# Models for construction of multivariate dependence: A comparison study

Daniel Berg

*University of Oslo & The Norwegian Computing Center.*
*E-mail: daniel@danielberg.no*

Kjersti Aas

*The Norwegian Computing Center.*
*E-mail: Kjersti.Aas@nr.no*
*(Address for correspondence)*

[Revised October 19th 2007]

**Abstract**. We review models for construction of higher-dimensional dependence that have arisen recent years. A multivariate data set, which exhibit complex patterns of dependence, particularly in the tails, can be modelled using a cascade of lower-dimensional copulae. We examine two such models that differ in their construction of the dependency structure, namely the nested Archimedean constructions and the pair-copula constructions (also referred to as vines). The constructions are compared, and estimation- and simulation techniques are examined. The fit of the two constructions is tested on two different four-dimensional data sets; precipitation values and equity returns, using state of the art copula goodness-of-fit procedures. The nested Archimedean construction is strongly rejected for both our data sets, while the pair-copula construction provides a much better fit. Through VaR calculations, we show that the latter does not overfit data, but works very well even out-of-sample.

*Keywords*: Nested Archimedean copulas, Pair-copula decompositions, Equity returns, Precipitation values, Goodness-of-fit, Out-of-sample validation

## 1. Introduction

A copula is a multivariate distribution function with standard uniform marginal distributions. While the literature on copulae is substantial, most of the research is still limited to the bivariate case. Building higher-dimensional copulae is a natural next step, however, this is not an easy task. Apart from the multivariate Gaussian and Student copulae, the set of higher-dimensional copulae proposed in the literature is rather limited.

The Archimedean copula family (see e.g. Joe (1997) for a review) is a class that has attracted particular interest due to numerous properties which make them simple to analyse. The most common multivariate extension, the exchangeable multivariate Archimedean copula (EAC), is extremely restrictive, allowing the specification of only one generator, regardless of dimension. There have been some attempts at constructing more flexible multivariate Archimedean copula extensions, see e.g. Joe (1997); Nelsen (1999); Embrechts et al. (2003); Whelan (2004); Morillas (2005); Savu and Trede (2006). In this paper we discuss one group of such extensions, the nested Archimedean constructions (NACs). For the $d$-dimensional case, all NACs allow for the modelling of up to $d-1$ bivariate Archimedean copulae.

For a $d$-dimensional problem there are in general $d(d-1)/2$ parings of variables. Hence, while the NACs constitute a huge improvement compared to the EAC, they are still not rich enough to model all possible mutual dependencies amongst the $d$ variates. An even more flexible structure, here denoted the pair-copula construction (PCC) allows for the free specification of $d(d-1)/2$ copulae. This structure was originally proposed by Joe (1996), and later discussed in detail by Bedford and Cooke (2001, 2002), Kurowicka and Cooke (2006) (simulation) and Aas et al. (2007) (inference). Similar to the NACs, the PCC is hierarchical in nature. The modelling scheme is based on a decomposition of a multivariate density into a cascade of bivariate copulae. In contrast to the NACs, the PCC is not restricted to Archimedean copulae. The bivariate copulae may be from any family and several families may well be mixed in one PCC.

This paper has several contributions. In Section 2 we compare the two ways of constructing higher dimensional dependency structures, the NACs and the PCCs. We examine properties and estimation- and

simulation techniques, focusing on the relative strengths and weaknesses of the different constructions. In Section 3 we apply the NAC and the PCC to two four-dimensional data sets; precipitation values and equity returns. We examine the goodness-of-fit and validate the PCC out-of-sample with respect to one day value at risk (VaR) for the equity portfolio. Finally, Section 4 provides some summarizing comments and conclusions.

## 2. Constructions of higher dimensional dependence

### 2.1. The nested Archimedean constructions (NACs)

The most common multivariate Archimedean copula, the exchangeable Archimedean copula (EAC), is extremely restrictive, allowing the specification of only one generator, regardless of dimension. Hence, all $k$-dimensional marginal distributions ($k < d$) are identical. For several applications, one would like to have multivariate copulae which allows for more flexibility. In this section, we review one group of such extensions, the nested Archimedean constructions (NACs). We first review two simple special cases, the FNAC and the PNAC, in Sections 2.1.2 and 2.1.3, and then we turn to the general case in Section 2.1.4. However, before reviewing NACs, we give a short description of the EAC in Section 2.1.1, since this structure serves as a baseline.

#### 2.1.1. The exchangeable multivariate Archimedean copula (EAC)

The most common way of defining a multivariate Archimedean copula is the EAC, defined as

$$C(u_1, u_2, \ldots, u_d) = \varphi^{-1}\left\{\varphi(u_1) + \ldots + \varphi(u_d)\right\}, \tag{1}$$

where the function $\varphi$ is a decreasing function known as the generator of the copula and $\phi^{-1}$ denotes its inverse (see e.g. Nelsen (1999)). Here, we assume that the generator has only one parameter, $\theta$. There are however cases in which the generator have more parameters, see e.g. Joe (1997). For $C(u_1, u_2, \ldots, u_d)$ to be a valid $d$-dimensional Archimedean copula, $\phi^{-1}$ should have an analytical property known as d-monotonicity. Se McNeil and Neslehova (2007) for details. One usually also assumes that $\phi(0) = \infty$, i.e. that the Archimedean copula is strict. Consider for example the popular Gumbel (Gumbel, 1960) and Clayton (Clayton, 1978) copulae. The generator functions for these two copulae are given by $\varphi(t) = (-\log(t))^\theta$ and $\varphi(t) = (t^{-\theta} - 1)/\theta$, respectively.

#### 2.1.2. The fully nested Archimedean construction (FNAC)

A simple generalization of (1) can be found in Joe (1997) and is also discussed in Embrechts et al. (2003), Whelan (2004), Savu and Trede (2006) and McNeil (2007). The structure, which is shown in Figure 1 for the four-dimensional case, is quite simple, but notationally cumbersome. As seen from the figure, one simply adds a dimension step by step. The nodes $u_1$ and $u_2$ are coupled through copula $C_{11}$, node $u_3$ is coupled with $C_{11}(u_1, u_2)$ through copula $C_{21}$, and finally node $u_4$ is coupled with $C_{21}(u_3, C_{11}(u_1, u_2))$ through copula $C_{31}$. Hence, the copula for the 4-dimensional case requires three bivariate copulae $C_{11}$, $C_{21}$, and $C_{31}$, with corresponding generators $\varphi_{11}, \varphi_{21}$, and $\varphi_{31}$:

$$\begin{aligned} C(u_1, u_2, u_3, u_4) &= C_{31}(u_4, C_{21}(u_3, C_{11}(u_1, u_2))) \\ &= \varphi_{31}^{-1}\left\{\varphi_{31}(u_4) + \varphi_{31}(\varphi_{21}^{-1}\{\varphi_{21}(u_3) + \varphi_{21}(\varphi_{11}^{-1}\{\varphi_{11}(u_1) + \varphi_{11}(u_2)\})\})\right\}. \end{aligned}$$

For the $d$-dimensional case, the corresponding expression becomes

$$\begin{aligned} C(u_1, \ldots, u_d) \quad = \quad & \varphi_{d-1,1}^{-1}\{\varphi_{d-1,1}(u_d) + \varphi_{d-1,1} \circ \varphi_{d-2,1}^{-1}\{\varphi_{d-2,1}(u_{d-1}) + \varphi_{d-2,1} \\ \circ \ldots \circ \quad & \varphi_{11}^{-1}\{\varphi_{11}(u_1) + \varphi_{11}(u_2)\}\}\}. \end{aligned} \tag{2}$$

In this structure, which Whelan (2004) refers to as fully nested, all bivariate margins are themselves Archimedean copulae. It allows for the free specification of $d-1$ copulae and corresponding distributional parameters, while the remaining $(d-1)(d-2)/2$ copulae and parameters are implicitly given through the construction. More specifically, in Figure 1, the two pairs $(u_1, u_3)$ and $(u_2, u_3)$ both have copula $C_{21}$ with dependence parameter $\theta_{21}$. Moreover, the three pairs $(u_1, u_4)$, $(u_2, u_4)$ and $(u_3, u_4)$ all have copula $C_{31}$ with dependence parameter $\theta_{31}$. Hence, when adding variable $k$ to the structure, we specify the relationships between $k$ pairs of variables.

The FNAC is a construction of partial exchangeability and there are some technical conditions that need to be satisfied for (2) to be a proper $d$-dimensional copula. The consequence of these conditions for the FNAC is that the degree of dependence, as expressed by the copula parameter, must decrease with the level of nesting, i.e. $\theta_{11} \geq \theta_{21} \geq \ldots \geq \theta_{d-1,1}$, in order for the resulting $d$-dimensional distribution to be a proper copula.

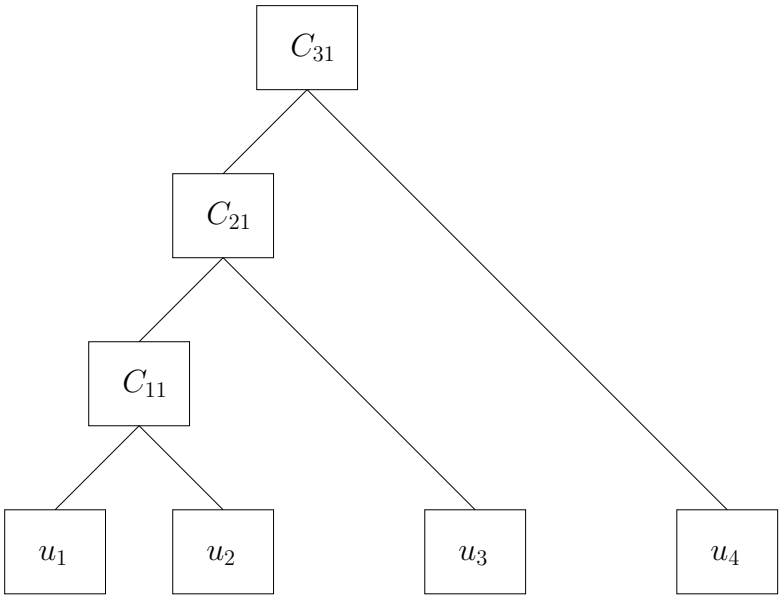


**Figure 1.** Fully nested Archimedean construction.

*2.1.3. The partially nested Archimedean construction (PNAC)*
An alternative multivariate extension is the PNAC. This structure was originally proposed by Joe (1997) and is also discussed in Whelan (2004), McNeil et al. (2006) (where it is denoted partially exchangeable) and McNeil (2007).

The lowest dimension for which there is a distinct structure of this class is four, when we have the following copula:

$$\begin{aligned} C(u_1,u_2,u_3,u_4) &= C_{21}(C_{11}(u_1,u_2),C_{21}(u_3,u_4)) \\ &= \varphi_{21}^{-1}\{\varphi_{21}(\varphi_{11}^{-1}\{\varphi_{11}(u_1)+\varphi_{11}(u_2)\})+\varphi_{21}(\varphi_{12}^{-1}\{\varphi_{12}(u_3)+\varphi_{12}(u_4)\})\}. \end{aligned} \tag{3}$$

Figure 2 illustrates this structure graphically. Again the construction is notationally cumbersome although the logic is straightforward. We first couple the two pairs $(u_1,u_2)$ and $(u_3,u_4)$ with copulae $C_{11}$ and $C_{12}$, having generator functions $\varphi_{11}$ and $\varphi_{12}$, respectively. We then couple these two copulae using a third copula $C_{21}$. The resulting copula is exchangeable between $u_1$ and $u_2$ and also between $u_3$ and $u_4$. Hence, it can be understood as a composite of the EAC and the FNAC.

For the PNAC, as for the FNAC, $d-1$ copulae and corresponding distributional parameters are freely specified, while the remaining copulae and parameters are implicitly given through the construction. More specifically, in Figure 2, the four pairs $(u_1,u_3)$, $(u_1,u_4)$ $(u_2,u_3)$ and $(u_2,u_4)$ will all have copula $C_{21}$, with dependence parameter $\theta_{21}$. Similar constraints on the parameters are required for the PNACs as for the FNACs.

*2.1.4. The general case*
The generally nested Arcimedean construction was originally treated by Joe (1997, Chapter 4.2), and is also mentioned in Whelan (2004). However, Savu and Trede (2006) were the first to provide the notation for arbitrary nesting, and to show how to calculate the $d$-dimensional density in general.

Savu and Trede (2006) use the notation hierarchical Archimedean copula for the generally nested case. The idea is to build a hierarchy of Archimedean copulas. Assume that there are $L$ levels. At each level

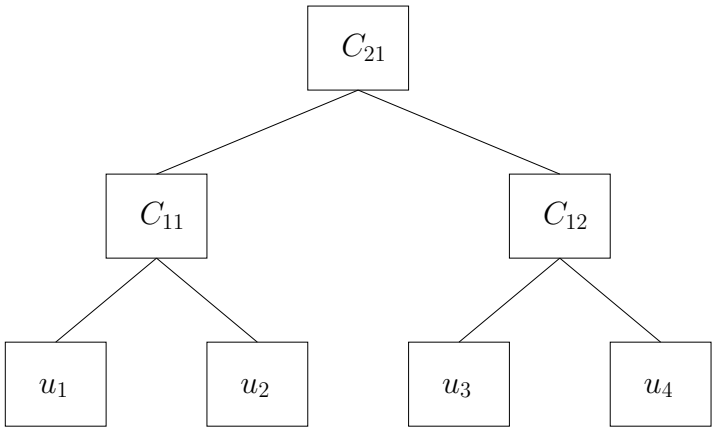


**Figure 2.** Partially nested Archimedean construction.

$l$, there are $n_l$ distinct objects (an object is either a copula or a variable). At level $l = 1$ the variables $u_1, \ldots, u_d$ are grouped into $n_1$ exchangeable multivariate Archimedean copulae. These copulae are in turn coupled into $n_2$ copulae at level $l = 2$, and so on.

Figure 3 shows an example. The 9-dimensional copula in the figure is given by

$$C(u_1, \ldots u_9) = C_{41}(C_{31}(C_{21}(C_{11}(u_1, u_2), u_3, u_4), u_5, u_6), C_{22}(u_7, C_{12}(u_8, u_9))).$$

At level one, there are two copulae. Both are two-dimensional EACs. The first, $C_{11}$ joins the variables $u_1$ and $u_2$, while the other, $C_{12}$, joins $u_8$ and $u_9$. At the second level, there are also two copulae. The first, $C_{21}$, joins the copula $C_{11}$ with the two variables $u_3$ and $u_4$, while the other, $C_{22}$ joins $C_{12}$ and $u_7$. At the third level there is only one copula, $C_{31}$, joining $C_{21}$, $u_5$ and $u_6$. Finally, at level four, the copula $C_{41}$ joins the two copulae $C_{31}$ and $C_{22}$.

There are a number of conditions to ensure that the resulting structure is a valid Archimedean copula (Savu and Trede, 2006). The number of copulas must decrease at each level, the top level may only contain one copula and all the inverse generator functions must be completely monotone. Further, we must have that the degree of dependence must decrease with the level of nesting. For example in Figure 3 we must have that $\theta_{11} \geq \theta_{21} \geq \theta_{31} \geq \theta_{41}$ and $\theta_{12} \geq \theta_{22} \geq \theta_{41}$. If mixing copula generators belonging to different Archimedean families, even this requirement might not be sufficient. Two Archimedean copulas from families 1 and 2 can only be nested if the derivative of the product $\varphi_1 \circ \varphi_2^{-1}$ is completely monotonic (McNeil, 2007). The issue of which copula families that can be mixed has been considered in some detail in Joe (1997), but it is still not fully explored. Hence, here we only work with structures for which all the generators are from the same family.

Unfortunately, it is not possible to obtain a simple expression for the density of a hierarchical Archimedean copula. Due to the complex structure of this construction, one has to use a recursive approach. One differentiates the $d$-dimensional top level copula with respect to its arguments using the chain rule. See Savu and Trede (2006) for more details.

### *2.1.5. Parameter estimation*

For the NACs, as for the EAC, the parameters may be estimated by maximum likelihood. However, not even for the EAC it is straightforward to derive the density in general for all parametric families. For instance, for the Gumbel family, one has to resort to a computer algebra system, such as *Mathematica*, or the function `D` in *R*, to derive the $d$-dimensional density.

Savu and Trede (2006) give the density expression for a general NAC. The density is obtained using a recursive approach. Hence, the number of computational steps needed to evaluate the density increases rapidly with the complexity of the copula, and parameter estimation becomes very time consuming in high dimensions.

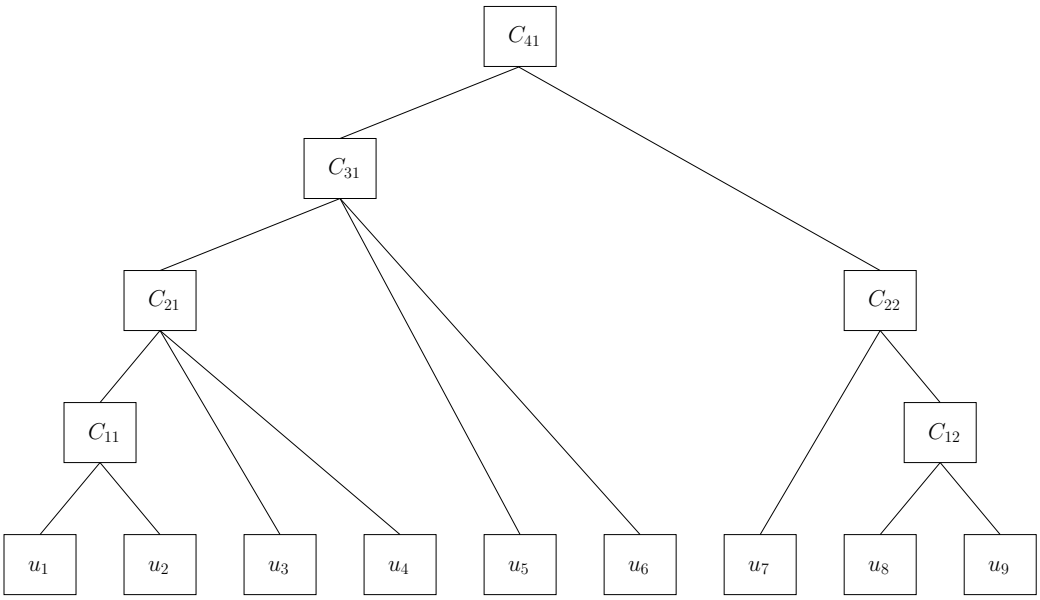


**Figure 3.** Hierarchically nested Archimedean construction.

*2.1.6. Simulation*

Simulating from the higher-dimensional constructions is a very important and central practical task. Simulating from an EAC is usually rather simple, and several algorithms exist. A popular algorithm utilizes the representation of the Archimedean copula generator using Laplace transforms (see e.g. Frees and Valdez (1998)). McNeil (2007) shows how to use the Laplace-transform method also for the NACs, in the case where all generators are taken from either the Gumbel- or the Clayton family. A problem with the Laplace transform method is that it is limited to copulae for which we can find a distribution that equals the Laplace transform of the inverse generator function, and from which we can easily sample. For some copulae, e.g. Frank, there is, as of now, no alternative to the conditional inversion method described in e.g. Embrechts et al. (2003). This procedure involves the $d-1$ first derivatives of the copula function and, in most cases, numerical inversion. The higher-order derivatives are usually extremely complex expressions (see e.g. Savu and Trede (2006)). Hence, simulation may become very inefficient for high dimensions.

### *2.2. The pair-copula constructions (PCC)*

While the NACs constitute a large improvement compared to the EAC, they still only allow for the modelling of up to $d-1$ copulae. An even more flexible structure, the PCC, allows for the free specification of $d(d-1)/2$ copulae. This structure was orginally proposed by Joe (1996), and it has later been discussed in detail by Bedford and Cooke (2001, 2002), Kurowicka and Cooke (2006) (simulation) and Aas et al. (2007) (inference). Similar to the NAC's, the PCC's are hierarchical in nature. The modelling scheme is based on a decomposition of a multivariate density into $d(d-1)/2$ bivariate copula densities, of which the first $d-1$ are unconditional, and the rest are conditional.

While the NACs are defined through their distribution functions, the PCCs are usually represented in terms of the density. Two main types of PCCs have been proposed in the literature; canonical vines and D-vines (Kurowicka and Cooke, 2004). Here, we concentrate on the D-vine representation, for which the density is (Aas et al., 2007):

$$f(x_1,\dots x_d) = \prod_{k=1}^{d} f(x_k) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{j,i}\left\{F(x_i|x_{i+1},\dots,x_{i+j-1}), F(x_{i+j}|x_{i+1},\dots,x_{i+j-1})\right\}. \tag{4}$$

The conditional distribution functions are computed using (Joe, 1996)

$$F(x|\boldsymbol{v}) = \frac{\partial\, C_{x,v_j|\boldsymbol{v}_{-j}}\{F(x|\boldsymbol{v}_{-j}).F(v_j|\boldsymbol{v}_{-j})\}}{\partial F(v_j|\boldsymbol{v}_{-j})}, \tag{5}$$

where $C_{ij|\boldsymbol{k}}$ is a bivariate copula distribution function. To use this construction to represent a dependency structure through copulas, we assume that the univariate margins are uniform in [0,1]. One 4-dimensional case of (4) is

$$\begin{aligned} c(u_1,u_2,u_3,u_4) &= c_{11}(u_1,u_2)\cdot c_{12}(u_2,u_3)\cdot c_{13}(u_3,u_4) \\ &\cdot\ c_{21}(F(u_1|u_2),F(u_3|u_2))\cdot c_{22}(F(u_2|u_3),F(u_4|u_3)) \\ &\cdot\ c_{31}(F(u_1|u_2,u_3),F(u_4|u_2,u_3)), \end{aligned}$$

where $F(u_1|u_2) = \partial C_{11}(u_1,u_2)/\partial u_2$, $F(u_3|u_2) = \partial C_{12}(u_2,u_3)/\partial u_2$, $F(u_2|u_3) = \partial C_{12}(u_2,u_3)/\partial u_3$, $F(u_4|u_3) = \partial C_{13}(u_3,u_4)/\partial u_3$, $F(u_1|u_2,u_3) = \partial C_{21}(F(u_1|u_2),F(u_3|u_2))/\partial F(u_3|u_2)$ and $F(u_4|u_2,u_3) = \partial C_{22}(F(u_4|u_3),F(u_2|u_3))/\partial F(u_2|u_3)$. Figure 2.2 illustrates this structure.

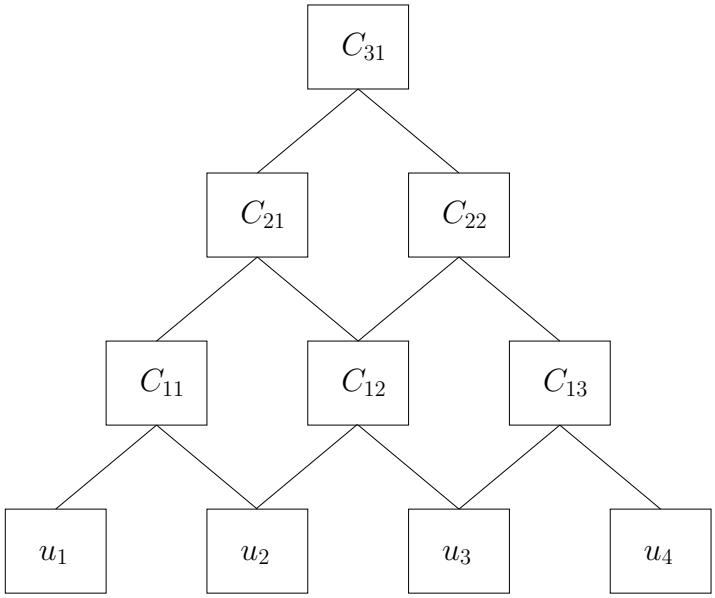


**Figure 4.** Pair-copula construction.

The copulae involved in (4) do not have to belong to the same family. In contrast to the NACs they do not even have to belong to the same class. The resulting multivariate distribution will be valid even if we choose, for each pair of variables, the parametric copula that best fits the data. As seen from (4) the PCC consists of $d(d-1)/2$ bivariate copulae of known parametric families, of which $d-1$ are copulae of pairs of the original variables, while the remaining $(d-1)(d-2)/2$ are copulae of pairs of variables constructed using (5) recursively. This means that in contrast to the NACs, the unspecified bivariate margins will belong to a known parametric family in general. However, it can be shown, that e.g. upper (lower) tail dependence on the bivariate copulae at the lowest level is a sufficient condition for all bivariate margins to have upper (lower) tail dependence†.

*2.2.1. Parameter estimation*

The parameters of the PCC may be estimated by maximum likelihood. In contrast to the NACs, the density is explicitly given. However, also for this construction, a recursive approach is used (see Aas et al. (2007, Algorithm 4)). Hence, the number of computational steps to evaluate the density increases rapidly with the complexity of the copula, and parameter estimation becomes time consuming in high dimensions.

*2.2.2. Simulation*

The simulation algorithm for a D-vine is straightforward and simple to implement, see Aas et al. (2007, Algorithm 2). Like for the NACs, the conditional inversion method is used. However, to determine each

†Personal communication with Harry Joe.

**Table 1.** Summary of construction properties for the EAC, NAC and PCC constructions.

| *Construction* | *Max no. of copulae freely specified* | *Parameter constraints* | *Copula mixing* |
|---|---|---|---|
| EAC | 1 | None | Only one copula |
| NAC | $d-1$ | Dependence must decrease with level of nesting | May combine different Arch. families but under complete monotonicity restrictions |
| PCC | $d(d-1)/2$ | None | May combine any copula families from any class |

**Table 2.** Computational times in sec. for different constructions and copulae, fitted to the equity data in Section 3.3.

| *Method* | *Likelihood evaluation* | *Estimation* | *Simulation* |
|---|---|---|---|
| *Gumbel* | | | |
| NAC | 0.32 | 34.39 | 0.02 |
| PCC | 0.04 | 5.09 | 7.56 |
| *Frank* | | | |
| NAC | 0.12 | 5.34 | 64.83 |
| PCC | 0.02 | 1.22 | 5.82 |

of the conditional distribution functions involved, only the first partial derivative of a bivariate copula needs to be computed (see Aas et al. (2007)). Hence, the simulation procedure for the PCC is in general much simpler and faster than for the NACs.

### *2.3. Comparison*

In this section we summarize the differences between the NACs and the PCCs with respect to ease of interpretation, applicability and computational complexity.

First, the main advantage of the PCCs is the increased flexibility compared to the NACs. While the NACs only allow for the free specification of $d-1$ copulae, $d(d-1)/2$ copulae may be specified in a PCC. Next, for the NACs there are restrictions on which Archimedean copulae that can be mixed, while the PCCs can be built using copulae from different families and classes. Finally, the NACs have another even more important restriction in that the degree of dependence must decrease with the level of nesting. When looking for appropriate data sets for the applications in Section 3, it turned out to be quite difficult to find real-world data sets satisfying this restriction. Hence, this feature of the NACs might prevent them from being extensively used in real-world applications. For the PCCs, on the other hand, one is always guaranteed that all parameter combinations are valid. Table 1 summarizes these properties.

It is our opinion that another advantage of the PCCs is that they are represented in terms of the density and hence easier to handle than the NACs that are defined through their distribution functions. The PCCs are also in general more computationally efficient than the NACs. Table 2 shows computational times (s) in *R* ‡ for likelihood evaluation, parameter estimation and simulation for different structures. The parameter estimation is done for the data set described in Section 3.3, and the simulation is performed using the parameters in Table 6 (based on 1000 samples). The values for NAC were computed using density expressions found in Savu and Trede (2006). However, general expressions may also easily be obtained symbolically using e.g. the function `D` in *R*. The estimation times in Table 2 are only indicative and included as examples since they are very dependent on size and structure of the data set. It is more appropriate to study the times needed to compute one evaluation of the likelihood given in the leftmost column. As can be seen from the table, the PCC is superior to the NAC for likelihood evaluation in both the Gumbel and the Frank case. Moreover, it is much faster for simulation in the Frank case, since one in this case must use the general conditional inversion algorithm with numerical inversion for the NAC. In the Gumbel case, however, one can perform much more efficient simulation from the NAC using the algorithms given in McNeil (2007). Hence, in this case, the NAC is superior to the PCC.

The multivariate distribution defined through a NAC will always by definition be an Archimedean copula (assuming that all requirements are satisfied), and all bivariate margins will belong to a known parametric family. This is not the case for the PCCs, for which neither the multivariate distribution nor

‡The experiments were run on a Intel(R) Pentium(R) 4 CPU 2.80GHz PC.

the unspecified bivariate margins will belong to a known parametric family in general. However, we do not view this as a problem, since both might easily be obtained by simulation.

Finally, it should be noted that for both structures, an important part of the full estimation problem is how to select the ordering of the variables. For smaller dimensions (say 3 and 4), one may estimate the parameters of all possible orderings and compare the resulting log-likelihoods. This is in practice infeasible for higher dimensions, since the number of different orderings increases very rapidly with the dimension of the data set. One may instead determine which bivariate relationships that are most important to model correctly and let this determine which ordering to choose. Very recently, there has been some attempts of formalising this procedure, both for the NACs (Okhrin et al., 2007) and for the PCCs (Min and Czado, 2007).

## 3. Applications

The fit of the NAC and the PCC is assessed for two different four-dimensional data sets; precipitation values and equity returns. Appropriate modelling of precipitation is of great importance to insurance companies which are exposed to growth in damages to buildings caused by external water exposition. Modelling precipitation and valuing related derivative contracts is also indeed a frontier in the field of weather derivatives, see e.g. Musshoff et al. (2006). The dependencies within an equity portfolio can have enormous impacts on e.g. capital allocation and the pricing of collateralized debt obligations. Before these two applications are further treated, we describe the tests used for goodness-of-fit in our study.

### 3.1. Goodness-of-fit

To evaluate whether a copula or copula construction appropriately fit the data at hand, goodness-of-fit testing is called upon. Lately, several procedures have been proposed, see e.g. Berg (2007) for an overview and power comparison. These power comparisons show that no procedure is always the best. However, the procedure that showed to have the overall best performance in the study referred to above, was one based on the empirical copula $C_n$ introduced by Deheuvels (1979),

$$C_n(\mathbf{u}) = \frac{1}{n+1}\sum_{j=1}^{n} \mathbf{1}(U_{j1} \leq u_1, \ldots, U_{jd} \leq u_d), \quad \mathbf{u} = (u_1, \ldots, u_d) \in (0,1)^d, \tag{6}$$

where $\mathbf{U}_j = (U_{j1}, \ldots, U_{jd})$ are the $U(0,1)^d$ pseudo-observations, defined as normalized ranks. This procedure is based on the process $\mathcal{C}_n = \sqrt{n}\{C_n - C_{\hat{\theta}_n}\}$ where $\hat{\theta}_n$ is some consistent estimator of $\theta$. Basing a goodness-of-fit procedure on $\mathcal{C}_n$ was originally proposed by Fermanian (2005), but there dismissed due to poor statistical properties. However, it has later been shown that it has the necessary asymptotic properties to be a justified goodness-of-fit procedure (Quessy, 2005; Genest and Rémillard, 2005). Moreover, Genest et al. (2007) and Berg (2007) have examined the power of $\mathcal{C}_n$ and concluded that it is a very powerful procedure in most cases.

We use the Cramér-von Mises statistic, defined by:

$$S_n = n\int_{[0,1]^d} \{C_n(\mathbf{u}) - C_{\theta_n}(\mathbf{u})\}^2 \, \mathrm{d}C_n(\mathbf{u}) = \sum_{j=1}^{n} \{C_n(\mathbf{U}_j) - C_{\theta_n}(\mathbf{U}_j)\}^2 . \tag{7}$$

Large values of $S_n$ means a poor fit and leads to the rejection of the null hypothesis copula. In practice, the limiting distribution of $S_n$ depends on $\theta$. Hence, approximate $p$-values for the test must be obtained through a parametric bootstrap procedure. We adopt the procedure in Appendix A in Genest et al. (2007), setting the bootstrap parameters $m$ and $N$ to 5000 and 1000, respectively. The validity of this bootstrap procedure was established in Genest and Rémillard (2005).

It will be shown in Section 3.2 that for the precipitation data set, $S_n$ leads to the rejection of all the different NAC and PCC structures that are investigated. Hence, to be able to compare the two structures for this data set, we also use another goodness-of-fit procedure based on the process $\mathcal{K}_n = \sqrt{n}\{K_n - K_{\hat{\theta}_n}\}$, where

$$K_n(t) = \frac{1}{n+1}\sum_{j=1}^{n} \mathbf{1}(C_n(\mathbf{U}_j) \leq t),$$

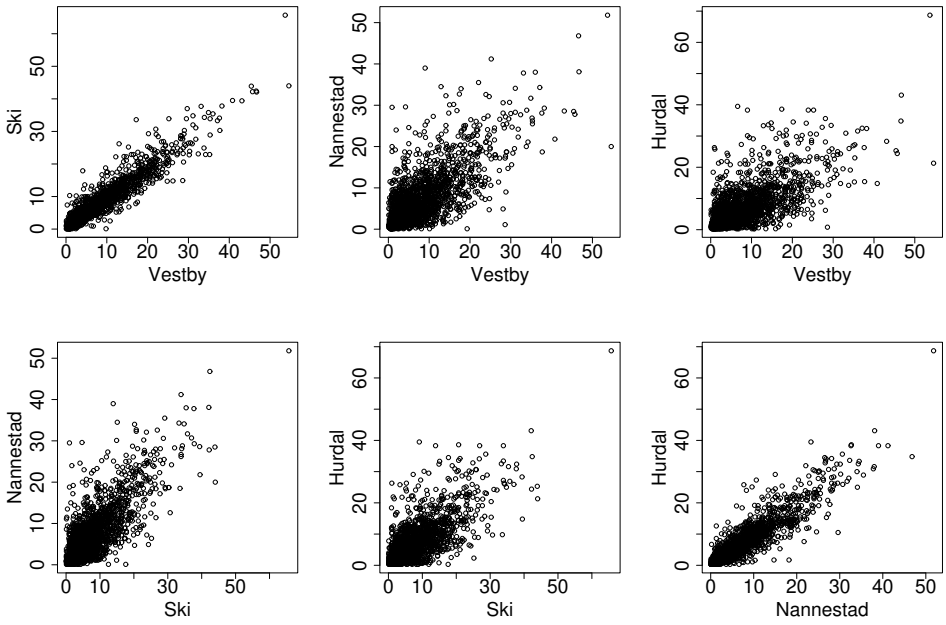


**Figure 5.** Daily precipitation (mm) for pairs of meteorological stations for the period 01.01.1990 to 31.12.2006, zeros removed.

is the empirical distribution function of $C_n(\mathbf{u})$. See Genest et al. (2006) for details. Also for this procedure we use the Cramér-von Mises statistic, i.e.:

$$T_n = \int_{[0,1]^d} \left\{K_n(\mathbf{u}) - K_{\hat{\theta}}(\mathbf{u})\right\}^2 \mathrm{d}K_n(\mathbf{u}) = \sum_{j=1}^{n} \left\{K_n(\mathbf{U}_j) - K_{\hat{\theta}_n}(\mathbf{U}_j)\right\}^2, \tag{8}$$

and parametric bootstrap to obtain the $p$-values.

For both procedures, we use a 5% significance level for all experiments in this section.

### *3.2. Application 1: Precipitation data*

In this section we study daily precipitation data (mm) for the period 01.01.1990 to 31.12.2006 for 4 meteorological stations in Norway; Vestby, Ski, Nannestad and Hurdal, obtained from the Norwegian Meteorological Institute. According to Musshoff et al. (2006), the stochastic process of daily precipitation can be decomposed into a stochastic process of "rainfall"/"no rainfall", and a distribution for the amount of precipitation given that it rains. Here, we are only interested in the latter. Hence, before further processing, we remove days with non-zero precipitation values for at least one station, resulting in 2065 observations for each variable. Figures 5-6 show the daily precipitation values and corresponding copulae for pairs of meteorological stations. Since we are mainly interested in estimating the dependence structure of the stations, the precipitation vectors are converted to uniform pseudo-observations before further modelling. In light of recent results due to Chen and Fan (2006), the method of maximum pseudo-likelihood is consistent even when time series models are fitted to the margins.

Based on visual inspection and preliminary goodness-of-fit tests for bivariate pairs (the copulae taken into consideration were the Student, Clayton, survival Clayton, Gumbel and Frank copulae), we decided to examine Gumbel and Frank NACs and Gumbel, Frank and Student PCCs for the precipitation data.

#### *3.2.1. Hierarchically nested Archimedean construction*

The most appropriate ordering of the variates in the decomposition is found by comparing Kendall's tau values for all bivariate pairs. These are shown in Table 3. They confirm the intuition that the degree of dependence between the variables corresponds to the distances between the stations. Ski and Vestby are closely located, and so is Hurdal and Nannestad, while the distance from Ski/Vestby to Hurdal/Nannestad

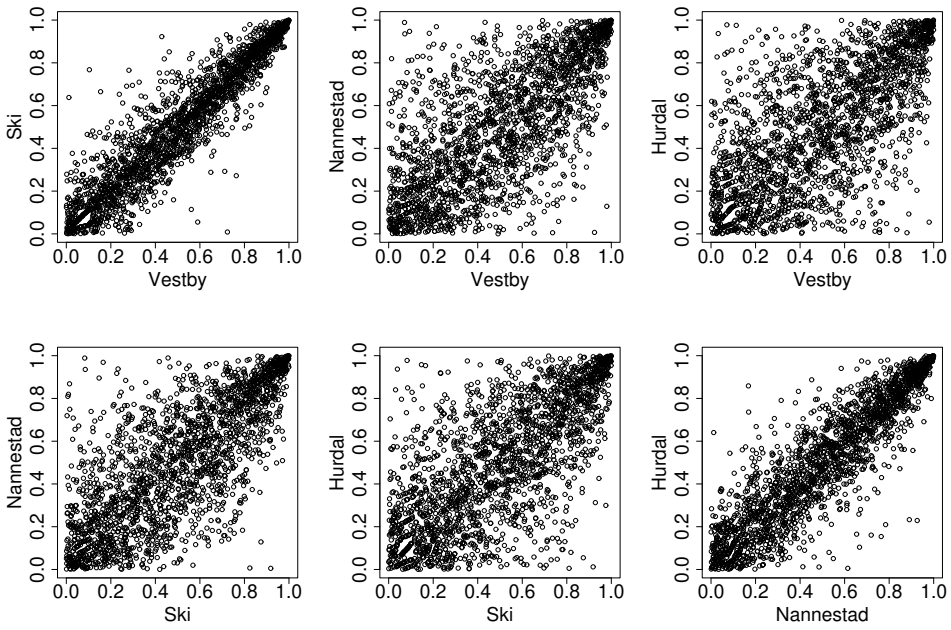


**Figure 6.** Pseudo-observations corresponding to Figure 5.

**Table 3.** Estimated Kendall's tau for pairs of variables.

| *Location* | Ski | Nannestad | Hurdal |
|---|---|---|---|
| Vestby | 0.79 | 0.49 | 0.47 |
| Ski | | 0.56 | 0.53 |
| Nannestad | | | 0.71 |

is larger. Hence, we choose $C_{11}$ and $C_{12}$ to be the copulae of Vestby and Ski and Nannestad and Hurdal, respectively, while $C_{21}$ is the copula of the remaining pairs.

The two leftmost columns of Table 4 show the estimated parameter values, resulting log-likelihoods, and estimated $p$-values for the Gumbel and Frank NACs, fitted to the precipitation data. We see that both goodness-of-fit procedures strongly reject the two NAC constructions. Hence we conclude that the NACs considered are not flexible enough to fit the precipitation data appropriately.

### *3.2.2. Pair-copula construction*

Also for the PCCs, the variables are ordered such that the copulae fitted at level 1 in the decomposition are those corresponding to the three largest Kendall's tau values. Hence, $C_{11}$ is the copula of Vestby and Ski, $C_{12}$ is the copula of Ski and Nannestad, and $C_{13}$ is the copula of Ski and Hurdal. The parameters of the PCC are estimated using Algorithm 4 in Aas et al. (2007). The three rightmost columns of Table 4 show the estimated parameter values, resulting log-likelihoods and $p$-values for the Gumbel, Frank and Student PCCs. We see that, as for the NACs, all considered PCCs are strongly rejected by $S_n$. Hence, from the $S_n$-results, it is not possible to determine which of the two constructions that best fit the precipitation data and we therefore also used $T_n$. This procedure also rejects both NACs, but it fails to reject the Gumbel and Student PCCs, with the Gumbel PCC seemingly the best. Hence, we conclude that the Gumbel PCCs provides the best fit, but that there is a need for further research to find copula types that better captures the properties of the precipitation data.

## *3.3. Application 2: Equity returns*

In this section, we study an equity portfolio. The portfolio is comprised of four time series of daily log-return data from the period 14.08.2003 to 29.12.2006 (852 observations for each firm). The data set was downloaded from `http://finance.yahoo.com`. The firms are British Petroleum (BP), Exxon Mobile

**Table 4.** Estimated parameters, log-likelihood and estimated $p$-values for NACs and PCCs fitted to the precipitation data.

| *Parameter* | *NAC* | | *PCC* | | |
|---|---|---|---|---|---|
| | *Gumbel* | *Frank* | *Gumbel* | *Frank* | *Student* |
| $\theta_{11} \setminus \nu_{11}$ | 4.32 | 16.69 | 4.34 | 16.78 | $0.93 \setminus 3.6$ |
| $\theta_{12} \setminus \nu_{12}$ | 3.45 | 13.01 | 2.24 | 7.10 | $0.78 \setminus 6.7$ |
| $\theta_{13} \setminus \nu_{13}$ | - | - | 3.45 | 12.98 | $0.90 \setminus 5.5$ |
| $\theta_{21} \setminus \nu_{21}$ | 1.97 | 5.96 | 1.01 | 0.08 | $0.01 \setminus 9.6$ |
| $\theta_{22} \setminus \nu_{22}$ | - | - | 1.02 | 0.61 | $0.09 \setminus 14.5$ |
| $\theta_{31} \setminus \nu_{31}$ | - | - | 1.03 | 0.27 | $0.04 \setminus 17.3$ |
| Log-likelihood | 4741.05 | 4561.72 | 4842.25 | 4632.19 | 4643.38 |
| $p$-value of $S_n$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| $p$-value of $T_n$ | 0.002 | 0.000 | 0.089 | 0.013 | 0.070 |

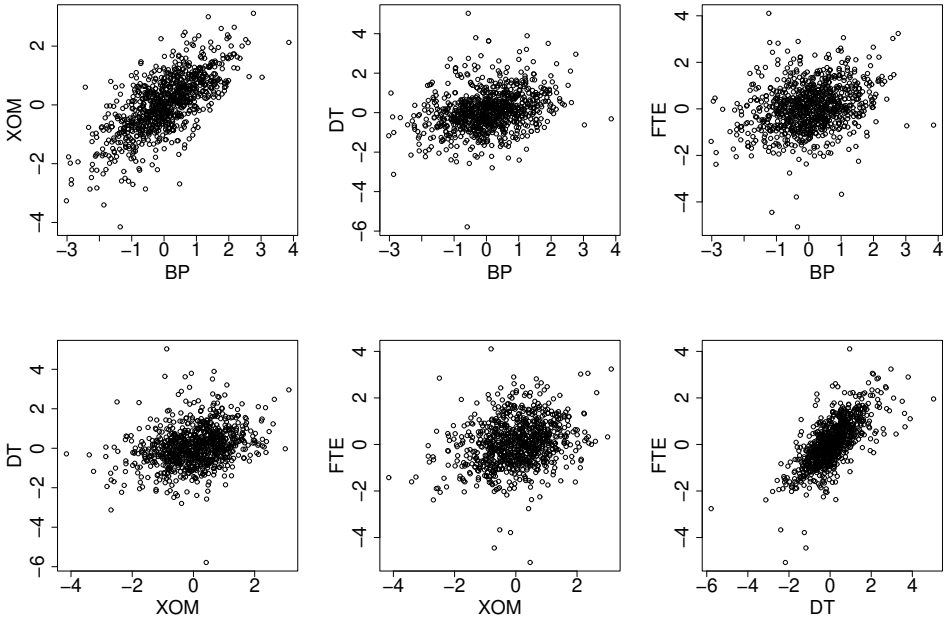


**Figure 7.** GARCH-filtered daily log-returns for our four stocks for the period from 14.08.2003 to 29.12.2006.

Corp (XOM), Deutsche Telekom AG (DT) and France Telecom (FTE). Financial log-returns are usually not independent over time. Hence, the original vectors of log-returns are processed by a GARCH filter before further modelling. We use the GARCH(1,1)-model (Bollerslev, 1986):

$$\begin{aligned} r_t &= c + \sigma_t \, z_t \\ \mathrm{E}[z_t] = 0 \text{ and } \mathrm{Var}[z_t] &= 1 \\ \sigma_t^2 &= a_0 + a \, \epsilon_{t-1}^2 + b \, \sigma_{t-1}^2. \end{aligned} \tag{9}$$

It is well recognised that GARCH models, coupled with the assumption of conditionally normally distributed errors are unable to fully account for the tails of the distributions of daily returns. Hence, we follow Venter and de Jongh (2002) and use the Normal Inverse Gaussian (NIG) distribution (Barndorff-Nielsen, 1997) as the conditional distribution. In a study performed by Venter and de Jongh (2004) the NIG distribution outperforms a skewed Student's t-distribution and a non-parametric kernel approximation as the conditional distribution of a one-dimensional GARCH process. After filtering the original returns with (9) (estimated parameter values are shown in Appendix A), the standardised residual vectors are converted to uniform pseudo-observations. Figures 6-8 show the filtered daily log-returns and pseudo-observations for each pair of assets.

Based on visual inspection and preliminary goodness-of-fit tests for bivariate pairs (like for the precipitation data, the copulae taken into consideration were the Student, Clayton, survival Clayton, Gumbel

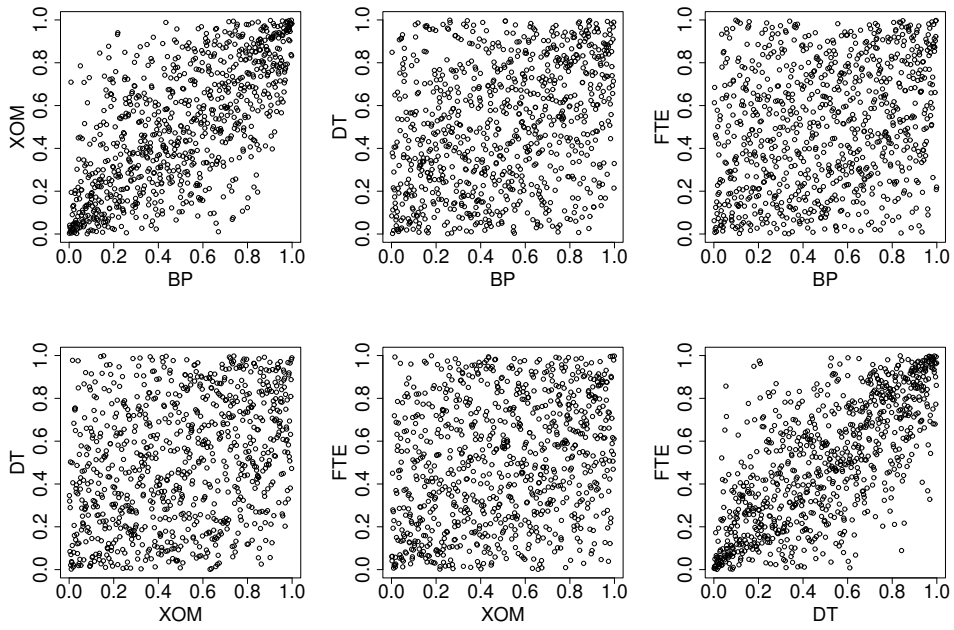


**Figure 8.** Pseudo-observations corresponding to Figure 7.

**Table 5.** Estimated Kendall's tau for pairs of variables for our four stocks.

| *Firm* | XOM | DT | FTE |
|---|---|---|---|
| BP | 0.45 | 0.19 | 0.20 |
| XOM | | 0.23 | 0.17 |
| DT | | | 0.48 |

and Frank copulae), we decided to examine a Frank NAC and Frank and Student PCC's for this data set.

### *3.3.1. Hierarchically nested Archimedean construction*

Also for this data set, the most appropriate ordering of the variates in the decomposition is found by comparing Kendall's tau values for all bivariate pairs. The Kendall's tau values are shown in Table 5. As expected, stocks within one industrial sector are more dependent than stocks from different sectors. Hence, we choose $C_{11}$ as the copula of $BP$ and $XOM$, $C_{12}$ as the copula of $DT$ and $FTE$, and $C_{21}$ as the copula of the remaining pairs. The leftmost column of Table 6 shows the estimated parameter values, resulting log-likelihood and $p$-value for the Frank HNAC. Even though this structure is not rejected by $T_n$, the strong rejection by $S_n$ suggests that the fit is not very good. Hence, we conclude that the Frank NAC is not able to appropriately fit the equity data.

### *3.3.2. Pair-copula construction*

Again, the most appropriate ordering of the variates in the decomposition is determined by the size of the Kendall's tau values. Hence, we choose $C_{11}$ as the copula of $BP$ and $XOM$, $C_{12}$ as the copula of $XOM$ and $DT$, and $C_{13}$ as the copula of $DT$ and $FTE$. The parameters of the PCC are estimated by maximum likelihood, see Algorithm 4 in Aas et al. (2007). The two rightmost columns of Table 6 shows the estimated parameter values, resulting log-likelihood and estimated $p$-values for the Frank and Student PCCs. We see that the Frank PCC is rejected by $S_n$. Moreover, the p-value of $T_n$ is equal to the one for the Frank NAC. The Student PCC, on the other hand, provides a very good fit and is not even rejected by $S_n$. Hence, we conclude that it fits the equity data very well.

**Table 6.** Estimated parameters, log-likelihood and estimated $p$-values for NAC and PCCs fitted to the filtered equity data.

| *Parameter* | *NAC* | *PCC* | |
|---|---|---|---|
| | *Frank* | *Frank* | *Student* |
| $\theta_{11} \setminus \nu_{11}$ | 5.57 | 5.56 | $0.70 \setminus 13.8$ |
| $\theta_{12} \setminus \nu_{12}$ | 6.34 | 1.89 | $0.32 \setminus 134.5$ |
| $\theta_{13} \setminus \nu_{13}$ | - | 6.32 | $0.73 \setminus 6.4$ |
| $\theta_{21} \setminus \nu_{21}$ | 1.78 | 0.91 | $0.14 \setminus 12.0$ |
| $\theta_{22} \setminus \nu_{22}$ | - | 0.30 | $0.06 \setminus 20.6$ |
| $\theta_{31} \setminus \nu_{31}$ | - | 0.33 | $0.07 \setminus 17.8$ |
| Log-likelihood | 616.45 | 618.63 | 668.49 |
| $p$-value of $S_n$ | 0.006 | 0.008 | 0.410 |
| $p$-value of $T_n$ | 0.385 | 0.385 | 0.697 |

### *3.4. Validation*

With the increasing complexity of models there is always the risk of overfitting the data. To examine whether this is the case for the PCC, we validate it out-of-sample for the equity portfolio. More specifically, we use the GARCH-NIG-Student PCC described in Section 3.3.2 to determine the risk of the return distribution for an equally weighted portfolio of BP, XOM, DT, and FTE over a one-day horizon. The equally-weighted portfolio is only meant as an example. In practice, the weights will fluctuate unless the portfolio is rebalanced every day.

The model estimated from the period 14.03.2003 to 29.12.2006 is used to forecast 1-day VaR at different significance levels for each day in the period from 30.12.2006 to 11.06.2007 (110 days). The test procedure is as follows: For each day $t$ in the test set:

(a) For each variable $j = 1, \ldots, 4$, compute the one-step ahead forecast of $\sigma_{j,t}$, given information up to time $t$.
(b) For each simulation $n = 1, \ldots, 10,000$
   - Generate a sample $u_1, \ldots u_4$ from the estimated Student PCC.
   - Convert $u_1, \ldots u_4$ to NIG(0,1)-distributed samples $z_1, \ldots, z_4$ using the inverses of the corresponding NIG distribution functions.
   - For each variable $j = 1, \ldots, 4$, determine the log-return $r_{j,t} = c_{j,t} + \sigma_{j,t}\, z_j$. (Here $c_{j,t}$ is computed as the mean of the last 100 observed log-returns.)
   - Compute the return of the portfolio as $r_{p,t} = \sum_{j=1}^{4} \frac{1}{4} r_{j,t}$.
(c) For significance levels $q \in \{0.005, 0.01, 0.05\}$
   - Compute the 1-day $\mathrm{VaR}_t^q$ as the $q$th-quantile of the distribution of $r_{p,t}$.
   - If $\mathrm{VaR}_t^q$ is greater than the observed value of $r_{p,t}$ this day, a violation is said to occur.

Figure 9 shows the actual log-returns for the portfolio in the period 30.12.2006 to 11.06.2007 and the corresponding VaR levels obtained from the procedure described above. Further, the two upper rows of Table 7 gives the number of violations $x$, of VaR for each significance level and with the expected values, respectively. To test the significance of the differences between the observed and the expected values, we use the likelihood ratio statistic by Kupiec (1995). The null hypothesis is that the expected proportion of violations is equal to $\alpha$. Under the null hypothesis, the likelihood ratio statistic given by

$$2\ln\left(\left(\frac{x}{N}\right)^x \left(1 - \frac{x}{N}\right)^{N-x}\right) - 2\ln\left(\alpha^x (1-\alpha)^{N-x}\right),$$

where $N$ is the length of the sample, is asymptotically distributed as $\chi^2(1)$. We have computed p-values of the null hypothesis for each quantile. The results are shown in the lower row of Table 7. If we use a 5% level for the Kupiec LR statistic, the null hypothesis is not rejected for any of the three quantiles. Hence, the GARCH-NIG-Student PCC seems to work very well out-of-sample.

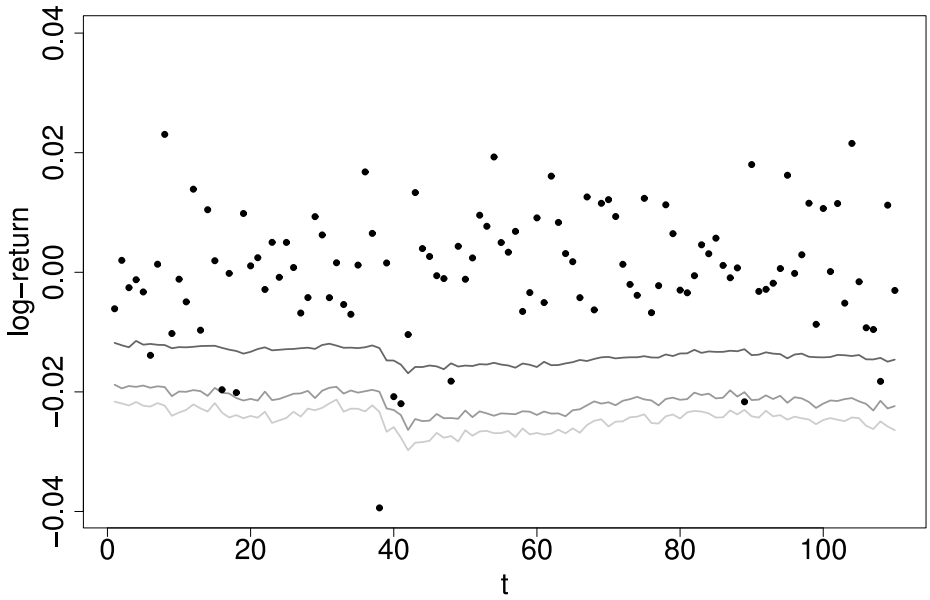


**Figure 9.** Log-returns for the equity portfolio for the period 30.12.2006 - 11.06.2007 along with 0.5%, 1%, 5% VaR simulated from the estimated GARCH-NIG-Student PCC.

**Table 7.** Number of violations of VaR, expected number of violations and p-values for the Kupiec test.

| *Significance level* | 0.005 | 0.01 | 0.05 |
|---|---|---|---|
| *Observed* | 1 | 2 | 9 |
| *Expected* | 0.55 | 1.1 | 5.5 |
| *P-value* | 0.13 | 0.44 | 0.16 |

## 4. Summary and Conclusions

In this paper we have reviewed two classes of structures for construction of higher-dimensional dependence; the nested Archimedean constructions (NACs) and the pair-copula constructions (PCCs). For both structures, a multivariate data set is modelled using a cascade of lower-dimensional copulae. They differ however in their construction of the dependence structure, the PCC being more flexible in that it allows for the free specification of $d(d-1)/2$ copulae, while the NAC's only allow for $d-1$.

Simulation and estimation techniques for the two structures have been examined, and we have shown that the PCCs in general are more computationally efficient than the NACs. The fit of the two constructions has been tested on two different four-dimensional data sets; precipitation values and equity returns, using state of the art copula goodness-of-fit procedures. The NACs considered are strongly rejected for both our data sets. For the precipitation data the Gumbel PCC provides a better fit. However, since even this structure is rejected by one of the goodness-of-fit tests used, one should look for other copula types that might capture the properties of the precipitation data even better than the Gumbel copula does. For the equity data, the Student PCC provides a good fit, and through VaR calculations we have shown that it does not overfit the training data, but works very well also out-of-sample.

Based on the properties presented and the results from the two applications we recommend in general the PCC over the NAC for the following reasons. First, the NAC has an important restriction in that the degree of dependence must decrease with the level of nesting. When looking for appropriate data sets for the applications in this paper, it turned out to be quite difficult to find real-world data sets satisfying this restriction. In addition, the NAC is restricted to the Archimedean class, and there are even restrictions on which Archimedean copulae that can be mixed. There might be real-world situations where there are natural hierarchy groupings of variables. In such cases the NAC's may come into consideration. However, the technical restrictions of the NAC might prevent extensive use.

The PCC, on the other hand, can be built using copulas from any class and there are no restrictions on the parameters of the structure. As far as we are concerned, the only potential disadvantage of the PCC compared to the NAC is that neither the unspecified bivariate margins nor the multivariate distribution

in general will belong to a known parametric family. However, we do not view this as a significant problem since these distributions might easily be obtained through simulation.

## Acknowledgements

Daniel Bergs work is supported by the Norwegian Research Council, grant number 154079/420 and Kjersti Aas' part is sponsored by the Norwegian fund Finansmarkedsfondet. We are very grateful to Cornelia Savu, Institute for Econometrics, University of Münster, Germany, for providing us with her code for the NAC's along with helpful comments. In addition we would like to express our deep gratitude for assistance on the GARCH-NIG filtration to Professor J.H. Venter, Centre of Business Mathematics and Informatics, North-West University, Potchefstroom, South Africa. Finally, we thank participants at the conference on copulae and multivariate probability distributions at Warwick Business School in September 2007 for their valuable comments, in particular Professor Alexander McNeil.

## References


Aas, K., C. Czado, A. Frigessi, and H. Bakken (2007). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics 42*. In press.

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics 24*, 1–13.

Bedford, T. and R. M. Cooke (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence 32*, 245–268.

Bedford, T. and R. M. Cooke (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics 30*, 1031–1068.

Berg, D. (2007). Copula goodness-of-fit testing: An overview and power comparison. Working paper, University of Oslo & Norwegian Computing Center.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Chen, X. and Y. Fan (2006). Estimation and model selection of semi-parametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics 135*, 125–154.

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

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés: Un test non paramétrique d'indépendence. *Acad. Royal Bel., Bull. Class. Sci.,* $5^e$ *série 65*, 274–292.

Embrechts, P., F. Lindskog, and A. McNeil (2003). Modelling dependence with copulas and applications to risk management. In S.T.Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance*. North-Holland: Elsevier.

Fermanian, J.-D. (2005). Goodness-of-fit tests for copulas. *Journal of Multivariate analysis 95*, 119–152.

Frees, E. W. and E. A. Valdez (1998). Understanding relationships using copulas. *North American Actuarial Journal 2*(1), 1–25.

Genest, C., J.-F. Quessy, and B. Rémillard (2006). Goodness-of-fit procedures for copula models based on the probability integral transform. *Scandinavian Journal of Statistics 33*, 337–366.

Genest, C. and B. Rémillard (2005). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. Technical Report G-2005-51, GERAD, Montreal, Canada.

Genest, C., B. Rémillard, and D. Beaudoin (2007). Omnibus goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics 42*. In press.

Gumbel, E. (1960). Distributions des valeurs extrémes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris 9*, 171–173.

Joe, H. (1996). Families of *m*-variate distributions with given margins and *m(m-1)/2* bivariate dependence parameters. In L. Rüschendorf and B. Schweizer and M. D. Taylor (Ed.), *Distributions with Fixed Marginals and Related Topics*.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. London: Chapman & Hall.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of derivatives 2*, 173–184.

Kurowicka, D. and R. M. Cooke (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. and R. M. Cooke (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. New York: Wiley.

McNeil, A. J. (2007). Sampling nested archimedean copulas. *To appear in Journal of Statistical Computation and Simulation*.

McNeil, A. J., R. Frey, and P. Embrechts (2006). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.

McNeil, A. J. and J. Neslehova (2007). Multivariate archimedean copulas, d-monotone functions and l1-norm symmetric distributions. *To appear in Annals of Statistics*.

Min, A. and C. Czado (2007). Bayesian estimation of multivariate copulas using pair-copula constructions. Working paper, Technische Universität München.

Morillas, P. M. (2005). A method to obtain new copulas from a given one. *Metrika 61*, 169–184.

Musshoff, O. M., M. Odening, and W. Xu (2006, July 23-26). Modeling and hedgning rain risk. In *Annual Meeting of the American Agricultural Economics Association (AAEA), Long Beach, USA*.

Nelsen, R. (1999). *An Introduction to Copulas*. New York: Springer.

Okhrin, O., Y. Okhrin, and W. Schmid (2007, May 30-31). 2007. In *Radon Workshop on Financial and Actuarial Mathematics for Young Researchers, Linz, Austria*.

Quessy, J.-F. (2005). *Théorie et application des copules : tests d'adéquation, tests d'indépendance et bornes pour la valeur-à-risque*. Ph. D. thesis, Université Laval.

Savu, C. and M. Trede (2006). Hierarchical Archimedean copulas. In *International Conference on High Frequency Finance, Konstanz, Germany, May*.

Venter, J. H. and P. J. de Jongh (2002). Risk estimation using the normal inverse Gaussian distribution. *Journal of Risk 4*(2), 1–24.

Venter, J. H. and P. J. de Jongh (2004). Selecting an innovation distribution for garch modelsto improve efficiency of risk and volatility estimation. *Journal of Risk 6*(3), 27–53.

Whelan, N. (2004). Sampling from Archimedean copulas. *Quantitative Finance 4*, 339–352.

**Table 8.** Estimated GARCH and NIG parameters for our four stocks.

| *Parameter* | *BP* | *XOM* | *DT* | *FTE* |
|---|---|---|---|---|
| $a_0$ | 1.598e-06 | 1.400e-06 | 1.801e-06 | 1.231e-06 |
| $a$ | 0.010 | 0.023 | 0.025 | 0.028 |
| $b$ | 0.978 | 0.968 | 0.963 | 0.966 |
| $\beta$ | -0.357 | -0.577 | 0.105 | 0.037 |
| $\psi$ | 3.686 | 2.293 | 1.173 | 1.670 |

## A. Parameters for GARCH-NIG model

Table 8 shows the estimated parameters for the GARCH-NIG model used in Section 3.3. For further details of the estimation procedure see Venter and de Jongh (2002).