MULTIVARIATE COPULA MODELS AT WORK: DEPENDENCE STRUCTURE OF ENERGY PRICES

MATTHIAS FISCHER†, CHRISTIAN KÖCK♮

† *Department of Computational Statistics*
Helmut-Schmidt University of Hamburg, Germany
Email: mfischer@hsu-hh.de

♮ *Department of Statistics and Econometrics*
University of Erlangen-Nürnberg, Germany
Email: Christian.Koeck@wiso.uni-erlangen.de

This research was kindly supported by the Hans Frisch-Stiftung.

SUMMARY

Since the pioneering work of Embrechts and co-authors in 1999, copula models enjoy steadily increasing popularity in finance. Whereas copulas are well-studied in the bivariate case, the higher-dimensional case still offers several open issues and it is by far not clear how to construct copulas which sufficiently capture the characteristics of financial returns. For this reason, elliptical copulas (i.e. Gaussian and Student-t copula) still dominate both empirical and practical applications. On the other hand, several attractive construction schemes appeared in the recent literature promising flexible but still manageable dependence models. The aim of this work is to empirically investigate whether these models are really capable to model differen sorts of exchange-traded energy prices.

Keywords and phrases: KS-copula; Hierarchical Archimedian; Product copulas; Pair-copula decomposition

1 Introduction

The increasing linkages between countries, markets and companies require an accurate and realistic modeling of the underlying dependence structure. This applies to financial markets and, in particular, to the financial assets traded there-on. For a long time both practitioners and theorists rely on the multivariate normal (Gaussian) distribution as statistical fundament, seemingly ignoring that this model assigns too less probability mass to extremal events. In order to remove this drawback but still maintain many of the attractive properties, elliptical distributions (e.g. multivariate Student-t or multivariate generalized hyperbolic distribution) occasionally found its way into financial literature. Though being able to model heavy tails, elliptical distributions fail to capture asymmetric dependence structures. The copula concept, in contrast, which originally dates back to Sklar (1959) but was made popular to finance through the pioneering work of Embrechts and co-authors (1999) provides a flexible tool to capture different patterns of dependence. Within this work we assume that the reader is already familiar with the notion of copulas. Otherwise, we refer

to Nelsen (2006) or Joe (1997). Whereas copulas are well-studied in the bivariate case, construction schemes for higher dimensional copulas are not. Recently, several publications on high-dimensional copulas appeared (e.g. Morillas, 2005, Palmitesta & Provasi, 2005, Savu & Trede, 2006, Liebscher, 2006, Aas et al., 2006). Each of them claims to provide a flexible dependence model, but there is no comprehensive comparison among these approaches, as far as we know. In particular, no references are found to the Student-t copula (i.e. the copula associated to the multivariate Student-t distribution) which is sometimes termed as "desert island copula" by Paul Embrechts on account of its excellent fit to multivariate financial return data.
The outline of this work is as follows: Section 2 overviews and connects several recent construction schemes of multivariate copulas. A short digression on goodness-of-fit measures can be found in section 3. Section 4 is dedicated to the description of the underling data sets, whereas the empirical results are summarized and discussed in section 5.

2 Multivariate copula models

2.1 Copulas - A short review

Representing the dependence structure of two or more random variables, the popularity of copulas is steadily increasing in many statistical disciplines. Let $[a,b]^d \subseteq \mathbb{R}^d$. A function $K : [a,b]^d \to \mathbb{R}$ is said to be *d-increasing* if its K-volume

$$V_K \equiv \sum_{i_1=1}^{2} \dots \sum_{i_d=1}^{2} (-1)^{i_1+\dots+i_d} C(u_{1i_1}, \dots, u_{di_d}) \geq 0 \tag{2.1}$$

for all $a \leq u_{i1} \leq u_{i2} \leq b$ and $i = 1, \dots, d$. If, additionally, $[a,b] = [0,1]$ and K satisfies the boundary conditions

$$K(u_1, \dots, u_{j-1}, 0, u_{j+1}, \dots, u_d) = 0 \quad \text{and} \quad K(1, \dots, 1, u, 1, \dots, 1) = u \tag{2.2}$$

for arbitrary $u \in [0,1]$, K is termed as copula and we write C, instead. Putting a different way, let $X_1, \dots, X_d$ denote d random variables with joint distribution $F(\mathbf{x}) = F(x_1, \dots, x_d)$ and continuous marginal distribution functions $F_1(x), \dots, F_d(x)$. According to Sklar's (1959) fundamental theorem, there exists a unique decomposition

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$$

of the joint distribution into its marginal distribution functions and the so-called copula

$$C(u_1, \dots, u_d) = P(U_1 \leq u_1, \dots, U_d \leq u_d), \quad U_i \equiv F_i(X_i)$$

on $[0,1]^d$ which comprises the information about the underlying dependence structure (For details on copulas we refer to Nelsen, 2006 and Joe, 1999). Finally, if C has dth order derivatives, the d-increasing condition is equivalent to

$$\frac{\partial^d C}{\partial u_1 \dots \partial u_d} \geq 0. \tag{2.3}$$

2.2 Elliptical copulas

The multivariate Gauss copula and t copula are examples of copulas which can be extracted from well-known multivariate distributions. Since they are widely used in practice they serve as benchmark models in this empirical analysis of energy prices. The d-variate Gauss copula is the copula of $\mathbf{X} = (X_1, \ldots, X_d) \sim N_d(\mathbf{0}, R)$ and defined as follows

$$C^{Ga}(u_1 \ldots, u_d) = \mathbf{\Phi}_R(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)) \ , \tag{2.4}$$

where $\mathbf{\Phi}_R$ denotes the d-variate standard normal cdf with correlation matrix R and Φ denotes the standard univariate normal cdf.
The d-dimensional t copula ist defined by

$$C^t(u_1, \ldots, u_d) = \mathbf{t}_{\nu,R}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d)) \ , \tag{2.5}$$

where t_ν is the cdf of a standard univariate t distribution, $\mathbf{t}_{\nu,R}$ is the joint df of the vector $\mathbf{X} \sim t_d(\nu, \mathbf{0}, R)$ and R is a correlation matrix.

2.3 Classical multivariate Archimedean copulas

Given a strict generator $\varphi : [0,1] \to [0,\infty]$, bivariate Archimedean copulas can be extended to the d-dimensional case. For every $d \geq 2$ the function $C : [0,1]^d \to [0,1]$ defined as

$$C(u_1, \ldots, u_d) = C(\mathbf{u}) = \varphi^{-1}\Big(\varphi(u_1) + \varphi(u_2) + \cdots + \varphi(u_d)\Big) \tag{2.6}$$

is a d-dimensional Archimedean copula if and only if φ^{-1} is completely monotonic on $\mathbb{R}_+$, i.e. if $\varphi^{-1} \in \mathcal{L}_\infty$ with

$$\mathcal{L}_m \equiv \Big\{\phi : \mathbb{R}_+ \to [0,1] \Big| \, \phi(0) = 1, \, \phi(\infty) = 0, \, (-1)^k \phi^{(k)}(t) \geq 0 \ , \, k = 1, \ldots, m, \Big\} .$$

The d-variate Clayton copula arises from $\varphi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$ and is given by

$$C^{Cl}(\mathbf{u}) = \Big(u_1^{-\theta} + \cdots + u_d^{-\theta} - d + 1\Big)^{-1/\theta}, \ \theta > 0. \tag{2.7}$$

The d-dimensional Gumbel copula

$$C^{Gu}(\mathbf{u}) = \exp\left\{-\Big[(-\ln u_1)^\theta + \cdots + (-\ln u_d)^\theta\Big]^{1/\theta}\right\}, \ \theta \geq 1, \tag{2.8}$$

derives from the generator $\varphi(t) = (-\ln t)^\theta$.
For an overview of further Archimedean copulas and the properties of the aforementioned ones, we refer the reader to the monographs by Nelson (2006) and Joe (1997).

2.4 Hierarchical Archimedean copulas

The basic idea of this approach (see, e.g. Savu & Trede, 2006) is to build a hierarchy of Archimedean copulas. Let there be L hierarchy levels indexed by l. At each level $l = 1, \ldots, L$ one has n_l distinct objects with index $j = 1, \ldots, n_l$. The $u_1, \ldots, u_d$ are located at the lowest level, $l = 0$. At level $l = 1$ the $u_1, \ldots, u_d$ are grouped into n_1 ordinary multivariate Archimedean copulas $C_{1,j}$, $j = 1, \ldots, n_1$, of the form

$$C_{1,j}(\mathbf{u}_{1,j}) = \varphi_{1,j}^{-1}\left(\sum \varphi_{1,j}(\mathbf{u}_{1,j})\right)$$

where $\varphi_{1,j}$ denotes the generator of copula $C_{1,j}$. Let $\mathbf{u}_{1,j}$ denote the set of elements of $u_1, \ldots, u_d$ belonging to copula $C_{1,j}$ for $j = 1, \ldots, n_1$. The copulas $C_{1,1}, \ldots, C_{1,n_1}$ might belong to different Archimedean families. All copulas of level $l = 1$ are in turn aggregated into copulas at level $l = 2$. The n_2 copulas $C_{2,j}$, $j = 1, \ldots, n_2$ are generalized Archimedean copulas, whose dependence structure is only of partial exchangeability. They consist of copulas from the previous level (as elements) and can be represented as

$$C_{2,j}(\mathbf{C}_{2,j}) = \varphi_{2,j}^{-1}\left(\sum_{\mathbf{C}_{2,j}} \varphi_{2,j}(\mathbf{C}_{2,j})\right),$$

where $\varphi_{2,j}$ denotes the generator of copula $C_{2,j}$, and $\mathbf{C}_{2,j}$ represents the set of all copulas from level $l = 1$ entering copula $C_{2,j}$ for $j = 1, \ldots, n_2$. We can proceed in this manner until attaining level L with the hierarchical Archimedean copula $C_{L,1}$ as single object.

2.5 Generalized multiplicative Archimedean copulas

In this section we focus on methods recently proposed by Morillas (2005) and Liebscher (2006). Both approaches are based on a second functional representation of Archimedean copulas via so called multiplicative generators (see Nelsen, 2006). Setting $\vartheta(t) \equiv \exp(-\varphi(t))$ and $\vartheta^{[-1]}(t) \equiv \varphi^{[-1]}(-\ln t)$, equation (2.6) can be rewritten as

$$C(u_1, \ldots, u_d) = \vartheta^{[-1]}\Big(\vartheta(u_1) \cdot \vartheta(u_2) \cdot \ldots \cdot \vartheta(u_d)\Big). \tag{2.9}$$

The function ϑ is called multiplicative generator of C.Equation (2.9) can also be expressed using the independence copula $C^{\perp}(\mathbf{u}) = \prod_{i=1}^{d} u_i$. Morillas (2005) substitutes $C^{\perp}$ by an arbitrary d-copula C in order to obtain

$$C_\vartheta(u_1, \ldots, u_d) = \vartheta^{[-1]}\Big(C(\vartheta(u_1), \vartheta(u_2), \ldots, \vartheta(u_d))\Big) \tag{2.10}$$

and proves that C_ϑ is a d-copula if $\vartheta^{[-1]}$ is *absolutely monotonic of order* d on $[0, 1]$, i.e. if $\vartheta^{[-1]}(t)$ satisfies $(\vartheta^{[-1]})^{(k)}(t) = \frac{d^k \vartheta^{[-1]}(t)}{dt^k} \geq 0$ for $k = 1, 2, \ldots, d$ and $t \in (0, 1)$.

Another way of generalizing Archimedean copulas is the method proposed by Liebscher (2006) who introduces the following copula representation

$$C(u_1, \dots, u_d) = \Psi\left(\frac{1}{m}\sum_{j=1}^{m} \psi_{j1}(u_1) \cdot \psi_{j2}(u_2) \cdot \ldots \cdot \psi_{jd}(u_d)\right), \tag{2.11}$$

where Ψ and $\psi_{jk} : [0,1] \to [0,1]$ are functions satisfying the following conditions: Firstly, it is assumed that $\Psi^{(d)}$ exist with $\Psi^{(k)}(u) \geq 0$ for $k = 1, 2, \dots, d$ and $u \in [0,1]$, and that $\Psi(0) = 0$. Secondly, ψ_{jk} is assumed to be differentiable and monotone increasing with $\psi_{jk}(0) = 0$ and $\psi_{jk}(1) = 1$ for all k, j. Thirdly, Liebscher's construction requires that $\Psi\left(\frac{1}{m}\sum_{j=1}^{m} \psi_{jk}(v)\right) = v$, for $k = 1, 2, \dots, d$ and $v \in [0,1]$. The three conditions guarantee that C defined in (2.11) is actual a copula. Proposals for Ψ and ψ are given in Liebscher (2006). A generalization which contains both the proposals by Morillas and Liebscher as special cases can be found in Fischer and Köck (2007).

2.6 Pair-copula decomposition of a copula

Originally, the pair-copula decomposition (PCD) decomposes the common density f of d random variables (Aas et al. 2006). Of course, one may also apply the pair-copula decomposition to the underlying copula density c, as we will show in this subsection. As an immediate consequence of Sklar's (1959) theorem, $c(F(x_1), F(x_2), F(x_3), F(x_4)) = \frac{f(x_1,x_2,x_3,x_4)}{f(x_1)\cdot f(x_2)\cdot f(x_3)\cdot f(x_4)}$. Substituting the common density by its PCD,

$$\begin{aligned} c(F(x_1), F(x_2), F(x_3), F(x_4)) &= c_{12}(F(x_1), F(x_2)) \cdot c_{23}(F(x_2), F(x_3)) \cdot c_{34}(F(x_3), F(x_4)) \\ &\quad \cdot c_{13|2}(F(x_1|x_2), F(x_3|x_2)) \cdot c_{24|3}(F(x_2|x_3), F(x_4|x_3)) \\ &\quad \cdot c_{14|23}(F(x_1|x_2, x_3), F(x_4|x_2, x_3)) \end{aligned}$$

with $c_{i|j}(\cdot,\cdot)$ being a pair-copula density and its indices i, j refer to x_i and x_j. Defining the function

$$h: \ (x, v, \theta\ ) \longmapsto \frac{\partial\, C_{x|v}(F_X(x), F_V(v))}{\partial\, F_V(v)}$$

with θ being the parameter vector of the copula $C_{x,v}$, the copula density decomposition can be written as follows: It is obvious that $F(x_1|x_2) = h(x_1, x_2, \theta_{12})$ with θ_{12} is the parameter (vector) of the of copula C_{12}. Analogously, $F(x_3|x_2) = h(x_3, x_2, \theta_{23})$, $F(x_2|x_3) = h(x_2, x_3, \theta_{23})$ and $F(x_4|x_3) = h(x_4, x_3, \theta_{34})$. $F(x_1|x_2, x_3)$, again, can be iteratively simplified. Finally, define $u_i = F(x_i), i = 1, \dots, 4$. The formula for the 4-dimensional PCD copula density now reads as

$$\begin{aligned} c(\mathbf{u}) &= c_{12}(u_1, u_2) \cdot c_{23}(u_2, u_3) \cdot c_{34}(u_3, u_4) \\ &\quad \cdot c_{13|2}(h(u_1, u_2, \theta_{12}), h(u_3, u_2, \theta_{23})) \cdot\ c_{24|3}(h(u_2, u_3, \theta_{23}), h(u_4, u_3, \theta_{34})) \\ &\quad \cdot c_{14|23}(h(h(u_1, u_3, \theta_{13}), h(u_2, u_3, \theta_{23}), \theta_{13|2}), h(h(u_4, u_3, \theta_{43}), h(u_2, u_3, \theta_{23}), \theta_{24|3})). \end{aligned}$$

2.7 Koehler-Symanowski (KS) copulas

Koehler & Symanowski (1995) introduce a multivariate distribution. The corresponding copula is defined as follows: With the index set $V = \{1, 2, \dots, d\}$, $\mathcal{V}$ being the power set of V and $\mathcal{I} \equiv \{I \in \mathcal{V} \text{ with } |I| \geq 2\}$ let $\mathbf{X}$ denote a d-dimensional random vector with univariate marginal distributions $F_i(x_i), i \in V$. For all subsets $I \in \mathcal{I}$ let $\alpha_I \in \mathbb{R}_0^+$ and $\alpha_i \in \mathbb{R}_0^+$ for all $i \in V$ such that $\alpha_{i+} = \alpha_i + \sum_{I \in \mathcal{I}} \alpha_I > 0$ for $i \in I$. Setting $u_i = F_i(x_i)$ for all $i \in V$, the KS copula is

$$C(u_1, \dots, u_d) = \frac{\prod_{i \in V} u_i}{\prod_{I \in \mathcal{I}} \left[\sum_{i \in I} \prod_{j \in I, j \neq i} u_j^{\alpha_{j+}} - (|I| - 1) \prod_{i \in I} u_i^{\alpha_{i+}} \right]^{\alpha_I}} . \tag{2.12}$$

In contrast to the cumulative distribution function the functional representation of the density is quite complicated due to complex factors with additive components. Koehler & Symanowski (1995) gave an explicit formula for the special case of a so called KS(2)-distribution, where all parameters α_I are set equal zero for $|I| > 2$. The corresponding copula will be termed as KS(2) copula henceforth. Palmitesta & Provasi (2005) apply this particular KS copula to weekly log-returns.
We enlarge the approach of Palmitesta & Provasi by setting the association parameter $\alpha_I \geq 0$ for $|I| = 2$ and $|I| = 4$ in (2.12), while all parameters α_I are set equal to zero for $|I| = 3$, i.e. we include a global dependence parameter. We refer to this copula as augmented KS (aKS) copula . Note that the aKS copula contains the Clayton copula as a special case.

2.8 Multiplicative Liebscher copulas

Liebscher (2006) discusses how to combine or connect a given set of k possibly different d-copulas $C_1, \dots, C_k$ to a new d-copula C in order to increase flexibility and/or introduce asymmetry. His proposal focusses on multiplicative connections of d-copulas of the form

$$C(u_1, \dots, u_d) = \prod_{j=1}^{k} C_j(g_{j1}(u_1), \dots, g_{jd}(u_d)) \tag{2.13}$$

with a set of $k \cdot d$ admissible functions $g_{11}, \dots, g_{1d}, \dots, g_{k1}, \dots, g_{kd}$, each of which being bijective, monotonously increasing or identically equal 1 satisfying $\prod_{j=1}^{k} g_{ji}(v) = v, \quad i = 1, \dots, d$. Possible choices for g can be found in Liebscher (2006). We use the following two specifications for g:

$$g_{ji}(v) \equiv v^{\theta_{ji}} \text{ with } \theta_{ji} > 0 \text{ and } \sum_{j=1}^{k} \theta_{ji} = 1 \text{ for } i = 1, \dots, d \tag{2.14}$$

$$g_{1i}(v) = f(v), \quad g_{2i}(v) \equiv v \cdot \frac{1}{f(v)}, \quad f(v) = \left(\frac{1 - e^{-\theta_i v}}{1 - e^{-\theta_i}} \right)^{\alpha}, \quad \theta > 0,\ \alpha \in (0, 1) \tag{2.15}$$

Setting $k = 2$, choosing $C_1(\mathbf{u}) = C(\mathbf{u})$ and $C_2(u_1, \ldots, u_d) = \prod_{i=1}^d u_i$, and defining $g_{1i}(v) \equiv v^{\theta_i}$ and $g_{2i}(v) \equiv v^{1-\theta_i}$, the last equation generalizes to the d-copula with $d+1$ dependence parameters ("Generalized Clayton of Liebscher type I", briefly L_1),

$$C(u_1, \ldots, u_d) = \prod_{i=1}^{d} u_i^{1-\theta_i} \left(1 + \sum_{i=1}^{d} (u_i^{-\gamma\theta_i} - 1)\right)^{-1/\gamma}, \quad \gamma > 0, \quad \theta_i \in (0, 1).$$

Applying (2.15) rather than (2.14), the Generalized Clayton of Liebscher type II (L_2) copula with $d+2$ dependence parameters

$$C(u_1, \ldots, u_d) = \prod_{i=1}^{d} \left(\frac{1 - e^{-\theta_i u_i}}{1 - e^{-\theta_i}}\right)^{\alpha} \left(1 - d + \sum_{i=1}^{d} u_i^{-\gamma} \left[\frac{1 - e^{-\theta_i u_i}}{1 - e^{-\theta_i}}\right]^{\alpha\gamma}\right)^{-1/\gamma}.$$

Similarly, combining two d-variate Clayton copulas with parameter γ and λ, respectively, and using g from (2.14) we obtain the d-dimensional copula family with $d+2$ parameters, termed as the Generalized Clayton of Liebscher type III (L_3) in the sequel. Finally, applying again (2.15) rather than (2.14), results in the Generalized Clayton of Liebscher type IV (L_4).

3 Goodness-of-fit measures

We now tackle the problem to compare the goodness-of-fit (GOF) of the different copula models from section 2, noting that most of them are not nested. As we apply maximum likelihood (ML) methods to obtain estimators for the unknown parameter vector, the first choice is the log-likelihood value ℓ or – in order to take the different numbers of parameters in account – the information criterion of Akaike $AIC = -2\ell + (2N(K+1))/(N-K-2)$, where K and N denote the number of parameters to be fitted and the number of observations, respectively. However, comparing log-likelihood values for non-nested models may produce misleading conclusions. Otherwise, one might compare the matrix of non-parametric dependency measures (e.g. Spearman's ρ, Kendall's τ or Blomberg's β) between assets pairs with the "theoretical ones" which can be derived from the parametric copula model, though keeping in mind that only bivariate dependencies are taken into consideration. Finally, certain GOF tests may come to application. Following Breymann, Dias & Embrechts (2003), Chen, Fan & Patton (2004) or recently Berg & Bakken (2006), the main idea is to project the multivariate problem into a set of independent and uniform $U(0,1)$ variables, given the multivariate distribution and to calculate the distance (e.g. Anderson-Darling, Kolmogorov-Smirnov, Cramér-von Mises, Kernel smoothing) between the transformed variables and the uniform distribution. In contrast to the authors above, we are not primarily interested whether the data stem from the specified copula model but we use these distances as citerion itself. The proceeding is roughly as follows:

By means of the Rosenblatt (1952) transformation the random vector $\boldsymbol{X} = (X_1, \ldots, X_d)'$ is mapped onto a random vector $\boldsymbol{Z}^* = (Z_1^*, \ldots, Z_d^*)'$ via

$$Z_1^* \equiv F_1(X_1) \text{ and } Z_i^* \equiv F_{X_i|X_1,\ldots,X_{i-1}}(X_i|X_1,\ldots,X_{i-1}), \quad i = 2,\ldots,d. \tag{3.1}$$

It can be shown that $\boldsymbol{Z}^*$ is uniformly distributed on $[0,1]^d$ with independent components $Z_1^*, \ldots, Z_d^*$. Assume that the cumulative distribution function of $\boldsymbol{X}$ admits the decomposition

$$F_{\boldsymbol{X}}(x_1,\ldots,x_d) = C(F_{X_1}(x_1),\ldots,F_{X_d}(x_d)),$$

where $C(\cdot)$ denotes a parametric copula which is the common distribution function of $\boldsymbol{U} = (U_1,\ldots,U_d)'$ with $U_i \equiv F_{X_i}(X_i)$. Define $C(u_1,\ldots,u_j) \equiv C(u_1,\ldots,u_j,1,\ldots,1)$ for $j \leq d$. Furthermore, the conditional distribution of $U_i|U_1,\ldots,U_{i-1}$ is given by

$$C_i(u_i) \equiv \frac{\partial^{i-1}C(u_1,\ldots,u_i)}{\partial u_1 \ldots \partial u_{i-1}} \Big/ \frac{\partial^{i-1}C(u_1,\ldots,u_{i-1})}{\partial u_1 \ldots \partial u_{i-1}}.$$

According to (3.1), the variables

$$Z_1 \equiv C(U_1) = U_1 \text{ and } \quad Z_i \equiv C_i(U_i), \quad i = 2,\ldots,d \tag{3.2}$$

are independent and uniform on $[0,1]$. Consequently, the sample $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_N$ from a parametric copula and with marginals given by $F_1,\ldots,F_d$ can be mapped onto an *iid* sample $\boldsymbol{Z}_1,\ldots,\boldsymbol{Z}_N$ from a uniform distribution on $[0,1]^d$.

Breymann et al. (2003) suggest to transform each random vector $\boldsymbol{Z}_i = (Z_{i1},\ldots,Z_{id})'$ in a (univariate) chi-square variable χ_j with d degrees of freedom through

$$\chi_j = \sum_{i=1}^{d} \Phi^{-1}(Z_{ji})^2, \quad j = 1,\ldots,N, \tag{3.3}$$

where $\Phi^{-1}(u)$ denotes the standard normal quantile function. If the margins are unknown, they may be replaced by the corresponding empirical counterparts. Breymann et al. (2003) state that "we do assume that the χ^2-distribution will not be significantly affected by the use of the empirical distribution functions used to transform the marginal data".

4 The data set

The data sets we used to compare the different copula models come from energy future markets. In particular, we selected natural gas (NG) futures, harbor unleaded (HU) gasoline futures, heating oil (HO) futures and light sweet crude oil (CL) futures. Instead of analyzing the prices themselves, we calculated and considered (percentual) continuously compounded returns ("log-returns")

$$R_t = 100(\log P_t - \log P_{t-1}), \quad t = 2,\ldots,N.$$

In order to account for possible time-dependencies (which are common to most financial return series), we fitted univariate GARCH models of the form

$$R_t = \mu + \gamma_1 R_{t-1} + \ldots + \gamma_k R_{t-k} + h_t \epsilon_t$$

with variance equations

$$h_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2 + \ldots + \alpha_1 R_{t-p}^2 + \beta_1 h_{t-1}^2 + \ldots + \beta_q h_{t-q}^2$$

to each of the series and considered standardized residuals ϵ_t rather the original returns R_t. Secondly, as we are primarily not interested in parametric models for the marginal distributions, all observations (i.e. returns or standardized residuals) were transformed into uniform ones by means of the (empirical) probability integral transform, i.e.

$$U_t = F_N(R_t) \text{ with } F_N(x) = \frac{\{\#R_t | R_t \leq x\}}{\#R_t} \text{ and } U_t^* = F_N(\epsilon_t),$$

where $\#$ denotes the number of observations. Figure 1 contains the series of prices and returns.

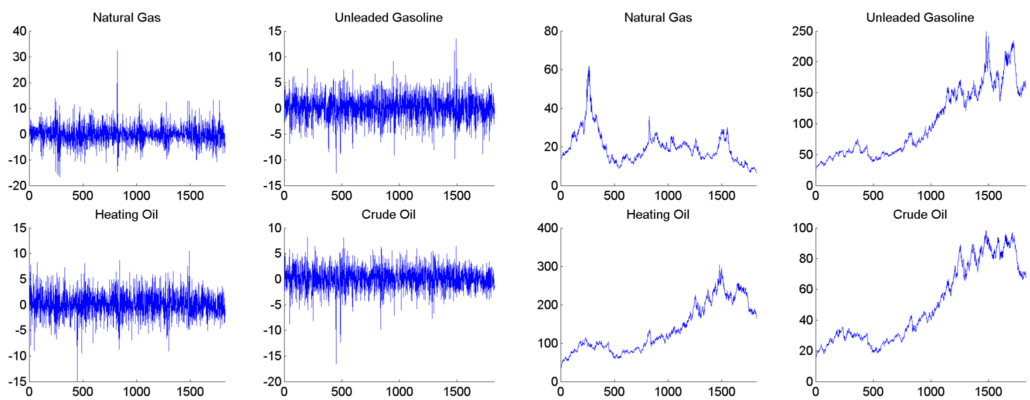

Figure 1: German stock prices and stock returns.

The following table 1 summarizes descriptive and inductive statistics. All series feature a certain amount of skewness and high kurtosis (measured by the third and fourth standardized moment $\mathbb{S}$ and $\mathbb{K}$). Moreover, there is no significant evidence for serial correlation but for GARCH effects as the Ljung-Box statistic $\mathcal{LB}$ and Engle's Lagrange multiplier statistic $\mathcal{LM}$ indicate.
Above that, the matrix plot in figure 2 illustrates the positive dependence between all series.

5 Empirical results

The 4-copulas under consideration are the following: Firstly, we selected the Clayton copula (CLA), the Gumbel copula (GUM) and its rotated version (roGUM) from the Archimedean

Start	End	N	Future	μ	s^2	$\mathbb{S}$	$\mathbb{K}$	$\mathcal{LB}$	$\mathcal{LM}$
02-01-00	27-12-06	1823	Natural Gas	−0.043	13.19	0.317	7.94	17.07 (18.307)	80.35 (18.307)
02-01-00	27-12-06	1823	Unleaded Gasoline	0.095	6.38	−0.12	4.25	6.18 (18.307)	32.16 (18.307)
02-01-00	27-12-06	1823	Heating Oil	0.079	5.77	−0.099	4.60	9.51 (18.307)	51.84 (18.307)
02-01-00	27-12-06	1823	Light sweet crude oil	0.075	4.77	−0.437	5.98	10.87 (18.307)	30.41 (18.307)

Table 1: Energy futures.

class. From the generalized Archimedean copula family, two hierarchical copula models (i.e. HA-CLA and HA-GUM) are included, based on the Clayton and the Gumbel copula, respectively. Moreover, six representatives of Morrillas' construction scheme (i.e. MO-CLA1, MO-CLA2, MO-CLA3, MO-GUM1, MO-GUM2, MO-GUM3) involving the Clayton, the Gumbel and different generator functions (no. 3, 2, 4 in Morillas, 2005) are included as well. In addition, two version of Liebscher's proposal (GMLF, GML2) are used. Above that, representing the "elliptical copula world", the Gaussian copula (NORM) and – as ultimate benchmark – the Student-t (T) copula are also included. From the pair-copula decomposition we chose five representatives (i.e. PC-NORM, PC-T, PC-CLA, PC-GUM) each of them derived from one single copula model (i.e. we used no decompositions based on different copulas). Finally, the KS(2)-copula and its augmented version (which is a generalized version of Palmitesta & Provasi, 2005 because we included a general dependence parameter) and four different types of multiplicative Liebscher copulas from example 2.8 (L_1, L_2, L_3, L_4) are considered.

The computer code for the ML-estimation was implemented in Matlab 7.1. For maximization purposes we used the line-search algorithm of Matlab. The following tables include the comprehensive results for the parameter estimates as well as the different goodness-of-fit measures only for GARCH-residuals and all copulas models mentioned above. As stated above, goodness-of-fit is measured by the log-likelihood value and Akaike's AIC criterion. Above that, two distance measures,

$$\mathrm{KS} = \sqrt{N} \max_{j=1,\ldots,N} \left| F_{\chi^2(d)}(\chi_j) - F_{N,\chi}(\chi_j) \right| \quad \text{and}$$

$$\mathrm{AKS} = \frac{1}{\sqrt{N}} \sum_{j=1,\ldots,N} \left| F_{\chi^2(d)}(\chi_j) - F_{N,\chi}(\chi_j) \right|$$

are calculated to quantify the distance after application of the Rosenblatt transformation (based on the different parametric copula models). A graphical illustration of this procedure is given in figure 3, where we contrasted the kernel density estimator (empirical cdf) of the projected data with the chi-squared density (cumulative distribution) for the Student-t and the Clayton copula.

The estimation results are are the following. As known from several empirical studies, the fit of the 4-variate Gaussian distributions may be considerably improved if the 4-variate

Student-t distribution is considered, instead. However, pair-copula decompositions based on bivariate Student-t copulas produce a better goodness-of-fit concerning the likelihood criteria and the minimal distance measure. Similar, PCD-decompositions based on Archimedean copulas may also be considered as possible alternatives. This also applies to the three copulas $L1$, $L3$ and $L4$, whereas all copulas based on Morillas' approach and, of course, the plain Archimedean copulas feature low goodness-of-fit measures. Considering hierarchical Archimedean copulas, instead, we found only slight improvement, at least for our data set. However, we have to confess that one might improve the results with another hierarchy which might be found on the basis of cluster algorithms.

The KS(2)-copula (recommended by Palmitesta and Provasi, 2005) provides only a poor fit to the return series. However, introducing an additional dependence parameter – which quantifies the overall dependence in the data set – clearly improves all goodness-fit measures.

To sum up, the 4-variate Student-t distribution still plays a predominant role. Some of the recently proposed construction schemes are partially competitive while others are more likely to be overestimated in the literature.

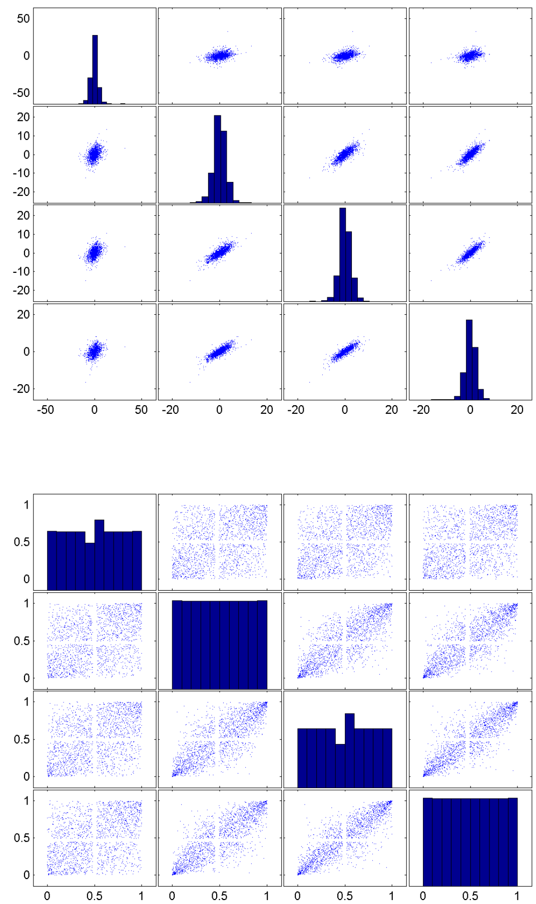

Figure 2: Scatter plots of the (transformed) GARCH residuals.

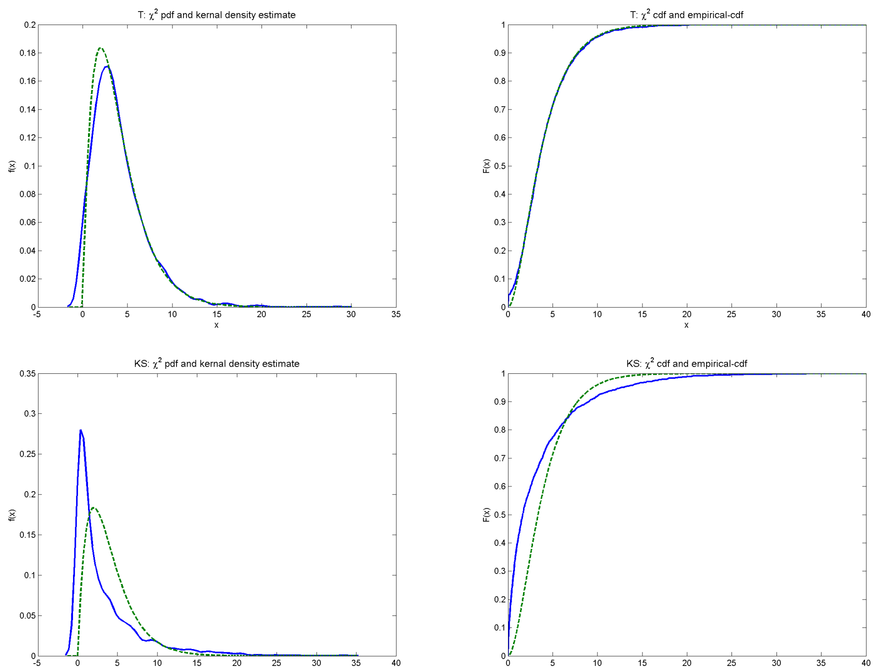

Figure 3: Goodness-of fit: Graphical representation.

Copula	l	AIC	KS	AKS
CLA	1469.38	-2934.76	4.908	0.7079
GUM	1546.30	-3088.59	4.099	0.7334
roGUM	1560.73	-3117.45	3.928	0.7088
NORM	2752.70	-5491.33	1.923	0.3602
T	2826.90	-5637.71	1.822	0.1236
PC-NORM	2755.79	-5497.52	1.947	0.3632
PC-T	2851.102	-5676.00	1.820	0.1173
PC-CLA	2274.32	-4534.58	4.095	0.6423
PC-GUM	2379.51	-4744.96	4.418	0.6344
GML2	1923.86	-3833.65	3.713	0.5339
GMLF	1923.86	-3835.67	3.713	0.5337
KS	226.67	-431.19	6.036	0.8146
aKS	2362.55	-4700.93	2.628	0.4256
MO-CLA1	1515.08	-3024.14	5.263	0.6564
MO-GUM1	1720.28	-3434.54	3.654	0.5036
MO-CLA2	1469.38	-2932.76	4.908	0.7079
MO-GUM2	1546.30	-3086.58	4.099	0.7334
MO-CLA3	1515.08	-3024.14	5.263	0.6564
MO-GUM3	1720.28	-3434.54	3.654	0.5036
L1	2279.35	-4546.64	3.153	0.3816
L3	2285.14	-4556.22	3.296	0.4124
L2	1777.60	-3541.15	3.737	0.5423
L4	2210.68	-4405.28	4.435	0.6047
HA-CLA	1852.58	-3697.14	4.364	0.6563
HA-GUM	2041.67	-4075.31	2.943	0.5583

Table 2: Goodness-of-fit measures: GARCH residuals

Copula	θ											
CLA	1.0014 (0.0235)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
GUM	1.6424 (0.0171)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
roGUM	1.6449 (0.0171)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
Copula	ρ_{12}	ρ_{13}	ρ_{14}	ρ_{23}	ρ_{24}	ρ_{34}	ν_1	ν_2	ν_3	ν_4	ν_5	ν_6
NORM	0.3972 (0.018)	0.456 (0.0167)	0.4162 (0.0175)	0.8199 (0.0057)	0.8368 (0.0054)	0.8758 (0.004)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
T	0.399 (0.0195)	0.4506 (0.0189)	0.4213 (0.019)	0.8245 (0.0064)	0.8433 (0.006)	0.8859 (0.0043)	10.7147 (1.1954)	– (–)	– (–)	– (–)	– (–)	– (–)
PC-NORM	0.397 (0.0176)	0.8193 (0.0059)	0.8759 (0.0041)	0.2483 (0.0215)	0.4354 (0.0187)	0.0245 (0.0238)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
PC-T	0.403 (0.0184)	0.8245 (0.0065)	0.8858 (0.0044)	0.2447 (0.0223)	0.4283 (0.0197)	0.0294 (0.023)	17.2145 (6.8706)	9.8544 (2.2578)	5.4978 (0.7184)	100 (110.1762)	7.8405 (1.4666)	35.54 (34.5133)
Copula	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6						
PC-CLA	0.4671 (0.0355)	2.0714 (0.0663)	2.8212 (0.0832)	0.2244 (0.0312)	0.5037 (0.0393)	0.0352 (0.0192)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
PC-GUM	1.2731 (0.1017)	2 (1.1936)	2 (1.0578)	1.1507 (0.0808)	1.5499 (0.6919)	1.027 (0.0281)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
Copula	δ_1	δ_2	δ_3	δ_4	θ_1	θ_2						
GML2	1 (1.639)	2 (5.2308)	2 (5.2364)	2 (5.2382)	1.0168 (13.7985)	0 (3.0911)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
GMLF	1 (1.0363)	2 (1.0075)	2 (1.0649)	2 (1.088)	4.1008 (0.5267)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)

Table 3: Energy prices (GARCH-Residuals): Parameter estimates and corresponding standard errors (in parenthesis)

Copula	a_{11}	a_{12}	a_{13}	a_{14}	a_{22}	a_{23}	a_{24}	a_{33}	a_{34}	a_{44}	a_{1234}	
KS	0.033 (9095.966)	0.0165 (6540.3683)	0.0291 (16906.5965)	0.0263 (25820.0807)	0.0734 (8746.73)	0.0121 (5413.0933)	0 (13758.2373)	0.0588 (14303.636)	0 (115.178)	0.5508 (765638.7764)	– (–)	– (–)
aKS	0.752 (0.4865)	0.004 (0.1002)	0.0083 (0.8775)	0 (0.8531)	0.0168 (0.0968)	0.0136 (0.0342)	0.0151 (0.2072)	0 (0.963)	0.0286 (0.182)	0.0014 (0.335)	0.2286 (0.5191)	– (–)
Copula	θ	r										
MO-CLA1	0.1741 (0.1241)	−31.3072 (8.8899)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
MO-GUM1	1.2546 (0.0061)	−7.4027 (0.0676)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
MO-CLA2	19.9807 (0.2751)	19.9533 (0.5583)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
MO-GUM2	1.6424 (0.0172)	2.0031 (1)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
MO-CLA3	0.1741 (0.1237)	0.969 (0.0086)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
MO-GUM3	1.2546 (0.0187)	0.881 (0.0172)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
Copula	γ	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6					
L1	3.247 (0.1364)	0.2487 (0.0148)	0.8792 (0.0144)	0.9335 (0.0113)	0.9529 (0.0095)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
L3	3.3771 (0.1527)	0.4771 (0.1819)	0.2354 (0.0149)	0.8668 (0.0154)	0.9115 (0.0153)	0.9383 (0.0126)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
L2	5.4703 (1.9839)	0.4751 (0.1403)	0 (0.9982)	1 (0.7172)	1 (0.9896)	1 (0.9702)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
L4	1.2645 (9.3339)	25.6291 (995.4507)	0.951 (2.533)	0 (7.6025)	1 (31.9871)	0.9865 (13.6176)	1 (3.9397)	– (–)	– (–)	– (–)	– (–)	– (–)
Copula	θ_{11}	θ_{12}	θ_{21}									
HA-CLA	0.8465 (0.9696)	2.8453 (0.1128)	0.8464 (0.3962)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)
HA-GUM	1.5492 (0.994)	2.9248 (0.2275)	1.5491 (0.2008)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)	– (–)

Table 4: Energy prices (GARCH-Residuals): Parameter estimates and corresponding standard errors (in parenthesis)

References

[1] Aas, K.; Czado, C.; Frigessi, A.; Bakken, H. (2006). Pair Copula Constructions of Multiple Dependence, Working Paper of the Norwegian Computing Center.

[2] Berg, D.; Bakken, H. (2006). A Copula Goodness-of-fit Test Based on the Probability Integral Transform, Working Paper of the Norwegian Computing Center.

[3] Breymann, W.; Dias, A.; Embrechts, P. (2003). Dependence structures for multivariate high-frequency data in finance, *Quantitative Finance* 1, 1-14.

[4] 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.

[5] Bedford, T.; Cooke, R.M. (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics* 30(4), 1031-1068.

[6] Caputo, A. (1998). Some properties of the family of Koehler Symanowski distributions, The Collaborative Research Center (SBF) 386, working paper No. 103, LMU, München.

[7] Chen, X.; Fan, Y.; Patton, A. (2004). Simple tests for models of dependence between multiple financial time series, with applications to U.S. equity returns and exchange rates. Working Paper.

[8] Crowson, P.; Sampson, R. (2001). Managing Metal Price Risk with the London Metal Exchange. London Metal Exchange, London.

[9] Embrechts, P.; McNeil, A.; Straumann, D. (1999). Correlation: Pitfalls and Alternatives. *Risk*, 5, 69-71.

[10] Fischer, M.; Köck, C. (2007): Constructing and Generalizing Given Multivariate Copulas: a unifying approach, discussion paper 80/2007, University of Erlangen-Nürnberg.

[11] Frees, E.W.; Valdez, E.A. (1998). Understanding relationship using copulas. *North American Actuarial Journal*, 2(1), 1-25.

[12] 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.

[13] Joe. H. (1997). Multivariate Models and Dependence Concepts. *Monographs on Statistics and Applied Probability* No. 37, London, Chapman & Hall.

[14] Köhler, K. J.; Symanowski, J. T. (1995). Constructing Multivariate Distributions with Specific Marginal Distributions, *Journal of Multivariate Distributions* 55, 261-282.

[15] Liebscher, E. (2006). Modelling and estimation of multivariate copulas, Working paper, University of Applied Sciences Merseburg.

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

[17] Nelsen, R.B. (2006). An introduction to copulas. Springer Series in Statistics, Springer.

[18] Palmitesta, P.; Provasi, C. (2005). Aggregation of Dependent Risks using the Koehler-Symanowski Copula Function, *Computational Economics* 25, 189-205.

[19] Rosenblatt, M. (1952). Remarks on multivariate Transformation, *The Annals of Statistics* 23, 470-472.

[20] Savu, C.; Trede, M. (2006). Hierarchical Archimedean Copulas, Working Paper, University of Münster.

[21] Sklar, A. (1959). Fonctions de Répartition á n Dimensions et Leurs Marges. *Publications of the Institute of Statistics*, 8, 229-231.

[22] Wang, S. (1998). Aggregation of correlated risk portfolios: models and algorithms, Working paper, www.casact.org.

[23] Whelan, N. (2004). Sampling from Archimedean copulas, *Quantitative Finance* 4, 339-352.