Prediction of Volatility and Value at Risk with Copulas for Portfolios of Commodities

Felix Mörée

Master's thesis
2016:E39

Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

Centrum Scientiarum Mathematicarum

Abstract

Value at Risk (VaR) is a popular measurement for valuing the risk exposure. Correct estimates of VaR are essential in order to properly be able to monitor the risk.

This thesis examines a copula approach for estimating VaR for portfolios of commodities. The predictions are made from a semi-parametric model with Monte Carlo methods. The underlying model is constructed by choosing the best fit from different (E)GARCH-models for margins and some of the most common Archimedean and Elliptical copulas for the dependence.

None of the copulas in the scope gave a good fit to the data for the dependence. However, the copula with the best fit was the t-copula, which later was compared with the normal copula, the variance-covariance method and the method with historical observations. The comparison was done with Kupiec's test for correct number of VaR breaks and Christoffersen's test for independent breaks.

The results showed that none of the models in scope performed well, although the copula approach followed the data better. Backtesting suggested that the copula models overestimated the risk, resulting in too few VaR-breaks that also were clustered.

The conclusion was that other copulas would be needed to appropriately model the dependence, or that more sophisticated modeling methods in general should be used.

Keywords: Commodities, Copula, GARCH, VaR

Acknowledgements

This thesis has been written in three different countries during my exchange studies. I would like to thank all the people that I have met along the way, thank you all for your patience and interest in the subject.

I would also like to thank my supervisor, Nader Tajvidi, for his valuable support during the process.

Contents

1 Introduction **1**
1.1 Background . . . 1
1.2 Purpose . . . 1
1.3 Problem . . . 1
1.4 Delimitations . . . 2
1.5 Previous research . . . 2
1.6 Outline . . . 3

2 Theory **4**
2.1 Financial data . . . 4
2.2 Modeling of margins . . . 4
2.2.1 General model . . . 4
2.2.2 Mean modeling . . . 5
2.2.3 Noise modeling . . . 5
2.2.4 Distributions . . . 6
2.3 Modeling of dependence . . . 7
2.3.1 Dependence theory . . . 7
2.3.2 Properties of copulas . . . 8
2.3.3 Elliptical copulas . . . 9
2.3.4 Archimedean copulas . . . 9
2.3.5 Random number generation . . . 10
2.4 Estimation of parameters . . . 11
2.5 Value at Risk . . . 12
2.5.1 Definition . . . 12
2.5.2 Historical method . . . 13
2.5.3 Variance-Covariance method . . . 13
2.5.4 Monte Carlo method . . . 14
2.6 Tests . . . 15
2.6.1 Goodness of fit of margins . . . 15
2.6.2 Goodness of fit of copulas . . . 15
2.6.3 Backtesting . . . 16

3 Data **17**

4 Method **21**

5 Results **22**
5.1 Estimation of parameters . . . 22
5.2 Prediction of VaR . . . 25

6 Discussion **29**
6.1 Conclusions 29
6.2 Recommendations 30

A Best fits **32**

B Diagnostic plots for margins **33**

C Pair plots **36**

1 Introduction

1.1 Background

Financial modeling is a field that has seen its fair share of praise and criticism. One thing is nevertheless true: the field is here to stay and is growing rapidly with the amount of data that is available.

A typical application of financial modeling is perhaps one of the most important - to estimate the risk connected to one's investments. One of the most well-known measures of risk is Value at Risk (VaR), which is a method dated back as long as to the beginning of the century (1922 [18]) and the use today is excessive. The popularity lies in that VaR gives a comprehensive value - a number of the smallest loss at a given scenario connected to a certain probability.

If estimating the risk of one asset is considered difficult, modeling a portfolio is more demanding. In a portfolio, except for the variance of each asset, the dependence between different assets need to be accounted for. For these situations, copulas could be implemented for modeling the dependence. The advantage of copulas is that the modeling can be divided into two parts, first the estimation of the variance for the separate assets followed by the estimation of the dependence with copulas.

1.2 Purpose

The purpose of this thesis is to use copulas to improve VaR estimations of portfolios. Especially the data will consist of common commodities with different degree of dependence, which also behave differently in bull and bear markets.

1.3 Problem

The problem statement for this thesis is to see if copulas can be used to improve the estimations of VaR for a portfolio. Especially, a portfolio with common commodities will be studied.

1.4 Delimitations

Every investigation needs boundaries. In this case this means that there will be a limitation of models in the estimations. For the estimations of the margins the limitations will lead to trying Autoregressive-moving average (ARMA) up to $a = b = 2$ and two different generalised autoregressive conditional heteroskedasticity (GARCH) models up to $p = q = 1$.

When it comes to copulas, only the most common Elliptical and Archimedean copulas will be used. The choice of this limitation is due to the frequency that those copulas are used, together with the fact that there is an implementation available in R.

As a last limitation, re-estimations of model parameters will be done in every step, but there will be no change of model during the predictions.

1.5 Previous research

Copulas are a popular tool for modeling dependence due to the possibility of modeling margins and dependence separately. The applications are many of which finance is one. In financial modeling copulas can be used to model a wide variety of financial contracts, see for example [5] for more about the topic.

The inspiration for this work is mostly from Jacobsson's [19] thesis. Jacobsson used the same commodities as in this thesis, together with a momentum strategy and copulas, in order to model a trading strategy. The modeling was proved to be a success.

Copulas saw its large break-through with Li's article [23] from the RiskMetrics group. Li described how risk could easily be modeled with the normal copula, a method praised and widely used until the financial crisis. Now, the formula is a subject of heavy criticism. Using a normal copula did not work out well due to assets' tendency of high correlations during crisis, something that was not captured in the Gaussian copula.

A similar work as this thesis has been written by Bob [3]. In his thesis he used copulas and extreme value theory in order to model VaR for a portfolio of four European indexes. The result showed that the copulas performed very well, especially the elliptical copulas.

1.6 Outline

The disposition of the report will be as follows: first the theoretical foundation will be laid. The theory includes financial data, the models, and the tests that will be used. Then the data will be presented, and after that the results will be walked through. In the end there is a discussion about the results.

2 Theory

2.1 Financial data

Financial data have many characteristics which are only briefly described here. For a more thorough presentation see for example [24]. To begin with, the autocorrelation is stronger in the absolute value of the returns than in the returns themselves. Furthermore, financial data shows the tendency of converging towards a normal distribution when more data is added. However, the variance is not constant and volatility clusters are created, indicating calmer and more busy periods.

When describing the data, the tails usually need to be heavier to capture more extreme events. Although, by definition, extreme events are rare and difficult to estimate when they will occur and of what scale the events will be. The consequence is hence that the volatility practically always will be incorrectly estimated.

As a last stylised fact, losses are generally larger and happen during a shorter time than profits. This is related to the leverage effect, meaning that the market reacts differently on positive and negative news.

2.2 Modeling of margins

2.2.1 General model

In this thesis, the following model will be used for modeling the marginal distributions:

$$r_t = \mu_t + \epsilon_t \tag{1}$$

where r_t denotes the daily log return, μ_t the mean model and ϵ_t is noise. In order to model the mean, ARMA models will be used, while GARCH models will be used to describe the noise. Below it is possible to find descriptions of the models used, but for a more thorough description see for example Madsen [25] or Lindström [24].

2.2.2 Mean modeling

In order to model the mean model an ARMA(a, b) model was used, for a more thorough description, see for example [20]. In the model; a denotes the number of AR terms, b the number of MA terms. The model is defined as

$$\left(1 - \sum_{i=1}^{a} \phi_i L^i\right) X_t = \delta + \left(1 + \sum_{i=1}^{b} \theta_i L^i\right) \epsilon_t \tag{2}$$

where X_t denotes the observation at time t and ϵ_t the noise. Furthermore, L denotes the lag operator, ϕ_i are parameters for the AR process and θ_i are parameters for the MA process.

2.2.3 Noise modeling

In order to model the noise, GARCH models will be used. The *Autoregressive Conditional Heteroskedasticity* (ARCH) model was introduced 1982 by Engle [13]. A generalisation of this model is GARCH(p, q) which was first introduced by Bollerslev [4] and is defined as

$$\epsilon_t = z_t \sigma_t \tag{3}$$

$$z_t \sim i.i.d.\ D(0, 1) \tag{4}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{5}$$

where ϵ_t denotes the noise in the time series, σ_t the conditional variance and $D(0, 1)$ is a distribution with mean 0 and variance 1. Furthermore, $p \geq 0$, $q > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$ och $\omega > 0$. If $p = 0$ the ARCH model is obtained and if $p = q = 0$ one has white noise.

The conditions above are sufficient to guarantee that σ^2 is non-negative, but not always necessary. The conditions are sometimes too restrictive and estimations without these constraints are often done. Nelson and Cao [28] mention that the restriction of positive parameters is not always necessary. They suggest that estimations should be done case-by-case by checking that the conditional variance is positive.

An alternative formulation of the model is the ARMA(max(p,q),q), formulated as

$$\phi(L)\epsilon_t^2 = \omega + [1+\beta(L)]\varphi_t \tag{6}$$

where $\phi(L) = [1-\alpha(L)-\beta(L)]$, which is of order $\max(p,q)-1$, $\varphi_t = \epsilon_t^2 - \sigma_t^2$, and L denotes the lag operator.

In order to account for the fact that the market reacts differently on positive and negative news, the leverage effect is included in the GARCH model. This is done by allowing asymmetrical innovations, which was first proposed by Nelson [27]. The name of the model is *Exponential GARCH* (EGARCH), and it is defined as

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^{q} g(z_{t-i}) + \sum_{i=1}^{p} \beta_i \ln(\sigma_{t-i}^2) \tag{7}$$

where $g(z_{t-i}) = (\alpha_i[|\epsilon_{t-i}| - E|\epsilon_{t-i}|] + \gamma_i\epsilon_{t-i})$ is suggested. In this model is $\alpha_i[|\epsilon_{t-i}|-E|\epsilon_{t-i}|]$ denotes the size effect, and $\gamma_i\epsilon_{t-i}$ is denoting the sign effect. The parameters ω, α_i, γ_i and β_i do not have any restrictions except for the model to be stationary.

2.2.4 Distributions

In the modeling of the marginal distributions, three distributions will be used. The first two are symmetrical and can be found in for example Gut [17]. To begin with, the normal distribution is defined as

$$f_N(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2} \tag{8}$$

where μ denotes the mean and σ the standard deviation. The distribution has expected value μ and variance σ^2.

A symmetrical distribution with heavier tails is the t-distribution, defined as

$$f_t(x;\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\,\Gamma(\frac{\nu}{2})}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2} \tag{9}$$

where ν is a positive parameter which denotes the degrees of freedom. The expected value is 0 and the variance is $\frac{\nu}{\nu-2}, \nu > 2$.

In order to be able to model asymmetrically distributed data, the asymmetrical t-distribution suggested by Fernandez and Steel [14] was used. The definition is

$$f_{skt}(x;\nu,\xi) = \frac{2\xi}{1+\xi^2}\left[f_t(\xi x,\nu)I(x<0) + f_t\left(\frac{x}{\xi},\nu\right)I(x\geq 0)\right] \tag{10}$$

which constitutes of two t-distributions with the same degrees of freedom, but are scaled in two different ways. The parameter $\xi > 0$ denotes the skewness and $I()$ is the indicator function. When $\xi = 1$, the symmetrical t-distribution is obtained. When $\xi < 1$, the distribution has a heavier tail to the left, while $\xi > 1$ leads to a heavier tail to the right.

2.3 Modeling of dependence

2.3.1 Dependence theory

Pearson's correlation coefficient is the measurement that is used on a daily basis to describe dependence. The mathematical definition for the correlation between X and Y is

$$\rho_{X,Y} = \frac{\mathcal{C}(X,Y)}{\sqrt{\mathcal{V}(X)\mathcal{V}(Y)}} \tag{11}$$

where $\mathcal{C}$ and $\mathcal{V}$ denotes covariance and variance respectively. However, this measurement is linear and does not perform well when non-linear transformations are needed, or when mean and variance do not exist.

Due to the limitations of Pearsons's ρ, another measure of dependence was introduced - concordance [29]. Given two observation pairs (x_i, y_i) and (x_j, y_j) from the vectors (X,Y), concordance means that either $x_i < x_j$ and $y_i < y_j$, or $x_i > x_j$ and $y_i > y_j$. Disconcordance means on the other side that either $x_i < x_j$ and $y_i > y_j$, or $x_i > x_j$ and $y_i < y_j$. In other words, concordance measures if large value from one vector tends to correspond with large values from another vector.

Knowing this, it is possible to define a measure for concordance. Kendall's

τ is estimated as

$$\tau_K = \frac{c-d}{c+d} \tag{12}$$

where c is the amount of concordant pairs between X and Y, and d is the amount of disconcordant pairs.

2.3.2 Properties of copulas

The theory of copulas was introduced with Sklar's theorem in 1959 [31]; below follows a summary of basic properties. For a more extensive introduction to this field, consult for example [29].

The idea with copulas is to find a function C that, together with known marginal distributions, creates the joint distribution. Mathematically, in the two-dimensional case this is formulated as

$$H(x,y) = C(F(x), G(y)) \tag{13}$$

where H is the joint distribution function, F and G are univariate distribution functions, and C is the copula. A d-dimensional copula C with a domain $[0,1]^d$ is grounded and d-increasing. In addition, the copula has the following two properties

$$\begin{cases} C(u_1,\dots,u_{i-1},0,u_{i+1},\dots,u_d) = 0 \\ C(1,\dots,1,u,1,\dots,1) = u \end{cases} \tag{14}$$

for uniform margins u_d. Furthermore, copulas fulfil the Fréchét-Hoeffding's limits which states the upper and lower boundaries of copulas. The limits are defined as

$$W(u_1,\dots,u_d) \le C(u_1,\dots,u_d) \le M(u_1,\dots,u_d) \tag{15}$$

where the lower limit W is defined as

$$W(u_1,\dots,u_d) = \max\left(1-d+\sum_{i=1}^{d} u_i, 0\right) \tag{16}$$

and the upper limit M is defined as

$$M(u_1,\dots,u_d) = \min(u_1,\dots,u_d). \tag{17}$$

2.3.3 Elliptical copulas

The first category of copulas that will be applied in this thesis is the elliptical copulas, copulas that have elliptical level curves, see for example [29] for a more thorough description. The copulas from this category to be studied are the t-copula and the Gaussian-copula. The first copula to be analysed is the Gaussian copula which is formulated as

$$C(u,v) = \boldsymbol{\phi}_{\boldsymbol{\Sigma}}(\boldsymbol{\phi}^{-1}(u),\, \boldsymbol{\phi}^{-1}(v)) \tag{18}$$

where $\boldsymbol{\phi}$ is the Gaussian distribution and $\boldsymbol{\Sigma}$ denotes the correlation matrix. The second one is the t-copula, which is defined as

$$C(u,v) = t_{\nu,\boldsymbol{\Sigma}}(t_\nu^{-1}(u),\, t_\nu^{-1}(v)) \tag{19}$$

where $t_{\nu,\boldsymbol{\Sigma}}$ is the t-distribution with ν degrees of freedom and the correlation matrix $\boldsymbol{\Sigma}$.

2.3.4 Archimedean copulas

The second category of copulas that will be applied is the Archimedean copulas. According to for example [29], these copulas fulfil the following criterion

$$C(u_1,\dots,u_d;\theta) = \phi^{[-1]}(\phi(u_1;\theta) + \cdots + \phi(u_d;\theta);\theta) \tag{20}$$

where $[-1]$ denotes the pseudo-inverse and θ is the parameter vector. Furthermore, ϕ is the so-called generator function which needs to be strictly decreasing, continuous, and convex. For $d > 2$, ϕ also needs to be monotonically increasing. In this thesis some of the most common Archimedean copulas will be used. Firstly, the Clayton copula is formulated as

$$C(u,v) = \max\left\{u^{-\theta} + v^{-\theta} - 1; 0\right\}^{-1/\theta} \tag{21}$$

where $-1 \leq \theta < \infty$ and $\theta \neq 0$. Secondly, the Frank copula is defined as

$$C(u,v) = -\frac{1}{\theta}\log\left[1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{\exp(-\theta) - 1}\right] \tag{22}$$

where $\theta \neq 0$. Thirdly, the Gumbel copula is defined as

$$C(u,v) = exp\left[-\left([-\log(u)]^{\theta} + [-\log(v)]^{\theta}\right)^{1/\theta}\right] \tag{23}$$

where $1 \leq \theta < \infty$. Lastly, the Joe copula is used, which is formulated as

$$C(u,v) = 1 - \left[(1-u)^{\theta} + (1-v)^{\theta} - (1-u)^{\theta}(1-v)^{\theta}\right]^{1/\theta} \tag{24}$$

where $1 \leq \theta < \infty$.

2.3.5 Random number generation

To obtain random numbers from a copula, the procedure described below is used. The method is further explained in [29]. In the case of a d-dimensional copula the following setting is studied. Let

$$C_k(u_1,\ldots,u_k) = C(u_1,\ldots,u_k,1,\ldots,1), \qquad k = 2,\ldots,d-1 \tag{25}$$

be the k-dimensional margins of C. With the cases $C_1(u_1) = u_1$ and $C_d(u_1,\ldots,u_d) = C(u_1,\ldots,u_d)$. Then the conditional distribution of U_k, given the values $U_1,\ldots,U_{k-1}$, is

$$C_k(u_k|u_1,\ldots,u_{k-1}) = \mathbb{P}\{U_k \leq u_k | U_1 = u_1,\ldots,U_{k-1} = u_{k-1}\} = \tag{26}$$

$$= \frac{\delta^{k-1}C_k(u_1,\ldots,u_k)}{\delta u_1\ldots\delta u_{k-1}}\Bigg/\frac{\delta^{k-1}C_k(u_1,\ldots,u_{k-1})}{\delta u_1\ldots\delta u_{k-1}} \tag{27}$$

where $\mathbb{P}$ denotes the probability. The conditions on the expression are existing denominator and numerator together with a non-zero denominator. A random variate from a d-dimensional can then be generated by using the following scheme:

- Simulate a random variate u_1 from the unison distribution $U(0, 1)$
- Simulate a random variate u_2 from $C_2(u_2|u_1)$
- Simulate a random variate u_3 from $C_3(u_3|u_1, u_2)$

  $\vdots$
- Simulate a random variate u_d from $C_d(u_d|u_1, \ldots, u_{d-1})$

2.4 Estimation of parameters

The estimation of the parameters is done in two steps, the method is called inference functions for margins (IFM) and was first proposed by Joe and Xu [21]. In the first step the parameters for the margins are estimated, then in the second step the parameters for the copula are estimated.

Firstly the margins are estimated by maximising the following maximum likelihood function

$$\hat{\boldsymbol{\delta}}_{MLE} = \arg\max_{\boldsymbol{\delta}} \mathcal{L}(\boldsymbol{\delta}) = \arg\max_{\boldsymbol{\delta}} \left\{ \sum_{i=1}^{n} \log f(x_i; \boldsymbol{\delta}) \right\} \tag{28}$$

where x_i are the observations, $\boldsymbol{\delta}$ the parameters of the margins and f is the marginal distribution.

Having estimated d marginal distributions, the dependence parameters $\boldsymbol{\theta}$ are estimated by maximising the following pseudo log-likelihood function

$$\mathcal{L}(\boldsymbol{\theta}, \hat{\boldsymbol{\delta}}_{\mathbf{1}}, \ldots, \hat{\boldsymbol{\delta}}_{\boldsymbol{d}}) = \sum_{i=1}^{n} \log C \left\{ F(x_{1,i}; \hat{\boldsymbol{\delta}}_{\mathbf{1}}), \ldots, F(x_{d,i}; \hat{\boldsymbol{\delta}}_{\boldsymbol{d}}); \boldsymbol{\theta}) \right\} \tag{29}$$

where C is a copula and F are the estimated marginal distributions with parameters $\hat{\boldsymbol{\delta}}_{\boldsymbol{d}}$.

2.5 Value at Risk

2.5.1 Definition

Value at Risk (VaR) is a measurement of risk which states, given a probability, the risk to lose at least a certain amount during a certain period of time. This measurement has gained popularity due to its simplicity of expressing the risk as just a number in percent or in absolute numbers. The popularity has led to that the model is currently the most preferred way for measuring market risk from the bank regulations Basel Accords [1].

The formal definition of VaR, given by for example Charpentier [8], is the minimal value l for the probability for the loss L to exceed l is maximally $(1-\alpha)$. In mathematical terms

$$\mathrm{VaR}_\alpha(L) = \inf\{l \in \mathbb{R} : P(l) \geq \alpha\} \tag{30}$$

where P denotes the probability function and $\alpha \in (0,1)$ is a chosen percentil. If the returns are studied, the conditional value of VaR is defined as

$$\mathbb{P}(r_{t+1} < -\mathrm{VaR}_{t+1|t}|\mathcal{F}_t) = \alpha \tag{31}$$

where r_t is the return at time t and $\mathcal{F}_t$ denotes the information until time t.

Knowing the definition and the popularity, it is also important to understand the limitations. The risk with VaR is that one do not know what could happen in case of extreme events because those are very difficult to model. Furthermore, it is hard to capture what the tails look like which can lead to incorrect estimations. It is also important to understand that VaR just shows the minimal amount an investor risks to lose, it does not answer the question of how much one could actually lose during a really bad day. See for example [9] for more criticism.

The risk for the investor depends on if a long or short position has been taken, see more in for example the description by Giot and Laurent [16]. A *long* position means that an investor buys an asset. The risk for the long position is a decline in value and therefore the left tail needs to be studied in the distribution. A *short* position on the other hand means that an asset has been sold. The risk for a short position is an increase in value for the underlying asset, hence the right tail needs to be studied. Given a certain α and the parametric model (1), VaR for the two positions are

$$\mathrm{VaR}_t^{short} = \mu_t + \sigma_t G_{1-\alpha}^{-1} \tag{32}$$

$$\mathrm{VaR}_t^{long} = \mu_t + \sigma_t G_{\alpha}^{-1} \tag{33}$$

where μ_t is the conditional mean, σ_t the conditional variance, and G^{-1} denotes the quantile function for a distribution.

2.5.2 Historical method

The simpliest method for calculating VaR is through the historical method which is a completely unparametric estimation of unconditional VaR. More theory about this model can be found in for example Dowd [12]. Given a set of n historical observations, all the log returns are sorted as

$$r_1 < r_2 < \cdots < r_{n-1} < r_n \tag{34}$$

where the sorting is done after *size* and not time. The estimate of VaR is the α quantile of this empirical distribution.

2.5.3 Variance-Covariance method

This model is one of the most basic ones, but remains one of the most used ones as well, see for example Bodie et al. [3]. The basic assumption is that the log returns are identically and independently distributed as a normal distribution. Let first

$$R_{PF} = \sum_{i=1}^{d} w_i r_i \tag{35}$$

be the portfolio return with d weights (w_i) and returns (r_i). The variance is then calculated as

$$\sigma_{PF}^2 = W^T \Sigma W \tag{36}$$

where W is a vector with the portfolio weights and Σ the covariance matrix. The VaR is then calculated as

$$\text{VaR}_\alpha = \mu_{PF} + \sigma_{PF}\boldsymbol{\phi}_\alpha^{-1} \tag{37}$$

where $\mu_{PF} = \mathbb{E}(R_{PF})$ and $\boldsymbol{\phi}_\alpha^{-1}$ is the quantile function of the normal distribution.

2.5.4 Monte Carlo method

The concept of Monte Carlo methods is, according to Dowd [12], to construct a model from which random numbers are generated in order to make inference. In its easiest form, the absolute error declines as $\frac{1}{\sqrt{n}}$, which follows the central limit theorem.

The Monte Carlo approach is based on semi-parametric models. The idea consists of four steps. Firstly, a model is selected for the stochastic variables of interests. Secondly, random numbers are drawn from the model. Thirdly, enough random numbers are drawn in order for convergence towards the real distribution. Lastly, from the distribution that is estimated, conclusions about the VaR can be made.

With the model presented in (1) together with the model for noise modeling, the forecast will be calculated as

$$r_{t+1} = \mu_{t+1} + \epsilon_{t+1} \tag{38}$$

$$\epsilon_{t+1} = z_{t+1}\sigma_{t+1} \tag{39}$$

$$z_{t+1} \sim i.i.d.\ D(0,1) \tag{40}$$

where the parameters μ, ϵ, σ are estimated as well as the parameters to the distribution $D(0,1)$. In this case $D(0,1)$ is an arbitrary distribution with mean 0 and unit variance. Random numbers are drawn from the distribution in order to calculate the returns and to create the empirical distribution of the returns. In order to draw conclusions about VaR, the α quantiles for the empirical distribution are decided.

2.6 Tests

2.6.1 Goodness of fit of margins

In order to choose the model that best fits the data, different goodness of fit tests are carried out. To begin with, Akaike's and Bayes' information criteria are used (AIC and BIC), which are described in for example Madsen [26]. The difference between the two is how extra model parameters are penalised. AIC is defined as

$$\text{AIC} = 2k - 2\ln(\mathcal{L}) \tag{41}$$

where k denotes the amount of parameters in the model and $\mathcal{L}$ is the value of the likelihood function that is received when estimating the parameters for the margins.

The second one, BIC, is defined as

$$\text{BIC} = k\ln(n) - 2\ln(\mathcal{L}) \tag{42}$$

where, in addition to the above notations, n denotes the number of observations. As it is possible to see, BIC penalises the number of parameters more.

2.6.2 Goodness of fit of copulas

In order to study how well an estimated copula fits to the data, Cramer-von Mises goodness of fit test is used. For a more thorough description of it, see for example [15]. The test statistic is formulated as

$$S_n = \int_{[0,1]} \mathbb{C}_n(\boldsymbol{u}^2) dC_n(\boldsymbol{u}) \tag{43}$$

where $\mathbb{C}_n$ is defined as

$$\mathbb{C}_n(\boldsymbol{u}) = \sqrt{n}(C_n(\boldsymbol{u}) - C_{\theta,n}(\boldsymbol{u})) \tag{44}$$

in which C_n is a parametric estimate of the distribution and $C_{\theta,n}$ is the theoretical distribution. The test evaluates the null hypothesis that the parametric estimate is indeed the theoretical distribution.

2.6.3 Backtesting

In order to study how well the calculated values of VaR corresponded to the true values, backtesting was used. Two things are tested: if there are a correct number of exceedances of VaR and if those exceedances are independent.

To see if there is a correct number of exceedances (henceforth called *VaR breaks*) the Kupiec test is used. The hypothesis tested is if a correct amount of values exceed the estimated VaR level. The test was first described by Kupiec [22] as

$$LN_{UC} = 2\left[\ln L(I, \hat{\pi}_1) - \ln L(I, p)\right] \sim \chi^2(1) \tag{45}$$

where $L(I, x) = x^{T_1}(1-x)^{T-T_1}$, and T is the number of independent, identically Bernouilli distributed values I_t. Furthermore, p is the significance level that is to be tested, and $\hat{\pi}_1 = T_1/T$. The total amount of observations is denoted T and T_1 is the number of observations that exceed the theoretical level of VaR. The distribution $\chi^2(1)$ is a chi-squared distribution with one degree of freedom. The null hypothesis that is evaluated is if there is a correct number of observations that exceed the estimated level of VaR.

To see if the exceedances are independent, Christoffersen's test is carried out. The test serves to validate that the VaR exceedances are not creating clusters. The test was formulated by Christoffersen [10] and is defined as

$$LR_{IND} = 2(\ln L(1, \hat{\pi}_{01}, \hat{\pi}_{11})) - 2(\ln L(I, \hat{\pi}_1)) \sim \chi^2(1) \tag{46}$$

where except the notations found above, $L(I, \pi_{01}, \pi_{11}) = (1-\pi_{01})^{T_0-T_{01}}\pi_{01}^{T_{01}}(1-\pi_{11})^{T_1-T_{11}}\pi_{11}^{T_{11}}$. Here T_{ij} is the number of observations j that follow i, $\hat{\pi}_{01} = T_{01}/T_0$ and $\hat{\pi}_{11} = T_{11}/T_1$. The distribution $\chi^2(1)$ is as well a chi square distribution with one degree of freedom. The null hypothesis that is tested is that the exceedances are independent.

3 Data

In this thesis, five of the most common commodities have been used. The five commodities are Brent oil ("Brent"), West Texas Intermediate ("WTI"), gold, aluminium and copper. The data consists of closing prices from 7 October 2009 until 8 October 2015 and was taken from Datastream [11]. To facilitate modeling, only dates when all the commodities were traded have been used. The number of days used was therefore 1567. The data was log transformed according to

$$r_t = \ln\left(\frac{p_t}{p_{t-1}}\right) = \ln(p_t) - \ln(p_{t-1}) \tag{47}$$

where p_t is the price on day t. The datasets were divided in two parts, the first 1000 observations were used to estimate the model parameters while the rest were used for validation. In Figure 1 it is possible to see the untransformed datasets, the validation sets are the coloured parts on the right hand side of the dashed lines. Furthermore, the log transformed data is plotted in Figure 2, the validation sets are once again the coloured datasets to the right of the dashed lines.

For the calculations of the portfolio returns, it is important to specify the weights of the different components. In the analysis, the assumption will be that the weights are constantly 0.2, meaning that the portfolio contains an equal value of every commodity. For further reference, the portfolio's log returns are plotted in Figure 3. As one can see from the validation set of the portfolio's log returns, there seems to be an unusual "calm" period in the beginning that later, after half of the validation set, turns into something more similar to the estimation set.

In order to get a better understanding of the data, descriptive statistics have been collected in Table 1 together with different test results. For a more thorough presentation of the tests used, see for example [6]. A short description of the tests and the results follow.

Two tests are first carried out in order to test stationarity. The first one is an Augmented Dicker-Fuller (ADF), which tests the null hypothesis that a unit root exists. The other test is a KPSS test which evaluates the null hypothesis that the data is stationary around a trend.

	Brent	WTI	Gold	Aluminium	Copper
Observations	1566	1566	1566	1566	1566
Mean	-0.0002	-0.0002	0.0001	-0.0001	-0.0001
SD	0.016	0.019	0.011	0.013	0.015
Kurtosis	5.577	6.620	10.241	4.867	5.720
Skewness	0.074	-0.080	-0.949	-0.107	-0.114
KPSS	0.035**	0.100	0.063*	0.100	0.100
ADF	0.010***	0.010***	0.010***	0.010***	0.010***
ARCH(4)	0.000***	0.000***	0.000***	0.005***	0.000***
ARCH(8)	0.000***	0.000***	0.001***	0.000***	0.000***
Q(10)	0.016**	0.297	0.955	0.161	0.164
Q(20)	0.085*	0.554	0.187	0.017**	0.098*
$Q^2(10)$	0.000***	0.000***	0.0001***	0.000***	0.000***
$Q^2(20)$	0.000***	0.000***	0.0002***	0.000***	0.000***

Table 1: Descriptive statistics of the data used. The notation *,**,*** denotes if the null-hypothesis can be rejected with the significance level 0.1, 0.05 and 0.01.

In order to investigate if the datasets have volatility clusters a Lagrange multiplier (ARCH) test is done. The ARCH test evaluates the null hypothesis that no volatility clusters are present. Lastly, to see if there is a time dependence, the Ljung-Box test is used. The null hypothesis tested is if the data is independent. Note that Q denotes a Ljung-Box test while Q^2 denotes the same test but for squared time series.

From the test results, it is possible to see that no time series has a unit root. Regarding stationarity, it can be seen that three out of five time series are stationary around a level. Gold and Brent seem to be on the border of not being stationary around a level, indicating that there might be a trend.

Lastly, the results of the ARCH tests show that all datasets have volatility clustering. One can also see that there is a dependence between the squared log returns in all datasets, but not necessarily a dependence between the log returns.

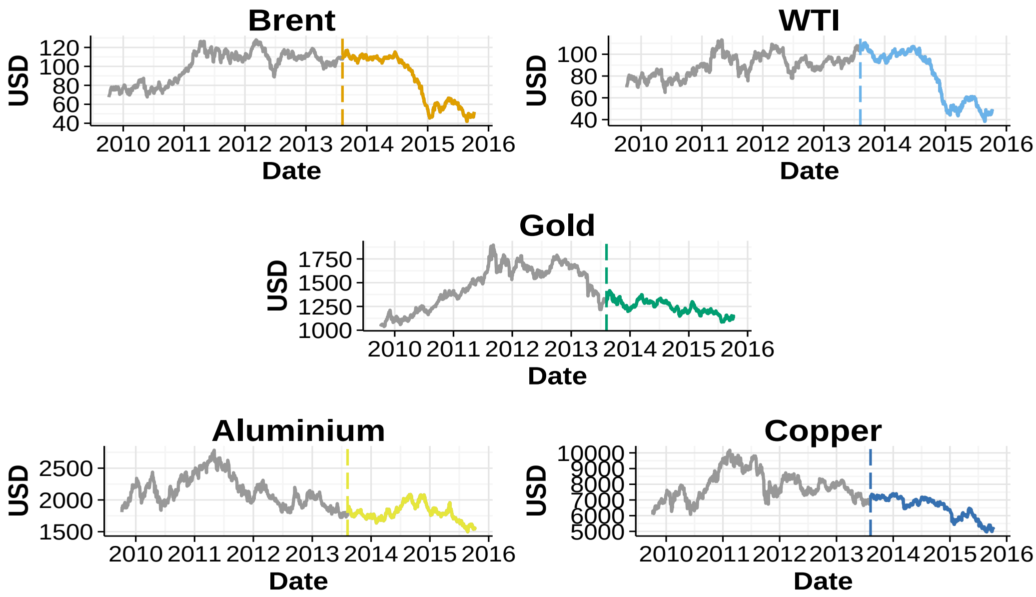

Figure 1: The data that was used. The plots show date against price in USD. The coloured part is the validation set. Note that the different y-axes have different scales.

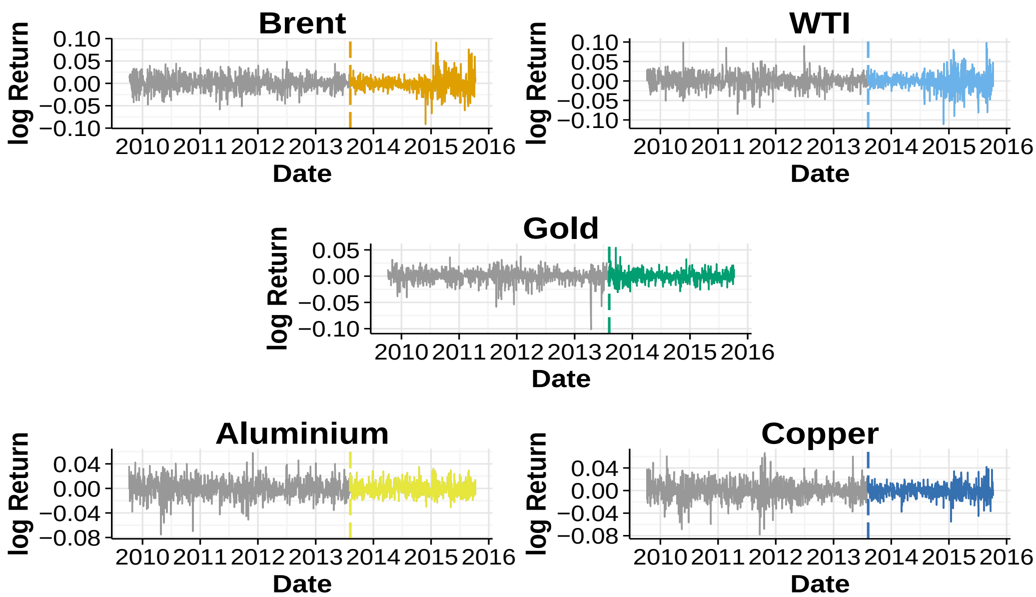

Figure 2: The log transformed data that was used. The plots show date against log return of the different datasets plotted in Figure 1. The coloured part is the validation set. Note that the y-axes have different scales.

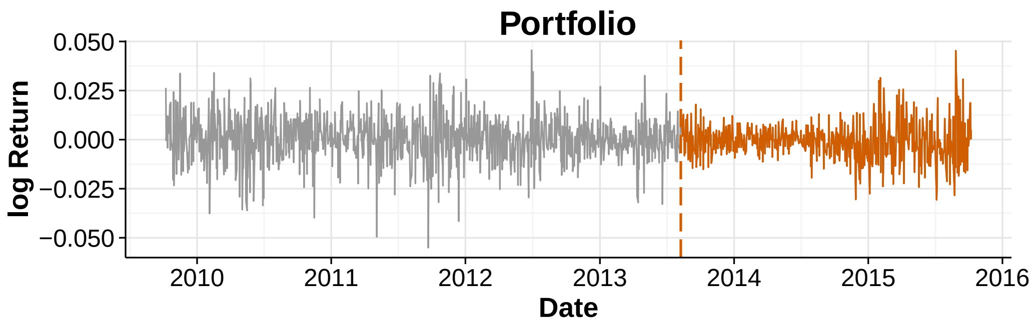

Figure 3: The portfolio return given equal weights. The data is plotted as date against log returns. The coloured part is the validation set.

4 Method

The method consisted of four steps. The first step was to model the margins, this was done by trying all possible combinations of ARMA and GARCH up until $p = q = 1$ and $a = b = 2$ with the distributions normal, Student's t, and skewed t. The different models were evaluated with AIC and BIC and if these crtieria indicated two different models, the smaller one was chosen as long as the residuals looked good. In order to determine what the residuals looked like, diagnostic plots were used (ACF, PACF and QQ-plot) together with different tests (ARCH, Ljung-Box).

The best model for each margin and the residuals were used to the second step - modeling the dependence. The dependence was modeled with the copulas which were presented in the section Theory. In order to decide which copula that best fitted the data, Cramér-von Mises goodness-of-fit test was used.

Thirdly, one step out-of-the-sample predictions of VaR were made for different α. Estimations were made with resampling every step together with a moving sample size of 1000. The best model from the previous steps was compared with three other models: historical method, variance-covariance method, and the best estimated margins with a normal copula. For the models with copulas, Monte Carlo methods were used to simulate random numbers, which later were used to estimate VaR.

Lastly, the results were evaluated with Kupiec and Christoffersen's tests in order to see how well the estimations performed. This was used to draw conclusions about the performance of the different methods.

For the quantitative part of this work, R has been used [30]. R is a software that is much used for statistical analysis, for this thesis the packages *rugarch* and *copulas* have been used. Furthermore, article [32] has been a large help for the work with copulas.

5 Results

5.1 Estimation of parameters

To start with, the margins and the parameters for the copula were to be found. The first step was to model the margins, which was done by finding the models that best fitted the estimation sets. This was done by fitting different combinations of ARMA and (E)GARCH models to the data. The size was up to $a = b = 2$ for ARMA, $p = q = 1$ for (E)GARCH, which were combined with the three different distributions mentioned above in the Theory. A total of 216 models were estimated for every dataset.

The three best fits, according to each information criterion and GARCH-model, are shown in Appendix A. The choice of models was done by examining the information criteria and the standardised residuals. In almost all cases the model with the largest BIC value was chosen, as long as the residuals were looking good. In the case of gold, the EGARCH model with the largest BIC was chosen, a decision based on the fact that it was that model that showed the best residuals.

The best fit for every dataset is shown in Table 2. It can be seen from the test results and the significance of the parameter that the fits are reasonable. In order to further investigate how well fitted the models were, diagnostic plots were plotted in Appendix B. ACF and PACF for standardised residuals and standardised squared residuals are shown together with a QQ-plot for the standardised residuals. From the diagnostical plots it is also possible to see that the models indeed seem to be good fits.

The second step, after having modeled the margins, was to model the dependence. To begin with, Kendall's τ and pair-wise data plots of the standardised residuals were studied. As one can see in Table 3 and in Figure 12, the largest dependence is between copper and silver, followed by brent and WTI. Except for that, the dependence is weak and vary in magnitude. From the pair plots, one can also see that the dependence is not very symmetrical in all cases.

Best estimations

	Brent	WTI	Au	Al	Cu
μ	0.0006	0.0006	0.0003	-0.0001	-0.0003
	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
ω	0.0000	-0.7977	-0.5586***	0.0000	-0.2242***
	(0.0000)	(0.6672)	(0.0689)	(0.0000)	(0.0006)
α	0.0467***	-0.1117***	-0.0190	0.0286***	-0.1003***
	(0.0096)	(0.0178)	(0.0289)	(0.0057)	(0.0202)
β	0.9491***	0.9025***	0.9373***	0.9660***	0.9730***
	(0.0106)	(0.0817)	(0.0075)	(0.0062)	(0.0006)
γ	-	0.1985***	0.1456***	-	0.1028***
		(0.048179)	(0.039310)		(0.010106)
df	10.4716***	6.5786***	3.8204***	7.3383***	7.5855***
	(3.0978)	(1.2962)	(0.5351)	(1.3366)	(1.6450)
$skew$	-	-	0.8917***	-	0.9439***
			(0.0382)		(0.0395)
$ARCH(4)$	0.4809	0.5037	0.0633*	0.4943	0.3285
$ARCH(8)$	0.3627	0.8038	0.2589	0.4431	0.3277
$Q(10)$	0.6556	0.8297	0.8096	0.3945	0.5146
$Q(20)$	0.9129	0.3824	0.5401	0.0729*	0.3374
$Q^2(10)$	0.4058	0.8277	0.4243	0.6690	0.4248
$Q^2(20)$	0.6953	0.8853	0.8984	0.6440	0.4008

Table 2: The parameters for the most suitable model as well as test result for the standardised residuals. The stars *,**,*** denotes the significances 0.1, 0.05 och 0.01.

In order to model the dependence, the different copulas mentioned in the theory part were fitted to the different time series. The data used for the fitting was the standardised residuals transformed to unit distributions. In order to be able to select the copula that fitted the data best, Cramér-von-Mises goodness-of-fit test was used.

In Table 4, the estimated parameters and the results from the goodness-of-fit tests are shown. As one can see, the p-values from the goodness-of-fit tests were in general very low. Furthermore, it is quite clear that the Elliptical copulas performed much better than the Archimedean copulas. The tests showed that none of the copulas had a very satisfactory goodness-of-fit, but the t-copula is the one that performed the least badly and was therefore used as the main model in the prediction of VaR.

	Brent	WTI	Au	Ag	Cu
Brent	1	0.449	0.217	0.275	0.314
WTI	0.449	1	0.183	0.305	0.329
Au	0.217	0.183	1	0.252	0.273
Ag	0.275	0.305	0.252	1	0.536
Cu	0.314	0.329	0.273	0.536	1

Table 3: Kendall's tau (τ_K) for the standardised residuals transformed to unit distributions.

Copula estimations

Copula	Estimated parameters	Test statisitc	p-value
Gaussian	$A = \begin{bmatrix} 1 & & \cdots & & \\ 0.726 & 1 & & \iddots & \\ 0.459 & 0.425 & 1 & & \vdots \\ 0.533 & 0.573 & 0.504 & 1 & \\ 0.576 & 0.595 & 0.558 & 0.802 & 1 \end{bmatrix}$	0.0698	0.0025***
t	$A = \begin{bmatrix} 1 & & \cdots & & \\ 0.711 & 1 & & \iddots & \\ 0.453 & 0.415 & 1 & & \vdots \\ 0.514 & 0.555 & 0.500 & 1 & \\ 0.557 & 0.584 & 0.551 & 0.789 & 1 \end{bmatrix}, \ \nu = 9.021$	0.0557	0.0155**
Clayton	0.798	0.9104	0.0005***
Frank	3.677	0.3587	0.0005***
Gumbel	1.559	0.7211,	0.0005***
Joe	1.778	1.9246	0.0005***

Table 4: The parameters for the most suitable model as well as test results from the goodness-of-fit tests. The stars *,**,*** denotes the significances 0.1, 0.05 och 0.01.

5.2 Prediction of VaR

After having found the best fits for the margins as well as for the copula, the chosen model could be used for predictions. The forecasts of the volatility (σ) and VaR were done as one step out-of-sample prediction with rolling re-estimation, with the constant estimation set size of 1000 for the re-estimations. The predictions were done by using four different models, representing three different methods. To begin with, VaR was estimated with Monte Carlo methods for the model with the best margins and the t-copula, the copula that gave best fit. Furthermore, as a comparison: VaR was estimated with Monte Carlo methods with the same margins and the normal copula, the historical method, as well as the variance-covariance method.

The calculations of the predictions for the historical method and the variance-covariance method are well described in the section Theory. For the parametric Monte Carlo method, estimation of every step was done with the following algorithm:

1. The models for the margins were fitted using the best models above to the latest 1000 data points.
2. The fitted models were used to decide the standardised residuals, which then were tranformed to unit distributions.
3. The copula was fitted to the unit distributed margins.
4. The estimated copula, with the estimated dependence and the models for the margins, was used to simulate 10000 random numbers.
5. The portfolio log return was calculated with the drawn random numbers. The empirical distribution of log returns was then decided.
6. VaR was determined from the empirical distribution.

The amount of random numbers in step 4 was chosen in a way that balances reliable estimates and computational power. Knowing from the theory that the error for this Monte Carlo method declines with $\frac{1}{\sqrt{n}}$, n was set to be 10000.

In Figure 4 and 5 it is possible to see different estimations for VaR levels for the different models. In the figures, dots indicate where there is a VaR break, meaning where the actual value over-/under-shoots the estimated value. The first thing that is noticable is that there are two periods: one with and one without a lot of volatility in the portfolio. Here it is possible to see that none of the models seem to perform well during the first half when the volatility is low. On the other hand, the VaR breaks in the second half seem to much more frequent.

In order to make the comparison clearer, Figure 6 can be consulted where different estimations of VaR from different models are shown in the same plot for $\alpha = (0.10, 0.05)$. Comparing the models show that the semiparametric methods follows the data better, but there seem to be too few VaR breaks, raising the question about whether the estimations are good or not. For the historical methods, there seem to be a lot of VaR breaks during the later part, but the estimations do not "follow" the data that well.

VaR comparison, $\alpha = (0.10, 0.05)$

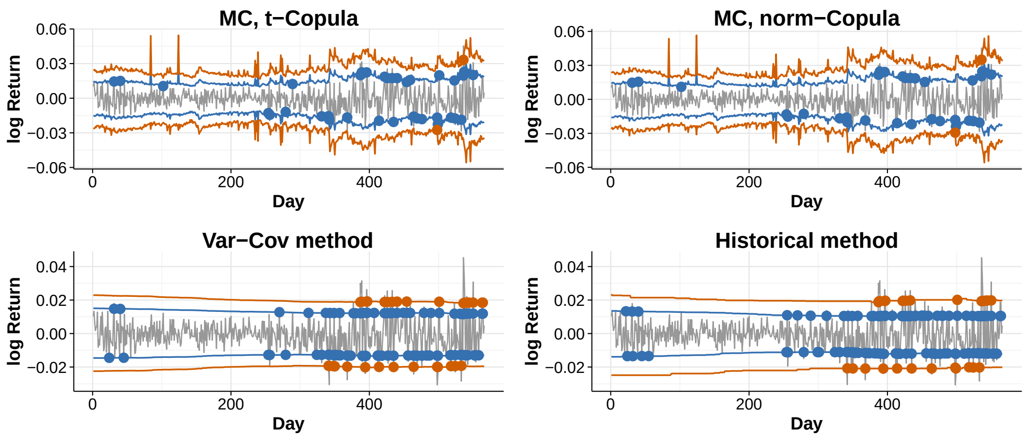

Figure 4: The estimations of different levels of VaR for the models used in the comparisons. The dots indicate VaR-breaks. As one can see, the VaR breaks are clustered in later part of the data set. Furthermore, the models with copulas have fewer VaR breaks than the other models.

VaR comparison, $\alpha = (0.025, 0.01)$

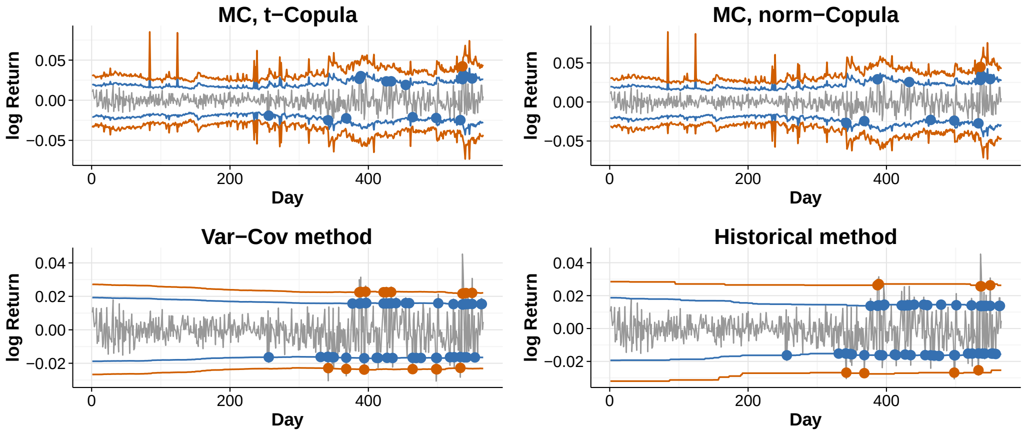

Figure 5: The estimations of different levels of VaR for the models used in the comparisons. The dots indicate VaR-breaks. As one can see, the VaR breaks are clustered in later part of the data set. Furthermore, the models with copulas have fewer VaR breaks than the other models.

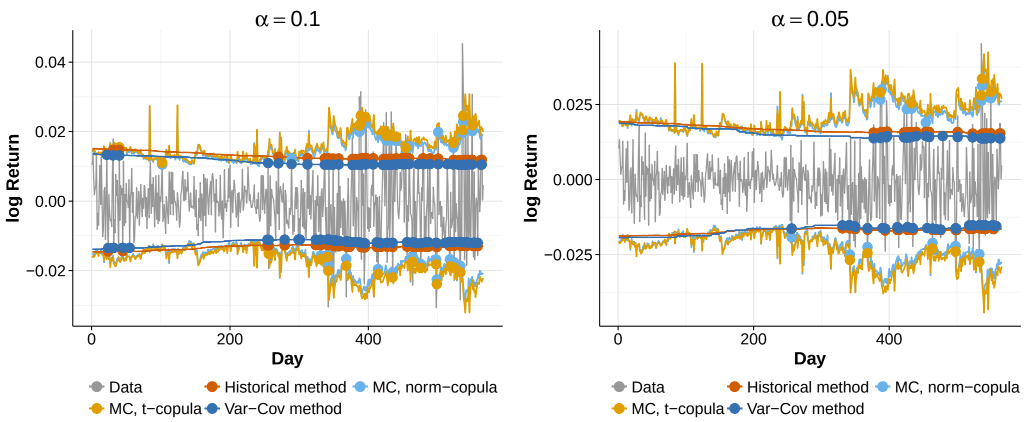

Figure 6: Comparisons of VaR for different α, using the different methods. As one can see, the methods with copulas "follow" the data better.

In Table 5, the results of the Kupiec and Christoffersen tests are shown. What Figures 4 and 5 showed is here confirmed, the estimations with copula and the Monte Carlo methods did not seem to give an adequate fit. Not for any position or level did the semi-parametric method give satisfactory results.

On the other hand, the non-parametric methods perform well in the Kupiec tests, especially for the long position. On the other hand, the non-parametric methods do not perform well in the Christoffersen tests, suggesting that the VaR breaks cluster. The test results suggest that also the non-parametric models give unsuccessful results.

Based on the test results, the best models are the ones that are non-parametrical, of which the historical observations method performed best. Among the parametric models, no conclusion could be drawn about which model performed the best between using normal- or t-copula. As a final remark, looking at the results as a whole, no model succeeded in giving satisfactory test results. Hence, none of the models in scope of this thesis should be relied upon for VaR predictions.

Backtesting

Method	α	**Long position**			**Short position**		
		$\hat{f}$	Kupiec	Chris.	$\hat{f}$	Kupiec	Chris.
MC with t-copula	0.100	0.028	0.000***	0.039**	0.034	0.000***	0.054*
	0.050	0.009	0.000***	0.624	0.012	0.000***	0.005***
	0.025	0.002	0.000***	1.000	0.002	0.000***	1.000
	0.010	0.000	0.000***	1.000	0.002	0.015**	1.000
MC with norm-copula	0.100	0.025	0.000***	0.079*	0.030	0.000***	0.006***
	0.050	0.009	0.000***	0.624	0.009	0.000***	0.056*
	0.025	0.002	0.000***	1.000	0.002	0.000***	1.000
	0.010	0.000	0.000***	1.000	0.002	0.015**	1.000
Var-Cov method	0.100	0.081	0.126	0.001***	0.069	0.009***	0.001***
	0.050	0.041	0.291	0.005***	0.041	0.291	0.000***
	0.025	0.025	0.968	0.003***	0.034	0.214	0.000***
	0.010	0.012	0.585	0.128	0.019	0.046**	0.000***
Historical method	0.100	0.104	0.738	0.001***	0.088	0.346	0.001***
	0.050	0.048	0.801	0.002***	0.042	0.395	0.000***
	0.025	0.021	0.552	0.028**	0.027	0.821	0.000***
	0.010	0.007	0.459	0.688	0.009	0.776	0.007***

Table 5: The frequency of VaR breaks, as well as the p-values for the Kupiec and Christoffersen tests, given different values of α. The notation *,**,*** denotes if the null-hypothesis can be rejected with the significance level 0.1, 0.05 and 0.01.

6 Discussion

6.1 Conclusions

Being aware of the risk exposure is crucial in order to properly plan capitalisation. This thesis examined if copulas can be used for estimating VaR for portfolios of commodities. This has been done using a semi-parametric approach, consisting of constructing a parametric model that was used in order to simulate VaR with Monte Carlo methods. The parametric model was made by fitting suitable time-series models for the margins and copulas for the dependence. Some of the most common time series models have been used to order the margins, while some of the most common copulas have been used to model the dependence.

The results showed that the margins could be well-fitted; however, the dependence could not be modeled satisfactorily with any of the suggested copulas. In the end, the VaR predictions of the semi-parametric models were not better than those of the non-parametric estimations according to the test results. Although the non-parametric models in scope of this thesis did not perform well either. At least, the parametric models followed the data better. But one cannot draw the conclusion that the semi-parametric models presented in this thesis gives satisfactory predictions.

The explication of the result is believed to mainly depend on the modeling. As seen in the results, the goodness-of-fit for the copulas in scope is not very good. The copula that best fitted the data was not significant with 95% certainty. This result shows that the dependence could not be modeled properly, and hence it most likely influenced the results. This shows that more effort is needed to find a suitable copula.

Another explication can also be that there was an unusually calm period in the beginning of the validation set, in comparison to the whole data set. This could be an explication to why no model in scope seems to have estimated well the VaR during that period - the period was just too calm. On the other hand, VaR is a tool for predicting losses in the future, a good model should not only be dependent on "standard" data. This touches one of the main problems of modeling rare movements, they are per definition rare and difficult, if at all possible, to model properly.

A problem related to this could be that too much or too little data was used in order to estimate the models. If fewer days were used to estimate the model, the model would be faster to adapt to periods which are more or less volatile. More data could have helped in a sense that more cases would be accounted for in the modeling and the model choice. Furthermore, from Figure 12 one can see that the distributions of the standardised residuals are not really unit distributions. With more data points, convergence towards a true unit distributions could be possible.

As a last remark, this thesis highlights fairy well the criticism raised about VaR and its limitations. It is difficult to predict the future, especially for stocks and commodities which are very volatile. The use of the same model to predict the whole estimation set can be a source of error. Allowing re-estimations of the model could be a way of improving the results.

Furthermore, the criticism raised about the difficulties of modeling dependence seems also to be justified, especially in using the normal copula method, due to the difficulties to actually properly model the dependence. As seen from the goodness-of-fit tests, the models in scope of this thesis performed poorly and it is not unreasonable to think that a portfolio of more assets would lead to a worse fit.

As a conclusion, this thesis highlights that in order to successfully model the risk more sophisticated methods for modeling are needed, above all it is important to better model the dependence.

6.2 Recommendations

As one can see from the results, the model in scope that best fitted the data gave an inadequate estimation. In order to improve the results some things could be done differently. During the work four things have been striking me as potential improvements.

To begin with, more care could be given to model the margins, perhaps include other models or allow re-estimation of models during the predictions.

Secondly, other copulas can be included and used in order to better estimate the dependence.

The following two suggestions would adjust the method: the third idea is to use extreme value theory in order to model the margins, for example a peak over threshold method. Fourthly, an idea could be to get inspiration from models used in the business and construct new methods from that.

A Best fits

Brent

Model	Dist.	AIC	Model	Dist.	BIC
ARMA(0,0)-GARCH(1,1)	std	5543.330	ARMA(0,0)-GARCH(1,1)	std	5518.791
ARMA(0,1)-GARCH(1,1)	std	5542.054	ARMA(0,0)-GARCH(1,1)	norm	5512.743
ARMA(1,0)-GARCH(1,1)	std	5541.894	ARMA(0,1)-GARCH(1,1)	std	5512.608
ARMA(0,1)-EGARCH(1,1)	std	5545.197	ARMA(0,1)-EGARCH(1,1)	std	5510.843
ARMA(0,2)-EGARCH(1,1)	std	5542.824	ARMA(0,0)-EGARCH(1,1)	std	5510.652
ARMA(0,0)-EGARCH(1,1)	sstd	5541.495	ARMA(0,0)-EGARCH(1,1)	norm	5510.469

Table 6: The top three models based on AIC and BIC for Brent.

WTI

Model	Dist.	AIC	Model	Dist.	BIC
ARMA(2,2)-GARCH(1,1)	std	5358.032	ARMA(0,0)-GARCH(1,1)	sstd	5328.287
ARMA(0,0)-GARCH(1,1)	sstd	5357.734	ARMA(1,0)-GARCH(1,1)	sstd	5322.360
ARMA(2,2)-GARCH(1,1)	sstd	5357.166	ARMA(0,1)-GARCH(1,1)	sstd	5321.994
ARMA(0,2)-EGARCH(1,1)	sstd	5371.682	ARMA(0,0)-EGARCH(1,1)	std	5335.521
ARMA(1,0)-EGARCH(1,1)	sstd	5369.692	ARMA(0,0)-EGARCH(1,1)	sstd	5334.798
ARMA(2,2)-EGARCH(1,1)	sstd	5369.620	ARMA(1,0)-EGARCH(1,1)	std	5331.139

Table 7: The top three models based on AIC and BIC for WTI.

Gold

Model	Dist.	AIC	Model	Dist.	BIC
ARMA(0,1)-GARCH(1,1)	sstd	6297.320	ARMA(0,0)-GARCH(1,1)	sstd	6267.268
ARMA(0,0)-GARCH(1,1)	sstd	6296.714	ARMA(0,1)-GARCH(1,1)	sstd	6262.965
ARMA(1,1)-GARCH(1,1)	sstd	6296.605	ARMA(1,0)-GARCH(1,1)	sstd	6262.104
ARMA(2,2)-EGARCH(1,1)	sstd	6297.860	ARMA(0,0)-EGARCH(1,1)	sstd	6254.080
ARMA(0,2)-EGARCH(1,1)	std	6290.670	ARMA(0,0)-EGARCH(1,1)	std	6253.377
ARMA(2,2)-EGARCH(1,1)	std	6288.768	ARMA(0,2)-EGARCH(1,1)	std	6251.408

Table 8: The top three models based on AIC and BIC for gold.

Aluminium

Model	Dist.	AIC	Model	Dist.	BIC
ARMA(1,1)-GARCH(1,1)	std	5688.121	ARMA(0,0)-GARCH(1,1)	std	5659.692
ARMA(2,0)-GARCH(1,1)	std	5687.881	ARMA(0,1)-GARCH(1,1)	std	5656.868
ARMA(1,2)-GARCH(1,1)	std	5687.233	ARMA(1,0)-GARCH(1,1)	std	5654.905
ARMA(2,2)-EGARCH(1,0)	std	5691.524	ARMA(0,0)-EGARCH(0,0)	std	5651.850
ARMA(2,2)-EGARCH(1,0)	sstd	5680.439	ARMA(1,0)-EGARCH(0,0)	std	5648.908
ARMA(2,2)-EGARCH(1,1)	std	5678.882	ARMA(0,1)-EGARCH(0,0)	std	5648.537

Table 9: The top three models based on AIC and BIC for aluminium.

Copper

Model	Dist.	AIC	Model	Dist.	BIC
ARMA(1,1)-GARCH(1,1)	std	5543.657	ARMA(0,0)-GARCH(1,1)	std	5515.880
ARMA(1,1)-GARCH(1,1)	sstd	5543.261	ARMA(1,0)-GARCH(1,1)	std	5511.684
ARMA(0,2)-GARCH(1,1)	sstd	5542.199	ARMA(0,1)-GARCH(1,1)	std	5511.557
ARMA(0,2)-EGARCH(1,1)	sstd	5560.018	ARMA(0,0)-EGARCH(1,1)	sstd	5524.608
ARMA(1,1)-EGARCH(1,1)	sstd	5559.331	ARMA(0,2)-EGARCH(1,1)	std	5519.509
ARMA(0,0)-EGARCH(1,1)	sstd	5558.962	ARMA(0,2)-EGARCH(1,1)	sstd	5515.848

Table 10: The top three models based on AIC and BIC for copper.

B Diagnostic plots for margins

Brent

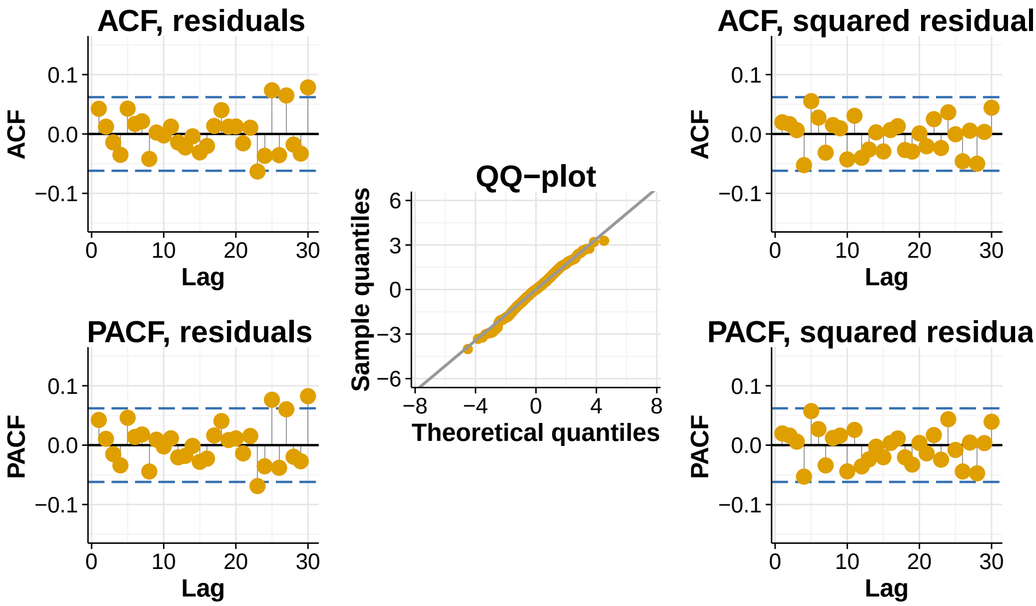

Figure 7: Diagnostic plots of standard residuals for Brent.

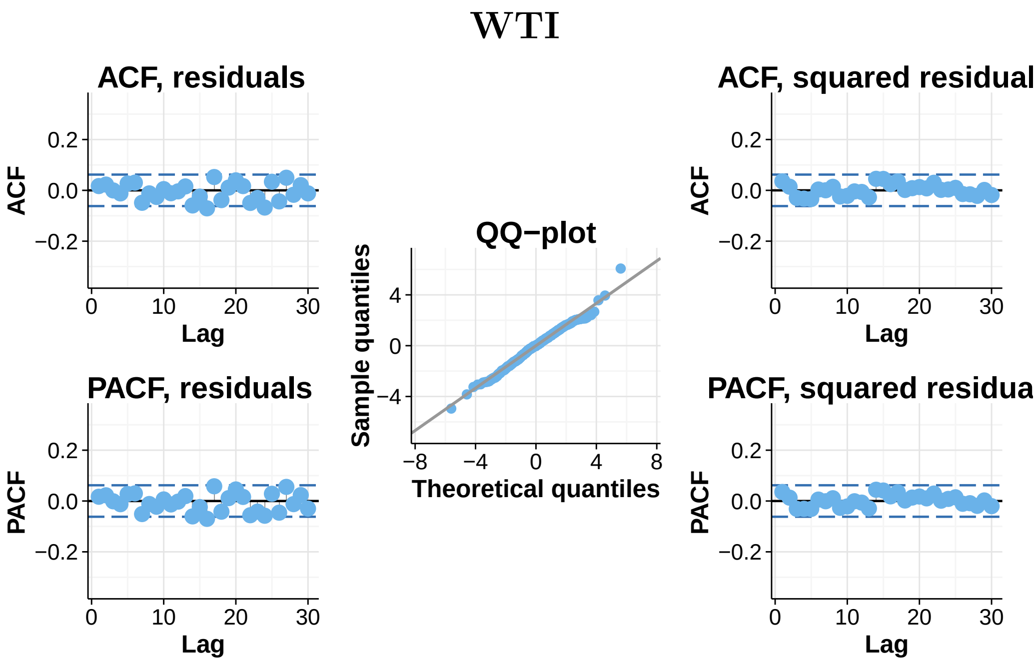

Figure 8: Diagnostic plots of standard residuals for WTI.

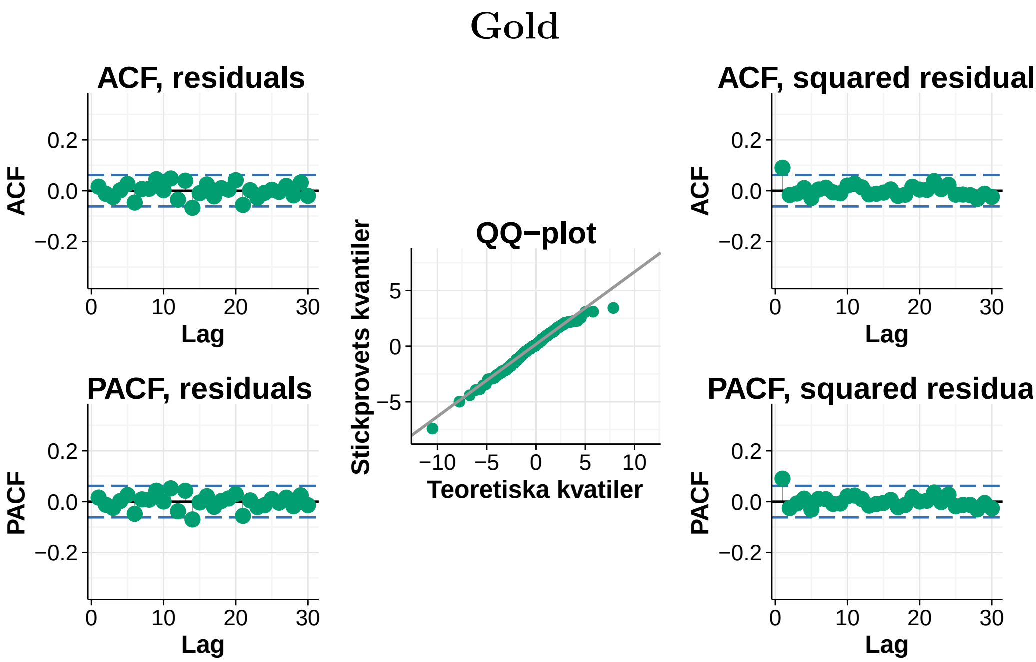

Figure 9: Diagnostic plots of standard residuals for gold.

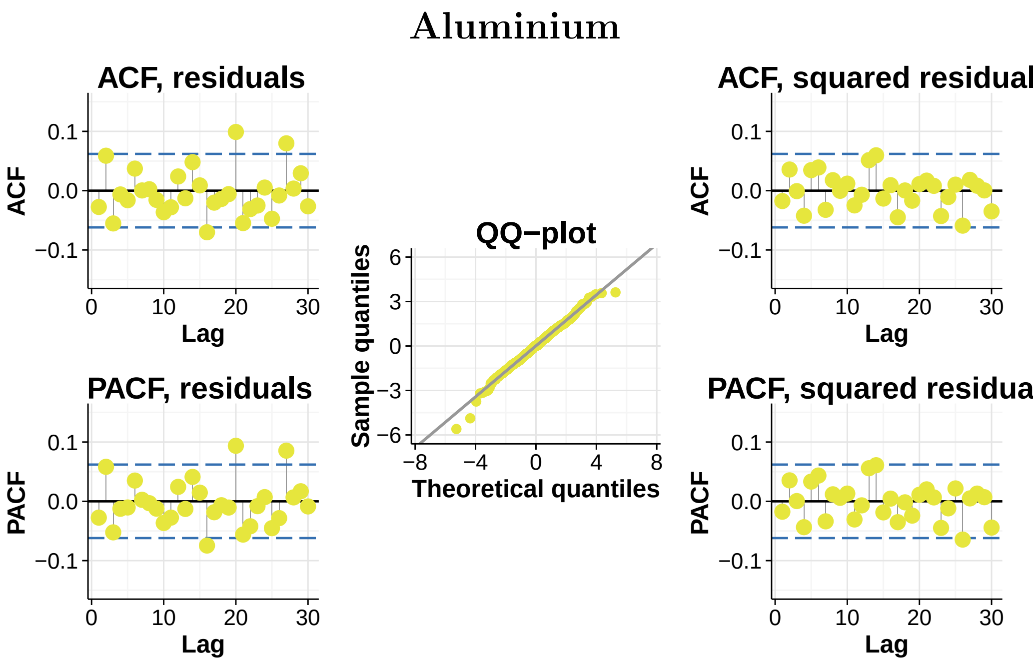

Figure 10: Diagnostic plots of standard residuals for aluminum.

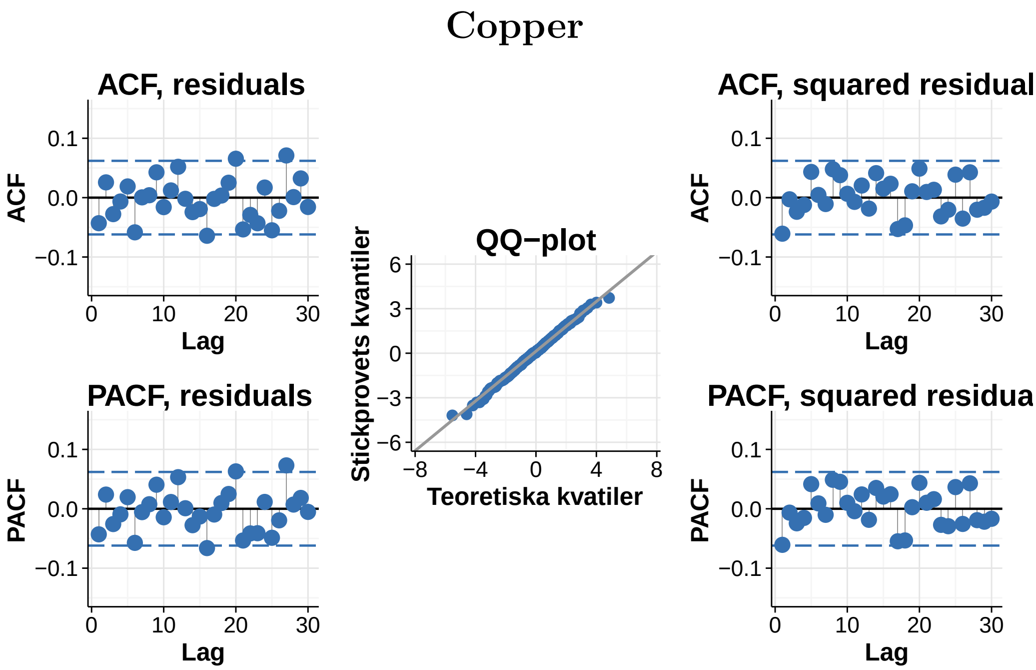

Figure 11: Diagnostic plots of standard residuals for copper.

C Pair plots

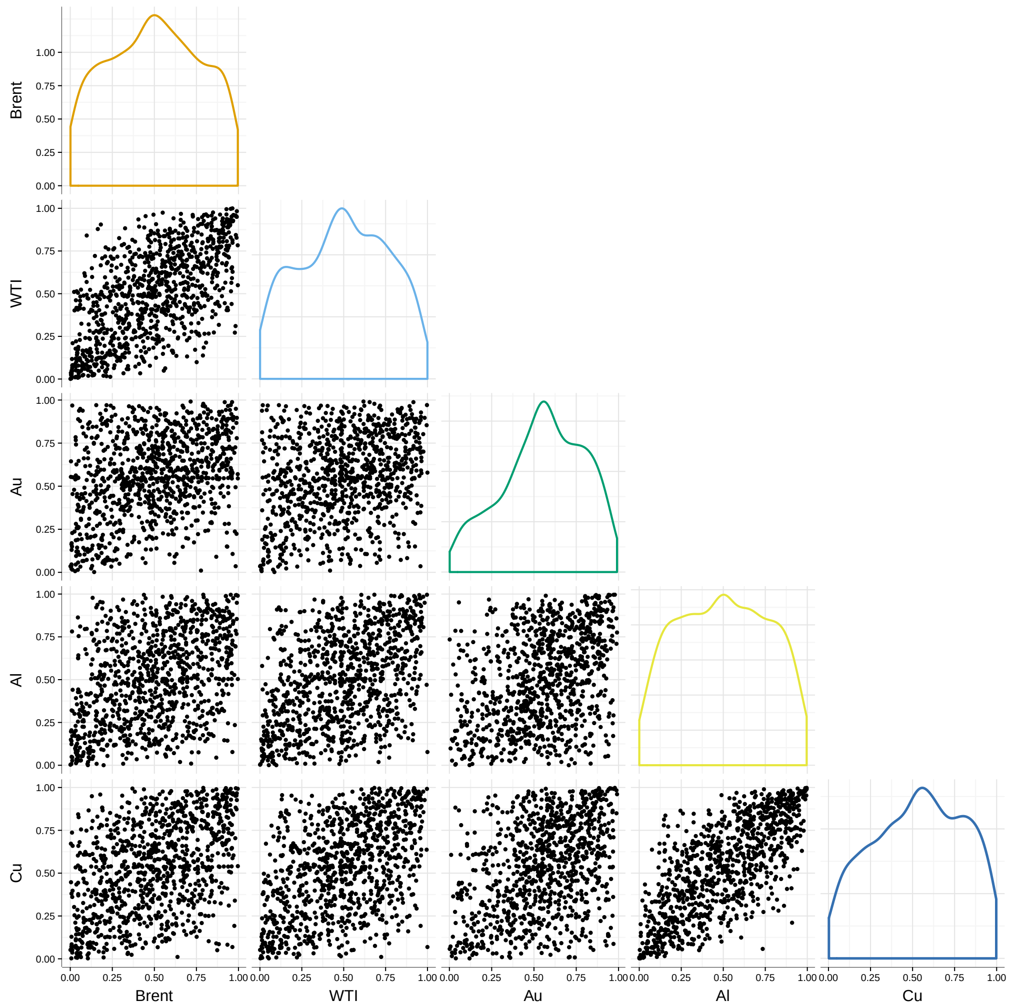

Figure 12: A pair plot between the standardised residuals transformed into unit distributions. As one can see, the dependence vary between the different commodities. Furthermore, the unit distriubtion can be seen. As one can see, the density is not completely uniform. Note that the values on the y-axis are in general not correct for the density plots.

References

[1] Basel Committee on Banking Supervision (2015).

[2] Bob, N. K. (2013) *Value at Risk Estimation, A GARCH-EVT-Copula Approach* http://www2.math.su.se/matstat/reports/master/2013/rep6/report.pdf

[3] Bodie, Z. et al (2013) *Investments*, McGraw Hill, Global Edition, 10th edition, ISBN 978-007-716114-9

[4] Bollerslev, T. (1986) *Generalized autoregressive conditional heteroskedasticity* pp. 307-327 in Journal of Econometrics 31

[5] Bouyé, E (2000) *Copulas for Finance A Reading Guide and Some Applications. Financial* Econometrics Research Centre, City Universe Business School, London

[6] Brooks, C. (2008) *Introductory Econometrics for Finance* Cambridge University Press, 2a upplagan, ISBN-13: 978-0-521-69468-1

[7] Campbell, J. Y. et al. (1997) *The Econometrics of Financial Markets* Princeton University Press, 2a upplagan, ISBN-13: 978-0-691-04301-2

[8] Charpentier, A. (2014), *Computational Actuarial Science with R.* CRC press, ISBN: 978-1-4665-9259-9

[9] Choudry, M. (2013), *An Introduction to Value-at-Risk*, Wiley 5th Edition, ISBN: 978-1-118-31672-6

[10] Christoffersen, P. F. (1998), *Evaluating Interval Forecasts.* International Economic Review, Vol. 39,No. 4, pp. 841-862

[11] Datastream. (2015) *Thomson Reuters Datastream.* (Accessed: October 2015)

[12] Dowd, K. (1998) *Beyond value at Risk. The new Science of risk management.* John Wiley and Sons, ISBN: 978-0-471-97622-6

[13] Engle, R. F. (1982). *Autoregressive Conditional Heteroscedasticity with Variance of United Kingdom Inflation.* Econometrica 50, no. 4: 987-1007.

[14] Fernandez, C.; Steel, M.F.J. (1998). *On bayesian modeling of fat tails and skewness.* Journal of the American Statistical Association Vol. 93, No. 441, pp. 359-371

[15] Genest, C. et al. *A goodness-of-fit test for bivariate extreme-value copulas.* Bernoulli 17(1), 2011.

[16] Giot, P.; Laurent, S. (2003) *Value-at-risk for long and short trading positions.* Journal of Applied Econometrics, 2003, 18 (6): pp. 641-664

[17] Gut, A. (2009). *An Intermediate Course in Probability.* Springer, ISBN: 978-1-4419-0162-0

[18] Holton, G.A. (2002) *History of Value-at-Risk: 1922-1998.* `http://stat.wharton.upenn.edu/~steele/Courses/434/434Context/RiskManagement/VaRHistlory.pdf`

[19] Jacobsson, M. (2015). *Forecasting commodity futures using Principal Component Analysis and Copula* `http://lup.lub.lu.se/student-papers/record/5426542`

[20] Jakobsson, A. (2013). *An Introduction to Time Series Modeling* Studentlitteratur, ISBN 9789144083742

[21] Joe, H. and Xu, J.J. (1996). *The estimation method of inference functions for margins for multivariate models.* Technical Report 166, Department of Statistics, University of British Columbia.

[22] Kupiec, P., (1995) *Techniques for verifying the accuracy of risk measurement models* Journal of Derivatives, 2, 174-184.

[23] Li, D. X. (2000). *On Default Correlation: A Copula Function Approach* The RiskMetrics Group, Working Paper Number 99-07

[24] Lindström, E. et al. (2015). *Statistics for Finance* Preprint, Chapman & Hall, ISBN-13: 978-1482228991

[25] Madsen, H. (2008) *Time Series Analysis.* Chapman & Hall, ISBN-13: 978-1420059670 0

[26] Madsen, H. ; Holst, J. (2006) *Modelling Non-Linear and Non-Stationary time series.* IMM-DTU

[27] Nelson, D. B. (1991) *Conditional Heteroskedasticity in Asset Returns: A New Approach* Econometrica Vol.59 Nr.2 pp. 347–370

[28] Nelson, D. B.; Cao, C.Q. (1992) *Inequality constraints in the univariate GARCH model.* Journal of Business & Economic Statistics, 10, 229-235.

[29] Nelson, R. B. (2006) *An Introduction to Copulas* Springer, ISBN-13: 978-0387-28659-4

[30] R Core Team (2016). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, `http://www.R-project.org/`

[31] Sklar, A. (1959) *Fonctions de répartition à n dimensions et leurs marges* Publ. Inst. Statist. Univ. Paris 8: 229–231

[32] Yan J. (2007) *Enjoy the Joy of Copulas: With a Package copula* Journal of Statistical Software, Vol. 21, Issue 4

Master's Theses in Mathematical Sciences 2016:E39
ISSN 1404-6342

LUTFMS-3306-2016

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lth.se/