]\to[0,1]$ given by (9) is a dependence function.

Observe that the formula (9) can be deduced by induction from the original formula given in [11], as well as seen as extension of Proposition 3 of [9] dealing with A_1,A_2 and $\alpha,\beta\in[0,1]$. Then the function $A\colon[0,1]^2\to[0,1]$ given by

$$\begin{aligned} A(t) &= (\alpha t+\beta(1-t))\,A_1\left(\frac{\alpha t}{\alpha t+\beta(1-t)}\right)\\ &+ ((1-\alpha)t+(1-\beta)(1-t))\,A_2\left(\frac{(1-\alpha)t}{(1-\alpha)t+(1-\beta)(1-t)}\right) \end{aligned}$$

is a dependence function. Moreover, if $(a_1,\dots,a_n)=(b_1,\dots,b_n)$, then the formula (9) turns into the standard convex sum $A(t)=\sum_{i=1}^n a_iA_i(t)$. Evidently, this method allows to introduce asymmetric Archimax copulas even if starting from symmetric Archimax copulas.

Example 3.2. Consider dependence functions A_1, A_2. Let $A_2 = A_{(2)}$, see (4), $a_1 = a_2 = 1/2$, $b_1 = 0, b_2 = 1$. Then the dependence function A given by (9) does not depend on A_1, and it holds

$$A(t) = \frac{t}{2} + \frac{(2-t)}{2} A_2\left(\frac{t}{2-t}\right) = \frac{t}{2} + \sqrt{\left(\frac{t}{2}\right)^2 + (1-t)^2}.$$

Observe that $A(1/3) = (1+\sqrt{17})/6 = 0.85385$ and $A(2/3) = (1+\sqrt{2})/3 = 0.80474$, proving the asymmetry of any relevant Archimax copula $C_{\phi,A}$ (recall that $C_{\phi,A}$ is symmetric if and only if $A(t) = A(1-t)$ for all $t \in [0,1]$).

Inspired by [1] where construction methods for generators of Archimedean copulas were discussed, we propose one more new construction method for dependence function. For a dependence function A, denote by B a $[0,1] \to [0,1]$ function given by $B(t) = A(t) - 1 + t$. Each such B is characterized by its convexity, non-decreasingness and boundary conditions

$$\max(0, 2t-1) \le B(t) \le t.$$

The pseudo-inverse $B^{(-1)}\colon [0,1] \to [0,1]$ of B is given by

$$B^{(-1)}(u) = \sup\{t \in [0,1] \,|\, B(t) \le u\},$$

and it is characterized by concavity, non-decreasingness and boundary conditions

$$u \le B^{(-1)}(u) \le \frac{u+1}{2}. \tag{10}$$

Consider dependence functions $A_1, \ldots, A_n$ and related functions $B_1^{(-1)}, \ldots, B_n^{(-1)}$. Then the convex combination $\sum_{i=1}^n \lambda_i B_i^{(-1)}$ is concave, non-decreasing and satisfy the boundary conditions (10), and thus there is a dependence function A such that its related function $B^{(-1)}$ is just equal to $\sum_{i=1}^n \lambda_i B_i^{(-1)}$. This fact proves our next construction method.

Proposition 3.3. Let $A_1, \ldots, A_n$ be dependence functions and let $(\lambda_1, \ldots, \lambda_n) \in [0,1]^n$ be a probability vector. Then the function $A\colon [0,1] \to [0,1]$ given by

$$A(t) = \left(\sum_{i=1}^n \lambda_i B_i^{(-1)}\right)^{(-1)}(t) + 1 - t \tag{11}$$

is a dependence function.

Example 3.4. Consider the extremal dependence functions $A_1 = A_*$ and $A_2 = A^* = 1$. Then $B_1(t) = \max(0, 2t-1) = B_*(t)$ and $B_2(t) = t = B^*(t)$. Moreover, $B_1^{(-1)}(u) = \frac{u+1}{2}$ and $B_2^{(-1)}(u) = u$. For a fixed $\lambda \in [0,1]$,
$\big(\lambda B_1^{(-1)} + (1-\lambda) B_2^{(-1)}\big)(u) = \big(1 - \frac{\lambda}{2}\big)(u) + \frac{\lambda}{2} = B_\lambda^{(-1)}(u)$, and thus $B_\lambda(t) = \big(B_\lambda^{(-1)}\big)^{(-1)}(t) = \max\big(0, \frac{2t-\lambda}{2-\lambda}\big)$ and $A_\lambda(t) = B_\lambda(t) + 1 - t = \max\big(1-t, \frac{2-2\lambda+\lambda t}{2-\lambda}\big)$. Note that both A_1 and A_2 are symmetric (wrt. axis $x = 1/2$), but not A_λ for any $\lambda \in \left]0,1\right[$.

4. ESTIMATION METHOD

Basically there are two main methods recently used for estimating one-parameter families. One uses various measures of dependence, such as Kendall's tau through formal relation with copula parameter θ, the another is based on maximization of a likelihood function [8]. For general multi-parameter copulas (not e. g. multivariate normal or pair-copulas) the first method is problematic and to our best knowledge no satisfactory study has been presented so far. Given a sample of n-dimensional random vectors $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m$, here we use pseudo-loglikelihood function

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{m} \log\big(c_{\boldsymbol{\theta}}\big(F_1(X_{1,i}), \ldots, F_n(X_{n,i})\big)\big)$$

employing copula density $c_{\boldsymbol{\theta}}$ (which is n-order mixed derivative with respect to all variables) with vector parameter $\boldsymbol{\theta}$ and marginal empirical distribution functions

$$F_j(x) = \frac{1}{m+1} \sum_{i=1}^{m} 1(X_{i,j} \le x), \quad j = 1, \ldots, n,$$

where $1(\cdot)$ is the indicator function which yields 1 whenever $\cdot$ is true and 0 otherwise. The marginal empirical distribution functions transform $\boldsymbol{X}_i$ into pseudo-observations $\boldsymbol{U}_i$, $i = 1, \ldots, m$. Goodness of fit can be checked by comparing (L_2-norm) squared distances

$$S_n = \sum_{i=1}^{m} \left(C_n(U_{i,1}, \ldots, U_{i,n}) - C_{\boldsymbol{\theta}}(U_{i,1}, \ldots, U_{i,n})\right)^2$$

between estimated parametric copulas $C_{\boldsymbol{\theta}}$ and empirical copula function

$$C_n(u_1, \ldots, u_n) = \frac{1}{m} \sum_{i=1}^{m} 1(U_{i,1} \le u_1, \ldots, U_{i,n} \le u_n), \quad (u_1, \ldots, u_n) \in [0,1]^n .$$

However because of computational intensity of simulation related to bootstrap method, we do not perform GOF test [7] here (unless one would be interested in classifying copulas according to their competence to describe particular data). Instead, for comparison purposes we employ model selection criterion, in particular the Bayesian information criterion defined

$$BIC = -2L(\boldsymbol{\theta}) + k\log(m)$$

where k denotes number of parameters. The model preference grows with decreasing BIC.

5. APPLICATION

To examine performance of new models we consider two kinds of bivariate ($n = 2$) hydrological data. One is constituted by monthly average flow rate of two rivers

– Danube at Nagymaros (Hungary) $\{X_{i,1}\}$, $i = 1, \dots m$, and Inn measured at Schärding (Austria) $\{X_{i,2}\}$ (Inn is tributary to Danube, Nagymaros lies about 570 km downstream) comprising $m = 660$ realisations recorded for 55 years until 1991, see [17]. Another sequence of $m = 113$ entries comes from annual summer term maxima of the Vltava river (Bohemia) flow rate $\{X_{i,1}\}$ (measured above the dam Kamyk until 2007) with corresponding flood volume $\{X_{i,2}\}$, which is total amount of water run within 8 days starting three days before the corresponding flow rate peak.

Due to temporal manner of the monthly river discharge, the data were found not being i.i.d. After filtering the lowest frequencies, seasonal component and applying AR(1) model, the residuals were tested by Ljung-Box test and test of serial independence (based on copulas and proposed by [6]) with positive result.

Both bivariate data transformed to unit square are shown in Figure 1.

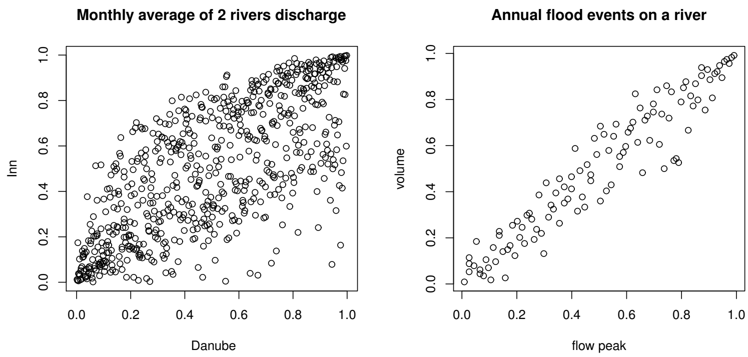

Fig. 1. Scatter plot of data after transformation by its marginal empirical distribution function.

Tables 2 and 3 summarize competition of new construction methods alongside well-established models (for overview see Table 1, [10, 15]) and related construction methods [1]. Besides parameters and maximized value of log-likelihood function we provide the corresponding estimation time and BIC criterion. Parameters were found by box-constrained optimisation (method L-BFGS-B implemented in R) which, if failed to find global maximum, was helped by pre-search over a grid. This happens mostly with more-parameter piece-wise construction of dependence function such as LPL. Parameters other than bounded by unit interval were rounded to one decimal place. Values in parentheses are fixed during estimation, square brackets indicate construction method of dependence function, in particular [bi] denotes biconvex combination given by Proposition 3.1 for $n = 2$, [li] represents special case of [bi] when $a_i = b_i$ $(i = 1, 2)$, and [inv] refers to Proposition 3.3. So far we implemented construction procedures for two dependence functions only and their individual parameters are estimated separately (in advance) from weighting parameters of their combination.

family	generator $\varphi_{\boldsymbol{\theta}}(t)$	parameter range	limiting case (Archimed.)
Gumbel	$(-\log(t))^{\theta_1}$	$[1,\infty]$	$\{1\}$ Π, $\{\infty\}$ M
Clayton	$t^{-\theta_1}-1$	$]0,\infty]$	$\{0\}$ Π, $\{\infty\}$ M
Frank	$-\log\left(\frac{e^{-\theta_1 t}-1}{e^{-\theta_1}-1}\right)$	$\Re$	$\{-\infty\}$ W, $\{0\}$ Π, $\{\infty\}$ M
Joe	$-\log\left(1-(1-t)^{\theta_1}\right)$	$[1,\infty]$	$\{1\}$ Π, $\{\infty\}$ M
BB1	$(t^{-\theta_1}-1)^{\theta_2}$	$]0,\infty]\times[1,\infty]$	$\{0,1\}$ Π, $\{\infty,\infty\}$ M
	dependence function $A_{\boldsymbol{\theta}}(t)$		limiting case (EV)
Mixed	$\theta_1 t^2-\theta_1 t+1$	$[0,1]$	$\{0\}$ Π
Gumbel (logistic)	$\left(t^{\theta_1}+(1-t)^{\theta_1}\right)^{1/\theta_1}$	$[1,\infty]$	$\{1\}$ Π, $\{\infty\}$ M
Hüsler Reiss	$t*\Phi\left(\frac{1}{\theta_1}+\frac{\theta_1}{2\log(t/(1-t))}\right)+$ $+(1-t)\Phi\left(\frac{1}{\theta_1}-\frac{\theta_1}{2\log(t/(1-t))}\right)$ Φ is CDF of standard normal	$[0,\infty]$	$\{1\}$ Π, $\{\infty\}$ M
Tawn (asymmetric logistic)	$1-\theta_2+(\theta_2-\theta_1)t+$ $+\left((\theta_1 t)^{\theta_3}+(\theta_2(1-t))^{\theta_3}\right)^{\frac{1}{\theta_3}}$	$[0,1]\times$ $[0,1]\times$ $[0,\infty]$	$\{0,0,1\}$ Π, $\{1,1,\infty\}$ M
LPL (linear-parabolic-linear)	$\begin{cases}1-\frac{1-b}{a}t & t\le a-c\\ \frac{b-a}{1-a}+\frac{1-b}{1-a}t & t\ge a+c\\ At^2+Bt+C & \text{otherwise}\end{cases}$ $A=\frac{(1-b)}{4(1-a)ac}$ $B=\frac{2(1-b)(2ac-a-c)}{4(1-a)ac}$ $C=\frac{2(1+b)ac+(1-b)c^2-(b+4c-1)a^2}{4(1-a)ac}$ $a=\theta_1$, $c=\theta_3\min(a,1-a)$ $b=\max(a,1-a)(1-\theta_2)+\theta_2$	$[0,1]\times$ $[0,1]\times$ $[0,1]$	$\{0,.,.\}$ $\{1,.,.\}$ $\{.,1,.\}$ Π $\{0.5,0,0\}$ M

Tab. 1. Overview of parametric families used to construct Archimax copula.

All procedures are implemented in R and freely available[1].

6. CONCLUSION

As seen from our results, given the two different data sets, the newly proposed construction methods do not give significantly better fit according to the selection criterion (which penalizes inclusion of additional parameters), however in case of dependence functions with roughly equal fitting performance they elevate the maximized likelihood. Note that the best results for fixed number of parameters are given by Archimax construction with both generator and dependence function, from which we may judge that the majority of well-established models in Archimedean and EV class capture mutually different dependence structure, in other words, they complement one another. The few exceptions that follow from Proposition 2.1 are equivalences of Archimedean copula with Gumbel generator and EV copula with Gumbel dependence function, or equivalence of BB1 and Archimax copula with Clayton generator and Gumbel dependence function.

In our software actually the estimation of Archimedean part is generally faster which may evoke a demand for some alternative to Proposition 2.1 in reverse order.

[1] www.math.sk/wiki/bacigal

generator		dependence function		log-lik	time	criterion
family	par.	family	par.	$L(\boldsymbol{\theta})$	[sec]	BIC
Gumbel	2.1			278.1	3	-549.8
Clayton	1.2			162.3	2	-318.1
Frank	6.6			255.2	3	-504.0
Joe	2.6			249.2	3	-492.0
BB1	0.1 2.1			278.5	11	-544.0
		Mixed	1.00	254.2	5	-502.0
		Gumbel	2.1	278.1	16	-549.8
		HüslerReiss	1.9	272.0	78	-537.7
		LL	0.56 0.70 (0.05)	66.9	1905	-120.9
		LPL sym.	(0.50) 0.05 0.80	274.0	1540	-528.5
		LPL	0.50 0.05 0.80	274.0	3122	-528.5
		Tawn	0.92 1.00 2.3	281.9	328	-544.3
Gumbel	2.1	Mixed	0.00	278.1	42	-543.3
Gumbel	1.5	Gumbel	1.5	278.1	36	-543.3
Gumbel	2.1	HüslerReiss	0.2	278.1	410	-543.3
Clayton	0.3	Mixed	1.00	269.1	84	-525.3
Clayton	0.1	Gumbel	2.1	278.5	55	-544.0
Clayton	0.9	HüslerReiss	1.8	272.9	190	-532.9
Frank	2.3	Mixed	0.97	280.8	86	-548.5
Frank	1.4	Gumbel	1.9	280.2	85	-547.4
Frank	1.8	HüslerReiss	1.5	276.1	255	-539.2
Joe	1.4	Mixed	0.98	273.9	81	-534.8
Joe	1.0	Gumbel	2.1	272.0	70	-531.2
Joe	1.0	HüslerReiss	1.9	272.1	348	-531.2
BB1	0.1 2.1	Mixed	0.00	278.5	71	-537.5
BB1	0.1 1.4	Gumbel	1.4	278.5	78	-537.5
BB1	0.1 2.1	HüslerReiss	0.1	278.5	40	-537.5
Gum–Cla	0.99			279.0	95	-538.5
Gum–Fra	0.46			278.7	53	-537.9
Gum–Joe	1.00			278.1	14	-536.7
Cla–Fra	0.00			256.2	335	-492.9
Cla–Joe	0.01			263.8	188	-508.1
Fra–Joe	0.74			271.8	64	-524.1
BB1–Gum	1.00			278.5	43	-537.5
BB1–Cla	1.00			278.5	14	-537.5
BB1–Fra	1.00			278.8	440	-537.5
BB1–Joe	1.00			278.5	16	-538.1
		[li] Mix–Gum	0.00	278.1	29	-536.7
		[inv]	0.00	278.1	322	-536.7
		[bi]	0.05 0.00	280.8	364	535.6
		[li] Mix–Hüs	0.10	276.3	314	-533.1
		[inv]	0.00	272.6	2049	-525.7
		[bi]	0.05 0.00	281.1	687	536.2
		[li] Gum–Hüs	0.71	278.5	125	-537.5
		[inv]	1.00	278.1	624	-536.7
		[bi]	0.92 0.99	279.2	1059	-532.4

Tab. 2. Estimation summary for **2 rivers flow rate**. Families denoted by [bi] and [li] (special case with $a_i = b_i$) refers to new construction method from Proposition 3.1 while [inv] to Proposition 3.3.

ACKNOWLEDGEMENT

This work was partially supported by the Slovak Research and Development Agency under the contract No. APVV-LPP-0004-07, SK-AT-0020-08 and APVV-0496-10, and by Program na podporu mladých výskumníkov, STU v Bratislave, projekt IMOPAR.

(Received August 31, 2010)

generator		dependence function		log-lik	GOF
family	par.	family	par.	$L(\boldsymbol{\theta})$	S_n
Gumbel	4.8			128.4	-252.1
Clayton	4.5			94.0	-183.3
Frank	18.4			123.9	-243.1
Joe	6.2			112.1	-219.5
BB1	0.3 4.3			129.5	-249.5
		Mixed	1.00	66.5	-128.3
		Gumbel	4.8	128.4	-252.1
		HüslerReiss	5.0	128.7	-252.7
		LL	0.50 0.50 (0.05)	57.3	-105.2
		LPL sym.	(0.50) 0.00 0.50	112.4	-215.3
		LPL	0.50 0.00 0.50	112.4	-210.6
		Tawn	1.00 1.00 4.8	128.4	-242.6
Gumbel	2.8	Mixed	1.00	128.6	-247.8
Gumbel	2.2	Gumbel	2.2	128.4	-247.2
Gumbel	3.1	HüslerReiss	1.3	128.9	-248.3
Clayton	2.0	Mixed	1.00	108.5	-207.5
Clayton	0.3	Gumbel	4.3	129.5	-249.6
Clayton	0.3	HüslerReiss	4.5	129.9	-250.5
Frank	10.0	Mixed	1.00	128.5	-247.6
Frank	4.0	Gumbel	3.2	131.3	-253.1
Frank	4.0	HüslerReiss	3.2	131.9	-254.4
Joe	3.4	Mixed	1.00	117.7	-225.9
Joe	1.0	Gumbel	4.8	128.4	-247.4
Joe	1.0	HüslerReiss	5.0	128.7	-247.9
BB1	0.3 2.5	Mixed	1.00	129.8	-245.4
BB1	0.3 1.7	Gumbel	2.5	129.5	-244.8
BB1	0.2 2.2	HüslerReiss	1.7	130.0	-245.8
Gum–Cla	1.00			128.4	-242.6
Gum–Fra	1.00			128.4	-242.6
Gum–Joe	1.00			128.4	-242.6
Cla–Fra	0.00			106.0	-197.8
Cla–Joe	0.00			112.1	-210.0
Fra–Joe	0.00			112.1	-210.0
BB1–Gum	1.00			129.4	-244.6
		[li] Mix–Gum	0.00	128.4	-242.6
		[inv]	0.00	128.4	-242.6
		[bi]	0.00 0.00	128.4	-237.9
		[li] Mix–Hüs	0.00	128.7	-243.2
		[inv]	0.00	128.7	-243.2
		[bi]	0.00 0.00	128.7	-238.5
		[li] Gum–Hüs	0.22	128.7	-243.2
		[inv]	0.26	128.6	-243.4
		[bi]	0.86 0.92	129.7	-240.5

Tab. 3. Estimation summary for **summer flood data**.

REFERENCES

[1] T. Bacigál, M. Juráňová, and R. Mesiar: On some new constructions of Archimedean copulas and applications to fitting problems. Neural Network World *1* (2010), 10, 81–90.

[2] P. Capéraá, A.-L. Fougéres, and G. Genest: Bivariate distributions with given extreme value attractor. J. Multivariate Anal. *72* (2000), 1, 30–49.

[3] B. De Baets, H. De Meyer, and R. Mesiar: Lipschitz continuity of copulas w.r.t. L_p-norms Nonlinear Anal., Theory, Methods Appl. *72(9-10)* (2010), 3722–3731.

[4] F. Durante and R. Mesiar: L^∞-measure of non-exchangeability for bivariate extreme value and Archimax copulas. J. Math. Anal. Appl. *369* (2010), 2, 610–615.

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

[6] C. Genest and B. Rémillard: Tests of independence or randomness based on the empirical copula process. Test *13* (2004), 335–369.

[7] C. Genest, B. Rémillard, and D. Beaudoin: Goodness-of-fit tests for copulas: A review and a power study. Insurance Math. Econom. *44* (2009), 199–213.

[8] C. Genest, K. Ghoudi, and L.-P. Rivest: A semi-parametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika *82* (1995), 543–552.

[9] C. Genest, K. Ghoudi, and L.-P. Rivest: Discussion of the paper by Frees and Valdez (1998). North Amer. Act. J. *2* (1998), 143–149.

[10] H. Joe: Multivariate Models and Dependence Concepts. Chapman and Hall, London 1997.

[11] A. Khoudraji: Contributions à l'étude des copules et à la modélisation des valeurs extrémes bivariées. PhD. Thesis, Université Laval, Québec 1995.

[12] E. P. Klement, R. Mesiar, and E. Pap: Triangular Norms. Kluwer, Dordrecht 2000.

[13] E. Liebscher: Construction of Asymmetric multivariate copulas. J. Multivariate Anal. *99* (2008), 10, 2234–2250.

[14] A. Mesiarová: k-l_p-Lipschitz t-norms. Internat. J. Approx. Reasoning *46* (2007), 3, 596–604.

[15] R. B. Nelsen: An Introduction to Copulas. Second edition. Springer–Verlag, New York 2006.

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

[17] C. J. Vorosmarty, B. M. Fekete, B. A. and Tucker: Global River Discharge, 1807-1991, V. 1.1 (RivDIS). Data set. 1998. Available on-line [http://www.daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, TN, U.S.A. doi:10.3334/ORNLDAAC/199.

[18] R. R. Yager: On a general class of fuzzy connectives. Fuzzy Sets and Systems *4* (1980), 3, 235–242.

Tomáš Bacigál, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, Bratislava. Slovak Republic.
e-mail: bacigal@math.sk

Vladimír Jágr, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, Bratislava. Slovak Republic.
e-mail: jagr@math.sk

Radko Mesiar, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, Bratislava, Slovak Republic, and Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mail: mesiar@math.sk