



## Mixture of D-vine copulas for modeling dependence

Daeyoung Kim<sup>a</sup>, Jong-Min Kim<sup>b,\*</sup>, Shu-Min Liao<sup>c</sup>, Yoon-Sung Jung<sup>d</sup>

<sup>a</sup> Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003, USA

<sup>b</sup> Statistic Discipline, Division of Science and Mathematics, University of Minnesota-Morris, Morris, MN 56267, USA

<sup>c</sup> Department of Mathematics, Amherst College, Amherst, MA 01003, USA

<sup>d</sup> Cooperative Agricultural Research Center, College of Agriculture and Human Sciences, Prairie View A&M University, Prairie View, TX 77446, USA

### ARTICLE INFO

#### Article history:

Received 24 June 2012

Received in revised form 13 February 2013

Accepted 14 February 2013

Available online 5 March 2013

#### Keywords:

Dependence

Multivariate data

Pair-copula

Vines

### ABSTRACT

The identification of an appropriate multivariate copula for capturing the dependence structure in multivariate data is not straightforward. The reason is because standard multivariate copulas (such as the multivariate Gaussian, Student-t, and exchangeable Archimedean copulas) lack flexibility to model dependence and have other limitations, such as parameter restrictions. To overcome these problems, *vine copulas* have been developed and applied to many applications. In order to reveal and fully understand the complex and hidden dependence patterns in multivariate data, a *mixture of D-vine copulas* is proposed incorporating D-vine copulas into a finite mixture model. As a D-vine copula has multiple parameters capturing the dependence through iterative construction of pair-copulas, the proposed model can facilitate a comprehensive study of complex and hidden dependence patterns in multivariate data. The proposed mixture of D-vine copulas is applied to simulated and real data to illustrate its performance and benefits.

© 2013 Elsevier B.V. All rights reserved.

### 1. Introduction

Interest in copulas has been growing as statistical tools for understanding the dependence between several random variables, and copula-based modeling has been extensively applied to many areas including actuarial sciences (Carrière, 2006), finance (Cherubini et al., 2004; McNeil et al., 2005), neuroscience (Onken et al., 2009), and weather research (Schölzel and Friederichs, 2008).

Copulas have several attractive properties. First, due to Sklar's theorem (Sklar, 1959), copulas allow us to model separately the marginal distributions and the joint dependence structure. Second, they are invariant under increasing and continuous transformations. Third, they do not require the marginals to be elliptically distributed, unlike correlation. Finally, copulas can be used to measure dependence in the tails of the joint distribution.

More importantly, the development of the vine copulas (Joe, 1996; Bedford and Cooke, 2001, 2002; Kurowicka and Cooke, 2006; Aas et al., 2009) has been accelerating the use of the copula-based modeling to depict the dependence for multivariate data. Vine copulas are a graphical model designed to overcome the limitations of existing standard multivariate copulas such as the multivariate Gaussian, Student-t, and exchangeable Archimedean copulas. It is hierarchical in nature in the sense that it can express a multivariate copula by using a cascade of bivariate copulas, the so-called pair-copulas. There are two popular classes of vine copulas that have been widely used by many researchers: the canonical (C-) vine copulas and the D-vine copulas (Kurowicka and Cooke, 2004).

Recently there has been research that combine copulas with a finite mixture model in various areas of application to not only fully understand the different dependence patterns between observed random variables, but also add more flexibility

\* Corresponding author. Tel.: +1 320 589 6341; fax: +1 320 589 6371.

E-mail addresses: [jongmink@morris.umn.edu](mailto:jongmink@morris.umn.edu), [kjonomi@hotmail.com](mailto:kjonomi@hotmail.com) (J.-M. Kim).

in terms of statistical modeling. By a finite mixture model, we mean a probabilistic model represented as a weighted sum of a few parametric densities. It is well known that the finite mixture model is very useful for uncovering hidden structures in the data (Lindsay, 1995; McLachlan and Peel, 2000). For instance, Vrac et al. (2005) developed a mixture model of Frank copulas to partition a global field of atmospheric profiles of temperature and humidity to characterize air masses while taking into account dependences between and among temperature and moisture through Frank copula functions. Cuvelier and Noirhomme-Fraiture (2005) proposed a mixture model of Clayton copulas for clustering in data mining, and showed the Clayton copula based-mixture model provided equivalent or better results in terms of clustering compared to a mixture of Normal copulas. Hu (2006) employed a mixture of three copulas, Gaussian, Gumbel and Gumbel-Survival, to model dependence of monthly returns between a pair of stock indexes and capture left and/or right tail dependence.

Although useful, the aforementioned approaches to model heterogeneous dependence structures are unfortunately limited to bivariate data, and thus it is difficult to apply them to multivariate data with diverse dependence structures. As will be shown in Section 4.2, the four-dimensional precipitation data measured from four municipalities (Vestby, Ski, Hurdal and Nannestad) in Akershus County of Norway has two different dependence structures: one describing the overall pattern of precipitation in the Akershus County while the other delineating dependence pattern of precipitation governed by each location of the four municipalities. Note that none of the existing methods are developed to reveal these two different types of dependence simultaneously.

In this paper we propose a *mixture of D-vine copulas* which includes multiple parameters for examining the different dependences inherent in multivariate data and can be extended to a multivariate copula function. As one of the most commonly used vine copulas, a D-vine copula utilizes  $d(d - 1)/2$  bivariate copula densities to express a  $d$ -dimensional multivariate density. By incorporating D-vine copulas into a finite mixture model, one can not only generate dependence structures that may not belong to existing copula families, but also conduct a comprehensive study of complex and hidden dependence patterns in multivariate data. Note that a finite mixture model can also be viewed as a semiparametric compromise between a fully parametric model (which often forces too much structure onto the data) and a nonparametric model (which leads to model estimates highly dependent on observed data). The reader will find from Section 4.2 that the proposed mixture of D-vine copulas successfully captures the diverse dependence structures hidden in multivariate precipitation data mentioned above.

This article is organized as follows: Section 2 briefly reviews the D-vine copula and its full inference. Section 3 proposes a finite mixture model of D-vine copulas, describes the model estimation using the EM algorithm (Dempster et al., 1977), and discusses the model selection process. Section 4 shows the performance of the proposed mixture model of D-vine copulas in simulated data and real data. A brief discussion is then followed in Section 5.

## 2. D-vine copula

This section reviews the pair-copula decomposition of a multivariate distribution, the D-vine copula and its inference presented in Aas et al. (2009) and Aas and Berg (2009).

### 2.1. The pair-copula construction

Let  $\mathbf{X} = (X_1, \dots, X_d)'$  be a  $d$ -dimensional random vector with the joint distribution function  $F = F(x_1, \dots, x_d)$  with marginals  $F_1 = F_1(x_1), \dots, F_d = F_d(x_d)$ . A copula is a function that allows us to represent the joint distribution in terms of its marginals and their dependence structure. Sklar's theorem (Sklar, 1959) states that the copula associated with  $F$  is a  $d$ -dimensional distribution function  $C: [0, 1]^d \rightarrow [0, 1]$  that satisfies

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d); \boldsymbol{\beta}), \quad (1)$$

where  $\boldsymbol{\beta}$  is the parameter vector measuring dependence between the marginals. For notational convenience, we will often suppress the parameter vector  $\boldsymbol{\beta}$ .

For an absolutely continuous  $F$  with strictly increasing, continuous marginal densities  $F_1, \dots, F_d$ , the joint density function  $f$  is obtained from Eq. (1):

$$f(x_1, \dots, x_d) = c_{1\dots d}(F_1(x_1), \dots, F_d(x_d)) \cdot \prod_{k=1}^d f_k(x_k), \quad (2)$$

where  $c_{1\dots d}(\cdot)$  is some uniquely identified  $d$ -variate copula density.

Under suitable regularity conditions, one can represent a  $d$ -dimensional multivariate density in Eq. (2) as a product of pair-copulas in an iterative manner, using the general form of a conditional marginal density in Eq. (3) (Aas et al., 2009):

$$f(x_k | \mathbf{v}) = c_{x_k, v_j | \mathbf{v}_{-j}}(F(x_k | \mathbf{v}_{-j}), F(v_j | \mathbf{v}_{-j})) \cdot f(x_k | \mathbf{v}_{-j}), \quad (3)$$

where  $\mathbf{v} = \mathbf{x}_{-k}$  is the  $(d - 1)$ -dimensional vector excluding  $x_k$ ,  $v_j$  is any one component arbitrarily chosen from  $\mathbf{v}$ , and  $\mathbf{v}_{-j}$  is the vector  $\mathbf{v}$  excluding  $v_j$ .

In the bivariate case ( $d = 2$ ), the unconditional bivariate density and conditional density are

$$f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2), \quad f(x_1 | x_2) = c_{12}(F_1(x_1), F_2(x_2))f_1(x_1),$$

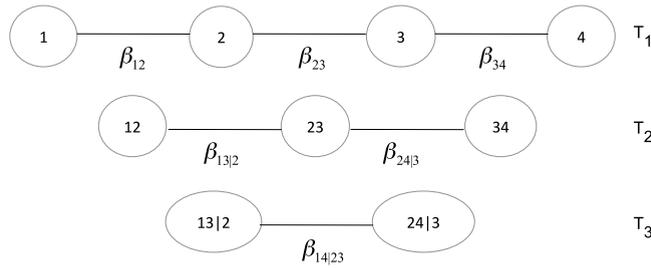


Fig. 1. A D-vine copula with four variables, 3 trees ( $T_1, T_2, T_3$ ), 9 nodes, and 6 edges.

where  $c_{12}(\cdot, \cdot)$  is the suitable pair-copula density for  $F_1(x_1)$  and  $F_2(x_2)$ . In the trivariate case ( $d = 3$ ) we have

$$\begin{aligned} f(x_1 | x_2, x_3) &= c_{13|2}(F(x_1 | x_2), F(x_3 | x_2))f(x_1 | x_2), \\ &= c_{13|2}(F(x_1 | x_2), F(x_3 | x_2))c_{12}(F_1(x_1), F_2(x_2))f_1(x_1), \end{aligned}$$

where  $c_{13|2}$  is the appropriate pair-copula using  $F(x_1 | x_2)$  and  $F(x_3 | x_2)$ . We can then express the trivariate density  $f(x_1, x_2, x_3)$  as a product of three pair-copulas,  $c_{12}$ ,  $c_{23}$ , and  $c_{13|2}$ :

$$\begin{aligned} f(x_1, x_2, x_3) &= f_3(x_3) \cdot f(x_2 | x_3) \cdot f(x_1 | x_2, x_3) \\ &= c_{12}(F_1(x_1), F_2(x_2))c_{23}(F_2(x_2), F_3(x_3))c_{13|2}(F(x_1 | x_2), F(x_3 | x_2)) \prod_{k=1}^3 f_k(x_k). \end{aligned}$$

In the four ( $d = 4$ ) dimensional case, the corresponding density can be expressed into two different forms according to different decompositions (Aas et al., 2009):

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= f_1(x_1) \cdot f(x_2 | x_1) \cdot f(x_3 | x_1, x_2) \cdot f(x_4 | x_1, x_3, x_2) \\ &= c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{13}(F_1(x_1), F_3(x_3)) \cdot c_{14}(F_1(x_1), F_4(x_4)) \\ &\quad \cdot c_{23|1}(F(x_2 | x_1), F(x_3 | x_1)) \cdot c_{24|1}(F(x_2 | x_1), F(x_4 | x_1)) \\ &\quad \cdot c_{34|12}(F(x_3 | x_1, x_2), F(x_4 | x_1, x_2)) \cdot \left[ \prod_{k=1}^4 f_k(x_k) \right] \\ &= f_2(x_2) \cdot f(x_3 | x_2) \cdot f(x_1 | x_3, x_2) \cdot f(x_4 | x_1, x_3, x_2) \\ &= c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \cdot c_{34}(F_3(x_3), F_4(x_4)) \\ &\quad \cdot c_{13|2}(F(x_1 | x_2), F(x_3 | x_2)) \cdot c_{24|3}(F(x_2 | x_3), F(x_4 | x_3)) \\ &\quad \cdot c_{14|23}(F(x_1 | x_2, x_3), F(x_4 | x_2, x_3)) \cdot \left[ \prod_{k=1}^4 f_k(x_k) \right]. \end{aligned}$$

## 2.2. D-vine copula

In the bivariate case, various copula families are available in practice, and are extensively studied in the literature (Joe, 1997; Nelsen, 2006). But the identification of an appropriate copula in higher dimensional problems is not straightforward, as standard multivariate copulas (such as the multivariate Gaussian, Student-t, and exchangeable Archimedean copulas) lack flexibility to model the dependence and have parameter restrictions. In order to overcome these problems, vine copulas are proposed by Joe (1996) and developed in more detail by Bedford and Cooke (2001, 2002), Kurowicka and Cooke (2003, 2004, 2006), Aas et al. (2009) and Aas and Berg (2009). Recent developments and applications of vine copula based models can be found in Bauer et al. (2012), Brechmann et al. (2012), Hobæk Haff et al. (2010), Joe et al. (2010), Panagiotelis et al. (2012) and Smith et al. (2010).

Vine copulas are graphical models that allow us to represent a  $d$ -dimensional multivariate density using  $d(d - 1)/2$  bivariate copula densities (pair-copulas) in a hierarchical manner: the first  $d - 1$  have dependence structures of unconditional bivariate distributions, and the remaining have dependence structures of conditional bivariate distributions. There are two popular types of vines that have been widely used by many researchers: the canonical (C-) vines and the D-vines (Kurowicka and Cooke, 2004). In the C-vines, one needs to specify in advance the relationships between one specific variable and the others, but in the D-vines, one is free to select which pairs for modeling the dependence. Since the D-vines are more flexible than the C-vines, in this paper we concentrate on the D-vine.

Fig. 1 shows the graphical specification for a 4-dimensional D-vine in the form of a nested set of tree structures. A D-vine with four variables has three trees  $T_j$ , and tree  $T_j$  has  $5 - j$  nodes and  $4 - j$  edges, where  $j = 1, 2, 3$ . Each edge is associated with a pair-copula density used for modeling dependence between two variables, and the edge label represents the dependence parameter in the associated pair-copula density. In the first tree,  $T_1$ , the dependences of the three pairs of variables, (1, 2),

(2, 3), (3, 4), are modeled using three pair-copulas,  $c_{12}(\cdot; \beta_{12})$ ,  $c_{23}(\cdot; \beta_{23})$ , and  $c_{34}(\cdot; \beta_{34})$ . In the second tree,  $T_2$ , two conditional dependences are modeled: (i) between the first and third variable given the second variable, the pair (1, 3 | 2), using associated pair-copula density  $c_{13|2}(\cdot; \beta_{13|2})$ , and (ii) between the second and fourth variable given the third variable, the pair (2, 4 | 3), using associated pair-copula density  $c_{24|3}(\cdot; \beta_{24|3})$ . Note that pairwise dependences of the two variables  $x_a$  and  $x_b$  where  $a, b = 1, \dots, d$  are modeled in subsequent trees conditioned on those variables between the variables  $x_a$  and  $x_b$  in  $T_1$ .

The  $d$ -dimensional D-vine density, given by Bedford and Cooke (2001), is

$$f(\mathbf{x}; \boldsymbol{\phi}) = \prod_{k=1}^d f_k(x_k) \times \prod_{i=1}^{d-1} \prod_{j=i}^{d-i} c_{j,j+i|(j+1):(j+i-1)}(F(x_j | x_{j+1}, \dots, x_{j+i-1}), F(x_{j+i} | x_{j+1}, \dots, x_{j+i-1}); \boldsymbol{\beta}_{j,j+i|(j+1):(j+i-1)}), \quad (4)$$

where  $f_k(x_k)$  are the marginal densities,  $c_{j,j+i|(j+1):(j+i-1)}$  are the bivariate copula densities with parameter(s)  $\boldsymbol{\beta}_{j,j+i|(j+1):(j+i-1)}$ , and  $\boldsymbol{\phi}$  is the set of all parameters in the D-vine density.

In the four-dimensional case ( $d = 4$ ), Eq. (4) then becomes

$$f(x_1, x_2, x_3, x_4; \boldsymbol{\phi}) = c_{12}(F(x_1), F(x_2); \beta_{12}) \cdot c_{23}(F(x_2), F(x_3); \beta_{23}) \cdot c_{34}(F(x_3), F(x_4); \beta_{34}) \\ \cdot c_{13|2}(F(x_1 | x_2), F(x_3 | x_2); \beta_{13|2}) \cdot c_{24|3}(F(x_2 | x_3), F(x_4 | x_3); \beta_{24|3}) \\ \cdot c_{14|23}(F(x_1 | x_2, x_3), F(x_4 | x_2, x_3); \beta_{14|23}) \cdot \prod_{k=1}^4 f_k(x_k),$$

where  $\boldsymbol{\phi} = (\beta_{12}, \beta_{23}, \beta_{34}, \beta_{13|2}, \beta_{24|3}, \beta_{14|23})$ .

### 2.3. Full inference for a D-vine copula

This section introduces the inference procedures of parameters in the D-vine density of Eq. (4). Assume that  $N$  observed samples, say  $\mathbf{x}_k = (x_{k,1}, \dots, x_{k,N})$ , where  $k = 1, \dots, d$ , are available. Assume further that each random variable  $X_k$  is assumed to be uniform in  $[0,1]$  (so that  $f_k(x_k) = 1$ ), which uses the normalized ranks of the original data for the parameter estimation. This is not a restrictive assumption because it is common that the distributions of the marginals are not known in practice, and the normalized ranks of the data, which are only approximately uniform and independent, keep the largest amount of information about the joint dependence between the variables (Oakes, 1982).

Since statistical inference for a D-vine copula is based on the normalized ranks of the data, the method of maximum pseudo-likelihood for parameter estimation (Oakes, 1994; Genest et al., 1995; Shih and Louis, 1995; Kim et al., 2007) is adopted to maximize the log-pseudo likelihood for the D-vine density over the parameters:

$$\ell(\boldsymbol{\phi}) = \sum_{i=1}^{d-1} \sum_{j=i}^{d-i} \sum_{n=1}^N \log c_{j,j+i|(j+1):(j+i-1)} \\ \times (F(x_{j,n} | x_{j+1,n}, \dots, x_{j+i-1,n}), F(x_{j+i,n} | x_{j+1,n}, \dots, x_{j+i-1,n}); \boldsymbol{\beta}_{j,j+i|(j+1):(j+i-1)}), \quad (5)$$

where  $\boldsymbol{\phi}$  is the set of all parameters in the D-vine density.

In the case of a four-dimensional variables ( $d = 4$ ), Eq. (5) becomes

$$\ell(\boldsymbol{\phi}) = \sum_{n=1}^N [\log c_{12}(x_{1,n}, x_{2,n}; \beta_{12}) + \log c_{23}(x_{2,n}, x_{3,n}; \beta_{23}) + \log c_{34}(x_{3,n}, x_{4,n}; \beta_{34}) \\ + \log c_{13|2}(F(x_{1,n} | x_{2,n}), F(x_{3,n} | x_{2,n}); \beta_{13|2}) + \log c_{24|3}(F(x_{2,n} | x_{3,n}), F(x_{4,n} | x_{3,n}); \beta_{24|3}) \\ + \log c_{14|23}(F(x_{1,n} | x_{2,n}, x_{3,n}), F(x_{4,n} | x_{2,n}, x_{3,n}); \beta_{14|23})],$$

where  $\boldsymbol{\phi} = (\beta_{12}, \beta_{23}, \beta_{34}, \beta_{13|2}, \beta_{24|3}, \beta_{14|23})$ .

Full inference of a D-vine copula in Eq. (4) consists of three steps (Aas et al., 2009):

- [Step 1] Choose which variables to include at the first tree  $T_1$  of a D-vine copula and decide the order of the variables using (tail) dependences (i.e., join the variables that have the strongest (tail) dependence).
- [Step 2] Specify the parametric shape of each pair-copula in an assumed D-vine.
- [Step 3] Estimate all parameters of an assumed D-vine by numerically maximizing the log-pseudo likelihood in Eq. (5).

In [Step 1], one can find the most appropriate ordering of the variables by comparing Kendall's tau values for all bivariate pairs (Aas et al., 2009; Aas and Berg, 2009). If there is information available on pair copula types for all pairs of a D-vine density, one can also choose the ordering of the variables by first fitting a specified bivariate copula to each pair, estimating dependence parameters, and then comparing them (Aas et al., 2009).

**Table 1**

Five copula functions, ranges of the dependent parameters and corresponding Kendall's tau. Note that  $x_1$  and  $x_2$  are uniform from  $[0,1]$ ,  $\Phi_G(x_1, x_2; \theta)$  is the standard bivariate normal distribution with correlation parameter  $\theta$ ,  $\Phi$  is the distribution function of the standard normal distribution, and  $D_1(\theta)$  represents the Debye function,  $D_1(\theta) = \frac{1}{\theta} \int_0^\theta \frac{t}{e^t-1} dt$ .

Copula	Function $C(x_1, x_2)$	Range of $\theta$	Kendall's tau
Gaussian	$\Phi_G[\Phi^{-1}(x_1), \Phi^{-1}(x_2); \theta]$	$-1 < \theta < 1$	$\frac{2}{\pi} \arcsin(\theta)$
FGM	$x_1 x_2 (1 + \theta(1 - x_1)(1 - x_2))$	$-1 \leq \theta \leq 1$	$\frac{2}{9} \theta$
Frank	$-\frac{1}{\theta} \log \left( 1 + \frac{(e^{-\theta x_1} - 1)(e^{-\theta x_2} - 1)}{e^{-\theta} - 1} \right)$	$[-\infty, \infty] \setminus \{0\}$	$1 - \frac{4}{\theta} [D_1(\theta) - 1]$
Clayton	$(x_1^{-\theta} + x_2^{-\theta} - 1)^{-1/\theta}$	$\theta > 0$	$\frac{\theta}{\theta+2}$
Gumbel	$\exp(-((-\log x_1)^\theta + (-\log x_2)^\theta)^{1/\theta})$	$\theta \geq 1$	$1 - \frac{1}{\theta}$

In [Step 2], one does not have to choose the pair-copula from the same family. One may then use the sequential estimation procedure proposed by Aas et al. (2009, p.189). Note that the sequential procedure does not guarantee a globally optimal fit.

In this paper, we consider the five commonly used bivariate copulas, namely Gaussian, Farlie–Gumbel–Morgenstern (FGM), Frank, Clayton and Gumbel copulas. Table 1 is a summary of properties of these five copulas.

The Frank copula and Gaussian copula are popular because it can cover strong negative or positive dependence between the marginals, and it is symmetric in both upper and lower tails. Furthermore, the range of the dependence parameter  $\theta$  includes both Fréchet bounds (lower bound at  $\theta = -\infty$  and upper bound at  $\theta = \infty$ ). Note that  $\theta = 0$  in the both copulas means independence between the two marginals. However, compared to the Frank copula, the dependence in the Gaussian copula is weaker in the center of the distribution and is stronger in both tails. The FGM copula is attractive because of its simple form, but is useful only when the dependence between the two margins is modest in magnitude. Note that the zero value of  $\theta$  in the FGM copula implies independence.

The Clayton copula and the Gumbel copula can account for only positive dependence, and attain only the Fréchet upper bound at  $\theta = \infty$  (i.e., they do not attain the Fréchet lower bounds for any value of  $\theta$  and do not allow negative dependence). Note that the marginals become independent as  $\theta$  approaches zero for the Clayton copula and when  $\theta = 1$  for the Gumbel copula. The Clayton copula exhibits strong left tail dependence and relatively weak right tail dependence. On the other hand, the Gumbel copula exhibits strong right tail dependence and relatively weak left tail dependence.

After the suitable parametric shapes for each pair copula are determined, the estimation of the parameters in an assumed D-vine copula can be done by numerically maximizing the log-pseudo likelihood in Eq. (5) over the parameters. Note that Aas et al. (2009) provides the algorithm designed to evaluate the log-pseudo likelihood in Eq. (5). One can employ the parameter estimates obtained from the sequential estimation procedure as a starting value for the method of maximum pseudo-likelihood.

In order to check if an estimated D-vine copula appropriately captures the dependence structure in the data, a goodness-of-fit (GOF) test is needed. There are a few GOF tests proposed in the literature that have been used for vine copulas. Aas et al. (2009) employed the probability integral transform of Rosenblatt (1952). Aas and Berg (2009) used two procedures, one based on the empirical copula (Deheuvels, 1979; Genest et al., 2009), and the other based on the empirical distribution function of the empirical copula (Genest et al., 2006).

### 3. Mixture of D-vine copulas

A mixture model is a powerful tool in (i) detecting hidden structure in the data (e.g., clustering the population into homogeneous subpopulations) and (ii) representing arbitrarily complex probability density functions. The mixture model is semiparametric in that it does not put too much structure to the data, unlike a fully parametric density, and does not produce model estimates highly dependent on the observed data, opposed to a fully non-parametric model (Lindsay, 1995; McLachlan and Peel, 2000).

The following subsections introduce a mixture of D-vine copulas that can detect diverse dependence structures hidden in multivariate data and describe an algorithm to estimate the parameters in the proposed model.

#### 3.1. Finite mixture model

A  $d$ -dimensional random vector  $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_d)$  is said to be generated from a mixture of  $M$ -component D-vine densities if its density function can be written as

$$g(\mathbf{x}; \boldsymbol{\theta}) = \sum_{m=1}^M \pi_m f(\mathbf{x}; \boldsymbol{\phi}_m), \tag{6}$$

where  $\pi_m$  is the mixing proportion of the  $m$ -th component satisfying  $0 < \pi_m < 1$  and  $\sum_{m=1}^M \pi_m = 1$ , and  $\boldsymbol{\phi}_m$  is the  $m$ -th component-specific parameter vector for the D-vine density  $f$  defined in Eq. (4). Note that  $\boldsymbol{\theta}$  in Eq. (6) is the set of parameters in the parameter space of dimension  $p$ , denoted by  $\Theta$ , which is the full product space (the simplex of  $\pi_m$ , and the cross

product space of  $\phi_m$ ):

$$\theta = \left[ \begin{pmatrix} \pi_1 \\ \phi_1 \end{pmatrix}, \dots, \begin{pmatrix} \pi_m \\ \phi_m \end{pmatrix}, \dots, \begin{pmatrix} \pi_M \\ \phi_M \end{pmatrix} \right], \quad (7)$$

with each column corresponding to the parameters of each component. Here  $p$  is the total number of free parameters to be estimated, which is equal to  $(M - 1) + \sum_{m=1}^M \dim(\phi_m)$  where  $\dim(a)$  denotes the dimension of a vector  $a$ .

Note that different column permutations of  $\theta$  in Eq. (7) will give the same density  $g(\mathbf{x}; \theta)$  in Eq. (6), which is called label-switching in mixture models. Therefore, we say that a mixture of  $M$ -component D-vine densities in Eq. (6) is identifiable up to a column permutation of  $\theta$ . See Lindsay (1995) and McLachlan and Peel (2000) for more details.

For the estimation of Eq. (6), both the number of components  $M$  and the parameters  $\theta$  need to be estimated, which will be addressed in the following two sections.

### 3.2. EM algorithm

In this section we describe the EM algorithm (Dempster et al., 1977) to obtain the estimates for the parameters  $\theta$  in a mixture of  $M$ -component D-vine densities, given the data set and the number of components  $M$ . The determination of  $M$  will be discussed later in Section 3.3.

Assume that  $N$  observations, say  $\mathbf{x}_k = (\mathbf{x}_{k,1}, \dots, \mathbf{x}_{k,N})$  where  $k = 1, \dots, d$ , drawn randomly from a  $M$ -component D-vine densities in Eq. (6), are available. Then the log-likelihood of  $\theta$ ,  $\ell(\theta)$ , is

$$\ell(\theta) = \log \left( \prod_{n=1}^N g(\mathbf{x}_n; \theta) \right) = \log \left( \prod_{n=1}^N \sum_{m=1}^M \pi_m f(\mathbf{x}_n; \phi_m) \right). \quad (8)$$

Let us denote latent variables  $\mathbf{z}_n = (z_{n1}, \dots, z_{nm}, \dots, z_{nM})$ , where  $z_{nm} = 1$  if  $\mathbf{x}_n$  comes from the  $m$ -th component and  $z_{nm} = 0$  otherwise. Assume that  $\mathbf{z}_n$  is independent and identically distributed from a multinomial distribution, that is,  $\mathbf{z}_n \sim \text{Mult}(M, \pi = (\pi_1, \dots, \pi_M))$ . The complete-data log likelihood for the complete data  $\mathbf{y}_n = (\mathbf{x}_n, \mathbf{z}_n)$ ,  $\ell_c(\theta)$ , is given by

$$\ell_c(\theta) = \log \prod_{n=1}^N \prod_{m=1}^M [\pi_m f(\mathbf{x}_n; \phi_m)]^{z_{nm}} = \sum_{n=1}^N \sum_{m=1}^M z_{nm} \log \pi_m + \sum_{n=1}^N \sum_{m=1}^M z_{nm} \log f(\mathbf{x}_n; \phi_m). \quad (9)$$

The EM algorithm for estimating  $\theta$  in a mixture of  $M$ -component D-vine densities is described as follows: Start with some initial guess for the parameters,  $\theta^{(0)}$ , and then repeat the E-step and M-step to compute successive estimates,  $\theta^{(s)}$  for  $s = 1, 2, \dots$ .

[E-step] Computes the conditional expectation of the complete-data log likelihood,  $\ell_c(\theta)$  in Eq. (9), given the observed data and current parameter estimates for  $\theta$ . This is equivalent to the calculation of the posterior probability that  $\mathbf{x}_n$  belongs to the  $m$ -th component, given the current values of the parameters:

$$\hat{z}_{nm}^{(s)} = E(z_{nm} | \mathbf{x}, \theta^{(s)}) = P(z_{nm} = 1 | \mathbf{x}, \theta^{(s)}) = \frac{\pi_m^{(s)} f(\mathbf{x}_n; \phi_m^{(s)})}{\sum_{\ell=1}^M \pi_\ell^{(s)} f(\mathbf{x}_n; \phi_\ell^{(s)})}. \quad (10)$$

[M-step] Compute the parameter estimates for each component independently,  $(\pi_1^{(s+1)}, \dots, \pi_m^{(s+1)}, \dots, \pi_M^{(s+1)})$  and  $(\phi_1^{(s+1)}, \dots, \phi_m^{(s+1)}, \dots, \phi_M^{(s+1)})$ , by maximizing the expected complete-data log likelihood from the E-step. Here we can obtain a closed-form solution for  $\pi_m^{(s+1)} : \pi_m^{(s+1)} = \frac{\sum_{n=1}^N \hat{z}_{nm}^{(s)}}{N}$ . Then the estimation of  $\phi_m^{(s+1)}$  in the  $m$ -th component D-vine density is equivalent to obtaining the parameter estimates weighted by  $\hat{z}_{nm}^{(s)}$  for the parameters in a D-vine density of Eq. (4). The `optim` function in the R software (R Development Core Team, 2012) is used to numerically maximize the second term in Eq. (9) with respect to  $\phi_m$ .

A nice property of the EM algorithm is that the log likelihood,  $\ell(\theta)$  in Eq. (8), is not decreased during the iteration:  $\ell(\theta^{(s+1)}) \geq \ell(\theta^{(s)})$ . Thus, the E-step and M-step are iterated until  $\ell(\theta^{(s+1)}) - \ell(\theta^{(s)})$  is smaller than a predetermined tolerance, say 0.000001. For a detailed treatment of the convergence properties of EM algorithm, please see Dempster et al. (1977) and Wu (1983).

It is well known that the likelihood function for a finite mixture model has multiple local maximizers. Thus, we run the EM algorithm from multiple starting values for  $\theta$  randomly selected from the range of the parameter space and choose a local maximizer with the highest log likelihood among the found multiple local maximizers.

### 3.3. Model selection

Section 3.2 described how to estimate the parameters  $\theta$  in Eq. (6). For full inference for a mixture model of D-vine densities, however, (a) the determination of the number of components  $M$  and (b) the selection of pair-copula types in each D-vine density are required. To these ends, we employ three widely used model selection criteria: Akaike's information

criterion (AIC) of Akaike (1973), the Bayesian information criterion (BIC) of Schwarz (1978), and the consistent AIC (CAIC) of Bozdogan (1987),

$$AIC = -2 \log L(\hat{\theta}) + 2p,$$

$$BIC = -2 \log L(\hat{\theta}) + p \log(n),$$

$$CAIC = -2 \log L(\hat{\theta}) + p(\log(n) + 1),$$

where  $\hat{\theta}$  is the estimate of a  $p$ -dimensional  $\theta$  in Eq. (6). Each criteria described above has two terms; the first term for measuring the goodness-of-fit, and the second term for penalizing model complexity.

In practice, the number of components  $M$  is often determined by the model selection techniques, as the simulations in the literature have shown that they can not only provide empirical evidence supporting the mixture model when the data are heterogeneous, but also that they can help us avoid using over-parameterized models (see McLachlan and Peel (2000) for detailed reviews of simulation results). Note that the AIC tends to select more components than necessary even when  $N$  is very large and a mixture model is correctly specified (Bozdogan, 1987; Celeux and Soromenho, 1996). The numerical examples shown in Section 4 will illustrate this potential issue of the AIC.

As to the specification of pair-copula types in each D-vine density of a mixture model, it is not straightforward to employ the sequential estimation procedure proposed by Aas et al. (2009, p.189) (such as constructing pairwise scatter plots of variables and/or applying a GOF test for each pair of variables). This is because, in the context of the mixture model, the heterogeneity in the dependence structure is unobserved, and thus there is no observable information available regarding the contribution of individual data points to unobserved heterogeneous dependence structures in the data. Thus, in this paper, we employ the five commonly used bivariate copulas described in Table 1 of Section 2.3 (Gaussian, Frank, Clayton, Gumbel and FGM) as candidates of copula types for all pairs in a mixture of D-vine copulas.

In the numerical examples of Section 4, we will consider several candidates of a mixture of D-vine densities (generated from five copula functions for all pairs and several values for the number of components) for given data, and identify a best-fit model by means of the model selection procedures described above.

We close this section by summarizing a procedure for the full inference of a mixture of D-vine copulas in Eq. (6):

- [Step 1] Obtain the normalized ranks of the  $d$ -dimensional observed data.
- [Step 2] Decide the ordering of the variables at the first level (tree 1) of each D-vine density in a mixture of D-vine densities by following a strategy proposed by Aas and Berg (2009): compute Kendall's tau values for all bivariate pairs, and the variables at tree 1 are ordered according to the  $(d - 1)$  largest Kendall's tau values.
- [Step 3] Consider various candidates of copulas for all pairs in an assumed model.
- [Step 4] Given a copula function chosen from the candidates of copula functions, fit a mixture of  $M$ -component D-vine densities, with  $M$  ranging from 1 to  $M^*$  to the normalized data. Note that when  $M = 1$ , a mixture of  $M$ -component D-vine densities is equivalent to the D-vine copula model without using mixtures. If the number of copula functions considered for the pair-copula is  $S$ , the number of mixture models fitted is  $S \times M^*$ . For the estimation of the parameters in each model, employ the EM algorithm described in Section 3.2.
- [Step 5] Choose the best fitted model (the final pair-copula type for all pairs and the value of  $M$ ) among the fitted models by finding the model with the smallest values of the model selection measures such as the AIC, BIC, and CAIC.

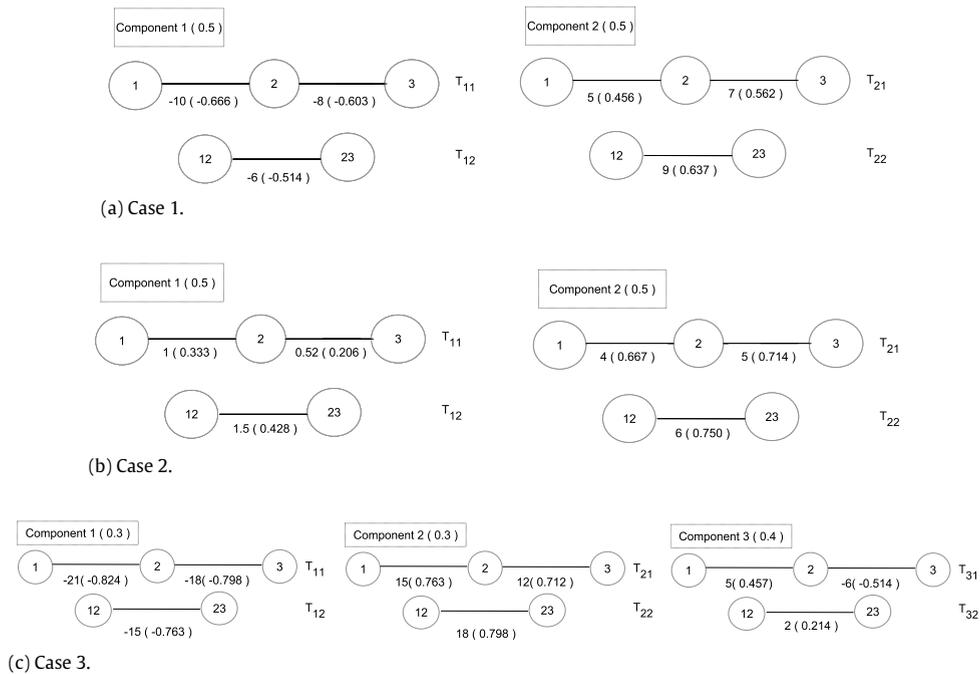
## 4. Numerical examples

In this section the performance of the proposed mixture of D-vine copulas is illustrated with numerous simulated examples and with a real data set studied in Aas and Berg (2009). The results of model estimation are presented in a form of the tree structure as shown in Fig. 1, with the dependence parameter in each pair-copula term at the edge and the corresponding Kendall's tau value in parenthesis.

### 4.1. Simulation studies

The results from the simulation examples are presented in three subsections. In the first subsection the focus is on the scenario where the data generating process (DGP) is a mixture of D-vine copulas in order to check whether the proposed model is able to capture different patterns of dependence structures hidden in the multivariate data. In the second and third subsection, the focus is on the model misspecification scenario. The second subsection is concerned with the case where the DGP is a single-component multivariate D-vine copula which does not have any meaningful mixture of dependences, but a mixture of D-vine copulas is fitted to the data. This scenario will investigate the effect of the misspecification on the number of components. The third subsection studies the performance of the proposed model when the DGP is misspecified; more specifically, we study the case when the data are generated from a mixture of multivariate skew normal distributions (Azzalini and Capitanio, 1999), rather than a mixture of D-vine copulas.

For the data simulation, the `CDVineSim` function in the R package **CDVine** (Brechmann and Schepsmeier, 2011) is used in the first two subsections and the `rmsn` function in the R package **sn** (Azzalini, 2011) is used in the third subsection.



**Fig. 2.** Simulation models for three cases: (a) Case 1 – a mixture of Frank based D-vine copulas with two components, the first one with only negative dependence and the second one with only positive dependence, (b) Case 2 – a mixture of Clayton based D-vine copulas with two components, the first one with weak positive dependence and the second one with strong positive dependence, and (c) Case 3 – a mixture of Frank based D-vine copulas with three components, the first one with only negative dependence, the second one with only positive dependence and the third one with both positive and negative dependences. Note that the values at each edge are the true value of the parameter (outside the parenthesis) and the corresponding Kendall tau value (inside the parenthesis).

#### 4.1.1. Cases with a mixture of D-vine copulas

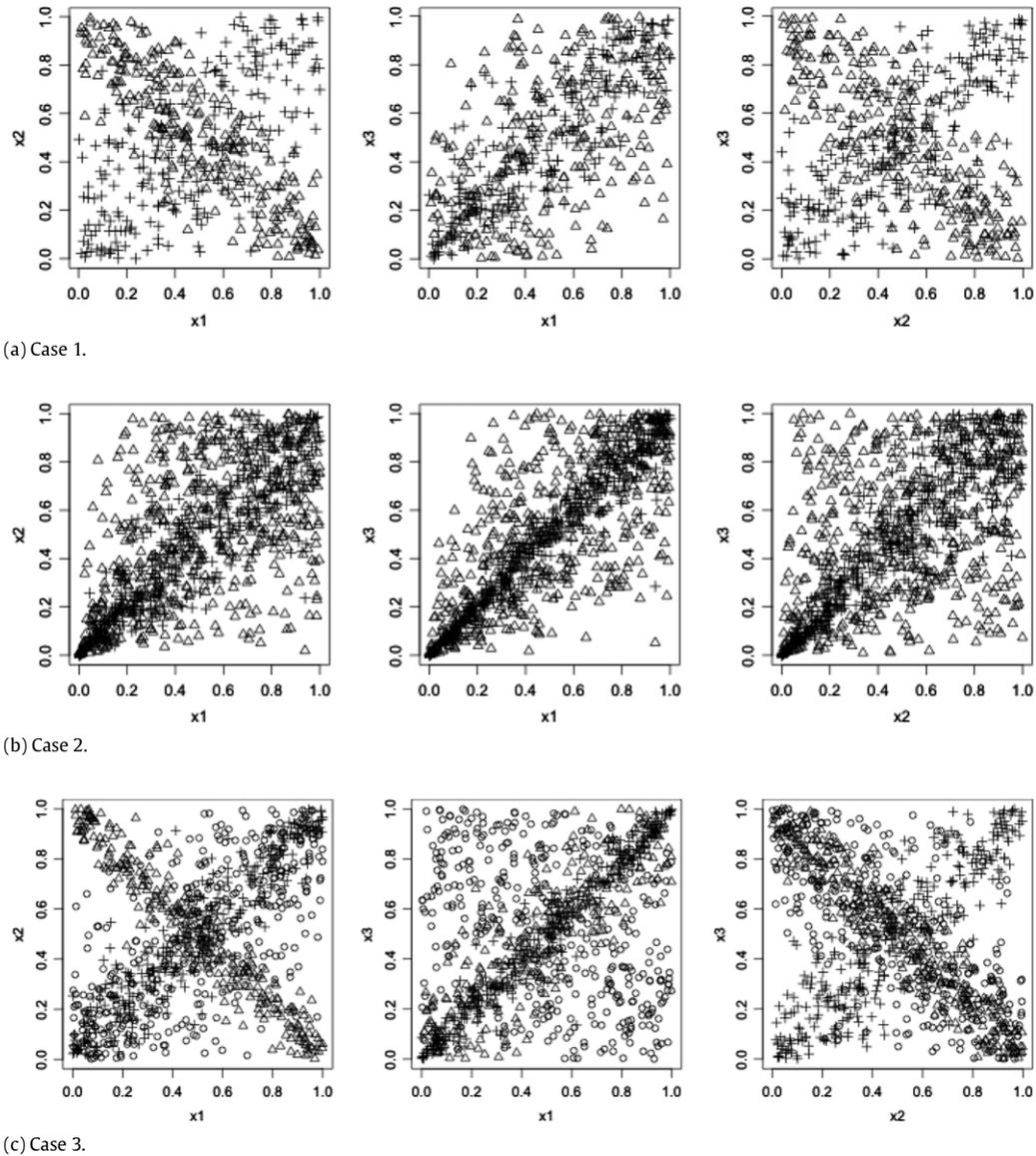
We here illustrate applications of a mixture of D-vine copulas to three sets of simulated data representing different types of dependence structure. The first two sets of simulated data, denoted by Case 1 and Case 2, are the cases where the observations are generated from a two-component ( $M = 2$ ) mixture of three-dimensional ( $d = 3$ ) D-vine densities with Frank and Clayton copulas for all pairs, respectively. The third set of simulated data, denoted by Case 3, is the case where there are three different dependence structures generated from a three-component ( $M = 3$ ) mixture of three-dimensional ( $d = 3$ ) D-vine densities with Frank copulas for all pairs.

Fig. 2 shows three simulation models (Case 1, Case 2, Case 3) in the form of tree structures. Fig. 2 (a) (Case 1) shows a two-component mixture of three dimensional Frank copula-based D-vine densities with equal mixing proportions ( $\pi_1 = \pi_2 = 0.5$ ), where the first component has only negative dependences and the second component has only positive dependences. Fig. 2 (b) (Case 2) is a mixture of Clayton-based D-vine densities with two components ( $\pi_1 = \pi_2 = 0.5$ ), one with weak positive dependence and the other with strong positive dependence. Fig. 2 (c) (Case 3) represents a mixture of Frank copula-based D-vine densities with three components ( $\pi_1 = \pi_2 = 0.3, \pi_3 = 0.4$ ), and has more challenging dependence structures than those of Case 1. The third component of this last simulation model contains both positive and negative dependence.

Fig. 3 gives pairwise scatter plots for three sets of data simulated from the three simulation models given in Fig. 2. The number of observations ( $N$ ) simulated from the three mixture models are 500 for Case 1, 1000 for Case 2 and 1000 for Case 3, respectively. Let  $T_{mj}$  be the  $j$ -th tree of the  $m$ -th component. The observations generated from the first component D-vine density ( $T_{11}, T_{12}$ ) and the second component D-vine density ( $T_{21}, T_{22}$ ) are represented by triangles ( $\Delta$ ) and pluses ( $+$ ), respectively. The observations generated from the third component ( $T_{31}, T_{32}$ ) in Case 3 are represented by circles ( $\circ$ ).

In order to find a best fitted model to each data set, we consider the five copula functions (Gaussian, FGM, Frank, Clayton, and Gumbel) as candidates for all pair-copulas in a mixture of D-vine densities, with the number of components ranging from 1 to  $M^*$ . For Case 1 and 2, we set  $M^* = 3$ , and for Case 3, set  $M^* = 4$ . Then we compute the model selection criteria (AIC, BIC, CAIC) at each fitted model. Note that we employ a multiple starting value strategy in the EM algorithm when estimating the parameters in an assumed model.

Tables 2–4 provide the results for the model selection in the three sets of simulated data. Note that **Boldface** indicates the number of components and pair-copula type selected by each criterion. As shown in Table 2 (for Case 1) and Table 3 (for Case 2), a two-component ( $M = 2$ ) mixture of Frank (for Case 1) and Clayton (for Case 2) based D-vine densities is better fitted to simulated data than the other models we considered because it has the smallest values for each of the model



**Fig. 3.** Pairwise scatter plots of three sets of simulated data generated from three simulation models shown in Fig. 2: (a) Case 1 – a mixture of Frank based D-vine copulas with two components, (b) Case 2 – a mixture of Clayton based D-vine copulas with two components, and (c) Case 3 – a mixture of Frank based D-vine copulas with three components. Note that the observations belonging to the first, second and third components are represented by triangles ( $\Delta$ ), pluses (+) and circles (o), respectively.

**Table 2**  
Model selection criteria for Case 1.

	AIC			BIC			CAIC		
	1	2	3	1	2	3	1	2	3
Gaussian	-251	-602	-657	-239	-572	-611	-236	-565	-600
Frank	-294	<b>-768</b>	-764	-281	<b>-739</b>	-717	-278	<b>-732</b>	-706
Clayton	-199	-294	-295	-186	-265	-249	-183	-258	-238
Gumbel	-223	-296	-293	-210	-267	-246	-207	-260	-235
FGM	-1591	-196	-188	-146	-167	-142	-143	-160	-139

selection criterion. Table 4 (for Case 3) shows that a three component ( $M = 3$ ) mixture of Frank-based D-vine densities fits better than the other models considered.

In Tables 2 and 4, the mixture models using Clayton and Gumbel copula showed poor model fit. This can be interpreted as evidence of copula misspecification caused by the use of Clayton and Gumbel copula, which cannot support negative

**Table 3**  
Model selection criteria for Case 2.

	AIC			BIC			CAIC		
	1	2	3	1	2	3	1	2	3
Gaussian	−1659	−2269	−2279	−1645	−2234	−2225	−1642	−2227	−2214
Frank	−18826	−2202	−2251	−1867	−2168	−2197	−1864	−2161	−2186
Clayton	−2134	−2727	−2724	−2119	−2692	−2670	−2116	−2685	−2659
Gumbel	−1262	−1918	−2014	−1247	−1883	−1960	−1244	−1876	−1949
FGM	−976	−968	−960	−962	−934	−906	−959	−927	−895

**Table 4**  
Model selection criteria for Case 3.

	AIC				BIC				CAIC			
	1	2	3	4	1	2	3	4	1	2	3	4
Gaussian	−458	−1352	−1842	−1916	−443	−1318	−1788	−1842	−440	−1311	−1777	−1827
Frank	−552	−1691	−2419	−2418	−537	−1657	−2365	−2344	−534	−1650	−2354	−2329
Clayton	−323	−774	−653	−877	−308	−740	−599	−803	−305	−733	−588	−788
Gumbel	−396	−914	−976	−1003	−381	−880	−922	−929	−378	−873	−911	−914
FGM	−307	−317	−309	−301	−293	−282	−255	−227	−290	−275	−244	−212

**Table 5**  
Estimated standard errors and biases for the parameter estimates in Case 1.

Parameter	Case 1	
	SE	Bias
$\pi_1$	0.02698	−0.00098
$\beta_{12}^1(\tau_{12}^1)$	0.64209 (0.01530)	−0.03778 (−0.00019)
$\beta_{23}^1(\tau_{23}^1)$	0.52939 (0.02008)	−0.02897 (−0.00020)
$\beta_{13 2}^1(\tau_{13 2}^1)$	0.62849 (0.03274)	−0.01013 (0.00130)
$\beta_{12}^2(\tau_{12}^2)$	0.50089 (0.02942)	0.01767 (−0.00027)
$\beta_{23}^2(\tau_{23}^2)$	0.61803 (0.02613)	0.04068 (0.00033)
$\beta_{13 2}^2(\tau_{13 2}^2)$	0.66773 (0.02459)	0.06072 (0.00087)

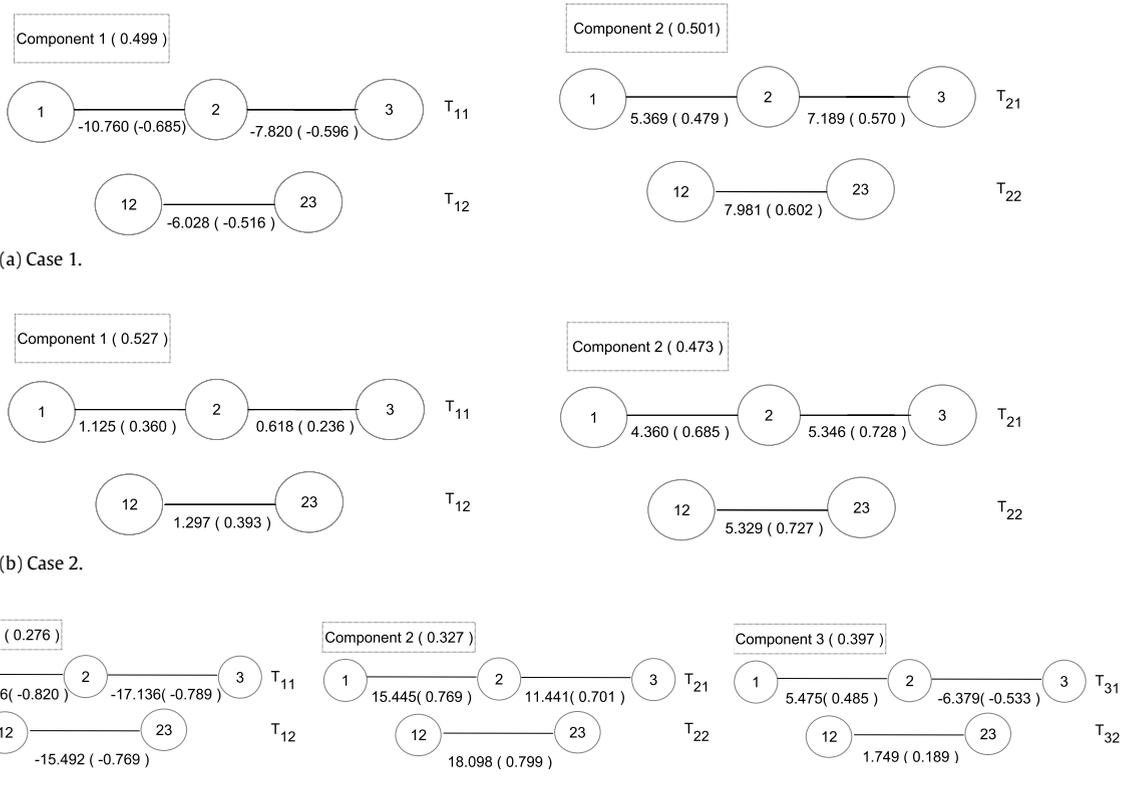
dependence in the simulated data. We also observe that the fit of a mixture model based FGM copula was the worst (and we found that the estimates for most dependence parameters in the models were on the boundary of the parameter space, either 1 or −1). This confirms a well-known property of FGM copula in that this copula is useful only when dependence between the variables is modest in magnitude.

Fig. 4 presents the final estimated models (selected by the model selection criteria) for the three sets of simulated data. From the comparison between the three estimated models (Fig. 4 (a), (b), (c)) and the corresponding true simulation models (Fig. 2 (a), (b), (c)), we can observe that the estimated dependence parameters and the corresponding estimated Kendall's tau values appear to be very close to the true values, and thus hidden dependences in the data are well captured.

To assess accuracy and precision of the parameter estimates in Fig. 4, we also obtain approximate standard errors and biases of the estimates by adopting a parametric bootstrap procedure suggested by Efron and Tibshirani (1993) and McLachlan and Peel (2000, pp. 68–70); see Tables 5–7. That is, given the estimates in each case, we first generate 1000 random samples of size  $N$  from the estimated model. Then, to get the bootstrap estimates of the parameters for each bootstrap sample, we take the parameter estimate calculated from the original data to be the initial value of the parameter when applying the EM algorithm to each bootstrap sample. Note that the parameter estimate from the original data corresponds to the true value of the parameter in the parametric bootstrap. Given the 1000 bootstrap estimates for the parameters, we compute the biases and standard errors for the estimates in Fig. 4. Note that  $\pi_m$ ,  $(\beta_{12}^m, \beta_{23}^m, \beta_{13|2}^m)$  and  $(\tau_{12}^m, \tau_{23}^m, \tau_{13|2}^m)$  in Tables 5–7 represent the mixing proportion, the dependence parameters (i.e., the component specific parameter vectors  $\phi_m$  in Eq. (6)), and the corresponding Kendall's tau values in a mixture of D-vine copulas, respectively.

Given the estimated mixture of D-vine copulas, we can explore the underlying structure of the data by clustering the data into clusters of data points such that the elements in the same cluster are as similar as possible, and the elements in different clusters are as dissimilar as possible. One of the advantages of using a mixture model is that it provides two natural ways to achieve the task of clustering, a “hard” clustering and a “soft” clustering. Suppose that the posterior probability that each point belongs to each cluster is given as Eq. (10). In the hard clustering, we divide data into distinct clusters, where each data point belongs to exactly one cluster corresponding to the highest estimated posterior probability.

However, when the clusters are not well-separated, the hard clustering may provide a poor presentation for the underlying structure of the data (in particular, for data points near cluster boundaries). Thus, one may consider the soft



**Fig. 4.** Fitted models for three cases generated from three simulation models shown in Fig. 2: (a) Case 1 – a mixture of Frank based D-vine copulas with two components, (b) Case 2 – a mixture of Clayton based D-vine copulas with two components, and (c) Case 3 – a mixture of Frank based D-vine copulas with three components. Note that the values at each edge are the estimated parameter (outside the parenthesis) and the corresponding Kendall tau value (inside the parenthesis).

**Table 6**  
Estimated standard errors and biases for the parameter estimates in Case 2.

Parameter	Case 2	
	SE	Bias
$\pi_1$	0.02386	-0.00077
$\beta_{12}^1 (\tau_{12}^1)$	0.08798 (0.01797)	0.00292 (0.00009)
$\beta_{23}^1 (\tau_{23}^1)$	0.06912 (0.02011)	0.00154 (-0.00008)
$\beta_{13 2}^1 (\tau_{13 2}^1)$	0.10815 (0.01976)	0.00181 (-0.00031)
$\beta_{12}^2 (\tau_{12}^2)$	0.27264 (0.01339)	0.00890 (-0.00013)
$\beta_{23}^2 (\tau_{23}^2)$	0.31926 (0.01178)	0.00797 (-0.00021)
$\beta_{13 2}^2 (\tau_{13 2}^2)$	0.44865 (0.01663)	0.03348 (0.00023)

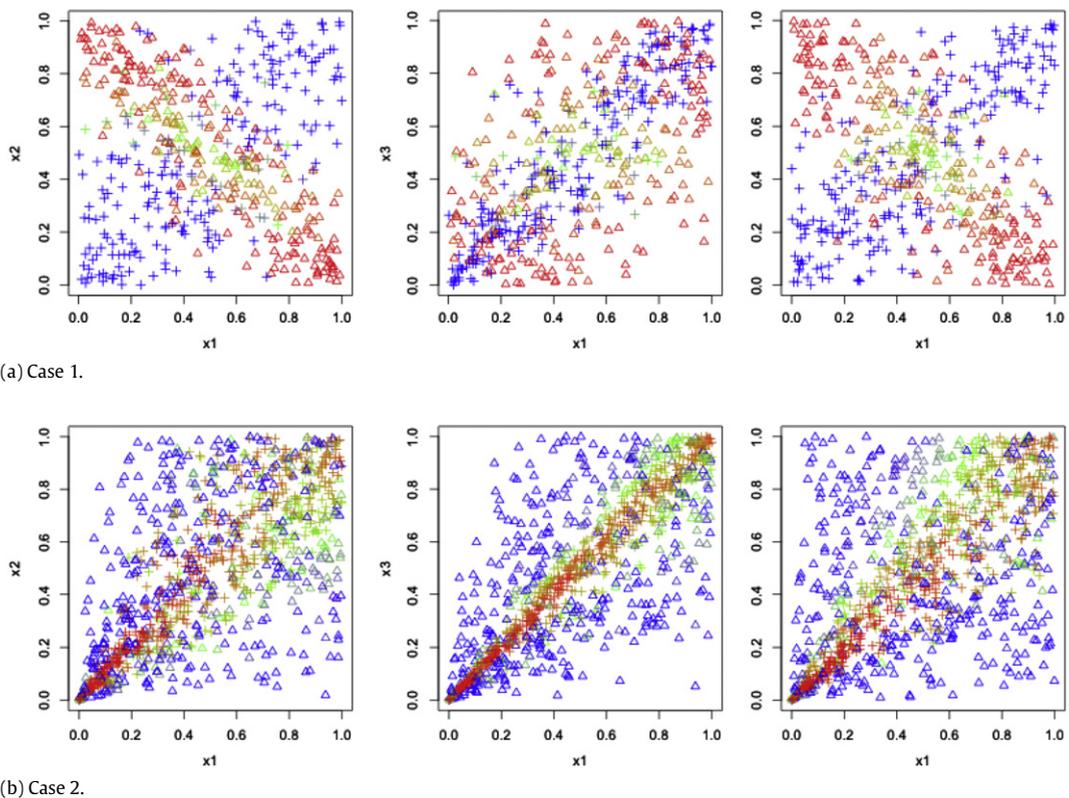
clustering that employs the estimated posterior probability because it measures the strength of the association between that data point and clusters. For visualization purposes, we construct the scatter plots where each point is colored by the magnitude of its estimated posterior probability.

Figs. 5 and 6 provide pairwise scatter plots of both hard clustering and soft clustering for the three cases given in Fig. 4. In Fig. 5(a) (Case 1), the observations belonging to the component with only positive dependence have plus signs (+) with blue color, and those belonging to the component with only negative dependence have triangles ( $\Delta$ ) with red color. Fig. 5(b) (Case 2) shows that the observations represented by triangles ( $\Delta$ ) with blue color belong in the first estimated D-vine component with weak positive dependence, and those represented by plus signs (+) with red color belong in the estimated D-vine component with strong positive dependence. Note that a few observations (represented by the remaining colors between red and blue, such as green colors) are positioned at the region where two clusters are merging; their estimated posterior probabilities were not close to either 0 or 1.

As to the hard clustering for Case 3, as shown in Fig. 6, we use triangles ( $\Delta$ ), pluses (+), and circles (o) to represent the observations belonging to the first, second, and third D-vine components, respectively. Thus, the pairwise scatter plots in

**Table 7**  
Estimated standard errors and biases for the parameter estimates in Case 3.

Parameter	Case 3	
	SE	Bias
$\pi_1$	0.01539	0.00020
$\pi_2$	0.01697	0.00005
$\beta_{12}^1 (\tau_{12}^1)$	0.83439 (0.00670)	-0.04610 (-0.00012)
$\beta_{23}^1 (\tau_{23}^1)$	0.64101 (0.00706)	0.00760 (0.00032)
$\beta_{13 2}^1 (\tau_{13 2}^1)$	1.12724 (0.01455)	-0.17388 (-0.00136)
$\beta_{12}^2 (\tau_{12}^2)$	0.99173 (0.01297)	0.06814 (0.00018)
$\beta_{23}^2 (\tau_{23}^2)$	0.72863 (0.01569)	0.06007 (0.00051)
$\beta_{13 2}^2 (\tau_{13 2}^2)$	1.24608 (0.01246)	0.11747 (0.00041)
$\beta_{12}^3 (\tau_{12}^3)$	0.41749 (0.02388)	0.02838 (0.00074)
$\beta_{23}^3 (\tau_{23}^3)$	0.47108 (0.02289)	-0.04818 (-0.00140)
$\beta_{13 2}^3 (\tau_{13 2}^3)$	0.34050 (0.03454)	-0.00571 (-0.00115)

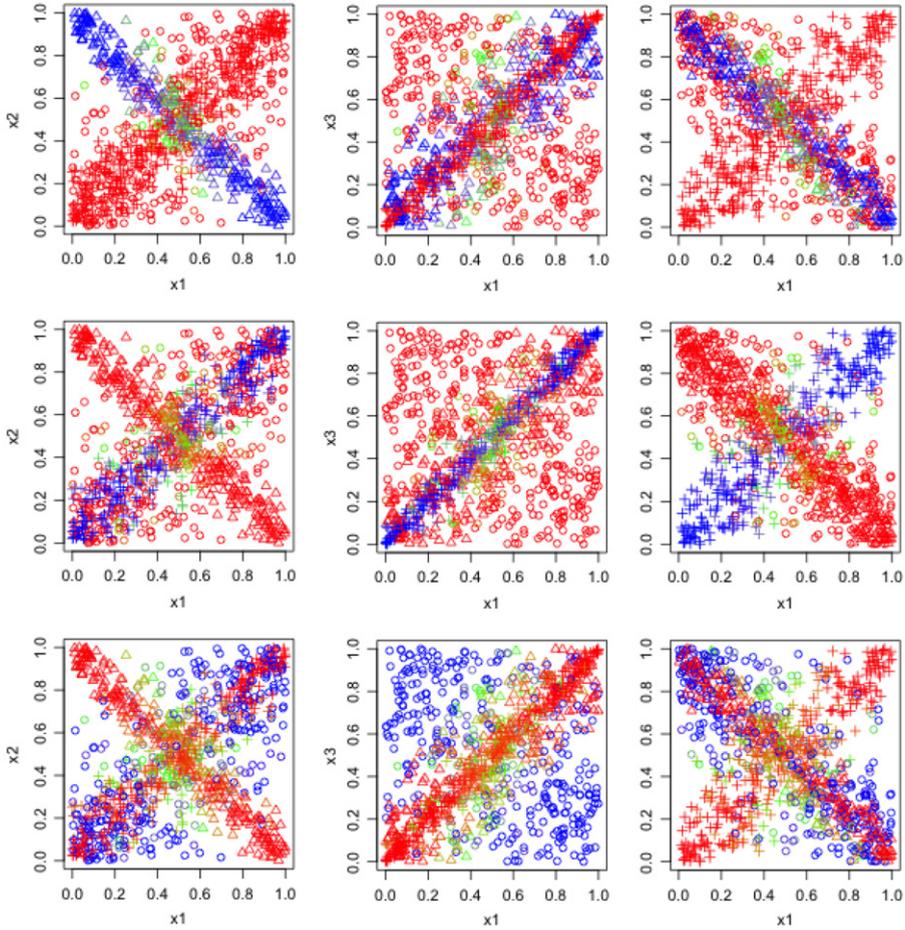


**Fig. 5.** Pairwise scatter plots of both hard clustering and soft clustering for two cases generated from two simulation models shown in Fig. 2: (a) Case 1 – a mixture of Frank based D-vine copulas with two components, and (b) Case 2 – a mixture of Clayton based D-vine copulas with two components.

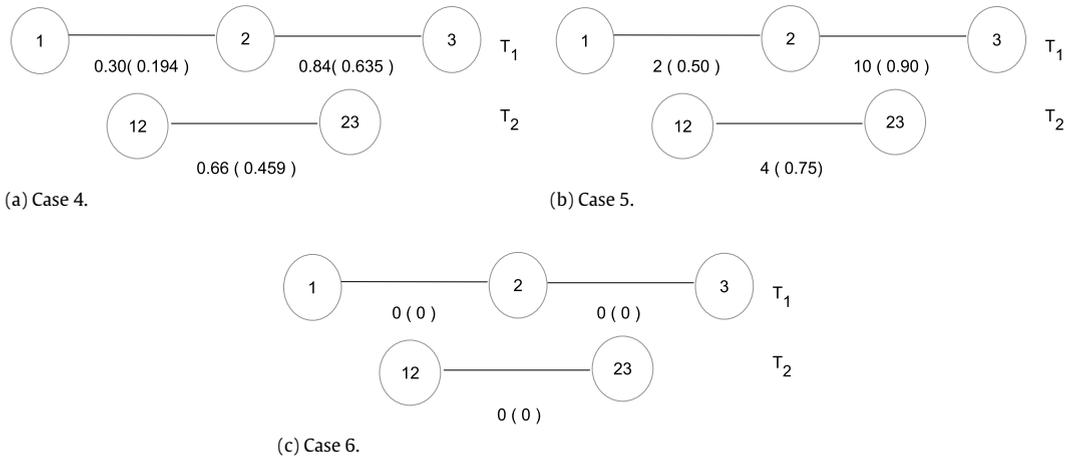
the three rows of Fig. 6 all show the same hard clustering results. For the soft clustering, on the other hand, each row of Fig. 6 focuses on one component because there are three components. That is, the pairwise scatter plots in the first row show that the first estimated D-vine component has observations represented by triangles ( $\Delta$ ) with blue color. From the second and third rows, we observe that the second D-vine component has observations represented by plus signs (+) with blue color and the third D-vine component has observations represented by circles (o) with blue color, respectively.

#### 4.1.2. Cases with a single-component D-vine copula

In this subsection we consider the scenario where a mixture of  $m$ -component D-vine copulas is fitted to the data generated from a single-component multivariate D-vine copula (without using a mixture model). We first generate the three data sets of size 1000 from a three-dimensional D-vine density using Gaussian, Gumbel and Independent copulas



**Fig. 6.** Pairwise scatter plots of both hard clustering and soft clustering for the third case generated from the third simulation model shown in Fig. 2: Case 3 – a mixture of Frank based D-vine copulas with three D-vine components.



**Fig. 7.** Simulation models: (a) Case 4 – a Gaussian based D-vine copula, (b) Case 5 – a Gumbel based D-vine copula, and (c) Case 6 – a D-vine copula using an independent copula. Note that the values at each edge are the true value of the parameter (outside the parenthesis) and the corresponding Kendall tau value (inside the parenthesis).

for all pairs and denote them as Case 4, Case 5 and Case 6, respectively. See Fig. 7 for these three simulation models. We then fit the proposed mixture of D-vine copulas to each simulated data and compute the model selection criteria at each fitted model. We here consider the five copula functions (Gaussian, FGM, Frank, Clayton, and Gumbel) as candidates for all pair-copulas in a mixture of D-vine densities, with the number of components ranging from 1 to 3. Notice that a mixture of D-vine densities with  $M = 1$  is equivalent to a single-component multivariate D-vine copula model.

**Table 8**

Model selection criteria for Case 4.

	AIC			BIC			CAIC		
	1	2	3	1	2	3	1	2	3
Gaussian	– <b>2002</b>	–1997	–1991	– <b>1987</b>	–1963	–1937	– <b>1984</b>	–1956	–1926
Frank	–1759	–1765	–1761	–1744	–1731	–1707	–1741	–1724	–1696
Clayton	–1456	–1530	–1542	–1441	–1495	–1488	–1438	–1488	–1477
Gumbel	–1753	–1770	–1765	–1738	–1736	–1711	–1735	–1729	–1700
FGM	–709	–701	–693	–694	–666	–639	–691	–659	–628

**Table 9**

Model selection criteria for Case 5.

	AIC			BIC			CAIC		
	1	2	3	1	2	3	1	2	3
Gaussian	–5823	–6083	–6120	–5809	–6049	–6066	–5806	–6042	–6055
Frank	–5511	–5669	–5692	–5496	–5634	–5638	–5493	–5627	–5627
Clayton	–3539	–4452	–4573	–3525	–4417	–4519	–3522	–4410	–4508
Gumbel	– <b>6583</b>	–6582	–6575	– <b>6569</b>	–6548	–6521	– <b>6566</b>	–6541	–6510
FGM	–1051	–1043	–1035	–1036	–1008	–981	–1033	–1001	–970

**Table 10**

Model selection criteria for Case 6.

	AIC			BIC			CAIC		
	1	2	3	1	2	3	1	2	3
Gaussian	1.716	–0.58	–1.59	16.44	33.78	52.40	19.44	40.78	63.40
Frank	2.15	0.74	– <b>6.53</b>	16.88	35.09	47.45	19.88	42.09	58.45
Clayton	0.98	7.84	15.52	<b>15.70</b>	42.19	69.51	<b>18.70</b>	49.19	80.51
Gumbel	2.60	6.09	13.48	17.33	40.45	67.47	20.33	47.45	78.47
FGM	2.09	2.31	8.13	16.82	36.66	62.12	19.82	43.66	73.12

The model-selection results in these three simulated data sets are summarized in Tables 8–10. We observe from Table 8 (for Case 4) and Table 9 (for Case 5) that all three model selection measures choose a mixture of D-vine with  $M = 1$ . That is, the model selection results support for a single-component three-dimensional D-vine copula without using mixtures. This makes sense because there is no available information in the simulated data for sensible clustering.

As shown in Table 10 for Case 6, the AIC choose a three-component mixture of Frank-based D-vine copulas which is more complex than the true model. This result is consistent with the potential issue of the AIC explained in Section 3.3. We also find that the estimated mixing proportions for the second and third components are in fact very close to zero ( $\hat{\pi}_2 = 0.026$  and  $\hat{\pi}_3 = 0.076$ ), implying that two out of three components are not practically meaningful. On the other hand, the BIC and CAIC both select a single-component Clayton-based D-vine density (i.e., a mixture of Clayton-based D-vine copula with  $M = 1$ ) for Case 6, and the estimated dependence parameters (and the corresponding estimated Kendall's tau) are found to be 0.022 (0.011), 0.05 (0.024) and 0.056 (0.027), all of which are very close to zero. This implies that a single-component D-vine copula model based on an independent copula for all pairs is appropriate.

#### 4.1.3. Case with a mixture of multivariate skew distributions

In this subsection, we fit a mixture of  $m$ -component D-vine copulas to the data generated from a mixture of multivariate skew normal distributions (Azzalini and Capitanio, 1999). Denoted as Case 7, a data set of size 1000 ( $=n$ ) is generated from a two-component ( $M = 2$ ) mixture of three-dimensional ( $d = 3$ ) skew normal distributions:

$$\pi_1 f(\mathbf{x}; \boldsymbol{\alpha}_1, \Omega_1) + (1 - \pi_1) f(\mathbf{x}; \boldsymbol{\alpha}_2, \Omega_2),$$

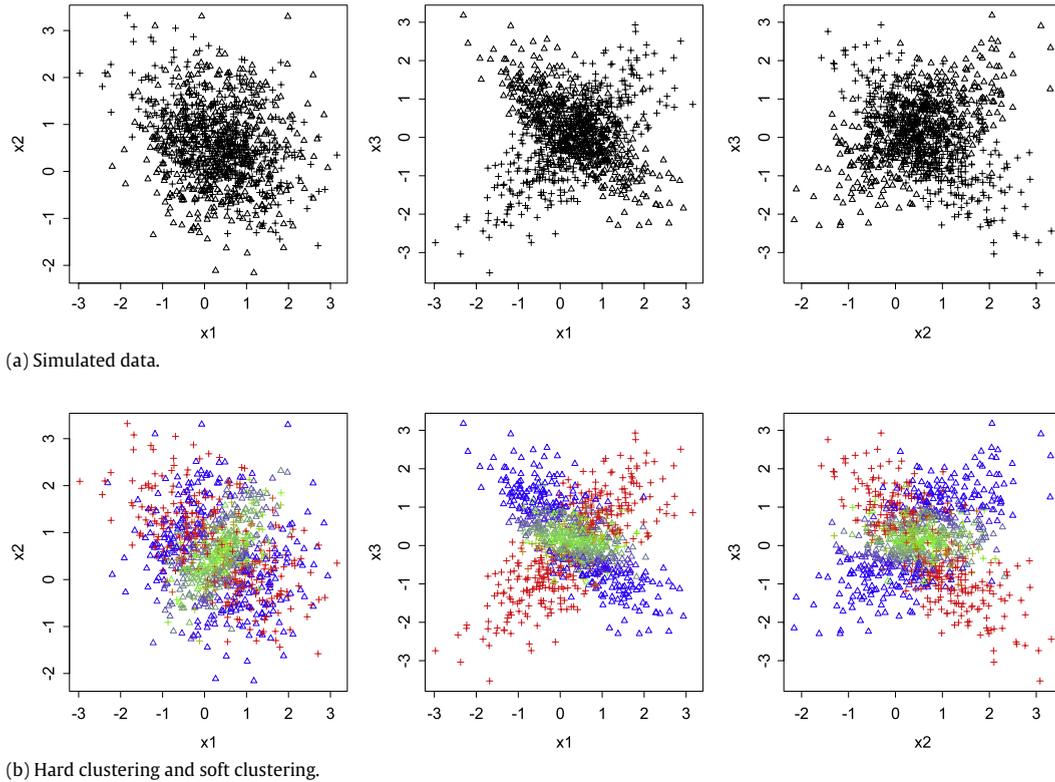
where  $f(\mathbf{x}; \boldsymbol{\alpha}, \Omega) = 2\phi_d(\mathbf{x}; \Omega)\Phi(\boldsymbol{\alpha}'\mathbf{x})$  is a  $d$ -dimensional skew normal density,  $\boldsymbol{\alpha}$  is a  $d$ -dimensional vector,  $\phi_d(\mathbf{x}; \Omega)$  is the  $d$ -dimensional normal density with zero mean and correlation matrix  $\Omega$  and  $\Phi(\cdot)$  is the standard normal distribution function. The true values of the parameters used in the simulation are  $\pi_1 = 0.5$ ,  $\boldsymbol{\alpha}_1 = (3, -1, 3)'$ ,  $\boldsymbol{\alpha}_2 = (1, 3, 1)'$ ,  $\text{vech}(\Omega_1) = (1, 0.6, -0.6, 1, 0.15, 1)$  and  $\text{vech}(\Omega_2) = (1, -0.6, 0.6, 1, -0.15, 1)$  where  $\text{vech}(A)$  vectorizes only the lower triangular part of a symmetric matrix  $A$ .

Fig. 8(a) shows the pairwise scatter plots for the simulated data of Case 7. The observations generated from the first and the second skew normal densities are represented by triangles ( $\Delta$ ) and pluses ( $+$ ), respectively. Notice that the two clusters are not well separated, especially in terms of their locations.

For the construction of a mixture of D-vine copulas appropriate to the data, we implement the five steps in Section 3.3. Here the three copula functions (Gaussian, FGM and Frank) are used as candidates for all pair-copulas in a mixture of D-vine densities, with the number of components ranging from 1 to 3. Note that the Clayton and Gumbel copula functions cannot be used here because some variables are negatively correlated.

**Table 11**  
Model selection criteria for Case 7.

	AIC			BIC			CAIC		
	1	2	3	1	2	3	1	2	3
Gaussian	−59	−758	−768	−44	−724	−714	−41	−717	−703
Frank	−50	−626	−636	−36	−592	−582	−33	−585	−571
FGM	−45	−179	−171	−30	−145	−117	−27	−138	−106



**Fig. 8.** Case 7 – pairwise scatter plots of (a) simulated data generated from a two-component mixture of three-dimensional skew normal distributions, and (b) hard clustering and soft clustering based on a two-component mixture of Gaussian-based D-vine densities. Note that the observations belonging to the first and second components are represented by triangles ( $\Delta$ ) and pluses ( $+$ ), respectively.

Table 11 provides the results for the model selection for Case 7. The BIC and CAIC suggest that a two-component ( $M = 2$ ) mixture of D-vine densities with Gaussian copulas is better fitted to the data than the other models, while the AIC tends to select more complex models than the other two criteria did. Fig. 8 (b) provides the pairwise scatter plots of both hard clustering and soft clustering based on a two-component mixture of Gaussian-based D-vine densities. The observations belonging to the first component have triangles ( $\Delta$ ) with blue color, and those belonging to the second component have pluses ( $+$ ) with red color. A group of points in the area where the clusters are overlapped are represented by the green colors. From Fig. 8 we observe that a fitted two-component ( $M = 2$ ) mixture of Gaussian-based D-vine densities provides the two estimated clusters similar to the true clusters and thus successfully captures the dependences between the variables.

#### 4.2. Data analysis: precipitation data

Introduced in Aas and Berg (2009), the real data we study in this section is the daily precipitation data (mm) for the period January 1, 1990 to December 31, 2006, measured from four meteorological stations in the four municipalities, one station for each municipality of Akershus County in southern Norway: Vestby, Ski, Nannestad, and Hurdal. Vestby and Ski are located in the southwestern part of the county, while Nannestad and Hurdal are situated in the northwestern part.

Modeling precipitation data and valuing related derivative contracts are very important not only in the insurance industry, but also in the field of weather derivatives (Muschhoff et al., 2006). As described in Aas and Berg (2009), the main interest here is to estimate the dependence structure between the amount of precipitations measured from four stations, conditional on that it rains. Thus, we remove days with zero precipitation values for at least one station. Note that the number of non-zero precipitation for all four stations during this period, i.e. the sample size, is 2065.

**Table 12**

Model selection results for precipitation data. The **Boldface** indicates the number of components ( $M$ ) and pair-copula type selected by each criterion.

	AIC			BIC			CAIC		
	1	2	3	1	2	3	1	2	3
Gaussian	−8599	−9495	−9548	−8565	−9422	−9435	−8559	−9409	−9415
Frank	−9234	−9443	−9498	−9200	−9370	−9385	−9194	−9357	−9365
Clayton	−5353	−7052	−7369	−5319	−6979	−7256	−5313	−6966	−7236
Gumbel	−9654	−9832	<b>−9835</b>	−9620	<b>−9759</b>	−9723	−9614	<b>−9746</b>	−9703
FGM	−3503	−3498	−3484	−3469	−3424	−3371	−3463	−3411	−3351

For the construction of a mixture of D-vine copulas, we follow the five steps summarized in Section 3.3. First, convert the precipitation values at each station to the normalized ranks. Note that the reader can refer to Aas and Berg (2009) for the pairwise scatter plots for daily precipitation data and the corresponding normalized ranks. Second, we decide the ordering of the variables at the first level of a D-vine density by computing Kendall's tau values for all bivariate pairs. The Kendall's tau values for six pairs {Vestby, Ski}, {Vestby, Nannestad}, {Vestby, Hurdal}, {Ski, Nannestad}, {Ski, Hurdal} and {Nannestad, Hurdal} are 0.785, 0.521, 0.484, 0.566, 0.519, and 0.732, respectively. Thus, the appropriate ordering of the variables is Vestby, Ski, Nannestad, and Hurdal. Note that the distance between Vestby and Ski (Nannestad and Hurdal) is shorter than the distance between Vestby/Ski and Nannestad/Hurdal.

In the third and fourth step, we consider each of the five copula functions, Gaussian, Frank, Clayton, Gumbel and FGM, as candidates for all pair-copulas in a mixture of D-vine copulas, and fit these five types of mixture of D-vine copulas, with the number of components ranging from 1 to 3, to the normalized ranks. Note that the total number of models fitted here is 15 (=five pair-copula types  $\times$  three values for the number of components). As the final step, we compute the three model selection criteria described in Section 3.3, AIC, BIC, and CAIC, at each fitted model.

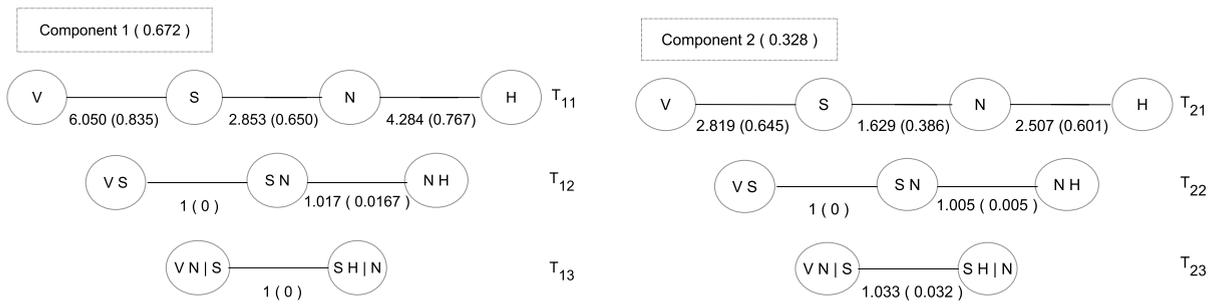
The results for the model selection are shown in Table 12. It appears that two-component ( $M = 2$ ) mixture of D-vine densities with Gumbel copulas for all pairs provides a better fit than other models because of the smallest values of the BIC and CAIC. Note that the AIC tends to select more complex models than necessary, which is consistent with the results found by many researchers. Table 12 also shows that the mixture models based on FGM and Clayton copulas are poorly fitted to the data, compared to those using Frank and Gumbel copulas. We should mention that the estimates for most dependence parameters in a mixture of FGM-based D-vine densities are in fact either 1 or  $-1$ , while a mixture of Clayton-based D-vine densities has a few estimates in the first level of D-vine trees that are very close to zero, the lower bound of the permissible range.

Aas and Berg (2009) fitted a D-vine copula (without using mixtures) based on three different copula families, Gumbel, Frank and Student- $t$ , to the same data. They found that a D-vine copula with Gumbel copulas for all pairs provided a better fit than the other two copulas they considered; its log likelihood was highest among three, and the goodness-of-fit test using the empirical distribution function of the empirical copula (Genest et al., 2006) gave a  $p$ -value larger than an  $\alpha$ -level of 0.05. We, however, find that the log-likelihood of a two-component mixture of Gumbel-based D-vine densities is much higher than that of a single-component Gumbel-based D-vine density without using mixtures ( $M = 1$ ).

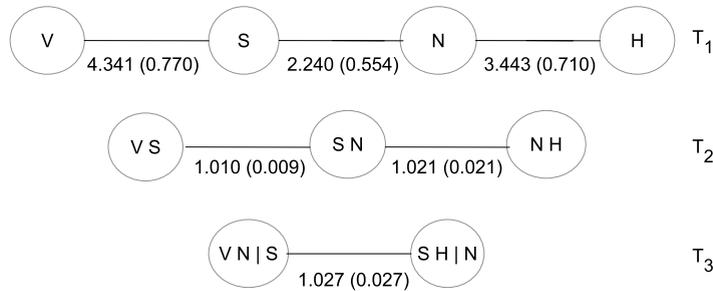
Fig. 9(a) shows the parameter estimation for a two-component mixture of Gumbel-based D-vine copulas in a tree structure. For the comparison purpose, Fig. 9(b) gives the estimates of the parameters for a single-component Gumbel-based D-vine copula without using mixtures. We observe from Fig. 9(a) that the data consists of two components with different types of dependence structures. The first component with 67.2% mixing proportion shows strong overall dependence between all four stations; the corresponding values of Kendall's tau in the first level are 0.835 for {Vestby, Ski}, 0.650 for {Ski, Nannestad}, and 0.767 for {Nannestad, Hurdal} in  $T_{11}$ . This strong dependence structure may result from the fact that all stations are located in the same county of Norway. The second component whose mixing proportion is 32.8%, on the other hand, describes a relatively weaker dependence pattern governed by the *locality* of the four stations. Note that the Kendall's tau value for {Ski, Nannestad} in the  $T_{21}$  is much smaller than the values for {Vestby, Ski} and {Nannestad, Hurdal} in the  $T_{21}$ .

To further visualize the underlying structure of the precipitation data explained above, we construct the pairwise scatter plots of both hard and soft clustering of daily precipitation data using the posterior probability in Eq. (10) computed from the two-component mixture of Gumbel-based D-vine copulas (see Fig. 10). In each pairwise plot the first cluster with  $\hat{\pi}_1 = 0.672$  (points with plus symbols and colors close to red) consists of the precipitation values densely located near the  $45^\circ$  line. On the other hand, the second cluster with  $\hat{\pi}_2 = 0.328$  (points with circle symbols and colors close to blue) has values relatively far from the  $45^\circ$  line. Notice that the points for {Vestby, Ski} and {Nannestad, Hurdal} belonging to the second cluster are closer to the  $45^\circ$  line, compared to those for the other pairs of stations in the second cluster.

In addition, we also check if the two clusters obtained from a mixture of Gumbel-based D-vine copulas correspond to particular periods of time by investigating the total precipitation of four stations for each cluster with respect to time. Fig. 11 shows the monthly plot of the total precipitation of four stations based on hard clustering. The pattern of the first cluster (represented by the pluses) appears to generally follow that of the whole data (represented by the squares) over the year, unlike that of the second cluster (represented by the circles). The contributions of the second cluster to the whole data are non-negligible only during the months of June to October.



(a) Two-component ( $M = 2$ ) mixture of Gumbel-based D-vine copulas.



(b) A single-component Gumbel-based D-vine copula without using mixture model ( $M = 1$ ).

**Fig. 9.** (a) Two fitted models for the precipitation data. The values outside and inside the parentheses at each edge are the estimated dependence parameter in each pair-copula term and the corresponding Kendall tau value. Note that the precipitation data measured from Vestby, Ski, Nannestad and Hurdal are denoted as V, S, N and H.

From the analysis performed in this section, the proposed mixture of D-vine copulas is useful for revealing the hidden dependence structure of the precipitation data, possibly related to both spatial and temporal factors that are not considered in the model.

### 5. Conclusion

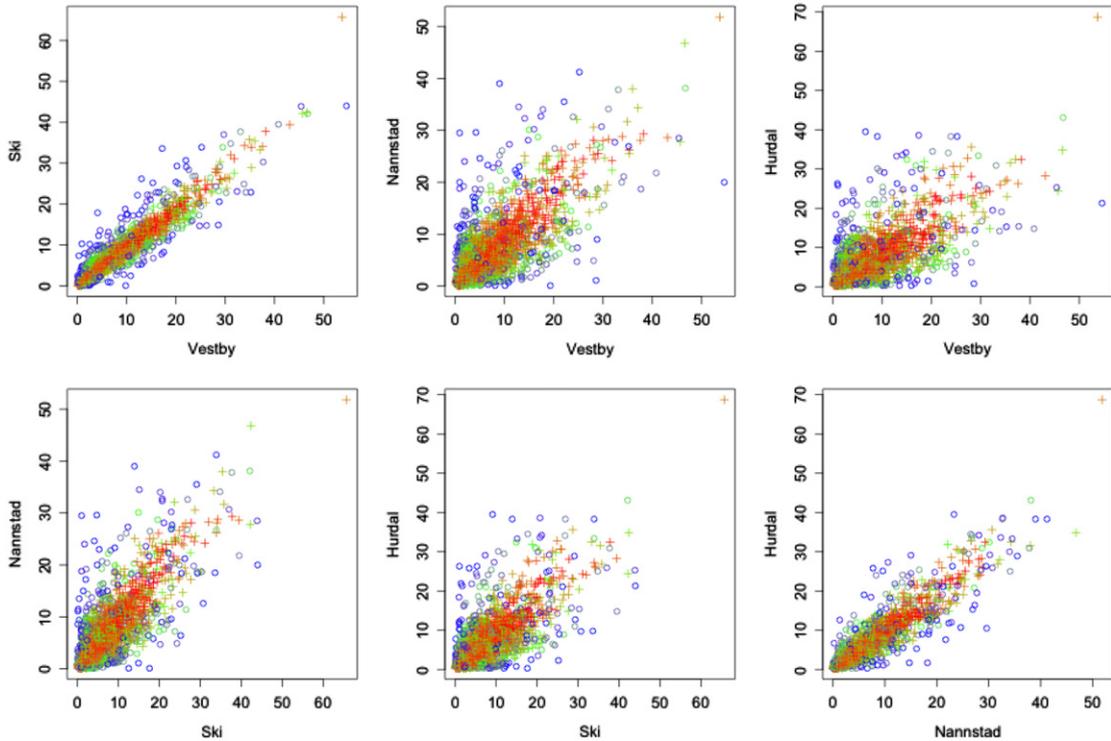
In this article we have presented a mixture of D-vine copulas for revealing complex and hidden dependence structures in multivariate data, and illustrated its performance on simulated data sets and real data. We have also shown that the EM algorithm enables the estimation of the dependence parameters in each D-vine density.

In applications of a mixture of D-vine copulas shown in this paper, we use the same parametric copula function for all pair-copulas. While the proposed methodology significantly outperforms a single-component multivariate D-vine copula model in revealing different hidden dependence structures, there is still space for further improvement. As one of the reviewers pointed out, the strength of the D-vine copulas relies, among other things, on the specification of different copula types for the pair-copulas. A valuable extension of this research would be to investigate the possibility of either specifying different copula types for all pair-copulas in a mixture of D-vine copulas or modeling each component of the mixture through a different D-vine copula model which uses the same copula for all pairs.

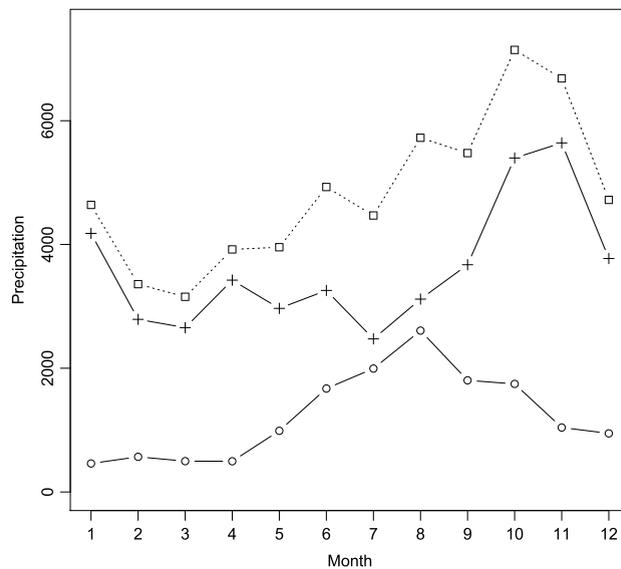
Finally, we employ the model selection criteria to determine the number of components and a suitable copula function for all pairs in a mixture of D-vine densities. One may wish to carry out a goodness-of-fit (GOF) test to check the suitability of the mixture model fitted to multivariate data. As Touboul (2011) pointed out, existing GOF tests for multivariate copulas present feasibility issues for high dimensional problems due to the curse of dimensionality. It appears that there are only few GOF tests for a mixture of multivariate copulas. Future work will consider development of a statistical tool to evaluate the appropriateness of a mixture of copulas to multivariate data.

### Acknowledgments

We appreciate both Dr. Kjersti Aas and Dr. Daniel Berg's help to provide us with the precipitation data and give us a detailed explanation about the precipitation data. We also appreciate the editor, an associate editor and two anonymous learned referees whose helpful suggestions and insightful comments greatly improved the quality of this paper.



**Fig. 10.** Pairwise scatter plots of both hard clustering and soft clustering of daily precipitation data using a two-component mixture of Gumbel-based D-vine copulas. For the hard clustering each data point is assigned to the cluster corresponding to the highest estimated posterior probability: the pluses and the circles represent the points belonging to the first cluster with  $\hat{\pi}_1 = 0.672$  and the second cluster with  $\hat{\pi}_2 = 0.328$ , respectively. For the soft clustering each point is colored by the magnitude of its estimated posterior probability: the observations belonging to the first cluster have the colors close to red, and those belonging to the second cluster have the colors close to blue. Note that the precipitation values positioned between the two clusters (their estimated posterior probabilities are not close to either 0 or 1) represented by colors close to green. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 11.** Monthly plot of the total precipitation of four stations based on hard clustering. Note that the pluses, circles and squares in each month represent the numbers computed from the first cluster with  $\hat{\pi}_1 = 0.672$ , the second cluster with  $\hat{\pi}_2 = 0.328$  and the whole data (consisting of both clusters), respectively.

## References

- Aas, K., Berg, D., 2009. Models for construction of multivariate dependence: a comparison study. *The European Journal of Finance* 15 (7), 639–659.
- Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics & Economics* 44 (2), 182–198.
- Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), 2nd International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 267–281.
- Azzalini, A., Capitanio, A., 1999. Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society Series B* 61, 579–602.
- Azzalini, A., 2011. R package *sn*: The skew-normal and skew-*t* distributions (version 0.4-17). <http://azzalini.stat.unipd.it/SN>.
- Bauer, A., Czado, C., Klein, T., 2012. Pair-copula constructions for non-Gaussian DAG models. *Canadian Journal of Statistics* 40 (1), 86–109.
- Bedford, T., Cooke, R.M., 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence* 32, 245–268.
- Bedford, T., Cooke, R.M., 2002. Vines — a new graphical model for dependent random variables. *Annals of Statistics* 30 (4), 1031–1068.
- Brechmann, E., Czado, C., Aas, K., 2012. Truncated and simplified regular vines in high dimensions with application to financial data. *Canadian Journal of Statistics* 40 (1), 68–85.
- Brechmann, E.C., Schepsmeier, U., 2011. Modeling dependence with C- and D-vine copulas: The R-package CDVine. R vignette of the R-package CDVine (version 1.1–4).
- Bozdogan, H., 1987. Model selection and Akaike information criterion (AIC): the general theory and its analytical extensions. *Psychometrika* 52 (3), 345–370.
- Carrière, J.F., 2006. Copulas. *Encyclopedia of Actuarial Science*.
- Celeux, G., Soromenho, G., 1996. An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification* 13 (2), 195–212.
- Cherubini, U., Luciano, E., Vecchiato, W., 2004. *Copula Methods in Finance*. John Wiley, Chichester.
- Cuvelier, E., Noirhomme-Fraiture, M., 2005. Clayton copula and mixture decomposition, in: Janssen, J., Lenca, P., (Eds.), In: *Applied Stochastic Models and Data Analysis (ASMDA 2005)*, Brest, France, 17–20 May, pp. 699–708.
- Deheuvels, P., 1979. La fonction de dépendance empirique et ses propriétés: Un test nonparamétrique d'indépendance. *Académie Royale de Belgique. Bulletin de la Classe des Sciences. 5e Série* 65, 274–292.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B* 39 (1), 1–38.
- Efron, B., Tibshirani, R., 1993. *An Introduction to the Bootstrap*. Chapman & Hall, London.
- Genest, C., Ghoudi, K., Rivest, L.-P., 1995. A semi-parametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 82, 543–552.
- Genest, C., Quessy, J.-F., Remillard, B., 2006. Goodness-of-fit procedures for copula models based on the probability integral transform. *Scandinavian Journal of Statistics* 33 (2), 337–366.
- Genest, C., Remillard, B., Beaudoin, D., 2009. Goodness-of-fit tests for copulas: a review and a power study. *Insurance: Mathematics & Economics* 44 (2), 199–213.
- Hobæk Haff, I., Aas, K., Frigessi, A., 2010. On the simplified pair-copula construction — simply useful or too simplistic? *Journal of Multivariate Analysis* 101 (5), 1296–1310.
- Hu, L., 2006. Dependence patterns across financial markets: a mixed copula approach. *Applied Financial Economics* 16 (10), 717–729.
- Joe, H., 1996. Families of *m*-variate distributions with given margins and  $m(m-1)/2$  bivariate dependence parameters, in: Rüschendorf, L., Schweizer, B., Taylor, M. D., (Eds.), *Distributions with fixed marginals and related topics*, pp. 120–141.
- Joe, H., 1997. *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
- Joe, H., Li, H., Nikoloulopoulos, A.K., 2010. Tail dependence functions and vine copulas. *Journal of Multivariate Analysis* 101, 252–270.
- Kim, G., Silvapulle, M.J., Silvapulle, P., 2007. Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis* 51 (6), 2836–2850.
- Kurowicka, D., Cooke, R.M., 2003. A parametrization of positive definite matrices in terms of partial correlation vines. *Linear Algebra and its Applications* 372, 225–251.
- Kurowicka, D., Cooke, R.M., 2004. *Distribution — free continuous bayesian belief nets*. in: *Fourth International Conference on Mathematical Methods in Reliability Methodology and Practice*, Santa Fe, New Mexico.
- Kurowicka, D., Cooke, R.M., 2006. *Uncertainty Analysis with High Dimensional Dependence Modelling*. Wiley, New York.
- Lindsay, B.G., 1995. *Mixture Models: Theory, Geometry and Applications*. In: *NSF-CBMS Regional Conference Series in Probability and Statistics, Vol. 5*. Institute of Mathematical Statistics and the American Statistical Association, Alexandria, Virginia.
- McLachlan, G.J., Peel, D., 2000. *Finite Mixture Models*. Wiley.
- McNeil, A.J., Frey, R., Embrechts, P., 2005. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.
- Musshoff, O., Odening, M., Xu, W., 2006. Modeling and hedging rain risk. in: *Annual Meeting of the American Agricultural Economics Association (AAEA)*, Long Beach, USA, July pp. 23–26.
- Nelsen, R.B., 2006. *An Introduction to Copulas*, second ed. Springer, New York.
- Oakes, D., 1982. A model for association in bivariate survival data. *Journal of the Royal Statistical Society. Series B* 44 (3), 414–422.
- Oakes, D., 1994. Multivariate survival distributions. *Journal of Nonparametric Statistics* 3, 343–354.
- Onken, A., Grünewälder, S., Munk, M.H.J., Obermayer, K., 2009. Analyzing short-term noise dependencies of spike-counts in macaque prefrontal cortex using copulas and the flashlight transformation. *PLoS Computational Biology* 5 (11), e1000577, 1–11.
- Panagiotelis, A., Czado, C., Joe, H., 2012. Pair copula constructions for multivariate discrete data. *Journal of the American Statistical Association* 107, 1063–1072.
- R Development Core Team, 2012. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, ISBN: 3-900051-08-9, <http://www.R-project.org>.
- Rosenblatt, M., 1952. Remarks on a multivariate transformation. *Annals of Mathematical Statistics* 23 (3), 470–472.
- Schwarz, G., 1978. Estimating the dimension of a model. *The Annals of Statistics* 6 (2), 461–464.
- Schölzel, C., Friederichs, P., 2008. Multivariate non-normally distributed random variables in climate research — introduction to the copula approach. *Nonlinear Processes in Geophysics* 15 (5), 761–772.
- Shih, J.H., Louis, T.A., 1995. Inference on the association parameter in copula models for bivariate survival data. *Biometrics* 51, 1384–1399.
- Sklar, A., 1959. Fonctions de répartition à *n* dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* 8, 229–231.
- Smith, M., Min, A., Almeida, C., Czado, C., 2010. Modelling longitudinal data using a pair-copula decomposition of serial dependence. *Journal of the American Statistical Association* 105, 1467–1479.
- Touboul, J., 2011. Goodness-of-Fit Tests For Elliptical and Independent Copulas through Projection Pursuit. *Algorithms* 4 (2), 87–114.
- Vrac, M., Chédin, A., Diday, E., 2005. Clustering a Global Field of Atmospheric Profiles by Mixture Decomposition of Copulas. *Journal of Atmospheric and Oceanic Technology* 22 (10), 1445–1459.
- Wu, C.F.J., 1983. On the convergence properties of the EM algorithm. *The Annals of Statistics* 11 (1), 95–103.