Ann Oper Res (2014) 219:333–357
DOI 10.1007/s10479-011-0900-9

Portfolio value-at-risk estimation in energy futures markets with time-varying copula-GARCH model

Xun Fa Lu · Kin Keung Lai · Liang Liang

Published online: 17 May 2011

Abstract This paper combines copula functions with GARCH-type models to construct the conditional joint distribution, which is used to estimate Value-at-Risk (VaR) of an equally weighted portfolio comprising crude oil futures and natural gas futures in energy market. Both constant and time-varying copulas are applied to fit the dependence structure of the two assets returns. The findings show that the constant Student t copula is a good compromise for effectively fitting the dependence structure between crude oil futures and natural gas futures. Moreover, the skewed Student t distribution has a better fit than Normal and Student t distribution to the marginal distribution of each asset. Asymmetries and excess kurtosis are found in marginal distributions as well as in dependence. We estimate VaR of the underlying portfolio to be 95% and 99%, by using the Monte Carlo simulation. Then using backtesting, we compare the out-of-sample forecasting performances of VaR estimated by different models.

Keywords Risk management · Copulas · Value-at-Risk · Time-varying models · Backtesting

Supported by the National Natural Science Founation of China (Grant No. 70821001).

X.F. Lu · L. Liang
School of Business, University of Science and Technology of China, 96 Jinzhai Road, Hefei, China

L. Liang
e-mail: lliang@ustc.edu.cn

X.F. Lu · K.K. Lai
Department of Management Sciences, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong

K.K. Lai
e-mail: mskklai@cityu.edu.hk

X.F. Lu (✉)
Zhishan Building, CityU-USTC Joint Advanced Research Center, 166 Renai Road, Dushu Lake Higher Education Town, Suzhou, China
e-mail: xflu@mail.ustc.edu.cn

1 Introduction

Value at Risk (VaR) model is one of the most widely used risk measure and has become a standard measure to aggregate risks across different factors. VaR reduces the (market) risk associated with any portfolio to just one monetary amount, i.e. the loss associated to a given probability. The key issue for VaR is to construct an estimation model which can provide accurate VaR of financial assets. The existing estimation models can be categorized into two broad groups: non-parametric and parametric (also including semi-parametric) methods. The former group mainly includes historical simulation method. The main advantage of this method is that it makes no assumptions about risk factor changes being from a particular distribution. The latter group usually makes specific distributional assumptions on returns, such as a normal distribution, and therefore calculates (analytically or by simulation) the corresponding VaR. The methods in this group assume that an adequately parametric distribution assumption can achieve accurate VaR (for a survey of VaR methodologies, please see Jorion 2007). In this study, we focus on the latter group only.

However, in reality, financial assets' returns usually exhibit non-normality (Ang and Chen 2002). Hence, many studies in empirical finance literature have found that the multivariate normal distribution does not provide adequate estimation of VaR, and often underestimates VaR of portfolios of financial assets. Bastianin (2009) summarized that the data of financial assets' returns show at least two kinds of non-normal features. The first is that the joint distribution of financial assets' returns includes two types of asymmetries (Patton 2004). Distribution of individual stock returns exhibits skewness or asymmetry (Harvey and Siddique 2000) and the dependence between financial assets' returns is also asymmetric (Patton 2006a). The second non-normal feature of financial assets' returns is excess kurtosis, which means probability distributions for assets' returns often exhibit fatter tails than the standard normal distribution. As defined, VaR concentrates on tails of the distribution. It will be underestimated on the assumption of multivariate normal distribution. Then, alternative distributions, such as Student t distribution focusing fat tails (Jorion 2007), or skewed Student t distribution focusing skewness and fat tails (Hansen 1994; Hull and White 1998), have been proposed and applied to build risk models of financial returns.

To sum up, when VaR is calculated using parametric models, the assumption of joint multivariate modeling is crucial. In order to overcome these problems resulting from the assumption of normal distribution, and allowing for the increasing body of empirical evidence of the non-normality of financial assets' returns and successful applications of the copula theory in finance and other disciplines in recent decades, we resort to the copula theory to improve VaR forecasts.

The copula theory has seen rapid development in recent times. This is first due to Sklar (1959), who showed that any n-dimensional joint distribution function may be decomposed into its n marginal distributions, which completely describe the location, scale and shape parameters of the n variables, and a copula, which completely describes the dependence between the n variables, and proves the well-known Sklar theorem. Based on the Sklar theorem, Nelsen (1998), and Joe (1997) further developed the copula theory. Embrechts et al. (1999) first introduced this concept to the finance literature. Copulas have already been widely applied in various fields of finance, such as risk management, derivative asset pricing, option valuation and so on. However, the copula methodology used in initial researches only dealt with unconditional distributions. In other words, the earlier applications using the copula methodology just focus on constant scenarios and do not include time-varying features. Patton (2001) expanded the constant copula into the conditional copula by allowing the first and second conditional moments to vary on time. After the methodological expansion of Patton (2001), the conditional copula began to be used in finance, e.g., Jondeau and Rockinger (2006). For further applications of the copula theory, readers can refer to McNeil et al. (2005), and Stoyanov et al. (2010), which provide the basic concept, detailed analysis and varied perspectives of copula's usage within the field of financial risk management.

In this paper, we focus on using copula functions to estimate VaR of a portfolio comprising crude oil futures and natural gas futures. The futures of energy commodities have also become an important financial instrument. In these energy commodities, crude oil and natural gas are the most important products. The relationship between prices of crude oil and natural gas is complex due to their constant interactions. Therefore, it is very useful and interesting to investigate the relationship between crude oil prices and natural gas prices for participants in energy market. In this paper, we have investigated the diversification benefits of a portfolio comprising crude oil and natural gas futures traded on the New York Mercantile Exchange (NYMEX), using VaR measure based on copula.

Copula has also been used in portfolio risk management, including estimation of VaR. For example, Embrechts et al. (2003), and Cherubini et al. (2004) have explicitly used constant copulas to measure portfolio VaR under unconditional distributions. Afterward, Fantazzini (2008) expanded the constant copula to time-varying copula for estimating VaR under conditional distributions. However, few researchers have considered applying copulas to estimate VaR in energy market. Therefore, it is necessary to find more appropriate multivariate distributions to fit the time-series data of energy market. Copula can be a solution to this problem. For example, Bastianin (2009) has explicitly used copulas to estimate VaR in energy market. The results find that asymmetric copula models such as symmetrized Joe-Clayton (also called SJC by Patton 2006a) copula with Student t marginals deliver the best VaR forecasts. His findings also confirm the importance of non-normalities and asymmetries of log-returns distributions in energy market.

What we do in this paper is to estimate and compare the conditional portfolio VaR using Copula-GARCH models with different innovation distributions in energy market. Firstly, we use models that can capture empirically observed time-varying mean values and variances of energy returns, such as GARCH-type models, which have been successfully used to fit the time-series data in financial markets, and to effectively capture the main observed characteristics in financial markets, e.g. autoregression in mean, volatility clustering in variances, and leverage effects from exogenous information (Bollerslev et al. 1994), in the presence of skewness and kurtosis (Jondeau and Rockinger 2006). Then, we use copulas to describe the dependence structure between crude oil and natural gas futures. Also, we allow copula parameters to be time-varying, but do not apply different copulas during bear markets and bull markets, and during in-sample dataset and out-of-sample dataset. After constructing joint bivariate distributions of both (crude oil and natural gas) assets' returns using in-sample dataset and combining copulas and GARCH-type models, we employ the selected appropriate models to forecast one-day-ahead rolling-over VaR of the portfolio using Monte Carlo simulation. Finally, backtests and loss functions are used to test which model provides better forecasts of VaR.

The contribution of this paper is threefold. First, we use time-varying copula-GARCH models to estimate portfolio VaR in energy markets. The results show that non-normality and asymmetry are significant in crude oil and natural gas returns, and the dependence structure is time-varying. Time-varying copula-GARCH models do have important implications for improving portfolio VaR forecasts. However, forecasting performances of time-varying copula-GARCH models are, at best, limited in comparison with the constant copula-GARCH models. We find that the constant Student t copula-GARCH model can better fit the time series in energy market, which is consistent with the usual findings that the Student t copula often provides a much better fit to multivariate financial return data. Second, there exists significant skewness in marginal distribution, as well as in dependence structure. Therefore, the skewed Student t distribution is better fitted to selected dataset than the normal or Student t distribution. Lastly, leverage effect is also found in crude oil returns, but not in natural gas returns. This implies that bad news can cause larger volatilities than good news in crude oil returns.

The rest of this paper is organized as follows. Section 2 introduces the theory of copulas and marginal distribution modeling. Section 3 illustrates how to use copulas to forecast VaR by Monte Carlo simulation. Empirical results are presented in Sect. 4. Section 5 concludes.

2 Multivariate modeling using copulas

Copulas can be very useful to model risk assets in financial markets, because they can help users flexibly construct many multivariate distributions to fit financial assets, without being subject to the curse of dimensionality. Additionally, copulas can easily capture extreme dependencies, such as tail dependence, while the normal distribution assumes zero extreme dependence. For example, extreme co-movement events are often observed (Mendes and Souza 2004). This means that good and bad extreme events will have some relationship and are not completely independent. However, the correlation coefficient based on multivariate normal distribution cannot capture this type of dependence. Another important property of copulas is that the copula of the underlying random variables is invariant under non-linear strictly increasing transformations, such as transforming assets returns into log-returns in financial time series. However, linear correlation based on multivariate elliptical, especially normal distribution most frequently used in practice, cannot remain invariant under non-linear strictly increasing transformations. Therefore, as Embrechts et al. (2003) said, linear correlation is often a misunderstood measure of dependence even though it is used popularly.

The following sub-section introduces the concepts and properties of copulas in brief. For exact depiction, we must first define the notations. The variables of interest are X and Y and the conditioning variable is W, which may be vector. According to Patton (2006a), F_{XYW} is the joint distribution of (X, Y, W), $F_{XY|W}$ is the conditional distribution of (X, Y) given W and let the conditional marginal distributions of $X|W$ and $Y|W$ be denoted by $F_{X|W}$ and $F_{Y|W}$, respectively. Recall that $F_{X|W}(x|w) = F_{XY|W}(x, \infty|w)$ and $F_{Y|W}(y|w) = F_{XY|W}(\infty, y|w)$. In this paper distribution function F_{XYW} is sufficiently smooth for all required derivatives to exist, and $F_{X|W}$, $F_{Y|W}$, and $F_{XY|W}$ are continuous. For unconditional distribution, the notation is similar. In this case we simply ignore the conditioning variable. Throughout the paper, we adopt the usual convention of denoting cumulative distribution function (c.d.f.) of a random variable using an uppercase letter, and the corresponding density (p.d.f.) using the lowercase letter. Also, we denote the extended real line as $\bar{\Re} = \Re \cup \{\pm\infty\}$, random variable in upper case, X_t, and Y_t, and the corresponding realizations in lower case, x_t, and y_t.

2.1 Copula theory

For simplicity purposes, throughout this paper we limit the copulas to being only 2-dimensional. According to Nelsen (1998) and Joe (1997), a 2-dimensional unconditional copula can be defined as follows.

Definition 2.1 A two-dimensional copula is a bivariate cumulative distribution function (c.d.f.), C, with uniform distribution margins in $I = [0, 1]$, and the following properties:

1. Dom $C = I^2 = [0, 1]^2$
2. For every u, v in I,

$$C(u, 0) = C(0, v) = 0, \qquad C(u, 1) = u, \quad \text{and} \quad C(1, v) = v;$$

3. For every u_1, u_2, v_1, v_2 in I such that $u_1 \leq u_2$ and $v_1 \leq v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

Following the above definition, if F_X and F_Y are univariate c.d.f.s of X and Y, U and V are the "probability integral transforms (PIT)[1]" of X and Y, that is, $U = F_X(X)$ and $V = F_Y(Y)$, then $C(F_X(x), F_Y(y))$ is a bivariate c.d.f. with margins F_X and F_Y. Sklar's (1959) theorem provides the theoretical proof and interprets why the margins and the dependence structure can be separated. Because this paper focuses on the conditional copula, now we move on to extension of the Sklar's theorem (1959) for conditional distribution (Patton 2006a).

Theorem 1 *Sklar's theorem for conditional distributions*: *Let $F_{X|W}(\cdot|w)$ and $F_{Y|W}(\cdot|w)$ be the conditional distributions of $X|W = w$ and $Y|W = w$, respectively, $F_{XY|W}(\cdot, \cdot|w)$ be the joint conditional distribution of $(X, Y)|W = w$, and Ω be the support of conditioning variable W. Assume that $F_{X|W}(\cdot|w)$ and $F_{Y|W}(\cdot|w)$ are continuous in x and y for all $w \in \Omega$. Then there exists a unique conditional copula $C(\cdot, \cdot|w)$ such that*

$$F_{XY|W}(x, y|w) = C(F_{X|W}(x|w), F_{Y|W}(y|w)|w)$$
$$\forall (x, y) \in \bar{\Re} \times \bar{\Re} \text{ and each } w \in \Omega \tag{1}$$

Conversely, if we let $F_{X|W}(\cdot|w)$ be the conditional distribution of $X|W = w$, $F_{Y|W}(\cdot|w)$ be the conditional distribution of $Y|W = w$, and $\{C(\cdot, \cdot|w)\}$ be a family of conditional copulas measurable in w, then the function $F_{XY|W}(\cdot, \cdot|w)$ defined by (1) is a conditional bivariate distribution with conditional marginal distributions $F_{X|W}(\cdot|w)$and $F_{Y|W}(\cdot|w)$.

As Patton (2006a) indicated, converse of Sklar's theorem is the most interesting for multivariate density modeling. It implies that any two univariate distributions, of any type (not necessarily from the same family), may be linked together via any copula to define a valid bivariate distribution as long as the information set used is unchanged. With Sklar's theorem, the set of parametric bivariate distributions available in econometric modeling is increased substantially. Therefore, we can flexibly use many other multivariate distributions different from the multivariate normal distribution for modeling financial series.

As found by Patton (2006b), the multi-stage maximum likelihood estimator (MSMLE) for copula parameters in conditional distributions is asymptotically not less efficient than

[1] Fisher (1932) and Rosenblatt (1952) showed that random variables of PIT have the Unif(0, 1) distribution, regardless of original distributions. In other words, a random variable X with c.d.f. F_X can be transformed into a variable with Unif(0, 1) distribution, viz. $U = F_X(X)$. Conversely, if U is uniformly distributed over the interval [0, 1], then $X = F_X^{-1}(U)$ has c.d.f. F_X. The PIT is also usually used to generate random variables given their distribution functions (Embrechts et al. 2003) and to test the goodness-of-fit of hypothesized model (Dias 2004). For extension of the PIT theory to the time series case see Diebold et al. (1998).

the usual one-stage maximum likelihood estimator under standard conditions, but the multi-stage likelihood estimator is easier to implement and to overcome the problem of heavy computational burden caused by the large number of parameters. The multi-stage maximum likelihood method is also known as the Inference Functions for Margins (IMF) method. We denote the conditional joint distribution as $F_{XY|W}(x_t, y_t; \theta|w_{t-1})$, the conditional marginal distributions as $F_{X|W}(x_t; \varphi|w_{t-1})$ and $F_{Y|W}(y_t; \gamma|w_{t-1})$, and the copula as $C(u_t, v_t; \kappa|w_{t-1})$. The same pattern is followed for different densities. According to Sklar's theorem, $F_{XY|W}(x_t, y_t; \theta|w_{t-1}) = C(F_{X|W}(x_t; \varphi|w_{t-1}), F_{Y|W}(y_t; \gamma|w_{t-1}); \kappa|w_{t-1})$, where $\theta \equiv [\varphi', \gamma', \kappa']'$ is the set of all parameters of both marginal distributions and copula to be estimated, and w_{t-1} is the information set until time $t-1$. To simplify the notation, we suppress the conditioning variable(s). Therefore, (A.3) can be rewritten as follows:

$$\begin{aligned} L_{XY}(\theta) &= L_X(\varphi) + L_Y(\gamma) + L_C(\kappa) \\ &= \sum_{t=1}^{T} \log f_X(x_t; \varphi) + \sum_{t=1}^{T} \log f_Y(y_t; \gamma) + \sum_{t=1}^{T} \log c(F_X(x_t; \varphi), F_Y(y_t; \gamma); \kappa) \end{aligned} \tag{2}$$

where $\varphi \in \text{int}(\Phi) \subseteq \Re^p, \gamma \in \text{int}(\Gamma) \subseteq \Re^q, \kappa \in \text{int}(K) \subseteq \Re^r$ and so $\theta \equiv [\varphi', \gamma', \kappa']' \in \text{int}(\Theta) \equiv \text{int}(\Phi) \times \text{int}(\Gamma) \times \text{int}(K) \subseteq \Re^{p+q+r} \equiv \Re^s$, where $\text{int}(A)$ is the interior of set A.

Now we introduce some important properties of copula. The first is the tail dependence of copula, which measures the probability of having a high (low) extreme value of random variable Y, given that a high (low) extreme value of random variable X has occurred. In portfolio risk management, investors will select assets having lower correlations, especially in extreme values, to hedge the relationship risk and to get diversification benefits. Tail dependence can help investors select assets for a portfolio. According to Nelsen (1998), the lower and upper tail dependence coefficients are defined as follows.

$$\lambda^L = \lim_{\alpha \to 0^+} P(Y < F_Y^{-1}(\alpha) | X < F_X^{-1}(\alpha)) = \lim_{\alpha \to 0^+} \frac{C(\alpha, \alpha)}{\alpha} \tag{3}$$

$$\lambda^U = \lim_{x \to 1^-} P(Y > F_Y^{-1}(\alpha) | X > F_X^{-1}(\alpha)) = \lim_{x \to 1^-} \frac{1 - 2\alpha + C(\alpha, \alpha)}{1 - \alpha} \tag{4}$$

where α is the probability and $F_X^{-1}(\alpha)$ and F_Y^{-1} are 100α-th percentiles of F_X and F_Y, respectively, λ^L and $\lambda^U \in [0, 1]$ are lower tail dependence coefficient and upper tail dependence coefficient, respectively. For unification of notations, we substitute τ^L (τ^U) for λ^L (λ^U). Another key property of copula is that copula of the underlying random variables is invariant under non-linear strictly increasing transformations. Financial data usually exhibit autoregression and volatility clustering. The second property makes sure that the copula of underlying assets does not change when we filter the financial assets' returns using some models, like GARCH models.

2.2 Copula selection: goodness-of-fit tests

As noted by Durrleman et al. (2000), the obtained results might be very different if the copula selected is not suitable. Thus, the choice of copula that is going to fit the data is very important. The initial approach to select the copula was proposed by Deheuvels (1979), who constructed the so-called "empirical copula", though it is neither a copula nor exactly the same (except asymptotically), and considered that the best copula is the one that minimizes the distance between the empirical copula and the hypothesized copula. The discrete L^2

norm was chosen by him to measure the distance. Another suggestion is to use a criterion like Akaike's information criterion (AIC) (Akaike 1973) and Schwarz's Bayesian information criterion (SBIC) (1978), which are defined as:

$$\mathrm{AIC}(M) = -\frac{2LL}{n} + \frac{2M}{n} \tag{5}$$

$$\mathrm{SBIC}(M) = -\frac{2LL}{n} + \frac{M\log(n)}{n} \tag{6}$$

where M is the number of parameters being estimated and LL is the value of maximum likelihood function when parameters are optimal, and n is the number of observations. When we select the optimal copula function, M is the number of copula parameters if models for marginal distributions are considered, as known, and LL can be obtained from $\log c(u, v; \kappa)$. Similarly, when we select marginal distributions, M is the number of marginal parameters, and LL can be obtained from $\log f_X(x; \varphi)$ or $\log f_Y(y; \gamma)$. These approaches to select the underlying true copula are helpful to identify the appropriate copula, but they are not able to provide any understanding about the power of the decision rule employed. On the other hand, Goodness-of-fit (GOF) tests are proposed by many academics. These tests are able to either reject or fail to reject a parametric copula and are thus preferred (Berg and Bakken 2006). Genest et al. (2009) briefly reviewed GOF tests of copula models. According to their findings, a good combination of power and conceptual simplicity is provided by the Cramér-von Mises (CVM) statistic:

$$S_n = \sum_{t=1}^{n} \{C_\kappa(u_t, v_t; \hat{\kappa}) - C_n(u_t, v_t)\}^2 \tag{7}$$

This statistic measures how close the fitted copula $C_\kappa(u_t, v_t; \hat{\kappa})$ is from the empirical copula C_n, as modified by Fermanian (2005). Because the definition of S_n involves $\hat{\kappa}$, the distribution of this statistic depends on the unknown value of copula parameter κ under the null hypothesis that C is from the class C_κ. Thus the P-value of the test must be computed using a parametric bootstrap procedure described by Genest et al. (2009).

2.3 Marginal distribution modelling

GARCH-type models have been applied by many authors to fit univariate variables, to analyze and forecast volatility of financial time series data, because they can effectively capture the main observed characteristics in financial markets. Taking into account characteristics of return series in energy futures market, we use the classical GARCH model and the Threshold GARCH (TGARCH, also called GJR by Glosten et al. 1993) model to model univariate variables. Let the log-returns of a given asset be given by $r_t = \log(P_t) - \log(P_{t-1})$, $t = 1, \dots, T$. P_t denotes the price of a given asset at time t. We model each marginal time series by the general AR(1)-TGARCH(1, 1) model and AR(1)-GARCH(1, 1) model.

$$r_t = \mu + \phi r_{t-1} + \varepsilon_t \tag{8}$$

$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \psi\varepsilon_{t-1}^2 I_{t-1} + \beta h_{t-1} \tag{9}$$

$$\varepsilon_t = \eta_t\sqrt{h_t},\ \eta_t \overset{i.i.d}{\sim} f(0, 1) \tag{10}$$

where h_t denotes the conditional variance. ε_t is the innovation, or the residual, and η_t is the standard residual with zero mean and unit variance. $I_t = 1$ if $\varepsilon_{t-1} < 0$, and 0 otherwise.

Good news $\varepsilon_{t-1} > 0$ and bad news $\varepsilon_{t-1} < 0$ have different effects on conditional variance in TGARCH model. That is, good news has an impact of α, while bad news has an impact of $\alpha + \psi$. If $\psi \neq 0$, we say that the leverage effect exists, while if $\psi = 0$, the TGARCH model degenerates into GARCH model.

We also try other autoregression models to fit the mean model, such as ARMA(p, q), using the pre-estimation analysis. The results show that AR(1) is sufficient to filter autoregression of the target time series. We estimate the AR(1)-TGARCH(1, 1) model and the AR(1)-GARCH(1, 1) model assuming three different density functions $f(0, 1)$ for η_t: Normal, Student t and skewed Student t (Skew-T). The density function of the skewed Student t distribution, which is generalized by Hansen (1994) to contemporarily capture excess kurtosis and excess skewness, is given by:

$$\text{Skew-T} \quad f(z; \upsilon, \lambda) = \begin{cases} bc(1 + \frac{1}{\upsilon-2}(\frac{bz+a}{1-\lambda})^2)^{-(\upsilon+1)/2} & \text{if } z < -a/b \\ bc(1 + \frac{1}{\upsilon-2}(\frac{bz+a}{1+\lambda})^2)^{-(\upsilon+1)/2} & \text{if } z \geq -a/b \end{cases} \tag{11}$$

where $2 < \upsilon < \infty$ and $-1 < \lambda < 1$ denote the degree of freedom parameter and the asymmetry parameter, respectively. The constants a, b and c are given by $a = 4\lambda c(\frac{\upsilon-2}{\upsilon-1})$,

$$b = \sqrt{1 + 3\lambda^2 - a^2}, \qquad c = \frac{\Gamma(\frac{\upsilon+1}{2})}{\sqrt{\pi(\upsilon-2)}\Gamma(\frac{\upsilon}{2})}, \quad \text{respectively.}$$

By construction, Skew-T density has zero mean and one variance. Moreover, it encompasses a large set of conventional densities. For instance, if $\lambda = 0$, the Skew-T distribution reduces to the traditional Student t distribution, which is not skewed. While if $\lambda = 0$ and $\upsilon \to \infty$, it reduces to the normal density. Similar to the Student t, given the restriction $\upsilon > 2$, this distribution is well defined and its second moment exists, while the skewness exists only for $\upsilon > 3$ and the kurtosis exists only for $\upsilon > 4$. The parameter λ controls the skewness of the density, which has a single mode at $-a/b$. If $\lambda > 0$, the distribution has a positive skewness, viz., skewed to the right, while if $\lambda < 0$, the distribution has negative skewness, viz., skewed to the left.

3 VaR forecasting with copula-GARCH models

3.1 Introduction: value at risk

VaR is defined as the maximum expected loss which may be incurred by a portfolio over a target horizon and at a given level of confidence under normal financial market conditions. Let q denote the confidence level, and L the expected profit and loss. We define VaR as a positive number (negative loss). Then, the VaR of a given portfolio at time t (return from $t - \Delta t$ to t), with confidence level $1 - q$, is defined as:

$$\text{VaR}_t(q) = -\inf\{x | F_t(x) \geq q\} \tag{12}$$

where $F_t(x)$ is the c.d.f. of the portfolio P&L at time t. From (12), it can be seen that the choice of the distribution function is paramount to VaR calculation. Traditionally, the distribution is usually assumed to be normal distribution. However, it underestimates the probability in the tail. Furthermore, it underestimates the VaR. Therefore, we use a fat-tailed distribution, such as Student t and skewed Student t, in order to capture fat-tailness and skewness, to model the distribution of P&L.

3.2 VaR estimation

Our objective is to estimate one-day-ahead VaR of an equally weighted portfolio composed of two assets, crude oil futures and natural gas futures. We denote r_t^X and r_t^Y as the time series of daily log-returns for futures of crude oil and natural gas, respectively. Then, the portfolio returns are approximately equal to $0.5r_t^X + 0.5r_t^Y$. To estimate the portfolio VaR, we need to investigate the joint distribution of the vector (r_t^X, r_t^Y). In this paper, we use the copula-GARCH model constructed by the aforementioned procedures to fit the time series (r_t^X, r_t^Y). Because there are no analytic and easy-to-use formulae to switch from conditional mean and volatility to VaR of the portfolio, Monte Carlo simulation is employed to forecast VaR based on copulas (Fantazzini 2008; Bastianin 2009). We denote L_t as the P&L function of this portfolio, expressed as:

$$\begin{aligned} L_t &= \frac{1}{2}P_t^X + \frac{1}{2}P_t^Y - \left(\frac{1}{2}P_{t-1}^X + \frac{1}{2}P_{t-1}^Y\right) \\ &= \frac{1}{2}P_{t-1}^X(\exp(r_t^X) - 1) + \frac{1}{2}P_{t-1}^Y(\exp(r_t^Y) - 1) \end{aligned} \tag{13}$$

The procedures we use to forecast one-day-ahead VaR based on copulas at 95% and 99% confidence level are the following:

1. AR-GARCH-type models are fitted, and marginal distributions are estimated for each return series using T observations in Sect. 2.3;
2. One-step return means and variances are forecasted in time $T+1$, and denoted as $\hat{r}_{T+1}^i$ and $\hat{h}_{T+1}^i$, for $i = X, Y$;
3. We simulate $N = 10{,}000$ Monte Carlo scenarios over the time horizon $[T, T+1]$, using the conditional bivariate distribution modeled by copula-GARCH models.
   (a) Estimate copula parameters $\hat{\kappa}$ by PITs u_t and v_t of standardized residuals η_t^X and η_t^Y of AR-GARCH-type models.
   (b) Simulate j random variables (u_{T+1}^j, v_{T+1}^j), where $j = 1, \ldots, N$, from the copula function estimated in step (a). See Cherubini et al. (2004), for a discussion about copula simulation.
   (c) Obtain the (simulated) standardized residuals $\eta_{T+1}^{i,j}$ by using the inverse functions of the estimated marginals.

$$\left(\eta_{T+1}^{X,j}, \eta_{T+1}^{Y,j}\right) = \left(F_{X,T+1}^{-1}(u_{T+1}^j; \hat{\varphi}), F_{Y,T+1}^{-1}(v_{T+1}^j; \hat{\gamma})\right)$$

   (d) Get the (simulated) asset log-returns by using standardized residuals in step (c) and the forecasted means and variances from step 2.

$$\left(r_{T+1}^{X,j}, r_{T+1}^{Y,j}\right) = \left(\hat{r}_{T+1}^X + \eta_{T+1}^{X,j} \cdot \sqrt{\hat{h}_{T+1}^X}, \hat{r}_{T+1}^Y + \eta_{T+1}^{Y,j} \cdot \sqrt{\hat{h}_{T+1}^Y}\right)$$

   (e) Repeat steps (b)–(d) for N times, and calculate the values of L_{T+1}^j using (13) for $j = 1, \ldots, N$.
   (f) Sort the 10,000 values of L_{T+1}^j in increasing order, and the 95%, 99% VaR is simply calculated as:
      (i) 95% VaR is the absolute value of $10{,}000 \times (1 - 95\%) = 500$th ordered scenario of L_{T+1};
      (ii) 99% VaR is the absolute value of $10{,}000 \times (1 - 99\%) = 100$th ordered scenario of L_{T+1}.

4. We repeat steps 1–3 M times by rolling over the daily returns for 2 year periods starting from Jan 2, 2008 to Dec 31, 2009 with one day increment. For example, we use the dataset from Jan 2, 1998 to Dec 31, 2007 to forecast the next trading day VaR (Jan 2, 2008). Then we add the price in Jan 2, 2008 to the dataset and delete the price on Jan 2, 1998, so that we assure that the dataset always has T observations. M is the number of out-of-sample instances. The results of this step are used for backtesting VaR.

Note that the simulation times N is a critical variable when using this procedure. Obviously, the larger N is, the more accurate the VaR will be. However, the simulation can be very time-consuming, especially when conducted in a recursive or rolling forecasting scheme (Bastianin 2009). Fantazzini (2008) suggests a choice of 100,000 simulations, while Bastianin (2009) suggests only 5,000. Considering the accuracy of VaR and CPU time, we select 10,000 simulations, which represent a good compromise between accuracy and speed.

3.3 VaR evaluation: backtesting

After forecasting VaR of each day in the out-of-sample (from the 2494 to 2999 observations) using the procedure described in Sect. 3.2, we compare these forecasted VaRs with the real observed portfolio P&L and then evaluate the performance of the constructed models using backtesting techniques. In this paper, we apply two statistical tests and three loss functions to backtest the performance of different VaR models.

3.3.1 Statistical tests

We first define $\{I_t\}_{t=1}^T$ as the hit series. $I_t = 1$ when the value of observed P&L function is less than the negative forecasted VaR threshold, and 0 otherwise. q is defined as the true probability coverage. Let $Z = \sum_{t=1}^T I_t$ be the number of exceptions in a sample of size T.

The first statistical test is the Kupiec's unconditional coverage test. This test, proposed by Kupiec (1995), tests the difference between the observed and the expected number of VaR exceptions of the effective portfolio profits and losses. The test of the null hypothesis that the observed exception frequency Z is equal to the expected exceptions is given by a likelihood ratio (LR) test statistic:

$$LR_{UC} = -2\log\bigl[(1-q)^{T-Z}q^Z\bigr] + 2\log\bigl[(1-Z/T)^{T-Z}(Z/T)^Z\bigr] \tag{14}$$

which, when T is large enough, is asymptotically distributed as $\chi^2(1)$ under H_0. This test can reject a model for both high and low failures. However, it ignores conditioning, or time variation in the data, and cannot cope with the case where observed exceptions cluster together in time (Jorion 2007).

To overcome the shortcomings of Kupiec's unconditional coverage test, Christoffersen (1998) developed the conditional coverage test, which tests the joint assumption of unconditional coverage and independence of failures. His statistic can be expressed as:

$$\begin{aligned} LR_{CC} &= LR_{UC} + LR_{IND} \\ &= -2\log\bigl[(1-q)^{T-Z}q^Z\bigr] + 2\log\bigl[(1-\pi_{01})^{n_{00}}\pi_{01}^{n_{01}}(1-\pi_{11})^{n_{10}}\pi_{11}^{n_{11}}\bigr] \end{aligned} \tag{15}$$

where n_{ij} is the number of observations with value i followed by j for $i, j = 0, 1$ and $\pi_{ij} = n_{ij}/\sum_j n_{ij}$ are corresponding probabilities. This test is distributed as $\chi^2(2)$ under H_0.

3.3.2 Loss functions

All statistical tests discussed above have focused on examining the behavior of the hit function. These tests choose acceptable models on the basis of number of exceptions, while information contained in the number of exceptions is limited. For example, one might be interested in the magnitude of the exception rather than simply whether or not an exception occurred (Campbell 2006). Moreover, previous statistical tests do not show any power in distinguishing among different, but close, alternatives. These problems in previous statistical tests result in construction of a general loss function. The first loss function, suggested by Lopez (1998) as an alternative to the approach that focuses exclusively on the hit series, can be written as:

$$C_t^L = \begin{cases} 1 + (|L_t| - \mathrm{VaR}_t)^2 & \text{if } L_t < -\mathrm{VaR}_t \\ 0 & \text{if } L_t \geq -\mathrm{VaR}_t \end{cases} \tag{16}$$

This measure includes an additional term based on the magnitude of an exception, except for the score of one when an exception occurs. A backtest using this loss function would typically be based on the sample average loss,

$$\hat{C}^L = \frac{1}{T} \sum_{t=1}^{T} C_t^L$$

Blanco and Ihle (1999) proposed an alternative loss function by focusing on the average size of exceptions:

$$C_t^{BI} = \begin{cases} \frac{|L_t| - \mathrm{VaR}_t}{\mathrm{VaR}_t} & \text{if } L_t < -\mathrm{VaR}_t \\ 0 & \text{if } L_t \geq -\mathrm{VaR}_t \end{cases} \tag{17}$$

The backtest using this loss function is similar to Lopez's loss function.

The third loss function is based on the quantile estimation by González-Rivera et al. (2004):

$$Q = T^{-1} \sum_{t=1}^{T} (q - I_t)(L_t - \mathrm{VaR}_t) \tag{18}$$

The criterion (using these three loss functions) to evaluate the performance of different VaR models is that a smaller $\hat{C}$ or Q indicates a better goodness-of-fit. When we backtest the performance of different VaR models using all of tests mentioned above, there is a general strategy to choose the best VaR model: first, we use the statistical tests to choose the best models; second, we apply the loss functions to compare the costs of different admissible choices.

4 Empirical application

In this section we apply the theory presented to forecast VaR of an equally weighted portfolio comprising crude oil futures and natural gas futures. By comparing performances of different models, we focus on influences of marginal distributions and copulas on VaR.

Table 1 Descriptive statistics

	Full sample		Estimation sample		Forecasting sample	
	Crude oil	Natural gas	Crude oil	Natural gas	Crude oil	Natural gas
Mean	0.000505	0.000317	0.000684	0.000500	−0.000377	−0.000584
Std. Dev.	0.026397	0.038411	0.023749	0.038193	0.036797	0.039497
5% VaR	0.040970	0.058684	0.037274	0.058373	0.059956	0.059962
1% VaR	0.076821	0.094051	0.057966	0.096303	0.103435	0.088520
Skewness	−0.123276	0.593258	−0.319427	0.479783	0.196153	1.105355
Kurtosis	6.817222	8.192984	6.170075	8.199213	5.476272	8.205871
Jarque-Bera	1828.387	3545.681	1086.710	2904.739	132.2643	673.0879
P-value	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
ADF	−41.2367	−59.0726	−49.9049	−52.9548	−23.7606	−26.2611
P-value	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
PP	−55.9781	−59.1067	−50.1489	−53.0229	−23.9058	−26.1164
P-value	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Correlation	0.283192		0.287260		0.287763	

4.1 Data description

We examine VaR of an equally weighted portfolio comprising crude oil and natural gas futures traded on the NYMEX. We collected closing futures prices covering the period from January 5, 1998 to December 31, 2009, with 3000 daily observations. Actually, contracts with different maturities are traded on the NYMEX. Here, we selected one-month maturity contracts (marked Contract 1 on the NYMEX; it is a futures contract specifying the earliest delivery date) as the target dataset. In what follows, returns are classically represented by changes in logarithms of price values, with $T = 2999$ log-returns. We use 2494 observations (from January 5, 1998 to December 31, 2007) to estimate the models, and reserve the last 505 observation (from January 2, 2008 to December 31, 2009) for out-of-sample evaluation of the models. The descriptive statistics of the two return series are presented in Table 1.

First, we can note that average returns of crude oil futures contracts are slightly higher than natural gas futures contracts over three periods, while volatilities are the opposite. But they are very close. This is different from the usual phenomenon that assets with high returns are accompanied by high risk. This somewhat implies that crude oil futures offer returns that are superior to natural gas futures. In this study, we focus on the relationship between these two assets. Meanwhile, notice that average returns in the full sample and the estimation sample are positive, while they are negative in the forecasting sample. The changes in comparative average returns of in-sample and out-of-sample periods suggest that allowing for structural breaks in the returns-generating process may improve portfolio decisions (Patton 2004). Due to computational constraints, we allow for no structural breaks in average returns. The skewness of crude oil in full sample and estimation sample is negative, while skewness is positive in the forecasting sample. Natural gas futures exhibit positive skewness in three periods. Both time series exhibit excess kurtosis. The Jarque-Bera statistic significantly rejects the null hypothesis of unconditional normality. Also, the unconditional correlation coefficient shows a positive degree of linear dependence. The empirical 5% and 1% VaRs are defined as the negative of the fifth and first empirical percentile of returns, that is, $\mathrm{VaR}(X; 0.05 \text{ or } 0.01) \equiv -\hat{F}_n^{-1}(0.05 \text{ or } 0.01)$, where $\hat{F}_n$ is the empirical distribution of

Table 2 Empirical exceedance correlation

	F-test[(1)]	P-value	F-test[(2)]	P-value	$\hat{\rho}(c)^- - \hat{\rho}(c)^+$			
					$c_1 = 0.0$	$c_2 = 0.5$	$c_3 = 1.0$	$c_4 = 1.5$
Full-sample	135.1631	0.0000	130.8226	0.0000	−0.010319	0.027300	0.033873	0.081916
In-sample	115.8545	0.0000	113.7120	0.0000	−0.344999	0.069498	0.058760	0.097934
Out-of-sample	23.8898	0.0000	23.1966	0.0000	−0.197679	0.003777	0.065353	

Notes: Columns (Rows) 1–4 display F-tests and the associated P-values based on the following regressions:

$^{(1)}r_t^X = \alpha_0 + \alpha_1 r_t^Y \cdot I_{(r_t^Y > 0)} + \alpha_2 r_t^Y \cdot I_{(r_t^Y \le 0)} + \xi_t$; $H_0 : \hat{\alpha}_1 = \hat{\alpha}_2$

$^{(2)}r_t^Y = \beta_0 + \beta_1 r_t^X \cdot I_{(r_t^X > 0)} + \beta_2 r_t^X \cdot I_{(r_t^X \le 0)} + \varepsilon_t$; $H_0 : \hat{\beta}_1 = \hat{\beta}_2$

returns using n observations. The ADF (Augmented Dickey-Fuller) and PP (Philips-Perron) are unit root tests used to test the null hypothesis of non-stationarity of time series. That P-values for all time series are less than 0.05 indicates that every series is stationary.

To examine the existence of this asymmetric dependence between crude oil futures and natural gas futures, we use the measures of Hong et al. (2007), called "exceedance correlation", which is defined to be between random variables X and Y as $\rho^e(q)$:

$$\rho^e(q) = \begin{cases} \rho^- = \operatorname{corr}(X, Y | X \le Q_x(q) \cap Y \le Q_y(q)) & \text{for } q \le 0.5 \\ \rho^+ = \operatorname{corr}(X, Y | X > Q_x(q) \cap Y > Q_y(q)) & \text{for } q \ge 0.5 \end{cases}$$

where $Q_x(q)$ and $Q_y(q)$ are q-th quantiles of X and Y. If X and Y are standardized returns, we can examine the degree of asymmetry in their unconditional copula using the exceedance correlation after having removed all asymmetries in marginal distributions. Then, the exceedance correlation is defined as:

$$\rho^e(c) = \begin{cases} \rho^- = \operatorname{corr}(X, Y | X \le c \cap Y \le c) & \text{for } c \le 0 \\ \rho^+ = \operatorname{corr}(X, Y | X > c \cap Y > c) & \text{for } c \ge 0 \end{cases}$$

Additionally, we use some regression equations to test the null hypothesis of symmetric correlation. Results are shown in Table 2. Obviously, both F-tests based on raw log-returns reject the null hypothesis of symmetric correlation. Also, exceedance correlations based on transformed standardized residuals for two indices are not equal at exceedance levels $c = 0, 0.5, 1.0, 1.5$. These results confirm that the asymmetry in the dependence exists even after removing all asymmetries in marginal distributions.

4.2 Marginal distributions modelling

In what follows, we multiply log-returns by 100 in order to get larger variances and parameters. As shown in Table 1, each series is stationary at 0.05 significance level. This implies that these series are appropriate to be fitted by AR-GARCH-type models. Using the correlogram analysis, it is found that autocorrelation and partial autocorrelation exist in both in-sample series. The Q-statistic of Ljung-Box test of crude oil futures on the 34th lag is about 47, with corresponding p-value 0.075. This implies that the null hypothesis is rejected at 0.1 level of significance and confirms the existence of autocorrelation. The Q-statistic of natural gas futures on the 1st lag is about 8.7, with corresponding p-value 0.003. This implies the autocorrelation coefficient is non-zero at the 0.05 level of significance. To fit the autocorrelation of both time series, we use the first order autoregressive model (AR(1)) for the conditional mean of log-returns of each series. Model test shows that the AR(1) model is sufficient to fit the mean of each series. After fitting the mean model, we use the Lagrange multiplier (LM) test of Engle (1982) to examine whether residuals of both in-sample series have heteroscedasticity. Obviously, ARCH effects of higher lags exist in both in-sample series. This implies that we can use GARCH-type models to capture the heteroscedasticity of the residuals of mean equations of the two in-sample series.

Models employed for marginal distributions are the AR(1)-TGARCH(1, 1) model and the AR(1)-GARCH(1, 1) model, assuming three different density functions $f(0, 1)$: Normal, Student T and Skew-T, given by (8), (9) and (10).

First, we allow for determining whether AR(1)-TGARCH(1, 1) or AR(1)-GARCH(1, 1) is more appropriate to both univariate time series under the same density functions of innovations. Results are listed in Table 3.

The KS tests reject the null hypothesis of normality, but do not reject the null hypothesis of Student t, and Skew-T distribution of either of the returns. This implies that the filtered standardized residuals are also non-normal. The LR tests show that the null hypothesis of no significance of restriction is rejected in case of the same innovation in crude oil. This implies that TGARCH models are superior to GARCH models with the same innovation. Therefore, we use TGARCH models to fit crude oil returns. The results are opposite in natural gas, where the null hypothesis of no significance of restriction can't be rejected in the same innovation case. Considering the no significance of leverage effect in natural gas, we use GARCH models to fit its returns. Also, the results of log-likelihood values AIC and SBIC show that the TGARCH model with Skew-T innovation is always superior to other models in crude oil, while GARCH model with Skew-T innovation performs the best in case of natural gas. Therefore, we select the AR(1)-TGARCH(1, 1) with Skew-T innovation to fit the marginal distribution of crude oil futures, and the AR(1)-GARCH(1, 1) with Skew-T innovation to fit the marginal distribution of natural gas futures. Parameter estimates for both marginal distributions are listed in Table 4.

Table 4 shows that the Ljung-Box test used to examine autocorrelation of standardized residuals of four GARCH models does not reject the null hypothesis of autocorrelation at lags 1, 5 and 10 at 5% significance level. The Engle-test, or LM test, is applied to the square of standardized residuals of all models. The p-values show that these models do not reject the null hypothesis of ARCH effects at lags 1, 5 and 10 at 5% significance level, except the TGARCH model with skewed-t innovations for crude oil futures. Although the TGARCH model with skewed-t innovations for crude oil futures rejects the null hypothesis of ARCH effects at lag 5 at 5% significance level, it does not reject the null hypothesis of ARCH effects at lag 5 at 1% significance level and at lag 10 at 5% significance level. Also, allowing for degrees of freedom ν and skewness λ is significant at 5% significance level, we consider the model to be adequate. The leverage effect is statistically significant in crude oil, but not in natural gas. This implies that bad news will cause larger volatility in crude oil, but news, bad or good, will bring an asymptotically symmetric effect on natural gas. In a word, we consider all the models are adequately fitted to marginal distributions of both time series.

4.3 Copulas modelling

After having estimated parameters of marginal distributions $\{F_{X,t}, F_{Y,t}\}$, using (A.4) and (A.5) in the first step, the following step estimates copula parameters using (A.6).

First, we use the seven constant copulas, Gaussian copula, Student t copula, Clayton copula, Rotated-Clayton copula, Gumbel copula, Rotated Gumbel copula, and SJC copula,

Table 3 Comparison between AR(1)-GARCH(1, 1) and AR(1)-TGARCH(1, 1)

Models	KS Test	LL	AIC	SBIC	LR Test	KS Test	LL	AIC	SBIC	LR Test
Crude oil for normal innovations						Natural gas for normal innovations				
GARCH	0.03156	−5629.4	4.52302	4.53181		0.04184	−6727.9	5.40145	5.41312	
	(0.013596)					(0.000310)				
TGARCH	0.02772	−5623.6	4.51636	4.53037	11.43	0.04230	−6727.5	5.40195	5.41596	0.7457
	(0.042556)				(0.000724)	(0.000256)				(0.387835)
Crude oil for Student t innovations						Natural Gas for Student t innovations				
GARCH	0.02063	−5578.6	4.48023	4.49424		0.01483	−66412.0	5.33332	5.34733	
	(0.236583)					(0.640331)				
TGARCH	0.02065	−5574.1	4.47739	4.49373	9.09	0.01510	−6641.9	5.33408	5.35043	0.0952
	(0.235637)				(0.002570)	(0.617289)				(0.757720)
Crude oil for Skew-T innovations						Natural gas for Skew-T innovations				
GARCH	0.01448	−5574.6	4.47778	4.49413		0.01179	−6638.5	5.33131	5.34766	
	(0.669716)					(0.877079)				
TGARCH	0.01425	−5569.2	4.47425	4.49293	10.80	0.01193	−6638.4	5.33206	5.35075	0.1273
	(0.689229)				(0.001013)	(0.868215)				(0.721239)

Notes: Values in parentheses are p-values. KS (Kolmogorov-Smirnov) test tests the null hypothesis that standardized residuals of GARCH-type models are from a specified distribution. LR (Likelihood Ratio) test compares specifications of nested models by assessing the significance of restrictions to an extended model with unrestricted parameters. LL is the log-likelihood value of specified model

Table 4 Parameter estimates for marginal distributions and statistic tests

Parameter	Crude oil: TGARCH-Skewed-T		Natural gas: GARCH-Skewed-T	
	Value	P-value	Value	P-value
μ	0.069705	0.131480	0.051831	0.196417
ϕ	−0.000191	0.498680	−0.059041	0.010578
ω	0.128463	0.004673	0.320922	0.000127
α	0.003943	0.291981	0.069986	0.000000
γ	0.043079	0.001197	/	/
β	0.949914	0.000000	0.908848	0.000000
υ	8.005057	0.000000	6.196872	0.000000
λ	−0.089351	0.001049	0.070105	0.003692
Ljung-Box test				
Lags	Q-stats.	P-value	Q-stats.	P-value
1	0.112582	0.737223	1.194217	0.274481
5	3.056530	0.691271	3.132940	0.679499
10	5.989665	0.816131	6.750912	0.748732
Engle test				
LM(1)	2.130321	0.144410	0.988648	0.320073
LM(5)	12.598090	0.027451	3.538283	0.617603
LM(10)	15.849887	0.104011	7.785774	0.649753

Notes: The Ljung-Box test is used to examine the autocorrelation of the standardized residuals of GARCH models. The Engle test is used to examine the ARCH effects of the square of standardized residuals series for all models

Table 5 Constant copula specification and estimation

Model	Parameter	LL	AIC	SBIC	Upper tail	Lower tail
Normal	0.330215	143.8992	−0.114640	−0.112305	0	0
Student's t: Corr.	0.334497	145.6411	−0.115236	−0.110565	0.000521	0.000521
Degree of freedom	29.2376					
Clayton	0.379802	102.5989	−0.081507	−0.079172	0	0.161055
Rotated Clayton	0.387761	107.1606	−0.085167	−0.082832	0.167226	0
Gumbel	1.239475	122.6218	−0.097571	−0.095236	0.250689	0
Rotated Gumbel	1.238084	121.5453	−0.096707	−0.094372	0	0.249589
SJC—Upper Tail	0.146384	131.0582	−0.102526	−0.098866	0.146384	0.130251
SJC—Lower Tail	0.130251					

Notes: The table shows the estimators of constant parameters of seven copulas, based on Skew-t marginals for crude oil and natural gas futures. LL is the copula log-likelihood at the optimum. Also presented are values of the Akaike information criteria (AIC) and the Schwarz's Bayesian information criteria (SBIC) at the optima

to fit standardized residuals of the best pair of marginal distributions obtained in Sect. 4.2 (in Table 5). For the functional forms of these seven copulas, readers can refer to Nelsen (1998), Patton (2006a, 2006b), and references therein.

Recall from Table 2 that most empirical lower exceedance correlations are slightly larger than the empirical upper exceedance correlations at exceedance levels $c = 0, 0.5, 1.0, 1.5$. However, to our surprise, the SJC copula tells something different, the coefficient of upper tail dependence is slightly larger than the coefficient of lower tail dependence. This is not consistent with common findings in equities markets that stock returns are more correlated with market downturns than market upturns. This may be determined by the unique characteristics of energy market. For instance, both crude oil and natural gas are used to generate electricity and for heating. Interestingly, it is found that copulas with larger upper tail dependence or symmetric tail dependence are always superior to those with larger lower tail dependence according to the maximized log-likelihood values, AIC and SBIC. The first copula is Student t copula, which is consistent with the usual findings that the Student t copula often provides a much better fit of multivariate financial return data (Grégoire et al. 2008). The Normal copula ranks the second order, followed by SJC and Gumbel copulas. The worst copulas are Clayton copula, ranked the last one, rotated Clayton copula, and rotated Gumbel copula. Based on these results, we select the four best copulas, Student t, Normal, SJC, and Gumbel copula, to forecast VaR of the portfolio composed of crude oil futures and natural gas futures.

As found by Patton (2006a), the dependencies among assets are usually time-varying. He assumes that the parameters of copulas can be modeled as the ARMA(1, 10) process, that is copula parameters are determined by past information such as previous parameters and historical asset log-returns. Therefore, the time-varying parameters of Normal copula and Student t copula are given by:

$$\rho_t = \tilde{\Lambda}\left(\omega_N + \beta_N \rho_{t-1} + \alpha_N \frac{1}{10}\sum_{j=1}^{10} \Phi^{-1}(u_{t-j}) \cdot \Phi^{-1}(v_{t-j})\right) \tag{19}$$

$$\rho_t = \tilde{\Lambda}\left(\omega_T + \beta_T \rho_{t-1} + \alpha_T \frac{1}{10}\sum_{j=1}^{10} T^{-1}(u_{t-j}; \upsilon) \cdot T^{-1}(v_{t-j}; \upsilon)\right) \tag{20}$$

where $\tilde{\Lambda}(x) \equiv (1 - e^{-x})(1 + e^{-x})^{-1} = \tanh(x/2)$ is the modified logistic transformation, designed to keep ρ_t in $(-1, 1)$ at all times (Patton 2006b). Φ^{-1} is the inverse c.d.f. of a standard normal distribution. T_υ^{-1} is the inverse c.d.f. of a Student t distribution. The degrees of freedom parameter υ in the Student t copula was assumed to be constant for simplicity, like the degrees of freedom were estimated in constant Student t copula.

For the Gumbel copula, the time-varying dependence processes are described as:

$$\delta_t = 1 + \left(\omega_G + \beta_G \delta_{t-1} + \alpha_G \frac{1}{10}\sum_{j=1}^{10} |u_{t-j} - v_{t-j}|\right)^2 \tag{21}$$

It is similar to time-varying parameters of the Normal copula, and Student t copula, while the forcing variable is the mean absolute difference between u_t and v_t over the previous 10 observations.

For SJC copula, the time-varying upper and lower tail dependencies are defined as:

$$\tau_t^U = \Lambda\left(\omega_U + \beta_U \tau_{t-1}^U + \alpha_U \frac{1}{10}\sum_{j=1}^{10} |u_{t-j} + v_{t-j}|\right) \tag{22}$$

$$\tau_t^L = \Lambda\left(\omega_L + \beta_L \tau_{t-1}^L + \alpha_L \frac{1}{10}\sum_{j=1}^{10} |u_{t-j} + v_{t-j}|\right) \tag{23}$$

where $\Lambda(x) = (1 + e^{-x})^{-1}$ is the logistic transformation to keep τ^U and τ^L in (0, 1).

We fit these four time-varying copulas to standardized residuals of AR(1)-TGARCH(1, 1) for crude oil futures, and AR(1)-GARCH(1, 1) for natural gas futures, both with skewed Student t innovations. The results are listed in Table 6.

According to the maximized log-likelihood values, AIC and SBIC, the best copula is the time-varying Student t copula. The time-varying Normal copula ranks No. 2. The third best copula is the time-varying SJC copula according to LL and AIC, while it is the fourth best according to SBIC. Comparing Table 5 with Table 6, it can be found that time-varying copulas always do better than their corresponding constant copulas. This implies that dynamics of copula parameters do exist and have important effects on the GOF of copulas to the two energy commodities. To see clearly this dynamics of copula parameters, we calculate the implied time path of conditional dependence between these two assets, and present the results in Figs. 1–4.

The four figures show that copula parameters significantly change over time. In Figs. 1 and 2, time paths of conditional linear correlation of Normal copula and Student t copula are very similar. Parameter δ of Gumbel copula also changes from 1.0001 to 1.5939 in Fig. 3. Figure 4 confirms that the change in linear correlation also takes place in tail dependence. This figure also shows that the upper tail dependence is mostly larger than the lower tail dependence. The difference between the upper and lower conditional tail dependence ranges from −0.0617 to 0.2854. Overall, the dynamics of copula parameters need to be considered when modeling portfolio risk in energy markets.

As noted above, the maximized log-likelihood values, AIC and SBIC, are helpful to identify the optimal copula, but they are not able to provide any understanding about the power of the decision rule employed. Therefore, we further use the GOF tests to identify the optimal copula. The Cramér-von Mises (CVM) statistic and its P-value are presented in Table 7.

In Table 7, the one-sided P-values show that all of these copula models are rejected at 0.05 significance level, but three models, constant Student t, and Normal, and time-varying Student t copulas, pass the goodness-of-fit tests at 0.01 significance level. In summary, the student t copulas, constant and time-varying, usually provide an adequate description of the dependence between crude oil and natural gas, which is consistent with results obtained from maximized log-likelihood values, AIC and SBIC. However, the differences among

Table 6 Time-varying copula specification and estimation

Model	ω	β	α	LL	AIC	SBIC
Normal	0.020695	1.947384	0.085363	154.7869	−0.113036	−0.106031
Student's t	0.021954	1.947513	0.082844	156.0728	−0.114433	−0.107428
Gumbel	1.017148	−0.096158	−1.547898	136.9449	−0.095966	−0.088961
SJC—Upper Tail	1.588172	−0.791271	−13.207845	146.8771	−1.003275	−0.086317
SJC—Lower Tail	0.752615	−1.856498	−9.381605			

Notes: The table shows the estimators of time-varying parameters of four copulas, based on skewed-t marginals for crude oil and natural gas futures

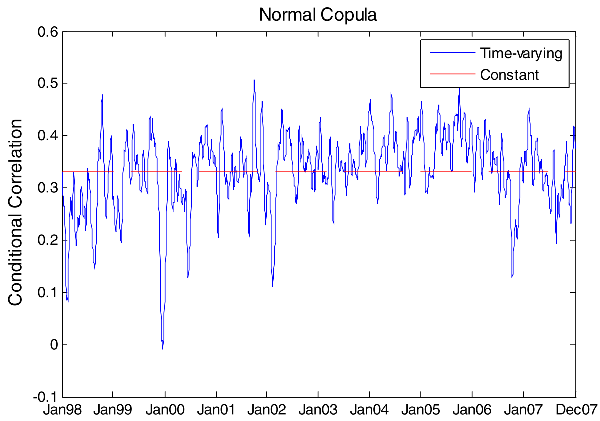

Fig. 1 Conditional correlation estimates from the Normal copula

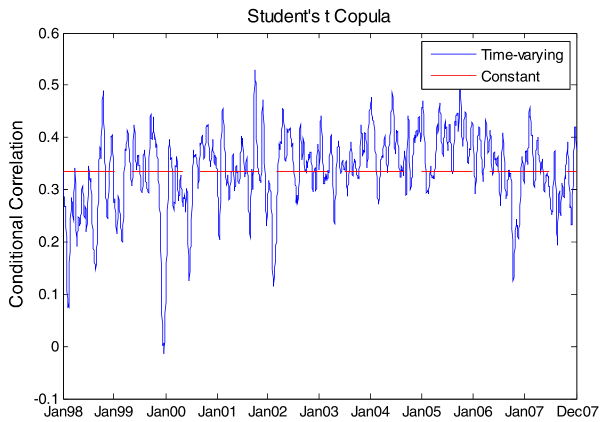

Fig. 2 Conditional correlation estimates from the Student *t* copula

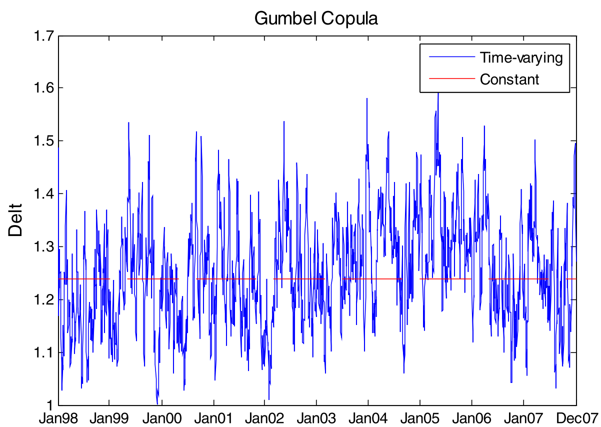

Fig. 3 Conditional Delt estimates from the Gumbel copula

these models are still limited, because the CVM statistics are very similar. Therefore, we take all of these copulas into account, and compare their performances in forecasting VaR of the portfolio concerned in the following section.

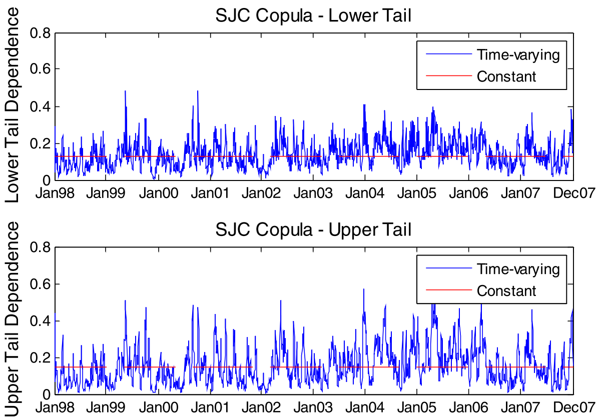

Fig. 4 Conditional lower and upper tail dependence estimates from the SJC copula

Table 7 The goodness-of-fit tests for different copula models

Constant copulas	Normal	Student's *t*	Gumbel	SJC
CVM stat.	0.053620	0.056076	0.143249	0.138600
P-value	0.014310	0.038246	0.000000	0.000000
Time-varying copulas	**Normal**	**Student's *t***	**Gumbel**	**SJC**
CVM stat.	0.167089	0.173276	0.349284	0.352775
P-value	0.002567	0.012931	0.000000	0.000573

4.4 Estimation of VaR

Tables 8 and 9 exhibit forecasting performances of different copula-GARCH models with Skew-T innovations using the out-of-sample dataset at 95% and 99% confidence levels. Table 8 shows the ratio of VaR exceedances Z/T, and two statistical tests. It first can be seen that the ratios of VaR exceedances are very close. Obviously, constant and time-varying Gumbel copula models fail to pass the unconditional coverage at 95% confidence level, although they pass all other statistical tests. This means Gumbel is not appropriate to fit the dependence structure between crude oil futures and natural gas futures. This finding is consistent with results obtained in Sect. 4.3. The remaining copula models pass all statistical tests. Using these statistical tests, it is difficult to determine which copula is better than other copula models. Thus we allow for loss functions listed in Table 9. Using the general strategy to choose the better VaR model mentioned in Sect. 3, we first omit Gumbel copula models. However, the results are not robust enough to choose the best copula model using the loss functions. We calculate the average score of each model based on rankings to evaluate forecasting performances of different models. That is, if a model ranks No. 1 using a certain loss function, we give it a score of 8; if it ranks No. 2, we give it a score of 7, and so on. Then the best model is the one that has the highest average scores. According to the average scores, the best model is the constant Normal copula, followed by the constant SJC copula and constant Student *t* copula. Notice that time-varying copulas usually perform worse than constant copulas. These results are opposite to those of GOF tests where time-varying copulas fit better than constant copulas to the dependence structure between crude oil and

Table 8 Backtesting of VaR forecasts: statistical test

Copulas	95%			99%		
	Z/T	LR_{UC}	LR_{CC}	Z/T	LR_{UC}	LR_{CC}
Constant:						
Normal	0.063366	1.757644	2.355868	0.005941	0.983739	1.031595
P-value		0.184919	0.307914		0.321278	0.597024
Student's *t*	0.065347	2.292848	2.442138	0.003960	2.413605	2.437486
P-value		0.129971	0.294915		0.120285	0.295602
Gumbel	0.073267	5.064949	5.250837	0.007921	0.237453	0.317374
P-value		0.024414	0.072409		0.626052	0.853263
SJC	0.065347	2.292848	2.442138	0.003960	2.413605	2.437486
P-value		0.129971	0.294915		0.120285	0.295602
Time-varying:						
Normal	0.063366	1.757644	2.355868	0.005941	0.983739	1.031595
P-value		0.184919	0.307914		0.321278	0.597024
Student's *t*	0.067327	2.892905	3.263401	0.007921	0.237454	0.317374
P-value		0.088970	0.195597		0.626052	0.853263
Gumbel	0.069307	3.556040	3.844677	0.007921	0.237453	0.317374
P-value		0.059329	0.146265		0.626052	0.853263
SJC	0.067327	2.892905	3.263401	0.005941	0.983739	1.031595
P-value		0.088970	0.195597		0.321278	0.597024

Notes: The table shows results of three statistical tests, based on skewed-t marginals for crude oil futures and natural gas futures, at 95% and 99% confidence level. Z/T denotes the ratio of VaR exceedances

natural gas futures. This implies that the time-varying parameters of copulas have, at best, limited effects on forecasting of portfolio VaR in energy market, although the dynamics of copula parameters do exist, and affect the choice of copulas by GOF tests. The constant SJC copula provides a better performance, as reported in many previous research results (Patton 2006a). This may be because it embodies symmetric as well as asymmetric dependencies. Additionally, results show that symmetric copulas perform better than asymmetric copulas. This means that asymmetry in the dependence has few effects on forecasting of portfolio VaR in energy market, although asymmetry exists even after removing all asymmetries in marginal distributions.

Considering the forecasting performances of copulas and their GOF tests, we conclude that the constant Student t copula should be a good compromise for effectively fitting the dependence structure between crude oil and natural gas futures. Also, the parameters of Student t copula can be time-varying, but the dynamics of parameters improve few forecasting performances.

5 Conclusions

In this paper we have presented a new approach to estimate the VaR of an equally weighted portfolio comprising crude oil and natural gas futures in energy market that combines constant or time-varying copulas with GARCH-type models to construct the joint conditional

Table 9 Backtesting of VaR forecasts: loss functions

Copulas	95%			99%			Average scores
	Lopez	BI	Q	Lopez	BI	Q	
Constant:							
Normal	0.119969	0.016113	0.130246	0.016089	0.001003	0.148758	6.666667
Student's t	0.124711	0.016307	0.130518	0.014053	0.001019	0.148656	5.500000
Gumbel	0.127889	0.016880	0.130635	0.017465	0.001059	0.147099	3.500000
SJC	0.123731	0.016269	0.130139	0.013825	0.001005	0.149270	6.333333
Time-varying:							
Normal	0.122718	0.016645	0.131094	0.015782	0.001015	0.148523	4.833333
Student t	0.125990	0.016403	0.130489	0.019181	0.001137	0.148929	3.333333
Gumbel	0.129191	0.016883	0.130548	0.018359	0.001082	0.147423	2.833333
SJC	0.126905	0.016550	0.130791	0.016119	0.001032	0.149326	3.000000

Notes: The table shows three loss functions, based on skewed-t marginals for crude oil futures and natural gas futures, at 95% and 99% confidence level. The average score of each model is calculated according to rankings

multivariate distributions. A distinguishing feature of our model is that it allows us to simultaneously take into account two kinds of non-normal features of financial data: asymmetry, such as skewness, and excess kurtosis, such as fat tails. The asymmetry and excess kurtosis in the marginals can be fitted by GARCH-type models, while the asymmetry and excess kurtosis in the dependence can be captured by copulas. Additionally, copulas provide the flexibility to estimate the parameters of joint multivariate distribution, because dependence structure is separated from modeling of marginal distributions. Therefore, parameters in the marginals and copulas can be estimated separately, without being subject to the curse of dimensionality.

One of our main findings is that the constant Student t copula is a good compromise for effectively fitting the dependence structure between crude oil and natural gas futures. The parameters of Student t copula can be time-varying, but the dynamics of parameters improve few forecasting performance. Although asymmetry is found in the dependence structure, it has few effects on performances in forecasting of portfolio VaR in energy market. Therefore, asymmetric copulas provide worse fitting performance to the dependence structure between crude oil and natural gas futures, and furthermore, worse performances in forecasting of VaR. As expected, the SJC copula provides a better performance in forecasting VaR. This is because it embodies symmetric and asymmetric dependencies. However, the GOF test of SJC copula shows that it fits the dependence structure between crude oil futures and natural gas futures worse than Student t and Normal copulas.

The empirical results affirm the importance of specifying correct marginals. Particularly, GARCH(1, 1)-type models with Skew-T innovation better fit the univariate returns of both time series. The skewness parameter of Skew-T distribution shows that asymmetry is significant in univariate returns. The degree of freedom parameter of Skew-T distribution also indicates the significant kurtosis in univariate returns. These imply that traditional models based on the normal assumption are not appropriate to fit the univariate returns of both crude oil futures and natural gas futures. Therefore, the GARCH(1, 1) with Normal innovation is rejected. Although the GARCH(1, 1) with Student t innovation is not rejected, it provides worse fitting performance to the univariate returns than the GARCH(1, 1) with Skew-T innovation according to the log-likelihood values, AIC and SBIC. More specifically, the

AR(1)-TGARCH(1, 1) with Skew-T innovation is selected to fit the marginal distribution for crude oil futures, because there exists significant leverage effect in crude oil futures, and the AR(1)-GARCH(1, 1) with Skew-T innovation is selected to fit the marginal distribution for natural gas futures.

This paper leaves possible extensions for future research. An interesting extension would be that the skewness parameters and the degrees of freedom of Skew-T marginal density can be modeled by conditional time-varying models (Jondeau and Rockinger 2006; Fantazzini 2008). We believe the dynamic model for skewness parameters and the degrees of freedom of Skew-T density can provide better fitting to the marginal distribution. Therefore, this can improve forecasting performances. Secondly, alternative techniques, such as the Bayesian method, can be considered to choose the most probable copula among a given set (Huard et al. 2006). Finally, using the conditional multivariate, beyond bivariate, copulas to model joint multivariate distributions is promising to evaluate VaR with multi-asset portfolios composed of commodities, exchange rates and equities.

Appendix: Estimation of multi-stage maximum likelihood

Provided $F_{X|W}(\cdot|w)$ and $F_{Y|W}(\cdot|w)$ are differentiable and $F_{XY|W}(\cdot,\cdot|w)$ and $C(\cdot,\cdot|w)$ are twice differentiable, the bivariate copula density $c(\cdot,\cdot|w)$ can be obtained by differentiating (1).

$$\begin{aligned} f_{XY|W(x,y|w)} &\equiv \frac{\partial^2 F_{XY|W}(x,y|w)}{\partial x \partial y} \\ &= \frac{\partial F_{X|W}(x|w)}{\partial x} \cdot \frac{\partial F_{Y|W}(y|w)}{\partial y} \cdot \frac{\partial^2 C(F_{X|W}(x|w), F_{Y|W}(y|w)|w)}{\partial u \partial v} \\ &= f_{X|W}(x|w) \cdot f_{Y|W}(y|w) \cdot c(u,v|w), \quad \forall (x,y,w) \in \bar{\Re} \times \bar{\Re} \times \Omega \end{aligned} \tag{A.1}$$

where $u = F_{X|W}(x|w)$ and $v = F_{Y|W}(y|w)$. Then,

$$c(u,v|w) = \frac{f_{XY|W(x,y|w)}}{f_{X|W}(x|w) \cdot f_{Y|W}(y|w)} \tag{A.2}$$

This procedure is very important to implement the maximum likelihood estimation. Let $L_{XY} \equiv \log f_{XY|W(x,y|w)}$, $L_X \equiv f_{X|W}(x|w)$, $L_Y = f_{Y|W}(y|w)$ and $L_C = \log c(u,v|w)$. So we can obtain the log-likelihood function of (A.1).

$$L_{XY} = L_X + L_Y + L_C \tag{A.3}$$

Especially, we denote the conditional joint distribution as $F_{XY|W}(x_t, y_t; \theta|w_{t-1})$, the conditional marginal distributions as $F_{X|W}(x_t; \varphi|w_{t-1})$ and $F_{Y|W}(y_t; \gamma|w_{t-1})$, and the copula as $C(u_t, v_t; \kappa|w_{t-1})$. The same pattern is followed for different densities. According to Sklar's theorem, $F_{XY|W}(x_t, y_t; \theta|w_{t-1}) = C(F_{X|W}(x_t; \varphi|w_{t-1}), F_{Y|W}(y_t; \gamma|w_{t-1}); \kappa|w_{t-1})$, where $\theta \equiv [\varphi', \gamma', \kappa']'$ is the set of all parameters of both marginal distributions and copula to be estimated, and w_{t-1} is the information set until time $t-1$. Let MSMLE of θ be denoted as $\hat{\theta}$. Then it can be estimated using IFM method by the following successive procedures:

1. Parameters φ and γ of univariate marginal distributions are estimated in the first stage by the following maximization:

$$\hat{\varphi} = \arg\max_{\varphi \in \Phi} \sum_{t=1}^{T} \log f_X(x_t; \varphi) \tag{A.4}$$

$$\hat{\gamma} = \arg\max_{\gamma \in \Gamma} \sum_{t=1}^{T} \log f_Y(y_t; \gamma) \tag{A.5}$$

2. Given $\hat{\varphi}$ and $\hat{\gamma}$ in Stage I, copula parameters κ can be estimated as:

$$\hat{\kappa} = \arg\max_{\kappa \in \mathrm{K}} \sum_{t=1}^{T} \log c(F_X(x_t; \hat{\varphi}), F_Y(y_t; \hat{\gamma}); \kappa) \tag{A.6}$$

Then, the IFM estimator is defined as:

$$\hat{\theta} = [\hat{\varphi}', \hat{\gamma}', \hat{\kappa}']' \tag{A.7}$$

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *The second international symposium on information theory* (pp. 267–281). Budapest: Akademiai Kiado.

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *The Review of Financial Studies*, *63*, 443–494.

Bastianin, A. (2009). *Modelling asymmetric dependence using copula functions: an application to value-at-risk in the energy sector*. FEEM Working Paper.

Berg, D., & Bakken, H. (2006). *Copula goodness-of-fit tests: a comparative study*. Working Paper.

Blanco, C., & Ihle, G. (1999). How good is your VaR? using backtesting to assess system performance. *Financial Engineering News*, 1–2.

Bollerslev, T., Engle, R. F., & Nelson, D. B. (1994). ARCH models. In R. F. Engle & D. L. McFadden (Eds.), *Handbook of econometrics* (Chapter 49, pp. 2959–3038). Amsterdam: Elsevier.

Campbell, S. D. (2006). A review of backtesting and backtesting procedures. *The Journal of Risk*, *9*, 1–18.

Cherubini, U., Luciano, E., & Vecchiato, W. (2004). *Copula methods in finance*. London: Wiley.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *Intermountain Economic Review*, *39*, 841–862.

Deheuvels, P. (1979). La Fonction de dépendance empirique et ses propriétés: un test non paramétrique d'indépendance. *Bulletin de la Classe Des Sciences. Académie Royale de Belgique*, *65*, 274–292.

Dias, A. (2004). *Copula inference for finance and insurance*. Unpublished PhD Thesis ETH, Swiss Federal Institute of Technology, Zurich.

Diebold, F. X., Gunther, T., & Tay, A. (1998). Evaluating density forecasts with applications to financial risk management. *Intermountain Economic Review*, *39*, 863–883.

Durrleman, V., Nikeghbali, A., & Roncalli, T. (2000). *Which copula is the right one*? Working Paper.

Embrechts, P., Lindskog, F., & McNeil, A. J. (2003). Modelling dependence with copulas and application to risk management. In S. T. Rachev (Ed.), *Handbook of heavy tailed distribution in finance*. Amsterdam: Elsevier.

Embrechts, P., McNeil, A. J., & Straumann, D. (1999). Correlation and dependency in risk management: properties and pitfalls. In M. Dempster & H. Moffatt (Eds.), *Risk management: value at risk and beyond*. Cambridge: Cambridge University Press.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. *Econometrica*, *50*, 987–1007.

Fantazzini, D. (2008). Dynamic copula modelling for value at risk. *Frontiers in Finance and Economics*, *5*, 72–108.

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

Fisher, R. A. (1932). *Statistical methods for research workers*. Edinburgh.

Genest, C., Rémillard, B., & Beaudoin, D. (2009). Goodness-of-fit for copulas: a review and power study. *Insurance. Mathematics & Economics*, *44*, 199–213.

Glosten, L., Jagannathan, R., & Runkle, D. (1993). On the relation between the expected value and the volatility on the nominal excess returns on stocks. *The Journal of Finance*, *48*, 1779–1801.

González-Rivera, G., Lee, T. H., & Mishra, S. (2004). Forecasting volatility: a reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of Forecasting*, *20*, 629–645.

Grégoire, V., Genest, C., & Gendron, M. (2008). Using copulas to model price dependence in energy markets. *Energy Risk*, *5*, 58–64.

Hansen, B. E. (1994). Autoregressive conditional density estimation. *Intermountain Economic Review*, *35*, 705–730.

Harvey, C. R., & Siddique, A. (2000). Conditional skewness in asset pricing tests. *The Journal of Finance*, *55*, 1263–1295.

Hong, Y., Tu, J., & Zhou, G. (2007). Asymmetries in stock returns: statistical tests and economic evaluation. *The Review of Financial Studies*, *20*, 1547–1581.

Huard, D., Évin, G., & Favre, A.-C. (2006). Bayesian copula selection. *Computational Statistics & Data Analysis*, *51*, 809–822.

Hull, J., & White, A. (1998). Value-at-risk when daily changes in market variables are not normally distributed. *The Journal of Derivatives*, *5*, 9–19.

Joe, H. (1997). *Multivariate models and dependence concepts*. London: Chapman and Hall.

Jondeau, E., & Rockinger, M. (2006). The copula-GARCH model of conditional dependencies: an international stock market application. *Journal of International Money and Finance*, *25*, 827–853.

Jorion, P. (2007). *Value at risk: the new benchmark for managing financial risk* (3rd ed.). New York: McGraw-Hill.

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

Lopez, J. (1998). *Methods for evaluating value-at-risk estimates*. Federal Reserve Bank of New York Research Paper, no. 9802.

McNeil, A., Frey, R., & Embrechts, P. (2005). *Quantitative risk management: concepts, techniques and tools*. New Jersey: Princeton University Press.

Mendes, B. V. M., & Souza, R. M. (2004). Measuring financial risks with copulas. *International Review of Financial Analysis*, *13*, 27–45.

Nelsen, R. B. (1998). *An introduction to copula*. New York: Springer.

Patton, A. J. (2001). *Applications of copula theory in financial econometrics*. Unpublished PhD Thesis, University of California, San Diego.

Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Economics*, *2*, 130–168.

Patton, A. J. (2006a). Modelling asymmetric exchange rate dependence. *Intermountain Economic Review*, *47*, 527–556.

Patton, A. J. (2006b). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics*, *21*, 147–173.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, *23*, 470–472.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464.

Sklar, A. (1959). Fonctions de repartition à n dimensions et leurs marges. *Publications de L'Institut de Statistique de L'Université de Paris*, *8*, 229–231.

Stoyanov, S. V., Racheva-Iotova, B., Rachev, S. T., & Fabozzi, F. J. (2010). Stochastic models for risk estimation in volatile markets: a survey. *Annals of Operation Research*, *176*, 293–309.