Depend. Model. 2021; 9:374–384 **DE GRUYTER**

Research Article **Open Access**
Special Issue in memory of Abe Sklar

Barry C. Arnold and Matthew Arvanitis*

On a general class of gamma based copulas

https://doi.org/10.1515/demo-2021-0117
Received May 25, 2021; accepted September 29, 2021

Abstract: A large family of copulas with gamma components is examined, and interesting submodels are defined and analyzed. Parameter estimation is demonstrated for some of these submodels. A brief discussion of higher-dimensional versions is included.

Keywords: copula, gamma components, bivariate beta, tail dependence, likelihood-free estimation, approximate Bayesian computation

MSC: 62H05, 62E10

1 Introduction

Arnold and Ng [3] introduced a bivariate second kind beta, or beta(2), distribution involving 8 independent random variables with gamma distributions (subsequently such random variables will be referred to as gamma components). It was identified as the most general bivariate model whose marginals are ratios of sums of independent gamma variables. The model involves 8 independent components $U_1, U_2, \ldots, U_8$ with $U_j \sim \Gamma(\alpha_j, 1),\ \ j = 1, 2, \ldots, 8$. The two-dimensional random vector (X, Y) is then defined by

$$X = \frac{U_1 + U_5 + U_7}{U_3 + U_6 + U_8}, \qquad Y = \frac{U_2 + U_5 + U_8}{U_4 + U_6 + U_7}. \tag{1}$$

This defines an 8-parameter family of bivariate distributions with beta(2) marginal distributions. If (X, Y) is defined as in (1) then we write: $(X, Y) \sim BB(2)(\underline{\alpha})$, to be read as (X, Y) has a bivariate second kind beta distribution with parameter vector $\underline{\alpha}$. Each α_i can assume any positive real value, so that the parameter space of the model is $(0, \infty)^8$. A corresponding family of bivariate distributions with beta marginals of the first or usual kind is obtained from (1) by defining

$$(V_1, V_2) = (X/(1 + X), Y/(1 + Y)).$$

Why there are 8 U_j's, and where they are located in the model (1), may require some explanation. There are four locations where a particular U_j may be placed. (1) In the numerator of X. (2) In the denominator of X. (3) In the numerator of Y and (4) In the denominator of Y. The variables U_1, U_2, U_3 and U_4 appear only once, and each one of them appears in only one of the four possible locations. A variable U_j cannot appear in both the numerator and denominator of X, nor of Y, since otherwise the independence of numerators and denominators, required for beta(2) marginals, would be destroyed. U_5 appears in the numerator of both X and Y. U_6 appears in the denominator of both X and Y. U_7 appears in the numerator of X and in the denominator of Y, while U_8 appears in the denominator of X and in the numerator of Y No U_j can appear in 3 or in 4 of

Barry C. Arnold: Department of Statistics, University of California, Riverside, USA, E-mail: barnold@ucr.edu
***Corresponding Author: Matthew Arvanitis:** USDA Forest Products Laboratory, Madison, Wisconsin, USA, E-mail: matthew.arvanitis@usda.gov

the possible locations, since that would destroy the required independence of at least one numerator and its corresponding denominator. If an additional independent gamma variable is introduced in one or two permissible locations in (1) then it can be combined with one of the existing 8 U_j's and no enrichment of the model will result. Thus for example, if U_9 is added to both numerators, then $U_5 + U_9$ will continue to play the role of U_5 with an adjusted shape parameter $\alpha_5 + \alpha_9$.

We adopt the convention that a random variable with a $\Gamma(\alpha, 1)$ distribution with $\alpha = 0$ will be defined to be a random variable that is degenerate at 0. By setting some of the α_j's in the Arnold-Ng model (1) equal to zero, simplified submodels (some of which have been discussed in the literature) will be obtained. Note that after setting certain α_j's equal to zero, we must retain $\alpha_1 + \alpha_5 + \alpha_7 > 0$, $\alpha_3 + \alpha_6 + \alpha_8 > 0$, $\alpha_2 + \alpha_5 + \alpha_8 > 0$, and $\alpha_4 + \alpha_6 + \alpha_7 > 0$, in order to continue to have beta(2) marginal distributions.

While this general bivariate beta model, and particularly its submodels, have demonstrated flexibility and usefulness in practice, the focus of this paper is on a specific class of submodels of the beta(1) containing only copulas, that is, distributions with uniform marginals. This paper is organized as follows. In Section 2, we construct this specific class of copulas by limiting the parameter space and discuss how further limitations of the parameter space can produce several familiar copula models discussed in the literature. In Section 3, natural symmetries of the copulas are considered. Section 4 discusses higher-dimensional versions. We demonstrate some examples of parameter estimation in Section 5, and finish with concluding remarks in Section 6.

2 Copulas

Our gamma based copulas are obtained by setting the values of the parameters in the Arnold-Ng(8) bivariate beta distribution so that the marginals are $Uniform(0, 1)$.

Recall that the Arnold-Ng 8-parameter bivariate beta model is of the form

$$(V_1, V_2) = \left(\frac{U_1 + U_5 + U_7}{U_1 + U_5 + U_7 + U_3 + U_6 + U_8}, \frac{U_2 + U_5 + U_8}{U_2 + U_5 + U_8 + U_4 + U_6 + U_7} \right), \tag{2}$$

where $U_1, U_2, U_3, \ldots, U_8$ are independent random variables with $U_i \sim \Gamma(\alpha_i, 1)$, $i = 1, 2, \ldots, 8$.

In order to have uniform marginals the α_i's must satisfy:

$$\alpha_1 + \alpha_5 + \alpha_7 = 1, \tag{3}$$
$$\alpha_3 + \alpha_6 + \alpha_8 = 1, \tag{4}$$
$$\alpha_2 + \alpha_5 + \alpha_8 = 1, \tag{5}$$

and

$$\alpha_4 + \alpha_6 + \alpha_7 = 1, \tag{6}$$

where all the α_i's are non-negative (and necessarily ≤ 1).

We may rewrite these constraints in the following form:

$$\begin{aligned} \alpha_1 &= 1 - \alpha_5 - \alpha_7; \\ \alpha_2 &= 1 - \alpha_5 - \alpha_8; \\ \alpha_3 &= 1 - \alpha_6 - \alpha_8; \\ \alpha_4 &= 1 - \alpha_6 - \alpha_7; \end{aligned} \tag{7}$$

so that the parameter space, the set of permissible values for $(\alpha_5, \alpha_6, \alpha_7, \alpha_8)$, may be defined by:

$$\begin{aligned} \alpha_5 &\in [0, 1]; \\ \alpha_6 &\in [0, 1]; \\ \alpha_7 &\in [0, 1 - \max\{\alpha_5, \alpha_6\}]; \end{aligned} \tag{8}$$

$$\alpha_8 \in [0, 1 - \max\{\alpha_5, \alpha_6\}].$$

It is convenient to use notation BB($i_1, i_2, \ldots, i_k$) to denote the model obtained from (2) by setting certain α_i's equal to 0 and retaining only $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}$ in the model. So, for $\alpha_5 = \alpha_6 = 1$ and $\alpha_7 = \alpha_8 = 0$, this model reduces to the co-monotone copula (BB(5,6)), that is, $V_1 = V_2$. Likewise, if $\alpha_5 = \alpha_6 = 0$ and $\alpha_7 = \alpha_8 = 1$, this model reduces to the counter-monotone copula (BB(7,8)), that is, $V_1 = 1 - V_2$. Also, if $\alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = 0$ (BB(1,2,3,4)), then V_1 is independent of V_2. Therefore, this model encompasses a continuous space of distributions with the co-monotone and counter-monotone copulas, also known as the Fréchet–Hoeffding bounds, and the independent copula as boundary points. Throughout this paper, correlations for this model or its submodels are expressed in terms of Spearman's Rank Correlation and are obtained by simulation when they cannot be determined analytically.

Submodels of this family are of particular interest. For example, the simplest is given when $\alpha_5 = \alpha$, and $\alpha_6 = \alpha_7 = \alpha_8 = 0$, so that it takes the form:

$$(V_1, V_2) = \left(\frac{U_1 + U_5}{U_1 + U_5 + U_3}, \frac{U_2 + U_5}{U_2 + U_5 + U_4} \right), \tag{9}$$

This family has positive correlations ranging from 0 to 0.478 (discussed in more detail later in the section), and the correlation has a monotone relationship with the parameter α. At $\alpha = 0$, this model reduces to the independent case. A larger submodel, the Magnussen model, is obtained by setting only $\alpha_7 = \alpha_8 = 0$ (see [5]). It thus is of the form:

$$(V_1, V_2) = \left(\frac{U_1 + U_5}{U_1 + U_5 + U_3 + U_6}, \frac{U_2 + U_5}{U_2 + U_5 + U_4 + U_6} \right), \tag{10}$$

where

$$\alpha_1 + \alpha_5 = 1, \tag{11}$$

$$\alpha_3 + \alpha_6 = 1, \tag{12}$$

$$\alpha_2 + \alpha_5 = 1, \tag{13}$$

and

$$\alpha_4 + \alpha_6 = 1, \tag{14}$$

where all the α_i's are non-negative (and necessarily ≤ 1). This is a two parameter family of copulas parameterized by $\alpha_5, \alpha_6 \in (0, 1)$. The other four α_i's are determined by (11)-(14). This model exhibits a full range of positive correlations, with the independent copula and the co-monotone copula appearing at the two extremes of the parameter space.

The Arnold-Ng 5 parameter model (introduced in [3]) is obtained by setting $\alpha_3 = \alpha_4 = \alpha_5 = 0$, and is therefore denoted by BB(1,2,6,7,8). The corresponding family of copulas includes those of the form:

$$(V_1, V_2) = \left(\frac{U_1 + U_7}{U_1 + U_7 + U_6 + U_8}, \frac{U_2 + U_8}{U_2 + U_8 + U_6 + U_7} \right), \tag{15}$$

where

$$\alpha_1 + \alpha_7 = 1, \tag{16}$$

$$\alpha_6 + \alpha_8 = 1, \tag{17}$$

$$\alpha_2 + \alpha_8 = 1, \tag{18}$$

and

$$\alpha_6 + \alpha_7 = 1, \tag{19}$$

where all the α_i's are non-negative (and necessarily ≤ 1). This is a one parameter family of copulas parameterized by $\alpha_1 \in (0, 1)$. The other four α_i's are determined by (16)-(19).

Note that if $(V_1, V_2) \sim BB(1, 2, 6, 7, 8)$ then

- $(1 - V_1, 1 - V_2) \sim BB(3, 4, 5, 7, 8)$
- $(1 - V_1, V_2) \sim BB(2, 3, 5, 6, 7)$
- $(V_1, 1 - V_2) \sim BB(1, 4, 5, 6, 8)$

We will denote these models by

AN(5A)=BB(1,2,6,7,8),
AN(5B)=BB(3,4,5,7,8),
AN(5C)=BB(2,3,5,6,7), and
AN(5D)=BB(1,4,5,6,8).

For these to be copulas we must impose the following constraints:

For AN(5A):

$$\alpha_1 + \alpha_7 = 1, \tag{20}$$
$$\alpha_6 + \alpha_8 = 1, \tag{21}$$
$$\alpha_2 + \alpha_8 = 1, \tag{22}$$
$$\alpha_6 + \alpha_7 = 1. \tag{23}$$

For AN(5B)

$$\alpha_5 + \alpha_7 = 1, \tag{24}$$
$$\alpha_3 + \alpha_8 = 1, \tag{25}$$
$$\alpha_5 + \alpha_8 = 1, \tag{26}$$
$$\alpha_4 + \alpha_7 = 1. \tag{27}$$

For AN(5C)

$$\alpha_5 + \alpha_7 = 1, \tag{28}$$
$$\alpha_3 + \alpha_6 = 1, \tag{29}$$
$$\alpha_2 + \alpha_5 = 1, \tag{30}$$
$$\alpha_6 + \alpha_7 = 1. \tag{31}$$

For AN(5D)

$$\alpha_1 + \alpha_5 = 1, \tag{32}$$
$$\alpha_6 + \alpha_8 = 1, \tag{33}$$
$$\alpha_5 + \alpha_8 = 1, \tag{34}$$
$$\alpha_4 + \alpha_6 = 1. \tag{35}$$

So we have four one parameter families of copulas.

But there are 45 more one parameter submodels that can be constructed by setting three of the five parameters to zero and applying the appropriate constraints. Four of the $\binom{8}{5} = 56$ choices must be excluded as they would produce zeros in one of the numerators or denominators in (1), and four additional choices all represent the same family as the following example shows. Take the BB(4,5,6,7,8) model for which we need to impose the following constraints:

$$\alpha_5 + \alpha_7 = 1, \tag{36}$$
$$\alpha_6 + \alpha_8 = 1, \tag{37}$$
$$\alpha_5 + \alpha_8 = 1, \tag{38}$$
$$\alpha_4 + \alpha_6 + \alpha_7 = 1. \tag{39}$$

We can take $\alpha_8 = \alpha \in (0, 1)$. Then $\alpha_5 = 1 - \alpha,\ \alpha_6 = 1 - \alpha,\ \alpha_7 = \alpha$, and, necessarily, $\alpha_4 = 0$, so our model is equivalent to the BB(5,6,7,8) with constraints. But how can one pick among all these models, and what properties do they have?

Let us look at the Olkin-Liu BB [7] model:

$$(V_1, V_2) = \left(\frac{U_1}{U_1 + U_6}, \frac{U_2}{U_2 + U_6} \right), \tag{40}$$

For this to be a copula, we need $\alpha_1 = \alpha_2 = \alpha_6 = 1$. Thus we have a 0-dimensional family of copulas. Three related copulas are obtainable by reflection about $1/2$. Clearly, when $\alpha_5 = 1$ for the simple family (9), it reduces to a reflected version of (40), and when the Magnussen parameters satisfy $\alpha_5 = 1 - \alpha_6$, it also reduces to (a possibly reflected version of) this model. As a copula, the Olkin-Liu BB model is also known as the Ali-Mikhail-Haq copula [1], a distribution that tends to appear in a vast array of copula families, including many of the Archimedean type (It should be noted that while this paper refers to the "Ali-Mikhail-Haq" copula multiple times, the copula that is being referenced is only one specific member of an entire family of copulas which is not a subset of the present family.) [6].

Let us return to consider the Arnold-Ng(5A) model of the form:

$$(V_1, V_2) = \left(\frac{U_1 + U_7}{U_1 + U_7 + U_6 + U_8}, \frac{U_2 + U_8}{U_2 + U_8 + U_6 + U_7} \right), \tag{41}$$

where

$$\alpha_1 + \alpha_7 = 1, \tag{42}$$

$$\alpha_6 + \alpha_8 = 1, \tag{43}$$

$$\alpha_2 + \alpha_8 = 1, \text{ and} \tag{44}$$

$$\alpha_6 + \alpha_7 = 1, \tag{45}$$

where all the α_i's are non-negative (and necessarily ≤ 1). This is a one parameter family of copulas parameterized by $\alpha_1 \in (0, 1)$. The other four α_i's are determined by (42)-(45). So we have $\alpha_1 = \alpha_2 = \alpha_6 = \alpha \in (0, 1)$ and $\alpha_7 = \alpha_8 = 1 - \alpha$. A plot of the correlation of the AN(5A) model, as a function of α, is given in Figure 1.

One immediate observation about the AN(5A) model is that its range of correlations is not full, that is, it is a proper subset of $[-1, 1]$. While this may at first thought appear to be a disappointing range of correlations, it should be noted that this family exhibits a different feature that may be desirable for many applications: it is a continuous collection of distributions ranging from the counter-monotone copula (having a correlation of -1) to an Ali-Mikhail-Haq copula, which favors strong upper tail (positive) dependence only, which coincides with the 0.478 correlation. Contour plots of the densities for various values of α are shown in Figure 2. It is applicable to phenomena that do not exhibit significant lower tail dependence, but do exhibit a trade-off between upper tail dependence and an overall negative correlation. For example, if X_1 represents a citizen's government entitlements, and X_2 represents the same citizen's overall wealth, then applying the AN(5A) to $V_1 = F_{X_1}(X_1)$ and $V_2 = F_{X_2}(X_2)$, α could be a measure of the government's plutocratic tendencies, since we might observe α nearer to 0 for socially democratic governments, while α may be much closer to 1 for highly plutocratic governments.

Remark Among the 4 avatars of the copula based on the Arnold-Ng 5 parameter bivariate beta model, two, namely AN(5A) and AN(5B), contain the counter-monotone copula but not the co-monotone copula. In contrast AN(5C) and AN(5D) contain the co-monotone copula but not the counter-monotone copula.

3 Symmetry considerations

There are three symmetry conditions that are on occasion deemed to be desirable properties for modeling purposes. It is therefore of interest to identify which members of our general copula family exhibit such symmetry features. A copula will be said to be marginally symmetric if $V_1 \stackrel{d}{=} 1 - V_1$ and $V_2 \stackrel{d}{=} 1 - V_2$, where $\stackrel{d}{=}$

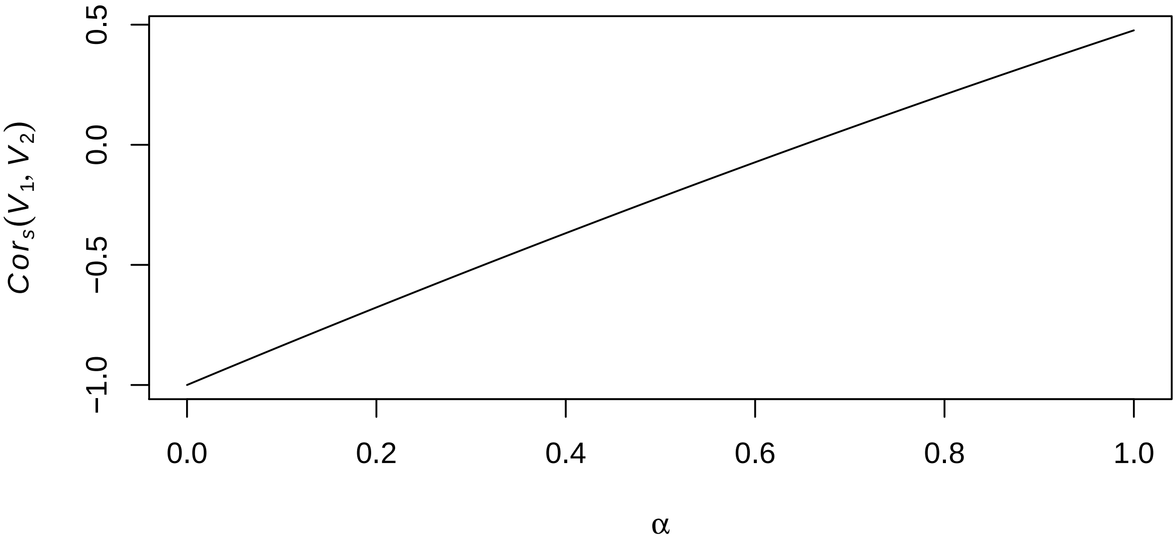

Figure 1: AN(5A) (Spearman) correlation as a function of α.

signifies equality of the distribution functions associated with the random variables. Clearly any BB copula with parameter vector $\underline{\alpha}$ satisfying the constraints (3)-(6) will exhibit marginal symmetry since V_1 and V_2 have $Uniform(0, 1)$ distributions.

Radial symmetry requires more restrictions. In order for a copula to be radially symmetric it must be the case that $(V_1, V_2) \stackrel{d}{=} (1 - V_1, 1 - V_2)$. For this to be true the parameter vector $\underline{\alpha}$ must be of the form

$$\underline{\alpha} = (\alpha, \alpha, \alpha, \alpha, \beta, \beta, 1 - \alpha - \beta, 1 - \alpha - \beta),$$

where $0 \leq \alpha + \beta \leq 1$. We thus have a two parameter subfamily of radially symmetric copulas.

For joint symmetry, even more is required, namely $(V_1, V_2) \stackrel{d}{=} (V_1, 1-V_2) \stackrel{d}{=} (1-V_1, V_2) \stackrel{d}{=} (1-V_1, 1-V_2)$. For this to occur we must have a copula with parameter vector of the form

$$\underline{\alpha} = (\alpha, \alpha, \alpha, \alpha, (1 - \alpha)/2, (1 - \alpha)/2, (1 - \alpha)/2, (1 - \alpha)/2).$$

where $0 \leq \alpha \leq 1$. We thus have a one parameter subfamily of jointly symmetric copulas.

There is another "symmetry" condition that can be considered. We will say that a copula is exchangeable if $(V_1, V_2) \stackrel{d}{=} (V_2, V_1)$. For this to be the case, the parameter vector of the copula must be of the form

$$(\alpha, \alpha, \beta, \beta, \gamma, \gamma, 1 - \alpha - \gamma, 1 - \beta - \gamma).$$

This is a three parameter family of copulas where the parameters satisfy the following constraints

$$0 \leq \gamma \leq 1, \quad 0 \leq \alpha \leq 1 - \gamma, \text{ and } \quad 0 \leq \beta \leq 1 - \gamma.$$

As examples, the AN(5C) and AN(5D) (among many others) are exchangeable.

4 On multivariate gamma based copulas

k-dimensional versions of the Arnold-Ng beta(2) distribution were mentioned in Arnold and Ghosh [2], in a context of copula models. First we consider the three dimensional case. It will then be evident how to deal with higher dimensions.

A three dimensional beta(2) distribution will be one whose structure is of a form which involves 26 independent gamma distributed U_j's. Thus, there are a total of 26 parameters in the model where U_j, j =

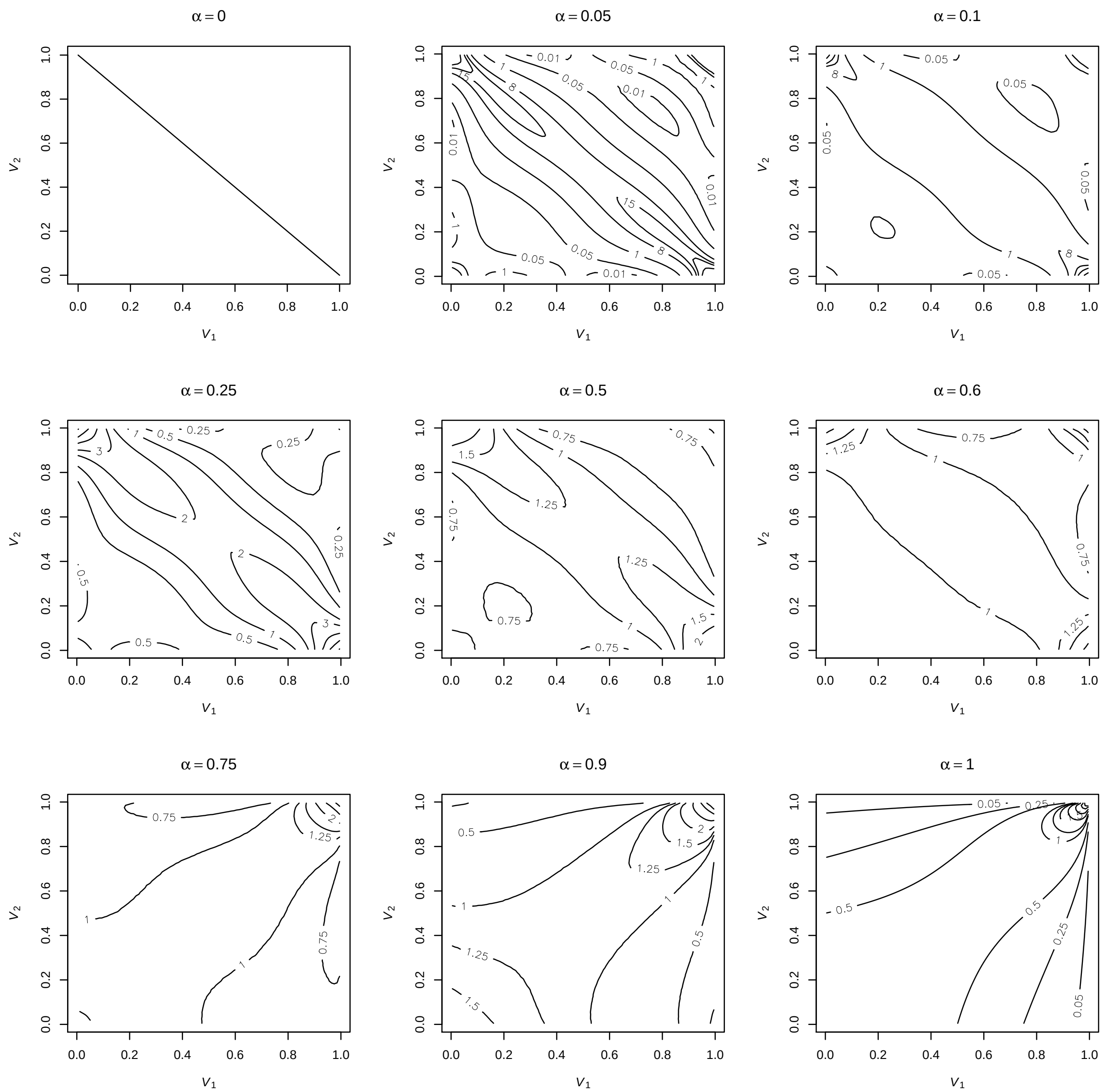

Figure 2: AN(5A) densities for various values of α.

$1, 2, \ldots, 26$ are independent variables with $U_j \sim \Gamma(\alpha_j, 1)$ for each j. The model can then be expressed in the following form.

$$X = \frac{U_1 + U_7 + U_8 + U_9 + U_{10} + U_{19} + U_{20} + U_{21} + U_{22}}{U_4 + U_{11} + U_{12} + U_{13} + U_{14} + U_{23} + U_{24} + U_{25} + U_{26}}, \tag{46}$$

$$Y = \frac{U_2 + U_7 + U_{11} + U_{15} + U_{16} + U_{19} + U_{20} + U_{23} + U_{24}}{U_5 + U_9 + U_{13} + U_{17} + U_{18} + U_{21} + U_{22} + U_{25} + U_{26}}, \tag{47}$$

$$Z = \frac{U_3 + U_8 + U_{12} + U_{15} + U_{17} + U_{19} + U_{21} + U_{23} + U_{25}}{U_6 + U_{10} + U_{14} + U_{16} + U_{18} + U_{20} + U_{22} + U_{24} + U_{26}}. \tag{48}$$

We must then impose 6 constraints to ensure that the associated trivariate beta distribution has uniform marginals.

The pattern for the dimensions of the parameter spaces of the multivariate copula models can now be recognized. The univariate model involves 2 U's with two constraints, i.e., a $3^1 - 1 - 2 = 0$ parameter model. The bivariate model involves 8 U's with 4 constraints, i.e., a $3^2 - 1 - 4 = 4$ parameter model. The trivariate case involves 26 U's with 6 constraints, i.e., a $3^3 - 1 - 6 = 20$ parameter model, and, in general, the k-dimensional model involves $3^k - 1$ U's with $2k$ constraints.

Use of fully parameterized k-dimensional copulas would almost never be recommended. Instead simplified sub-models, obtained by setting many of the α's equal to zero, can be expected to be adequate for many data sets.

5 Parameter estimation

In general, densities are not available in analytic form for the gamma-based copulas under discussion, except for special cases, such as the Ali-Mikhail-Haq case. So, if a sample is available from the bivariate gamma based copula then maximum likelihood is unavailable for parameter estimation. What we do have available are relatively simple expressions for mixed moments of the coordinate random variables. In principle, then, we could choose several sample mixed moments, equate them to their expectations and solve the resulting equations for the α_i parameters. This, in many cases, will prove to be a non-trivial exercise. However, it will typically be feasible for simplified sub-models involving only a few of the α_i's. With fewer equations to deal with, the method of moments approach is often not difficult to implement.

Example 1. *As a very simple example consider a model in which* $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$. *The model is thus of the form*

$$\begin{aligned} V_1 &= (U_5 + U_7)/(U_5 + U_7 + U_6 + U_8), \text{ and} \\ V_2 &= (U_5 + U_8)/(U_5 + U_8 + U_6 + U_7). \end{aligned} \tag{49}$$

In this one parameter family of copulas we can denote α_5 *by* α *and then we have* $\alpha_6 = \alpha$ *and* $\alpha_7 = \alpha_8 = 1 - \alpha$. *In this case, we can verify that*

$$E(V_1 V_2) = (\alpha + 1)/6.$$

If we define $M_{V_1 V_2} = (1/n)\sum_{i=1}^n V_{1i} V_{2i}$, *then a moment based estimate of* α *will be*

$$\widetilde{\alpha} = 6M_{V_{1i}V_{2i}} - 1.$$

Example 2. *Suppose that* (V_1, V_2) *is of the AN(5A) type. While it may be tempting to make use of the simple relationship between the Spearman correlation and the parameter* α *shown in Figure 1 in a method of moments approach, it is not the case that Spearman is the best option. Figure 3 depicts the three most well-known correlation measures for the AN(5A) model. In addition to the correlations, the figure shows 95% confidence ranges of values for the correlation measures, given observed samples of the shown sizes. In this light, Kendall's Tau appears to be the superior choice (at least among these three choices). But even using Kendall's Tau, it is clear that samples would need to be large in order to obtain reasonably accurate estimates of* α. *For example, at* $n = 50$, *a 95% confidence interval for* α *would have a width of about* 0.4, *while for* $n = 250$, *that width would reduce to about* 0.15, *depending on the actual value of* α.

Examples in which more of the α's are non-zero will generally require iterative solutions of the moment equations, but except for that, they can be expected to yield reasonable estimates of the parameters provided that sample sizes are adequate, and enough moments are included in the estimation procedure. More complex examples exist for which a simple method of moments approach may be inadequate.

Example 3. *Consider the two-parameter case where* $\alpha, \beta \in [0, 1]$, $\alpha_5 = \alpha\beta$, $\alpha_6 = \alpha$, *and* $\alpha_7 = (1 - \alpha)\beta$, *and* $\alpha_8 = (1-\alpha\beta)(1-\alpha)$, *and a sample,* $\mathbf{v}$, *of size n, is available. This is a two-parameter subfamily which contains both the co-monotone copula and the counter-monotone copula, and therefore can exhibit the full range of correlations. It also, unsurprisingly, includes two rotated versions of the Ali-Mikhail-Haq copula. Approximate Bayesian Computation, or ABC, can be applied, where the assumed prior for* (α, β) *can be any distribution deemed appropriate. Here, we will assume it to be a bivariate distribution with independent* $Uniform(0, 1)$ *marginals.*

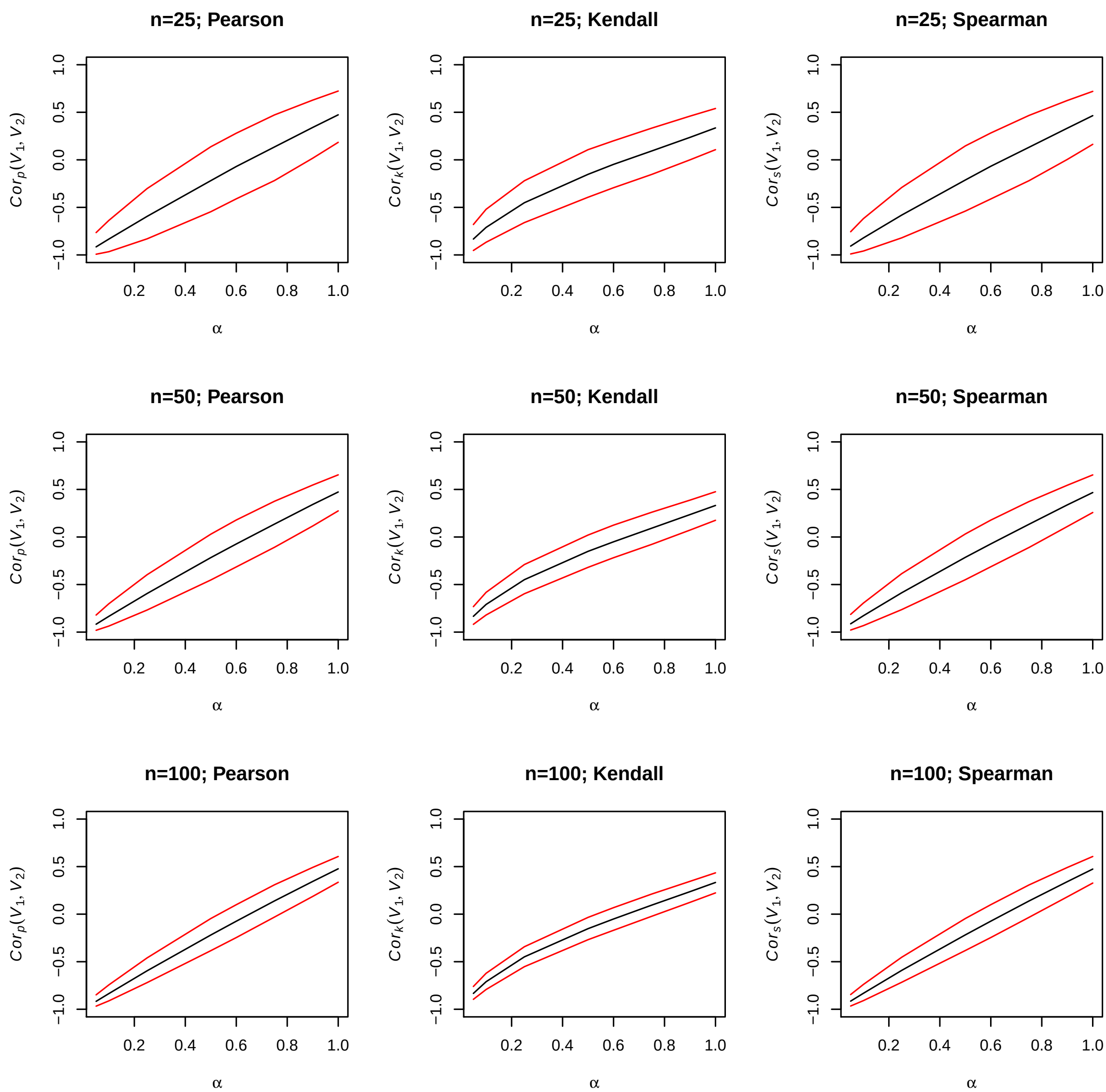

Figure 3: AN(5A) correlations, Pearson, Kendall's Tau, and Spearman, as a function of α, with 95% confidence bounds, obtained from simulation.

Now, given the behavior of this particular subfamily of distributions based on (α, β), we choose eight specific measures. The first four are simply proportions; partitioning the unit square into four equally-sized squares, we record the total number of observations in each, and divide by n. For the other four, we compute the four rotated Ali-Mikhail-Haq log-liklihood functions on the data contained in the same four regions (normalized).

$$S_1(\boldsymbol{v}) = \frac{n_{11}}{n}, \text{ where } n_{11} = \sum_{k=1}^{n} I\left\{v_{1k} \le \frac{1}{2} \& v_{2k} \le \frac{1}{2}\right\},$$

$$S_2(\boldsymbol{v}) = \frac{n_{12}}{n}, \text{ where } n_{12} = \sum_{k=1}^{n} I\left\{v_{1k} \le \frac{1}{2} \& v_{2k} > \frac{1}{2}\right\},$$

$$S_3(\boldsymbol{v}) = \frac{n_{21}}{n}, \text{ where } n_{21} = \sum_{k=1}^{n} I\left\{v_{1k} > \frac{1}{2} \& v_{2k} \le \frac{1}{2}\right\},$$

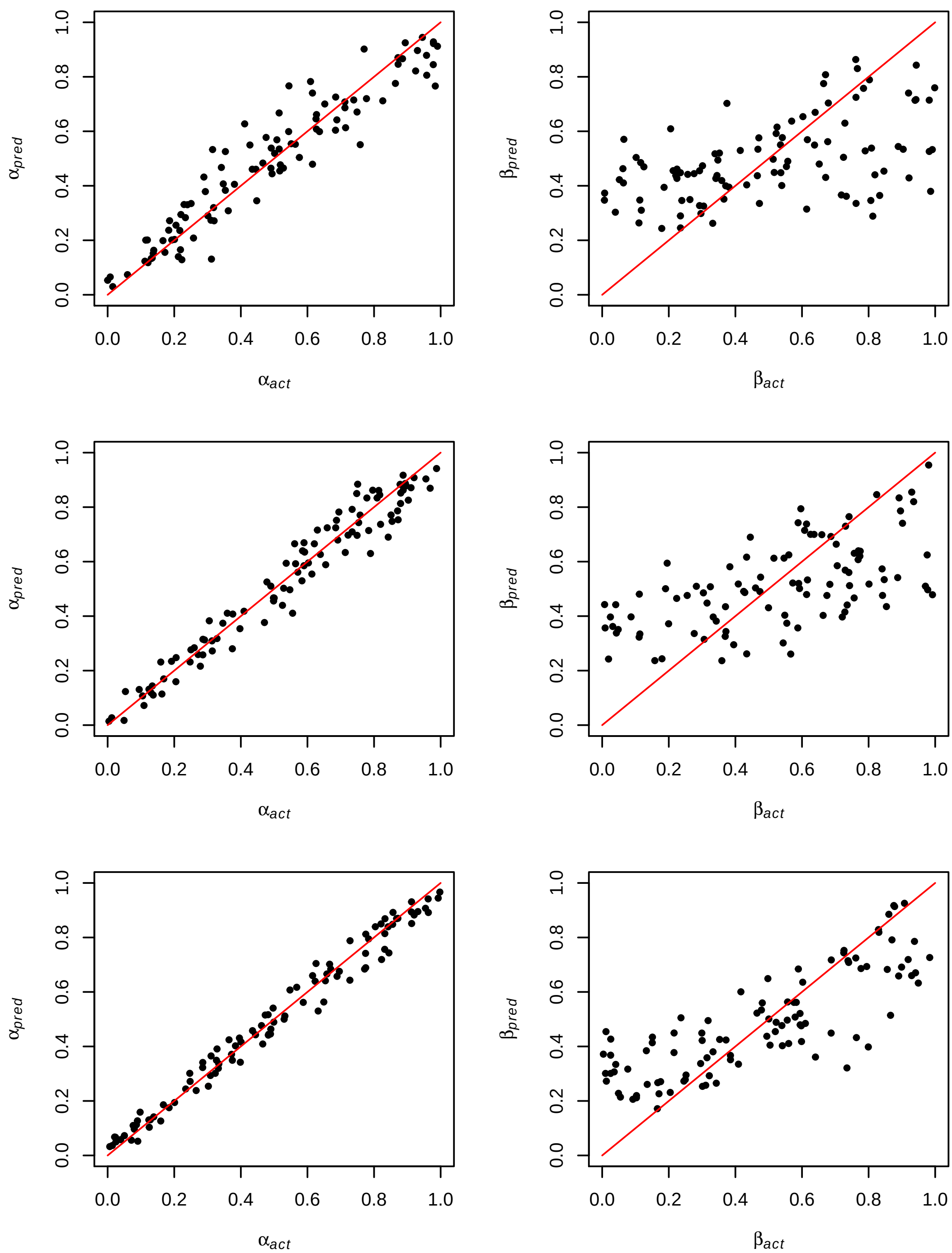

Figure 4: Comparisons of actual values of α and β with the predicted values by the ABC procedure. The three rows of plots represent sample sizes of 100, 250, and 1000, respectively.

$$S_4(\boldsymbol{v}) = \frac{n_{22}}{n}, \textit{ where } n_{22} = \sum_{k=1}^{n} I\left\{v_{1k} > \frac{1}{2} \& v_{2k} > \frac{1}{2}\right\}, \tag{50}$$

$$S_5(\boldsymbol{v}) = \frac{1}{n_{11}} \sum_{k: v_{1k} \le \frac{1}{2} \& v_{2k} \le \frac{1}{2}} \log\left[\frac{(1+v_{1k})(1+v_{2k}) - 2 + (1-v_{1k})(1-v_{2k})}{(1-(1-v_{1k})(1-v_{2k}))^3}\right],$$

$$S_6(\boldsymbol{v}) = \frac{1}{n_{12}} \sum_{k: v_{1k} \le \frac{1}{2} \& v_{2k} > \frac{1}{2}} \log \left[\frac{(1 + v_{1k})(2 - v_{2k}) - 2 + (1 - v_{1k})v_{2k}}{(1 - (1 - v_{1k})v_{2k})^3} \right],$$

$$S_7(\boldsymbol{v}) = \frac{1}{n_{21}} \sum_{k: v_{1k} > \frac{1}{2} \& v_{2k} \le \frac{1}{2}} \log \left[\frac{(2 - v_{1k})(1 + v_{2k}) - 2 + v_{1k}(1 - v_{2k})}{(1 - v_{1k}(1 - v_{2k}))^3} \right],$$

$$S_8(\boldsymbol{v}) = \frac{1}{n_{22}} \sum_{k: v_{1k} > \frac{1}{2} \& v_{2k} > \frac{1}{2}} \log \left[\frac{(2 - v_{1k})(2 - v_{2k}) - 2 + v_{1k}v_{2k}}{(1 - v_{1k}v_{2k})^3} \right],$$

We tested the ABC procedure for this family by arbitrarily selecting 100 values of (α, β) *from a* $Uniform\left((0,1) \times (0,1)\right)$ *distribution, simulating a sample for each, and applying the ABC procedure, repeating for three different sample sizes:* 100, 250, *and* 1000. *Figure 4 depicts the results. It can be observed that there is much more difficulty in estimating* β *than* α. *This is a common issue, and oftentimes with various subfamilies such as this, reasonable estimation may be assured with only very large sample sizes [4].*

6 Concluding Remarks

The present paper has provided an introduction to a broad spectrum of gamma based copula models and sub-models which can potentially be useful additions to the modeler's tool kit. These new flexible copula models can be expected to find application in cases in which the simpler well-known copula models prove to be inadequate to adapt to particular data sets, or where "big" data is encountered. It will be unlikely that use of the full 4 parameter model will be frequently deemed appropriate.

Acknowledgements: We are grateful for the helpful suggestions and dialogue provided by the referees which have resulted in an improved presentation.

Conflict of interest statement: Authors state no conflict of interest.

References

[1] Ali, M. M., N. Mikhail, and M. S. Haq (1978). A class of bivariate distributions including the bivariate logistic. *J. Multivariate Anal. 8*(3), 405–412.

[2] Arnold, B. C. and I. Ghosh (2014). Bivariate Kumaraswamy models involving use of Arnold-Ng copulas. *J. Appl. Stat. Sci. 22*(3/4), 227–240.

[3] Arnold, B. C. and H. K. T. Ng (2011). Flexible bivariate beta distributions. *J. Multivariate Anal. 102*(8), 1194–1202.

[4] Arvanitis, M. (2018). *Likelihood-free Estimation for Some Flexible Families of Distributions*. Ph.D. thesis, UC Riverside.

[5] Magnussen, S. (2004). An algorithm for generating positively correlated beta-distributed random variables with known marginal distributions and a specified correlation. *Comput. Statist. Data Anal. 46*(2), 397–406.

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

[7] Olkin, I. and R. Liu (2003). A bivariate beta distribution. *Statist. Probab. Lett. 62*(4), 407–412.