Modeling copper prices and risk management

Mikael Teern
Supervisor: Magnus Wiktorsson
Mats Tegnér

December 9, 2014

Contents

Chapter 1

Introduction

When I started to work at Cavotec Connectors (CC), Staffanstorp, Sweden, I quite soon figured that there might be some need of my expertise. I needed to come up with a thesis project and I came to the conclution that the firm are exposed to a few financial risks. My quick brainstorming concluded that the firm where exposed to risks from:

1. changes in the aluminum price,
2. changes in currency exchange rate EUR/SEK,
3. changes in the copper price,

in that order.

When the ideas where pitched to the Managing Director (MD)[1], Mats Tegnér, the first idea got dismissed immediately on the grounds that there where many other costs involved in the production that changes in the aluminum price do not make an impact on the total costs of the firm. I had found an imaginary risk. The two other risks are real though.

In the recent years we have seen the highest currency exchange rate EUR/SEK in the life time of the currency and also one of the lowest levels of since the currency got introduced and since CC is producing a lot of its products for the global market, changes in the copper price and changes in the currency exchange rate affects the total cost and thus also the profit of CC.

CC is, according to MD Mats Tegnér, sensitive to changes in the copper price, but not enough to start buying insurances on the copper price. However, Cavotec Sweden (CS) is producing copper cables and for them copper is a big part of their total cost.

I realized quite soon that the project got too big when trying to both model the currency exchange rate EUR/SEK and model the copper price,

[1]Since the start of this thesis Mats has changed position from MD to Business Development Director.

so after a discussion with Mats Tegnér we decided it was best to drop the currency exchange angle and focus on trying to model the copper price to estimate the risk, since he figured that it was more important to Cavotec.

Chapter 2

From copper to cable

When trying to get your head around how copper cables are made, I got serious headache. There are so many things to consider. Or if you just want the rough basics, then it is quite simple.

Cable manufacturers buys copper catodes from a mining company, the catodes are made into very thin wires which then are twined together into thicker wires and then once again. Now it is time to dress the cable with both inner and outer coating from some kind of material; for instance rubber or silicone etc.

If you are in the business of buying copper cables and want to know if you are getting a good deal, then here is where the headache comes in to play as you try to learn more about the business. Especially if you do not have a clue about the history of the business. So let us start there.

Today copper catodes are traded on a global market. There are a few places in the world where you can trade in copper and nowadays is possible to trade at any of these markets at any time of the day (almost at least) and that it is quite easy to do it. But it was not always like that. Earlier there were several different markets in Europe where copper (among other things) was traded and they each had their own way of specifying what a copper cable is.

When talking to people in the business about copper and copper cables you often hear them talk about the amount of copper per meter cable i.e. amount Cu/m. From the context of the conversation you are supposed to know about the thickness of the cable, or put in more mathematical terms, i.e. the cross sectional area in the unit[1] mm^2.You also need to know what kind of cable you are talking about since there are a lot of differences in the designs of the cables (coating, number of single conductors et cetera). From now on in this thesis we will talk about a single conductor since the arguments can be generalized easily.

Even though you are familiar with the context (mentioned above) in

[1]In the USA they use a measure called AWG, but let us not go there

which the cable (single conductor) is discussed, there are still things that are unclear. Such as: how much copper does one meter copper cable contain? You think that this would be a matter of simple arithmetic since we have been so thorough in our definitions. Well it is not. At best you will get an upper and lower bound for the amount of copper in one meter copper cable. And given todays high and volitile copper prices this has become a serious issue for the people in the industry[2].

But why is it a problem? The reason is that one meter copper cable, with some cross sectional area, is defined by its conductivity rather then the actual copper content which is different from one country to the next. Yet the producer is providing a measure of the copper amount in kilograms based on *the cross sectional area*.

When the price of copper was low and relativly stable, compared to todays prices, this was not of any concern since the potential loss or extra cost from not being informed of the true copper content was very low compared to the total cost of producing the cable. Today the copper price is four times greater than in the beginning of the 21st century and also more volatile.

[2]The volitility and the high prices of copper was one of the reasons for writing this thesis

Chapter 3

Price risks in copper and copper cables

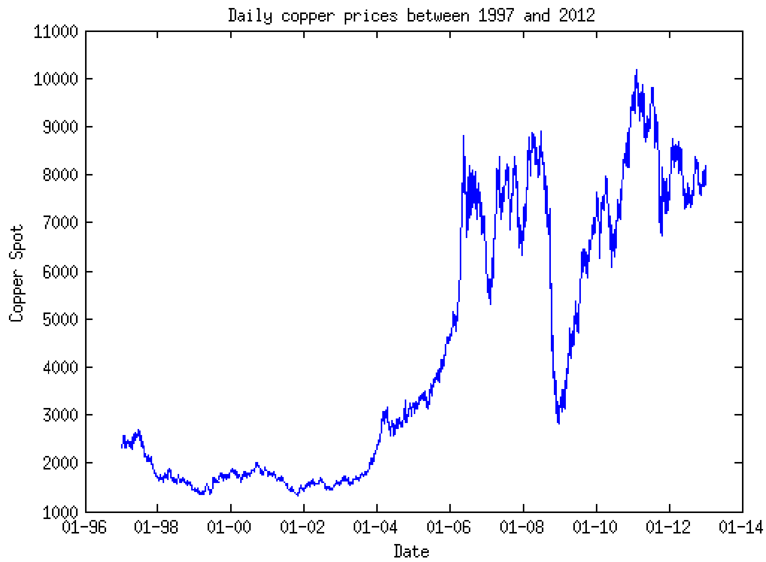

Figure 3.1: Daily copper price in US dollar per tonne. From 1997 to about 2004, the copper price remained fairly stable around 2000 US dollars per tonne, but then the copper price became both volitile and started to rise high.

For a long time, the copper price was constant. At least in the eyes of buyers and sellers of copper cables. Because of the low volatility and that the price seemed to vary around a fixed value, the copper price was thought of as being constant. Also important is that the price of copper was low compared to other costs in manufacturing of cables. Even if the copper price increased a lot (relative the previous value), that would not give a huge impact on the total cost of producing copper cables.

How the cable contracts are outlined can differ from deal to deal, but one typical way to set the contract is to split the total cost of the copper cable in two parts. The first part refers to the cost of production (labour, machinery etcetera) and the second part refers to the copper price. While the first part is fixed and set upon signing the contract, the second will be variable and the price is set on delivery. This means that neither the buyer nor the seller will know the the total price of the copper cable until the date of delivery. Add to this the fact that it can take up to five months to deliver a copper cable and we can conclude that the part of the total price depending on the copper price can change substantially and both the buyer and the seller take a financial risk.

If the buyer of the copper cable is also a seller of the copper cable, then he will experience both sides of the financial risk. But, likely is that the reseller of the copper cable needs to be able to provide a piece of the copper cable in a much shorter timespan than when the reseller bought it. Thus in order to be able to deliver on a short basis, the reseller needs to put a buffert of copper cable on stock and in turn introducing the risk of having the copper cables to long on stock.

Perhaps the most important aspect to producers (and buyers) of copper cables is that the volatility was very low. This meant that you could be sure that price of one meter copper cable would not change much from when you got (put) the order until delivery. And since the delivery time is up to five months with price set as to what the current price of copper is, low volatility is very much prefeered.

Since 2002 the behaviour of the copper price has changed dramatically. The price has become much more volatile and it has also risen a lot. The price took a dive back to pre 2002 levels during the financial crisis in 2008-2009 before it went skyhigh again. It is hard to say if the copper price fell because all the markets fell or if it had to do with the fact that we got a recession in the world and hence needed less copper. Disregarding the financial crisis, the copper price has been very high ever since 2002 (see Figure 3.1).

The copper price risk itself is not the only thing that traders in copper cable has to deal with. Not all copper goes into producing cables. This makes the supply of copper for cables lower than if you assume that all copper is available for cable production. Also the ratio between copper for cables and copper for other areas varies, making this a risk as well. Furthermore, there is an enormous amount of different types of cable available in the market, and the supply of each different type varies as well introducing yet another type of risk. Once a type of cable is made it is difficult to change it. Also cutting the cable too short is a risk that is worth considering.

In this thesis only the risk of the copper price has been considered. The effect of the rise in copper price alone is big enough, that the choice of simplifing the modeling to only looking at the copper price, is worth doing.

The risk metric used will Value at Risk (VaR). As a benchmark, VaR with time period of one day for daily copper prices will be modeled and then a more realistic time period of three months with monthly average copper prices will be modeled.

The price of copper is very much dependent on the economic situation in the world. If the world is in recession then the demand for copper is lower than if the world economy is flourishing. Copper is usually said to be an indicator of how the economy is going.

Copper is considered a leading indicator of economic trends. It is believed that a high price in copper indicates high demand which leads to high economic activity since copper is used in construction among other things. Generally when there is economic growth, roads are built, building are built etc. All of areas where copper is used. Copper is also considered an indicator of equity prices on the stock markets.

To investors in the financial markets copper is considered the only metal with a PhD, nicknamed 'Dr Copper'.
http://ravarumarknaden.se/kopparpriset-en-indikator-borskurserna-sp500/

Chapter 4

Copper price trading

4.1 Introduction to copper price trading

Copper is mainly traded on two exchanges in the world: The London Metal Exchange (LME) and the Comex division of the New York Mercantile Exchange (NYMEX). However, the Shanghai Metals Market (SMM) also provide the possibility to trade in copper (although the currency is Renminbi). There are some variations in the how copper (and other metals) are traded on the different exchanges (different: currencies, units, time to maturity). From a **academic/modeling** point of view, these differences are not really relevant. Many parts are just a question of a scaling parameter. `http://www.tradertech.com/information/coppertrading.asp`. LME was chosen for providing the data set for the analysis of the copper prices in this thesis. The data were acquired from the organization called LINC - Lund University's Finance Society during the author's membership period at said organization.

4.2 Contracts on London Metal Exchange

There are several different contracts to choose from at the London Metal Exchange (LME). Which contract to use depend on the particular desires of the investor.

Things get more complicated when the most common contracts that exist can be used to either hedge or speculate.

LME offers four types of contracts:

- LME Copper Futures.
- LME Copper Options.
- LME Copper TAPOS (Traded Average Prices Options).
- Copper LMEswap.

www.lme.com

Chapter 5

Theory

We need to establish some basic theory that we are going to use through out the thesis. The risk measure Value at Risk (VaR) will be used. As for the modeling of the copper price, different GARCH-models have been used. It has a nice property that it models the volatility as non-constant, thus making it possible to model the price as if it has varying volatility.

5.1 Log-transformation

One usual approach to model finance data is to first transform the data from daily prices into the logarithm of the returns, i.e. if the process representing the price at time t is denoted $\{P_t\}_{t=0}^{T}$, then the log returns, r_t, is found by

$$r_{t+1} = \ln\left(\frac{P_{t+1}}{P_t}\right) = \ln(P_{t+1}) - \ln(P_t). \tag{5.1}$$

From now on all modeling will be done on r_t.

From algebra and a neat trick, P_{t+1} can be expressed as

$$P_{t+1} = P_t \frac{P_{t+1}}{P_t} = P_t e^{\ln\left(\frac{P_{t+1}}{P_t}\right)} = P_t e^{r_{t+1}}, \tag{5.2}$$

which means that if we can estimate the log return one time step ahead, denoted $\hat{r}_{t+1}$, then the price one time step ahead can be estimated as

$$\hat{P}_{t+1} = P_t e^{\hat{r}_{t+1}}. \tag{5.3}$$

Furthermore, if the confidence interval of $\hat{r}_{t+1}$ is $I_{\hat{r}_{t+1}} = [c, d]$, for some c and d, then because of monotonicity of the exponential function, it follows that $I_{e^{\hat{r}_{t+1}}} = [e^c, e^d]$ and the confidence interval of P_{t+1} is $I_{P_{t+1}} = [P_t e^c, P_t e^d]$.

This can be generalized for k time steps ahead as

$$\hat{P}_{t+k} = P_t \frac{P_{t+1}}{P_t} \cdots \frac{P_{t+k}}{P_{t+(k-1)}} = P_t e^{\ln\left(\frac{P_{t+1}}{P_t}\right)+\ldots+\ln\left(\frac{P_{t+k}}{P_{t+(k-1)}}\right)} = P_t e^{\hat{r}_{t+1}+\ldots+\hat{r}_{t+k}} \tag{5.4}$$

and the if the confidence interval of $\hat{r}_{t+1} + \ldots + \hat{r}_{t+k}$ is $I_{\hat{r}_{t+1}+\ldots+\hat{r}_{t+k}} = [f, g]$, using the same reasoning as above, the the confidence interval of P_{t+k} is $I_{P_{t+k}} = [P_t e^f, P_t e^g]$.

5.2 Value at Risk

Financial risks come in many forms. Firms can default on their payments and even go bankrupt. If you have stocks or other claims in that firm, the value of your assets might diminish or even be wiped out altogether. Risks like this are called credit risks. There are several other financial risks as well. In the case of Cavotec, increases in the copper price is a significant one. The firm will suffer badly if copper prices increases too much too fast. But also a sudden drop in copper prices is bad. One particular way of measure financial risk is the measure called Value at Risk (VaR). It gives a measure of how much one stands to lose over a specific time period with a certain probability. Usually the time period is one day and the probability is usually in the range $[0.9, 1)$. From Jorion, P. (2001) it is defined at time t as the smallest number x such that the amount X_{t+1} one will stand to lose, at time $t + 1$, will fall below x with probability γ:

$$\begin{aligned} \mathrm{VaR}_\gamma(X_t) &= \inf\{x \in \mathbb{R} | \mathrm{P}(X_{t+1} \leq x) > \gamma\} \\ &= \inf\{x \in \mathbb{R} | \mathrm{P}(X_{t+1} > x) > 1 - q = \gamma\}. \end{aligned} \tag{5.5}$$

Sometimes the coverage rate $\gamma = 1 - q$ is used instead. In this thesis both will be used.

In the above defintion, the time period goes from t to $t + 1$ which is naturally thought of as a one day time period and this interpretation will be partly used in this thesis. The other interpretation in this thesis will be a time period of one month from t to $t + 1$. If the time period is known, then it is often excluded in the text and one simply writes 95%-VaR and from the context it should be understood which time period is assumed.

The above definition just looks at one side of the risk. In this thesis, as have been discussed above, the firm finds risks both for increase and decrease of the copper price. Therefore the definition of VaR needs to be expanded to cover both increases and decreases. By doing so a two-sided confidence interval is recived. The upper and lower bound, $\mathrm{VaR}^{\mathrm{up}}_\gamma(X)$) respectivly $\mathrm{VaR}^{\mathrm{lo}}_\gamma(X)$, of the confidence interval is defined from equation 5.6 below:

$$\mathrm{P}(\mathrm{VaR}_\gamma^{\mathrm{lo}}(X) < X < \mathrm{VaR}_\gamma^{\mathrm{up}}(X)) = \gamma \tag{5.6}$$

5.3 Volatility Models

In this thesis I have chosen to use models which assume non-constant variance or, using the terminology of finance, non-constant volatility. This because it is a well know fact in the finance industry that the volatility is not constant. Had it been constant, then there would not have been any reason to use other models then the Black-Scholes-Merton model[1]. We know that Black-Scholes-Merton model does not work very well and one flaw is its assumption of constant variance.

Looking at prices we see in general not only variations in day-to-day price changes but also that these variations are somewhat asymmetric. We see clusters of large volatility. To be able to model the prices one needs to take this into account. This is the reason for volatility models.

I will discuss how different volatility models work and compare them to see which is best suited for modeling the copper prices. The models are the GARCH(p,q) and a variation of it called EGARCH(p,q).

GARCH(p,q) Models

GARCH is an abbreviation for Generalized Autoregressive Condtional Heteroskedasticity and is a extension of the ARCH[2] model which was introduced by Engle (1982). GARCH(p,q) is defined as follows:

Let r_t define the error terms. In this thesis the log-returns of the data will be modeled as the error terms. These are assumed to be products of a time-dependent stochastic process describing the conditional deviation denoted σ_t and a stochastic process, denoted z_t consisting of white noise i.e. a process of independent identically distributed random variables with zero mean and variance equal to one.

The error terms are there described as:

$$r_t = \sigma_t z_t \tag{5.7}$$

This in itself is not very special. The interesting part is how the process σ_t is modeled. For GARCH(p,q) σ_t is modeled as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_p r_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_q \sigma_{t-q}^2 \tag{5.8}$$

[1]Not entirely true. Black-Scholes-Merton model has more flaws than the assumption of constant variance

[2]Autoregressive Condtional Heteroskedasticity

In this thesis three different distributions will be assumed when modeling the error terms, i.e. the process z_t will be assumed to follow three different distributions. The three distributions are: the standard normal distribution, student-T distribution and the normalized Generalized Error Distribution (GED) with zero mean and variance one. While the the standard normal distribution and student-T distribution are common distributions, the GED is not. The probability density function for normalized GED is defined as follows, according to pp.352-353 in Nelson (1991):

$$f_X(x) = \frac{\nu e^{-|\frac{x}{2\lambda}|^\nu}}{\lambda 2^{(1+1/\nu)}\Gamma(1/\nu)}, \; -\infty < x < \infty, \; 0 < \nu \leq \infty \tag{5.9}$$
$$\lambda = \sqrt{\frac{2^{-(2/\nu)}\Gamma(1/\nu)}{\Gamma(3/\nu)}}$$

and ν is a tail-thickness parameter. $\Gamma(\cdot)$ is gammafunction defined as follows:

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\mathrm{d}t. \tag{5.10}$$

This choice of λ implies $\mathbb{E}[X^2] = 1$. The gamma function can by analytic continuation be defined for all $z \in \mathbb{C}$ except the negative integers and zero. Choosing $\nu = 2$ we get the probability density function for the familiar standard normal distribution as[3]:

$$f_X(x) = \frac{2e^{-|\frac{x}{2\lambda}|^2}}{\lambda 2^{(1+1/2)}\Gamma(1/2)} = \left[\lambda = \sqrt{\frac{2^{-(2/2)}\Gamma(1/2)}{\Gamma(3/2)}} = \sqrt{\frac{2^{-1}\sqrt{\pi}}{\frac{\sqrt{\pi}}{2}}} = 1\right] \tag{5.11}$$
$$= \frac{2e^{-\frac{x^2}{2}}}{2^{(1+1/2)}\sqrt{\pi}} = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

This means that $\nu > 2$ gives thinner tails and $\nu < 2$ gives fatter tails than the normal distribution. For example, when $\nu = 1$ the function becomes:

$$f_X(x) = \frac{1 \cdot e^{-|\frac{x}{2\lambda}|^1}}{\lambda 2^{(1+1/1)}\Gamma(1/1)} = \left[\lambda = \sqrt{\frac{2^{-(2/1)}\Gamma(1/1)}{\Gamma(3/1)}} = \sqrt{\frac{0!}{2^2 \cdot 2!}} = \frac{1}{2\sqrt{2}}\right] \tag{5.12}$$
$$= \frac{e^{-\left|\frac{x}{2\frac{1}{2\sqrt{2}}}\right|}}{\frac{2^2 \cdot 0!}{2\sqrt{2}}} = \frac{\sqrt{2}}{2}e^{-|\sqrt{2}x|},$$

[3]In the calculations use the well known relations $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(z+1) = z\Gamma(z)$ to get the result.

which is known as the probability density function of the Laplace distribution with scale parameter $b = \frac{1}{\sqrt{2}}$ and location parameter $\mu = 0$.

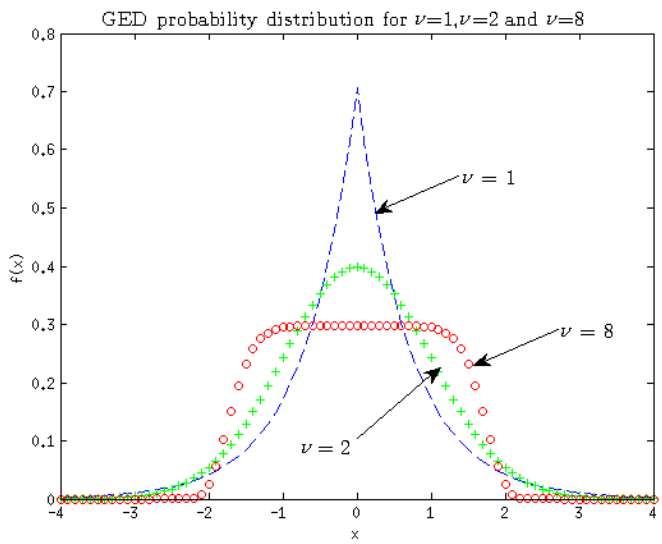

Figure 5.1: Probability density function for the normalized Generalized Error Distribution when $\nu = 1$ (Laplace distribution), $\nu = 2$ (standard normal distribution) and $\nu = 8$. Already when $\nu = 8$ the distribution is almost uniform which is expected when $\nu = \infty$.

EGARCH(p,q) Models

The EGARCH(p,q) model is similar to the GARCH(p,q) model. It was introduced by Nelson (1991), pp. 350-351. The conditional variance has to be positive and one way of ensuring this is to model the natural logarithm of the conditional variance. In this thesis the definition of the EGARCH(p,q) model used in Oxford MFE Toolbox is used (which can be found at `http://www.kevinsheppard.com/MFE_Toolbox`) and is defined as:

$$r_t = \sigma_t z_t \tag{5.13}$$

where σ_t is modeled as:

$$\log \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i g(Z_{t-i}) + \sum_{i=1}^{q} \beta_i \log \sigma_{t-i}^2 \tag{5.14}$$

and

$$g(Z_t) = \left| \frac{r_t}{\sigma_t} \right| - \sqrt{\frac{2}{\pi}} \tag{5.15}$$

As in the GARCH(p,q) model σ_t^2 is the conditional variance, $\{\alpha_i\}_{i=0}^p$, $\{\beta_i\}_{i=1}^p$, θ and λ are deterministic coefficients and .The distributions of z_t are the same as for GARCH(p,q)

Chapter 6

Results

The data used in this thesis are from daily copper prices from the London Metal Exchange during the period 1997-01-03 – 2013-01-03, with a total of 4175 data points. This is a fair amount of data. For the second section of the analysis the 4175 data points were transformed into 192 data points (the three data points in January of 2013 were not used since they were only three). The original data points were transformed in such a way that the new data points are the average monthly copper price. 192 data points is not much but it is what could be found and is still useful. Rare events, events happening with a probability of 0.01 or 0.001 might not even occur during this time period (possibly two events with a probability of 0.01 is to be expected), while in the first set of data points some forty events with probability of 0.01 are to be expected.

This is the trouble with trying to do analysis of average monthly data. The time span over which the data are being collected needs to be huge in order to get enough data to do a proper analysis.

As pointed out earlier, the aim of this thesis is to look at risks in the change in copper price. The risk will be measured with the risk metric called Value at Risk.

For the one day VaR estimations, the models described earlier (and again forthcoming) is sufficient to estimate the VaR quantiles. But for the three month predictions of VaR, the models are not sufficient and to be able to get the predictions, simulations of logreturns for future values will be used to predict the logreturns and thus the VaR for the logreturns.

As described earlier, the confidence intervals of estimated price $I_{P_{t+3}} = [P_t e^f, P_t e^g]$, where $I_{\hat{r}_{t+1}+\hat{r}_{t+2}+\hat{r}_{t+3}} = [f, g]$ and thus the

$$\mathrm{VaR}_\gamma(P_{t+3}) = P_t e^{\mathrm{VaR}_\gamma(\hat{r}_{t+1}+\hat{r}_{t+2}+\hat{r}_{t+3})}$$

.

In order to estimate the value at risk for the monthly copper price three months in the future, the $\hat{r}_{t+1}, \hat{r}_{t+2}$ and $\hat{r}_{t+3}$ will be estimated using simulations.

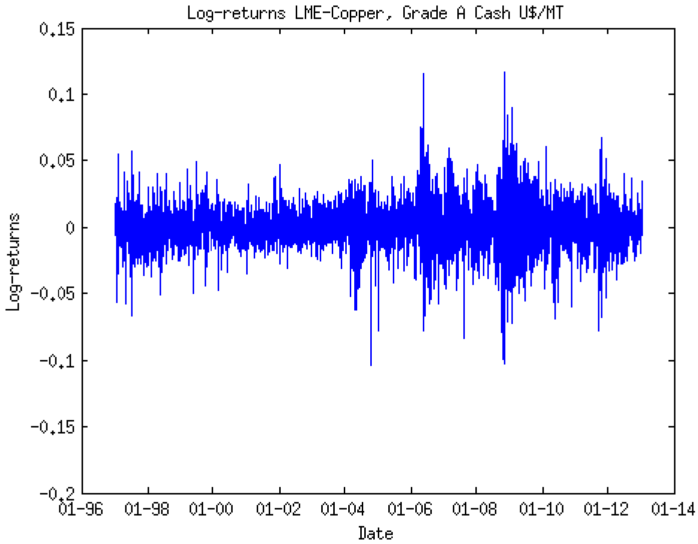

Figure 6.1: Log-returns of daily copper prices from LME from January 4th 1997 to January 3rd 2013.

When looking at the log-returns plot from the daily data from LME in figure 6.1, we can not see any trends and hence it appears to be stationary. The normplot, as can be seen in figure 6.2 gives a somewhat good fit (one normally does not expect log-returns to be normally distributed. Usually the tails are too fat). However, as expected the tails are too fat for the distribution being normal.

Finally looking at the sample autocorrelation plot of the log-returns in figure 6.3 we see some lags that are larger than what could be considered as white noise. There are some lags that are just outside of the 95% confidence interval of the standard deviation. The lags that lies outside of the 95% confidence interval are lag one, lag four, lag 40 and lag 49. The last two lags are however not really of any concern, when speaking of daily data points, as there are no real reason for their being any dependence between r_t and r_{t-50} at the same time as there are non for almost all the other time intervals. The correlations $Corr(r_t, r_{t-1})$ and $Corr(r_t, r_{t-4})$ are also small but more likely be real as opposed to $Corr(r_t, r_{t-50})$, considering that we believe that there is dependencies in the data. We see in figure 6.1 that there are volatility clusterings which indicate that there are dependence in the data.

This led to the conclusion to try and fit a volatility model to the data. In this thesis different GARCH -models where chosen and then compared to see which fits the best. The models are:

1. GARCH(p,q)-models

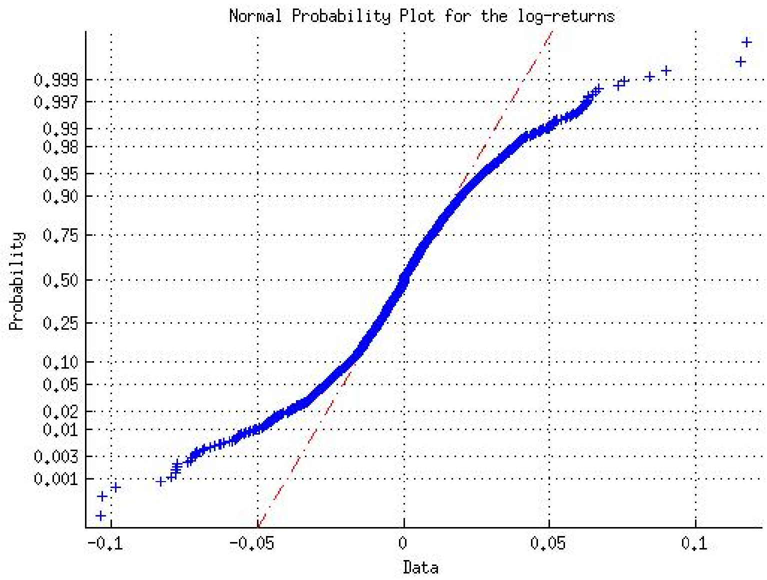

Figure 6.2: Normal probability plot of log returns daily copper prices from LME from January 4th 1997 to January 3rd 2013.

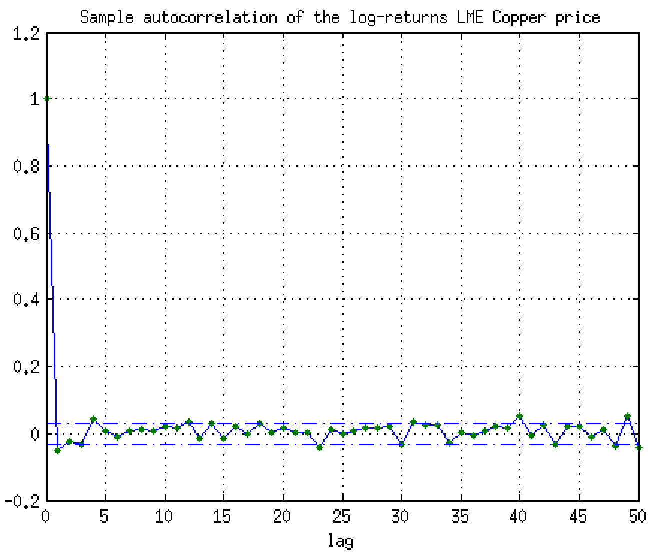

Figure 6.3: Sample autocorrelation of log-return of daily copper prices from LME from January 4th 1997 to January 3rd 2013.

2. EGARCH(p,q)-models

These models or families of models are very common to use. Both GARCH(p,q)-models and EGARCH(p,q)-models have the advantages of being relativly easy to estimate the parameters using Maximum Likelihood Estimation (MLE). A property which is normally difficult to find in a model.

6.1 GARCH(p,q)-models and EGARCH(p,q)-models

The Oxford MFE Toolbox, which can be found at `http://www.kevinsheppard.com/MFE_Toolbox`, was used for the estimations of the parameters for the different GARCH(p,q)-models and EGARCH(p,q)-models.

First out was the models from GARCH(1,1) to GARCH(10,10) with the assumption that the log-returns $r_t \in N(0, \sigma_t^2)$. Almost regardless of which model was used, the estimated parameter that got the largest value was one of the β-parameters.

Starting with GARCH(1,1)[1] we get the following model after parameter estimations[2]:

$$\hat{\sigma}_t^2 = \hat{\alpha}_1 r_{t-1}^2 + \hat{\beta}_1 \hat{\sigma}_{t-1}^2 = 0.046 r_{t-1}^2 + 0.947 \hat{\sigma}_{t-1}^2. \tag{6.1}$$

A quick look at the parameters gives $\hat{\alpha}_1 + \hat{\beta}_1 = 0.046 + 0.947 = 0.993$ which is very close to one. If $\alpha_1 + \beta_1 > 1$ the model is unstable.

Next, estimating the parameters for GARCH(2,1) through GARCH(10,1) gives that only three parameters were non-zero and it was $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\beta}_1$. which made the volatility process look like this:

$$\hat{\sigma}_t^2 = \hat{\alpha}_1 r_{t-1}^2 + \hat{\alpha}_2 r_{t-2}^2 + \hat{\beta}_1 \hat{\sigma}_{t-1}^2 = 0.02 r_{t-1}^2 + 0.03 r_{t-2}^2 + 0.94 \hat{\sigma}_{t-1}^2. \tag{6.2}$$

Also for this volatility process, the sum of parameters is very close to one, $\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\beta}_1 = 0.02 + 0.03 + 0.94 = 0.99$.

When checking the sum of the parameters of the other GARCH(p,q)-models (GARCH(1,1) to GARCH(10,10)), it is found that all of them sum close to one.

What is also remarkable is that for all the models the parameters connected to previous log-returns are all very low, $0 < \alpha_i < 0.1$ for $i = 1, \ldots, 10$, while the parameters connected to previous volatility terms are either zero or much larger than the α-parameters. Especially if the number of β-parameters are low i.e. GARCH(1,1) or GARCH(1,2) where the parameter connected

[1] the most common GARCH(p,q) model.

[2] $\hat{\alpha}_0$ was estimated to zero in all the estimation for daily data and has thus been left out.

to the earliest volatility term is the the highest and close to 0.9. The more parameters in the model, the less the earliest volatility term contributes to the updated volatility term, since the other terms in the process contributes somewhat.

In conclusion, previous terms in the volatility process contributes more to the updated volatility term than previous terms in the log-returns and among the volatility terms only one will make a big contribution.

After estimating the parameters it is time to test the model. To do that the distribution and the correlation structure must be analyzed. Remembering that the model is $r_t = \sigma_t z_t$ where z_t is said to be a zero mean i.i.d. process with unit variance, the log-returns needs to be divided by the square root of the volatility process which, if the model is correct, should give residuals $\hat{z}_t$ belonging to the correct distribution, i.e.

$$\hat{z}_t = \frac{r_t}{\hat{\sigma}_t}.$$

To find out if the model is correct, the quantile-quantile plot, QQ plot, of $\hat{z}_t$ is used to determine if the distribution is correct. The QQ plot for testing if $\hat{z}_t$ belongs to the normal distribution is called normplot. For evaluting this, if possible, built-in functions in matlab were used. If the residuals belongs to the distribution, the dots in the plot should follow a straight line.

There are no visible difference in the normality plots for any of the models. All of them look the same and they look fairly normal. There are tails in the plots, which is always expected, but these are a bit bigger than expected. See for instance figure 6.4

Next is the property of independence to be tested. For this a sample autocorrelation plot was used. This plot indicates if there exist any correlation between residuals at different times. As long as the points in the graph are within the confidence interval marked out by dashed lines, the autocorrelation is considered to be zero. For uncorrelated normal distributed random variables it is a well known fact that they are also independent random variables. For other distributions this is not the case, but in this thesis testing for correlation will be used as a approximation of independence.

The sample autocorrelation plots also looks identical and they look good. There are no visible autocorrelation among the terms in the process. See for example figure 6.5

One interesting aspect is that the volatility process does not change much if the GARCH-model is changed as can be seen in figure 6.6 where the ten different graphs looks like one graph with an extra solid line.

This leads to the conclusion that a GARCH(1,1) fits data pretty well, but still not good enough.

When doing the same analysis of GARCH(1,1) to GARCH(10,10) for Student's-t distribution, the same patterns appears as for normally distributed GARCH(p,q). It is no difference in the sample autocorrelation plots,

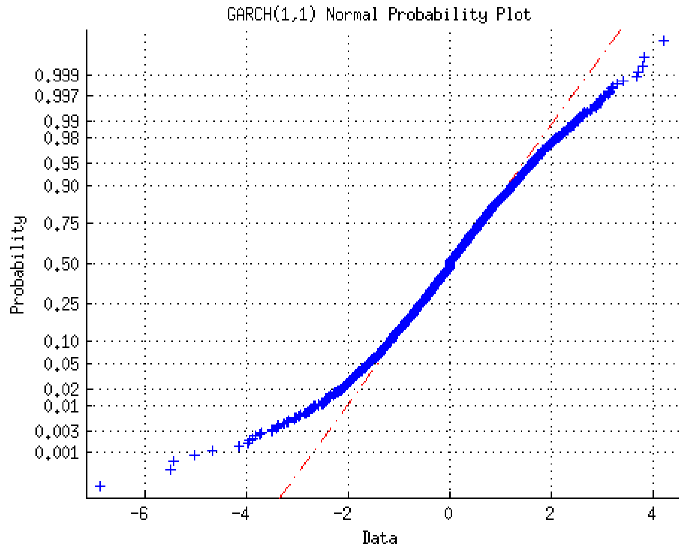

Figure 6.4: Normal probability plot of estimated residuals of the log-returns for GARCH(1,1) with normally distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013.

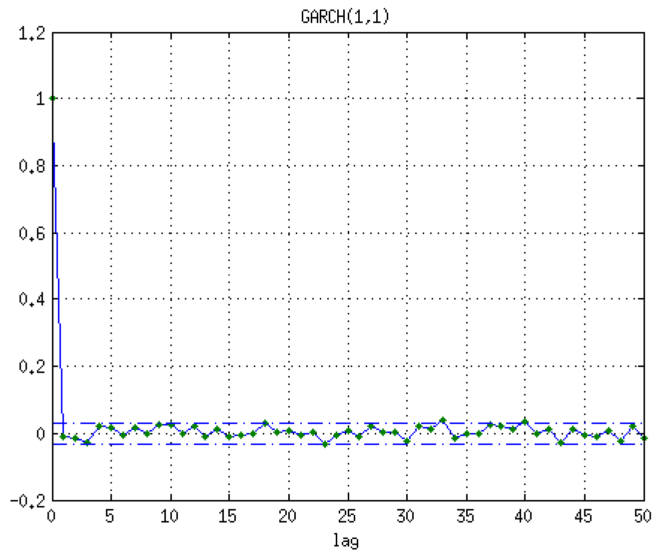

Figure 6.5: Sample autocorrelation of the estimated residuals of the log-returns for GARCH(1,1) with normally distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013.

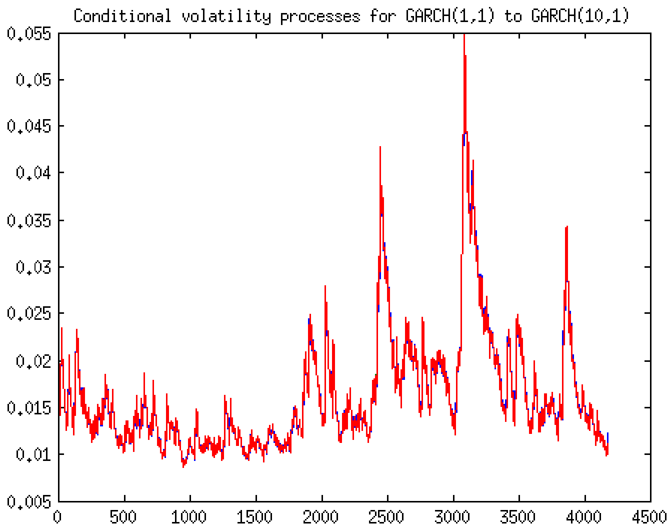

Figure 6.6: Conditional volatility for GARCH(1,1) to GARCH(10,1)

no difference in the Student's-t probability plots, and very little difference in the volatility processes from the different models.

Comparing the Student's-t probability plots to the normally distributed probability plots[3] shows that Student's-t distribution is preferred since it has a better fit with less heavy tails as can be seen in figure 6.7.

The degrees of freedom for the innovations with Student's-t distribution was estimated to $\nu = 6.5874$ for GARCH(1,1) and for the other GARCH models with innovations with Student's-t distribution, the degrees of freedom estimate was $6 < \nu < 7$.

Before leaving the GARCH(1,1) through GARCH(10,10), Generalized Error Distribution (GED) was also considered. As with the previous models, the QQ plots do not differ between the different models under the assumption of GED. Neither do sample autocorrelation plots. For example see in figure 6.8 the QQ plot of the residuals of the log-returns for GARCH(1,1) and in figure 6.9 the sample autocorrelation plot of the residuals of the log-returns for GARCH(1,1).

Since there are no matlab function implemented to create a QQ plot for GED, I had to create one myself by using existing matlab function qqplot(x,y) where x is a $n \times 1$ vector of log-return data and y is a $n \times 1$ vector of random numbers drawn from the Generalized Error Distribution.

Of the different GARCH(p,q)-models, the one with the assumption of GED had the best fit. Since there were no difference between the models with GED assumption, GARCH(1,1) is considered the best fit. There is no need

[3] or one of each since they look the same

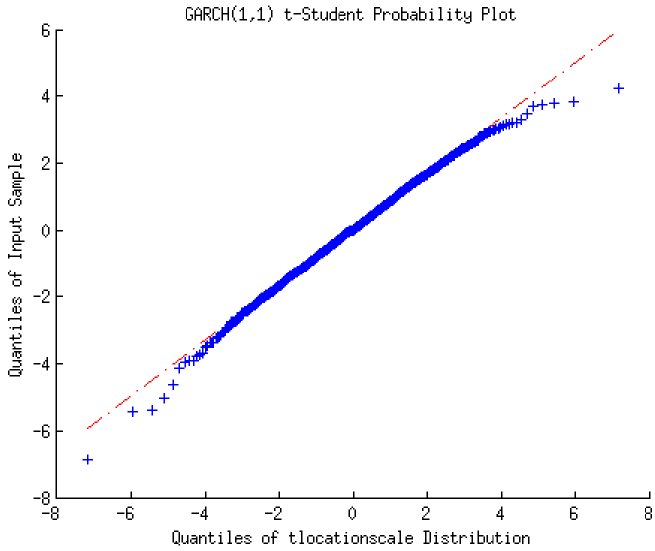

Figure 6.7: Student's-t QQ plot of the estimated residuals of the log-returns for GARCH(1,1) with Student's-t distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013.

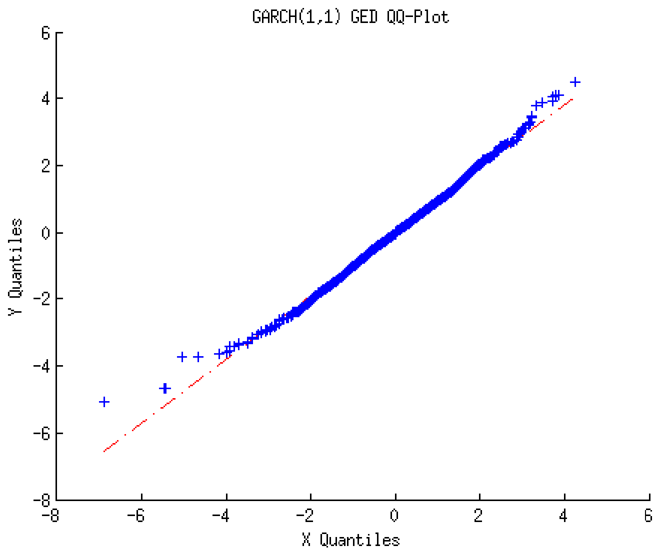

Figure 6.8: QQ plot of the estimated residuals of the log-returns for GARCH(1,1) with Generalized Error Distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013.

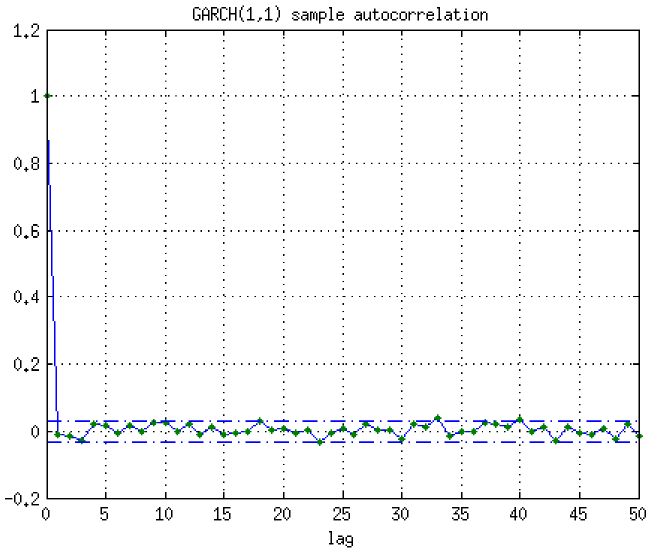

Figure 6.9: Sample autocorrelation the estimated residuals of the log-returns for GARCH(1,1) with Generalized Error Distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013.

to work with more complex models if they do not provide any improvements compared to simpler ones.This leads to the following model of the volatility process $\hat{\sigma}_t^2$:

$$\hat{\sigma}_t^2 = \hat{\alpha}_1 r_{t-1}^2 + \hat{\beta}_1 \hat{\sigma}_{t-1}^2 = 0.046 r_{t-1}^2 + 0.946 \hat{\sigma}_{t-1}^2 \tag{6.3}$$

and the full model becomes:

$$\hat{r}_t = \hat{\sigma}_t z_t, z_t \in GED(\hat{\nu}), \tag{6.4}$$

where $\hat{\nu} = 1.33$.

The 95% confidence intervals for the parameters are:

$\hat{\alpha}_1$	0.030	0.062
$\hat{\beta}_1$	0.928	0.965
$\hat{\nu}$	1.226	1.432

Also the EGARCH(p,q)-models behave as the regular GARCH(p,q)-models in the sense that it has little effect on the outcome if EGARCH(1,1) is used or if EGARCH(10,10) is used. There are no visible differences in the normal probability plots, no visible difference in the sample autocorrelation plots and very small differences in the conditional volatility processes. Only when putting two or more volatility processes in the same plot, there are visible differences and they are so small that it does not matter which one is used. As for the GARCH-case, the plot of ten different graphs of different EGARCH volatility processes, as can be seen in figure 6.10, looks like one graph with an extra solid line.

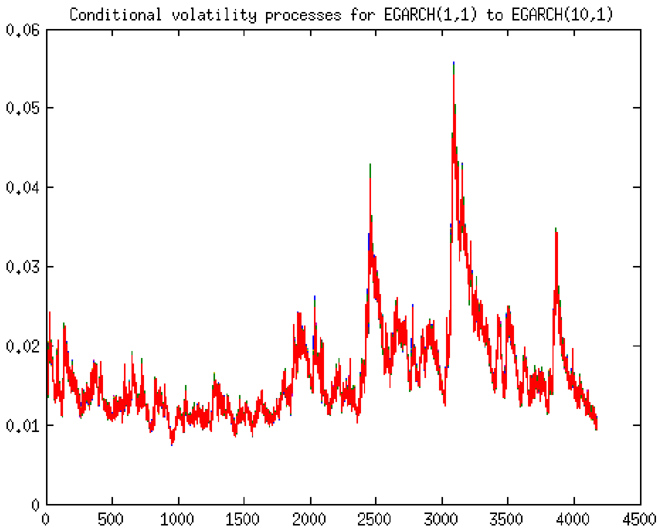

Figure 6.10: Conditional volatility for EGARCH(1,1) to EGARCH(10,1)

When comparing the EGARCH(p,q)-models to the previous models, GARCH(p,q) using Generalized Error Distribution is still the best. See the normal probability plot of the estimated residuals of the log-returns for EGARCH(1,1) with normal distributed innovations in figure 6.11 and the sample autocorrelation plot in figure 6.12. The EGARCH(p,q)-models are indistinguishable from the regular GARCH(p,q)-models, thus making a less involved model the best choice.

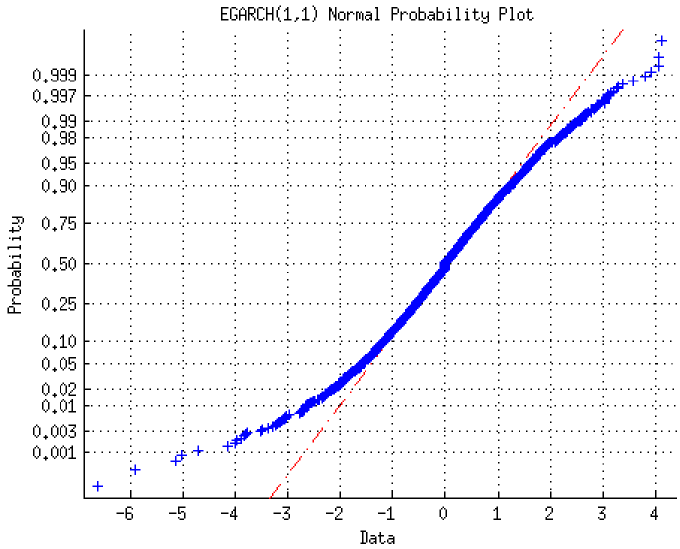

Figure 6.11: Normal probability plot of the estimated residuals of the log-returns for EGARCH(1,1) with normal distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013.

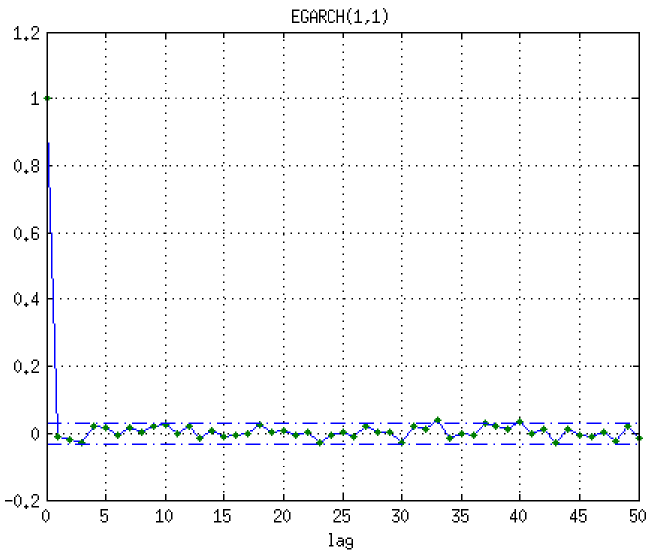

Figure 6.12: Sample autocorrelation the estimated residuals of the log-returns for EGARCH(1,1) with normal distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013

6.2 Value at Risk results

Now that the model is chosen for the data, it is time to compute the Value at Risk for the copper price. The time period will be one day and the confidence levels,γ, will be 95%, 99% and 99.9%. From previous chapters, the condfidence interval is defined from the following equation:

$$\mathrm{P}(\mathrm{VaR}_{\gamma}^{\mathrm{lo}}(X) < X < \mathrm{VaR}_{\gamma}^{\mathrm{up}}(X)) = \gamma \tag{6.5}$$

This definition is only good for a single outcome of the stochastic process describing the copper prices. Thus for the purposes here the definition will be expanded to

$$\mathrm{P}(\mathrm{VaR}_{\gamma}^{\mathrm{lo}}(\hat{r}_t) < \hat{r}_t < \mathrm{VaR}_{\gamma}^{\mathrm{up}}(\hat{r}_t)) = \gamma,\ 0 < t < T. \tag{6.6}$$

Now a confidence band is received instead of a confidence interval. In the plots below you will see the residuals of the log-returns, the confidence band and asterixes indicating that the residuals breached the confidence band at that time. The confidence bands is received from the $\frac{\gamma}{2}$-quantile and the $1-\frac{\gamma}{2}$-quantile as:

$$\begin{aligned} \mathrm{VaR}_{\gamma}^{\mathrm{lo}}(\hat{r}_t) &= \sigma_t F^{-1}(\frac{\gamma}{2};\hat{\nu}) \\ \mathrm{VaR}_{\gamma}^{\mathrm{up}}(\hat{r}_t) &= \sigma_t F^{-1}(1-\frac{\gamma}{2};\hat{\nu})) \end{aligned} \tag{6.7}$$

where $F^{-1}(\frac{\gamma}{2};\nu)$ is the quantile function of the GED with $\hat{\nu} = 1.33$ and $\gamma = (0.05, 0.01, 0.001)$.

Proceeding with VaR for GARCH(1,1) under the assumption of Generalized Error Distribution leads to the following figures which, together with the table below, shows that the one-day Value at Risk for GARCH(1,1) works pretty well.

	$\mathrm{VaR}_{0.05}$	$\mathrm{VaR}_{0.01}$	$\mathrm{VaR}_{0.001}$
Number of breaches	180	44	7
Breach ratio	0.0431	0.0105	0.0017

In the case of 95%-VaR it is expected that 5% of the residuals break through the confidence band. Looking in the table it is seen that breach ratio is 0.0431 which is fairly close to 0.05. If the model would be perfect, then the breach ratio would be exactly 0.05, but since no model is perfect 0.0431 is close enough for saying that the model works. Since the breach ratio is a bit less then 0.05 the model has made the the confidence bands a bit to wide and thus overestimating the losses.

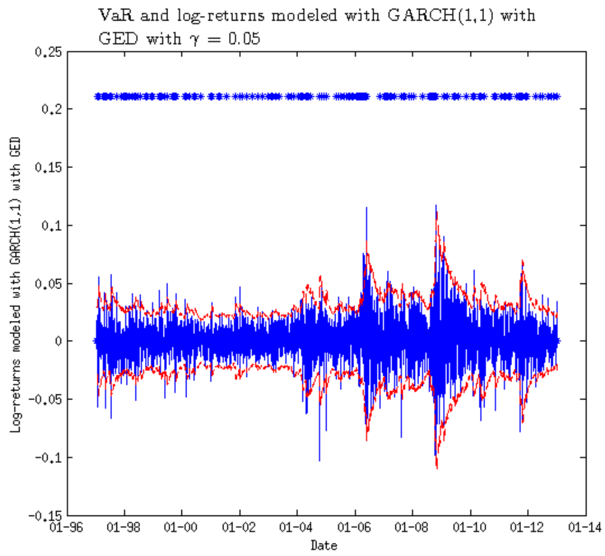

Figure 6.13: Confidence bands for the estimated residuals of the log-returns for GARCH(1,1) with Generalized Error Distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013, for $\text{VaR}_{0.05}$ and asterisks for every time the residuals broke the interval barriers.

Also the case of 99%-VaR looks good. Here the model has made the confidence bands a bit to narrow and thus underestimating the losses. But the error is small enough for it to be considered a good fit. When looking at the 99.9%-VaR the breach ratio is actully quite high and is exceeding the 0.1% level by 70%, but considering that there were only seven breaches in the dataset which spans roughly 16 years of daily data, then it is a good enough fit. With not that much data one extra breach will influence the breach ratio a lot. In order for the model to work better for the 99.9%-VaR it would have needed a lot more data. Only 16 years of daily data is clearly not enough. In this perspective the model works good enough.

To make sure that the VaR measure works, one can do a unconditional coverage test and an independence test. The unconditional coverage test tests if the proportions of failures of the VaR model is equal to γ. The independence test tests weather the failures are independent of one another. If the failures are not independent then the VaR metric is wrong since it says that at any given time the chance of failure is γ%. Thus if, as an example, the failures always comes in pairs, the probability of getting a second failure after the first one is not γ% as it should be.

When doing the unconditional coverage test proposed by Kupiec (1995) which is used in *A review of backtesting and backtesting procedures* by Sean D. Campbell in 2005, the following results were found:

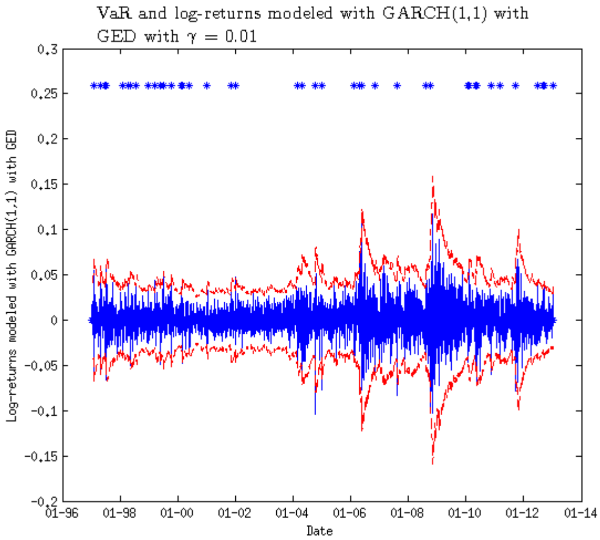

Figure 6.14: Confidence bands for the estimated residuals of the log-returns for GARCH(1,1) with Generalized Error Distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013, for $\mathrm{VaR}_{0.01}$ and asterisks for every time the residuals broke the interval barriers.

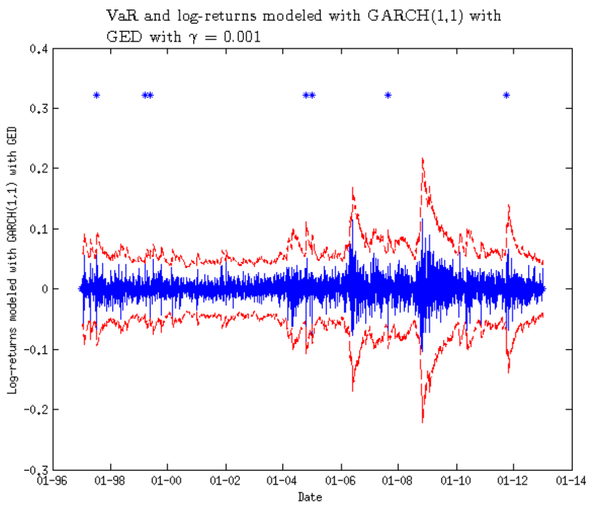

Figure 6.15: Confidence bands for the estimated residuals of the log-returns for GARCH(1,1) with Generalized Error Distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013, for $\mathrm{VaR}_{0.001}$ and asterisks for every time the residuals broke the interval barriers.

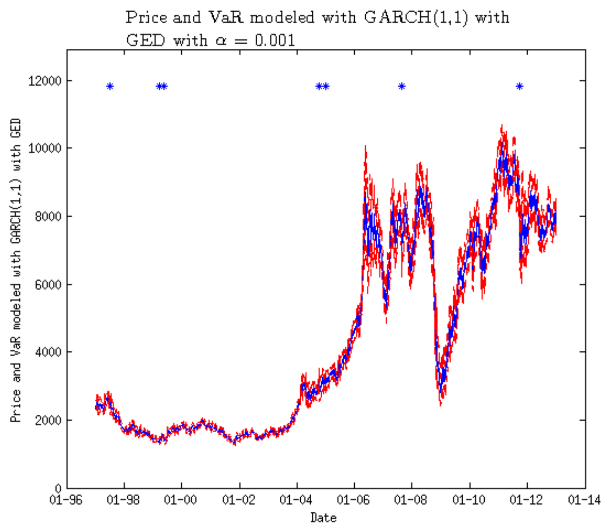

Figure 6.16: Confidence bands for the daily copper price for GARCH(1,1) with Generalized Error Distributed innovations, of the daily copper prices from LME from January 4th 1997 to January 3rd 2013, for $\text{VaR}_{0.001}$ asterisks for every time the residuals broke the interval barriers.

	$\text{VaR}_{0.05}$	$\text{VaR}_{0.01}$	$\text{VaR}_{0.001}$
POF	4.3485	0.1215	1.5884

where POF stand for proportion of failures and is defined as follows:

$$\text{POF} = 2\log\left(\left(\frac{1-\hat{\gamma}}{1-\gamma}\right)^{T-I(\gamma)}\left(\frac{\hat{\gamma}}{\gamma}\right)^{I(\gamma)}\right) \tag{6.8}$$
$$\hat{\gamma} = \frac{1}{T}I(\gamma)$$
$$I(\gamma) = \sum_{t=1}^{T} I_t(\gamma).$$

The $I_t(\gamma)$ is a so called 'hit' function defined as:

$$I_t(\gamma) = \begin{cases} 1 & \text{if} \quad x_{t,t+1} \leq -VaR_t(\gamma) \\ 0 & \text{if} \quad x_{t,t+1} > -VaR_t(\gamma) \end{cases} \tag{6.9}$$

where $x_{t,t+1}$ is the profit or loss on the portfolio over a fixed time interval from t to $t+1$.

If the proportions of failures are the same as the γ then the POF statistic will be zero.

The null hypothesis and the alternative hypothesis for the POF-test are

$$\begin{aligned} H_0 : & I(\gamma) \in \text{Bin}(T, \gamma) \\ H_1 : & I(\gamma) \in \text{Bin}(T, p),\ p \neq \gamma \end{aligned} \tag{6.10}$$

Under the null hypothesis that the model is correct, POF is asymptotically χ^2 distributed with one degree of freedom. The null hypothesis can be rejected, with 5% significance level, if POF is larger than the 5%-quantile of the χ^2 distribution which is equal to 3.84. Thus it can be seen that the POF for $\text{VaR}_{0.05}$ is larger than 3.84 and the null hypothesis can be rejected on the 5% level (while the null hypothesis for neither $\text{VaR}_{0.01}$ nor $\text{VaR}_{0.001}$ can be rejected on the 5% level).

The null hypothesis for $\text{VaR}_{0.05}$ cannot be rejected if the significance level is lower than 3.6%. Since the significance level of 3.6% is not to far from the 5% significance level, the overall performance of the model is good enough.

For testing the independence property, the Markov test suggested by Christoffersen (1998) through *A Review of Backtesting and Backtesting Procedures*, was used. It compares the value of the hit function at time t with the hit function at time $t-1$ and counts the number of time each event take place, i.e. the events are denoted $N_{00}, N_{01}, N_{10}, N_{11}$and the number of outcomes of each event are defined as:

$$\begin{aligned} N_{00} &= \#(I_t(\gamma) = 0 | I_{t-1}(\gamma) = 0) \\ N_{01} &= \#(I_t(\gamma) = 1 | I_{t-1}(\gamma) = 0) \\ N_{10} &= \#(I_t(\gamma) = 0 | I_{t-1}(\gamma) = 1) \\ N_{11} &= \#(I_t(\gamma) = 1 | I_{t-1}(\gamma) = 1) \\ N &= N_{00} + N_{01} + N_{10} + N_{11} \end{aligned}$$

	$I_{t-1}(\gamma) = 0$	$I_{t-1}(\gamma) = 1$	
$I_t(\gamma) = 0$	N_{00}	N_{10}	$N_{00} + N_{10}$
$I_t(\gamma) = 1$	N_{01}	N_{11}	$N_{01} + N_{11}$
	$N_{00} + N_{01}$	$N_{10} + N_{11}$	N

If $\frac{N_{00}}{N_{00}+N_{01}} = \frac{N_{10}}{N_{10}+N_{11}}$ then the events are considered independent.

In this thesis these results were received:

	$\text{VaR}_{0.05}$	$\text{VaR}_{0.01}$	$\text{VaR}_{0.001}$
N_{00}	3816	4085	4159
N_{01}	177	44	7
N_{10}	177	44	7
N_{11}	3	0	0
N	4173	4173	4173

which gives us these quotients:

	$\mathrm{VaR}_{0.05}$	$\mathrm{VaR}_{0.01}$	$\mathrm{VaR}_{0.001}$
$\frac{N_{00}}{N_{00}+N_{01}}$	0.9557	0.9893	0.9983
$\frac{N_{10}}{N_{10}+N_{11}}$	0.9833	1.0000	1.0000
$\frac{N_{00}+N_{10}}{N}$	0.9569	0.9895	0.9983

As can be seen in the table above the qoutients in each column seems to be of approximately the same values, and hence the one day VaR-measure is accurate.

6.3 GARCH(p,q)-models and EGARCH(p,q)-models three months predictions

For Cavotec, one-day forecasts is not really enough. The timespan is too short. Since it takes usually two to three months to get delivery of the cables and also that the final price is set at delivery, a three month forecast is more sought after. Another aspect is that the daily spot price is usually not used when buying copper. A more common approach is to base the purchase price on the average of the daily spot prices over one month time, from now on called the monthly average price or MAP. Even if the copper is not traded directly at the commodity exchanges, there are similar contracts offered at them. See for instance the contract called Traded Average Price Option (TAPO) at the London Metal Exchange's (LME) website `http://lme.com/trading/contract-types/tapos/`, which can be used to hedge against the fluctuations in what they call the Monthly Average Settlement Price (MASP).

In this thesis the monthly average spot price of copper was computed as the sum of the prices in a particular month divided by the number of days copper was traded during that month, e.g.

$$M_k = \frac{1}{N_k} \sum_{i=1}^{N_k} P_k^i \tag{6.11}$$

where k is the month, N_k the number of trading days of month, P_k^i is the copper price of the i-th trading day of the month.

This means that the new monthly data points will not be based on the same number of daily data points, since the number of traded days will differ from one month to another. One might think of other ways to create monthly data points. For instance one could use 30 consecutive data points and take the average[4] or, perhaps more realistic in this setting, use 21 consecutive

[4] A common way to estimate the number of days in a month.

data points and take the average[5]. The method in this thesis was chosen since it resemble the way the Monthly Average Settlement Price (MASP) at LME is computed. This to make the results of this thesis to be applicable to the trading environment at LME[6]. `https://www.lme.com/en-gb/metals/reports/averages/`

I decided to also look at the monthly median of the daily copper prices and compare the average to the median to see if there are any skewed data sets among the daily copper prices and make sure not very skewed data influence the monthly average copper price. As can be seen in the following figures, the difference between the monthly median price and the monthly average price is not that large. Thus the daily copper prices are not skewed. All in all, the average seems like a good measure to represent the overall performance of the copper price during one month.

The estimated standard deviation S_k where k is the month, N_k the number of trading days of month t is computed as

$$S_k = \sqrt{\frac{1}{N_k - 1} \sum_{i=1}^{N_k} (P_k^i - M_k)^2}. \tag{6.12}$$

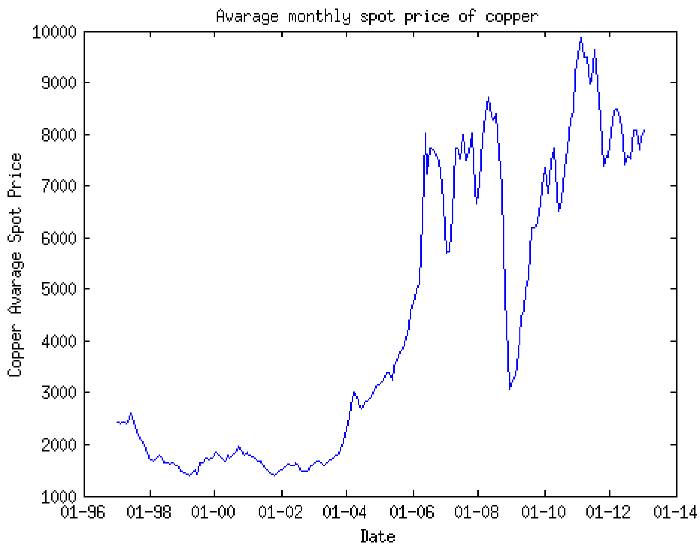

Figure 6.17: Monthly average copper price in US dollar per tonne. I would say that the curve look a lot like the one for the daily data, but with less 'noise'.

[5]There is usually on average 21 trading days in a month.

[6]One of the points of this thesis is to create support for Cavotec to hedge their risks on the financial markets.

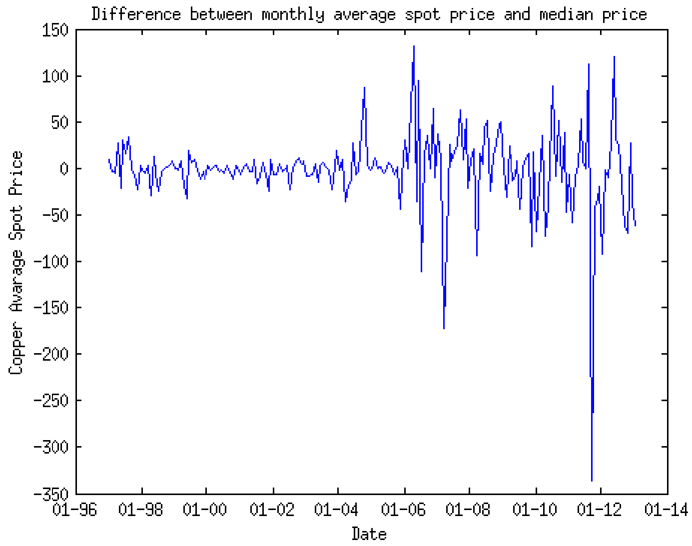

Figure 6.18: This plot shows the difference between the monthly average price and the monthly median price of copper.

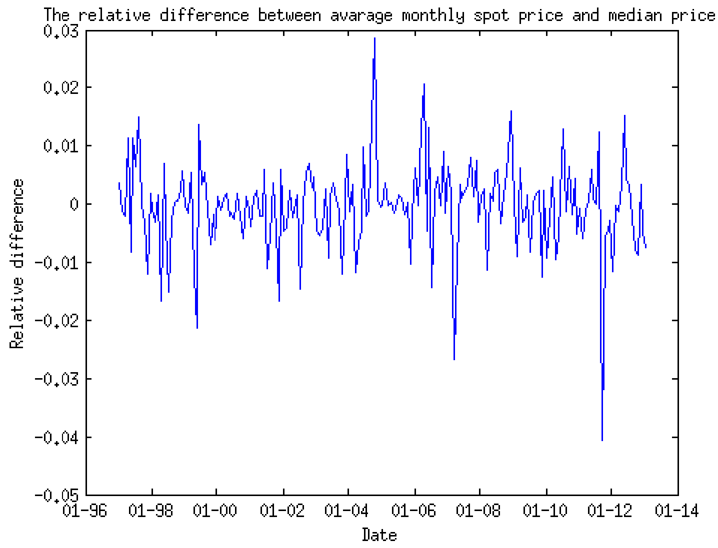

Figure 6.19: This plot shows the relative difference between the monthly average price and the monthly median price of copper. The average and the median only differs by at the most four percent.

The fact that figure 6.19 shows such low values in the relative metric of the monthly average copper price and the monthly median copper price, shows that the distribution of the monthly data is symmetric[7].

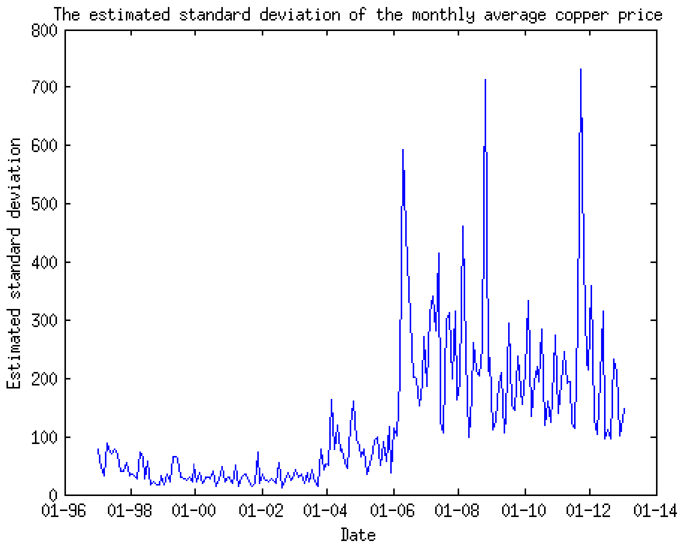

Figure 6.20: Estimated standard deviation of the monthly average copper price

The logreturns of the monthly average copper price r_k is defined as

$$r_k = \ln\left(\frac{M_{k+1}}{M_k}\right) = \ln(M_{k+1}) - \ln(M_k). \tag{6.13}$$

Since the data in figure 6.26 shows dependencies between r_k, r_{k-1} and r_k, r_{k-2}, these must first be modeled away. Looking at the plot of the sample autocorrelation function an AR(1) process or AR(2) process seems to be suitable. An AR(1) process is defined as a stochastic process $\{r_k\}_{k=1}^{T}$ according to the following:

$$\begin{aligned} r_k &= \phi r_{k-1} + \epsilon_t,\ \epsilon_k \in N(0, \sigma_k^2) \\ \sigma_k^2 &= \alpha_0 + \alpha\epsilon_{k-1}^2 + \beta\sigma_{k-1}^2. \end{aligned} \tag{6.14}$$

Similarly an AR(2) process is defined as

$$\begin{aligned} r_k &= \phi_1 r_{k-1} + \phi_2 r_{k-2} + \epsilon_k,\ \epsilon_k \in N(0, \sigma_k^2) \\ \sigma_k^2 &= \alpha_0 + \alpha\epsilon_{k-1}^2 + \beta\sigma_{k-1}^2. \end{aligned} \tag{6.15}$$

[7]or symmetric enough

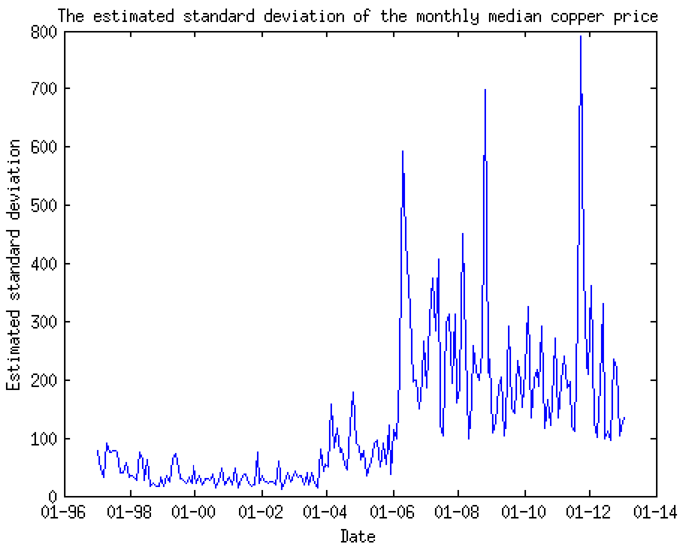

Figure 6.21: Estimated standard deviation of the monthly median copper price

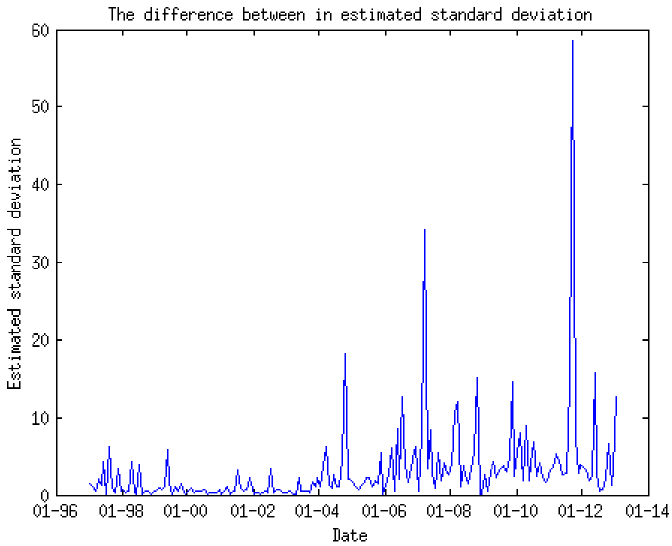

Figure 6.22: The difference between the estimated standard deviation of the monthly average copper price and the monthly median copper price.

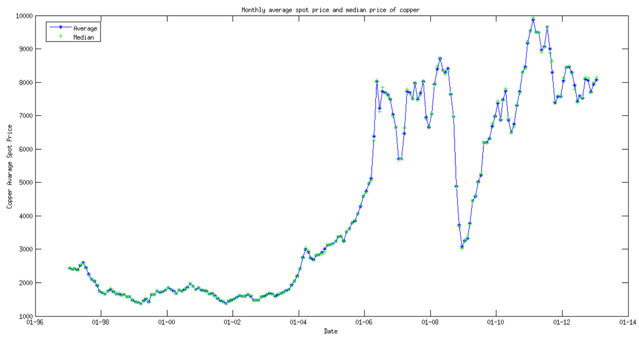

Figure 6.23: Here are the monthly average copper price and the monthly median copper price. There doesn't seem to be any large deviations between them in relative terms.

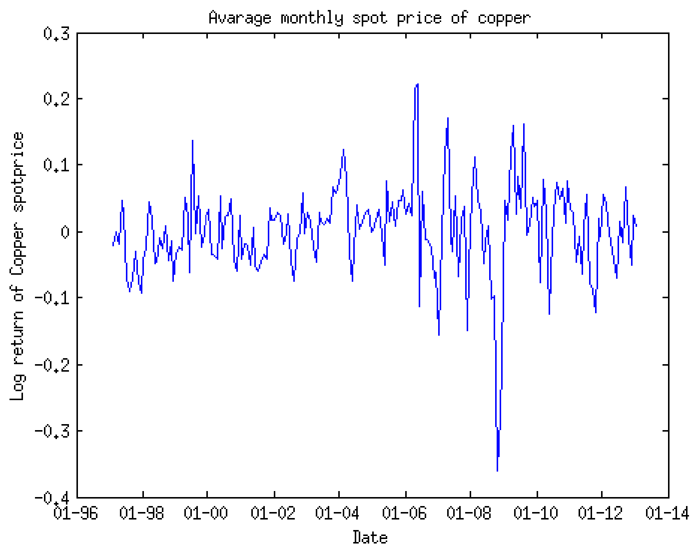

Figure 6.24: The log returns of the monthly average copper price. It doesn't look that well behaved.

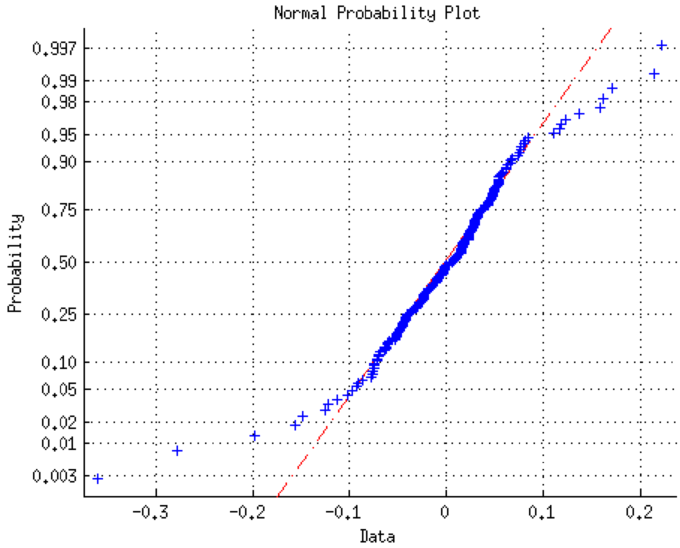

Figure 6.25: Normplot of the monthly average copper price. It has too many outliers too far away from the straight line, for the distribution of the sample to be considered normally distributed.

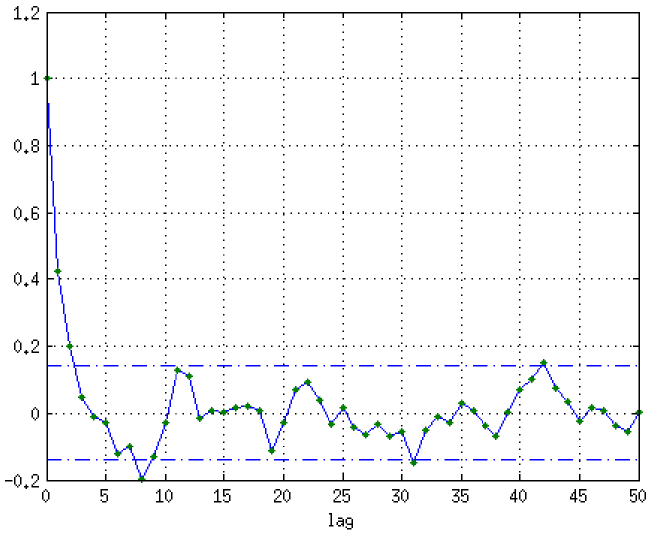

Figure 6.26: Plot of the sample autocorrelation function. If the sample were to show no correlation, all correlations should be inside the band marked by dashed lines. Here we see two lags of correlation that are above the dashed line, and the rest (almost) are inside the band. This suggests that the data should be modeled by a AR(1) or AR(2) process.

When testing if the AR(1) process is a suitable model, one must estimate the parameter (ϕ). The parameter, denoted $\hat{\phi}$ have been estimated by:

$$\hat{\phi} = \frac{\frac{1}{T-1}\sum_{k=1}^{T-1} r_{k+1}r_k}{\frac{1}{T-1}\sum_{k=1}^{T-1} r_k^2}$$

where r_k is the log returns of the data.

Similarly for the AR(2) process, the parameters, denoted $\hat{\phi}_1$, $\hat{\phi}_2$ were estimated by using the `arx.m` function in matlab. The estimated parameters became

	AR(1)
$\hat{\phi}$	0.4251

	AR(2)
$\hat{\phi}_1$	-0.4143
$\hat{\phi}_2$	-0.0257

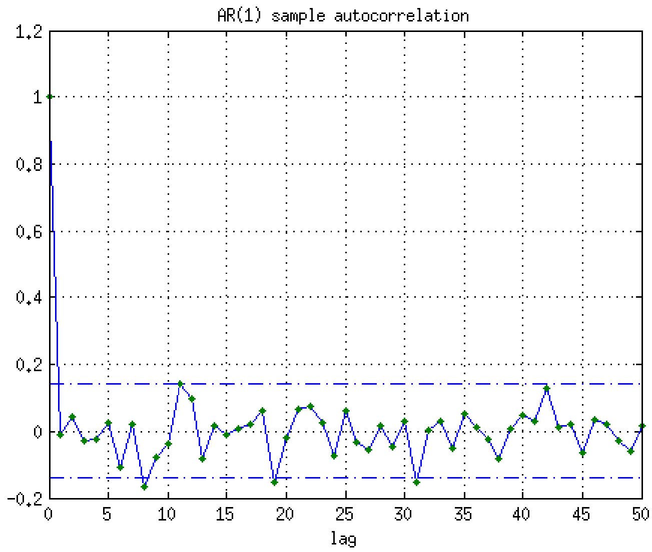

Figure 6.27: Sample autocorrelation plot of the log returns of the data when an AR(1) process is applied to it.

In figure 6.27 to figure 6.30 we can see that an AR(1) process fits the data better than an AR(2) process.

After concluding that the AR(1) fits best, the residuals are given as

$$\hat{\epsilon}_k = r_k - \phi r_{k-1}. \tag{6.16}$$

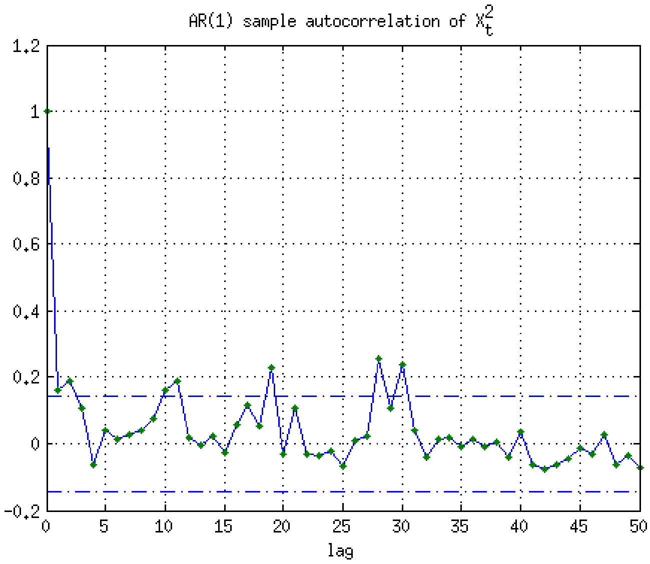

Figure 6.28: Sample autocorrelation plot of the squared log returns of the-data when an AR(1) process is applied to it.

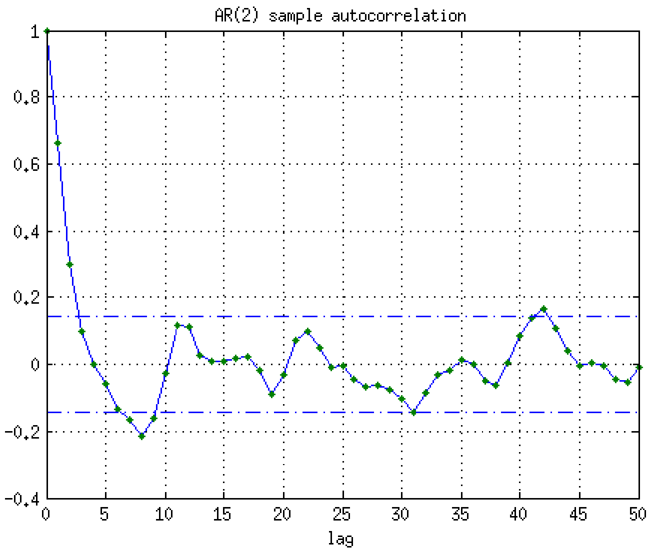

Figure 6.29: Sample autocorrelation plot of the log returns of the data when an AR(2) process is applied to it.

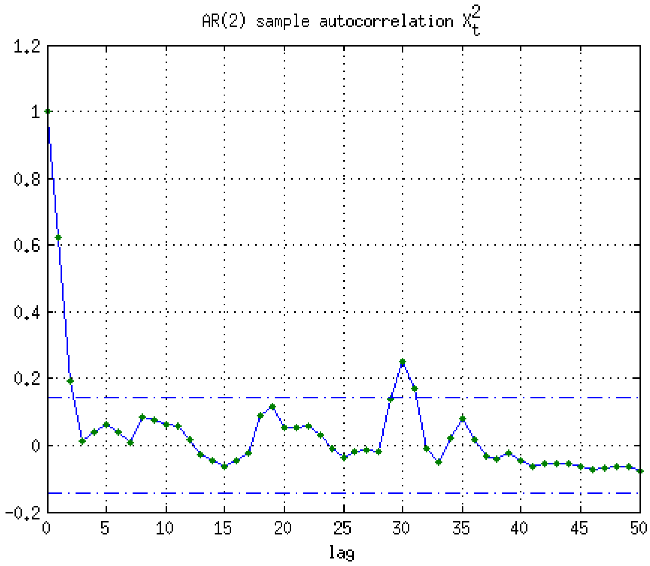

Figure 6.30: Sample autocorrelation plot of the squared log returns of the-data when an AR(2) process is applied to it.

When looking at figure 6.28 there still seems to be some correlation in the volatility, which suggests that the volatility needs to be modeled here as well as for the daily data.

The model now becomes

$$\hat{\epsilon}_k = \sigma_k z_k. \tag{6.17}$$
$$\sigma_k = \alpha_0 + \alpha_1 \hat{\epsilon}_{k-1}^2 + \beta \sigma_{k-1}^2.$$

The Oxford MFE Toolbox, which can be found at `http://www.kevinsheppard.com/MFE_Toolbox`, was again used for the estimations of the parameters for the different GARCH(p,q)-models and EGARCH(p,q)-models.

When estimating the parameters for GARCH(1,1), the following model after parameter estimations was received[8]:

$$\sigma_t^2 = \hat{\alpha}_1 \epsilon_{t-1}^2 + \hat{\beta}_1 \sigma_{t-1}^2 = 0.076\epsilon_{t-1}^2 + 0.892\sigma_{t-1}^2. \tag{6.18}$$

The 95% confidence intervals for the parameters are:

$\hat{\alpha}_1$	0.071	0.082
$\hat{\beta}_1$	0.884	0.899

The normalized residuals becomes

[8] α_0 was estimated to 0.00013889 which is very close to zero and has thus been left out.

$$\hat{z_k} = \frac{\hat{\epsilon}_k}{\sigma_k}. \tag{6.19}$$

Looking at the normplot in figure 6.31 and the sacfplot in figure 6.32 of the normalized residuals, the distribution is normal enough and with no correlations between the normalized residuals.

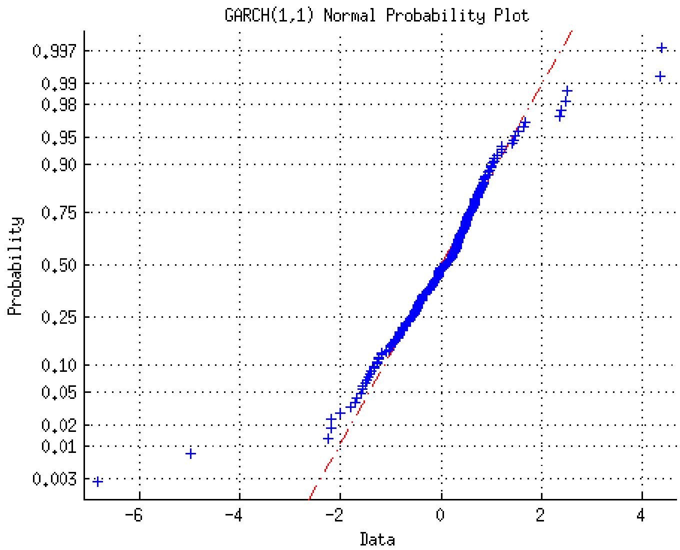

Figure 6.31: Normplot of the normalized residuals modeled by a GARCH(1,1) model.

Looking at its normplot in figure 6.33, we see that it is not yet normal enough.

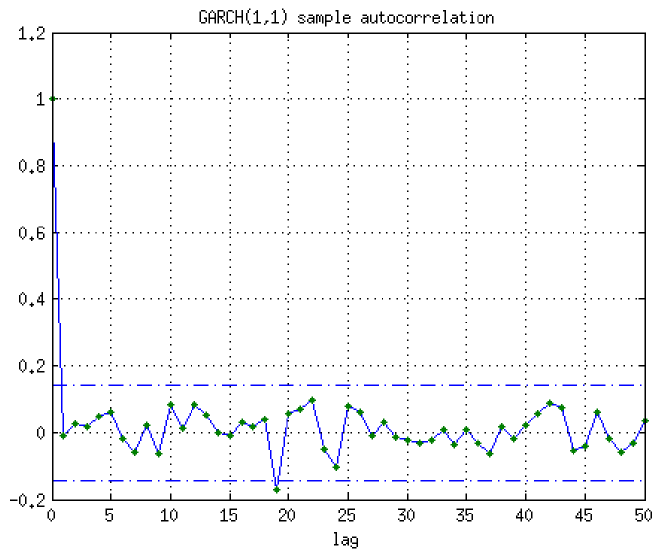

Figure 6.32: Sample autocorrelation plot of the normalized residuals modeled by a GARCH(1,1) model.

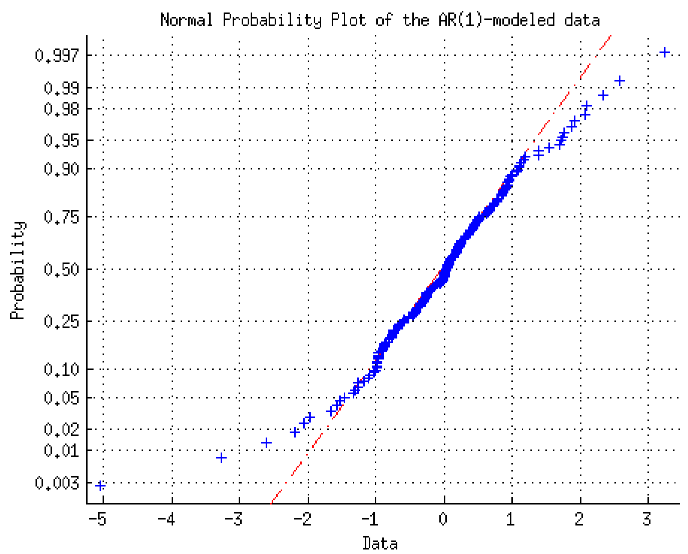

Figure 6.33: Normplot of the AR(1)-modeled data. Not yet normal enough.

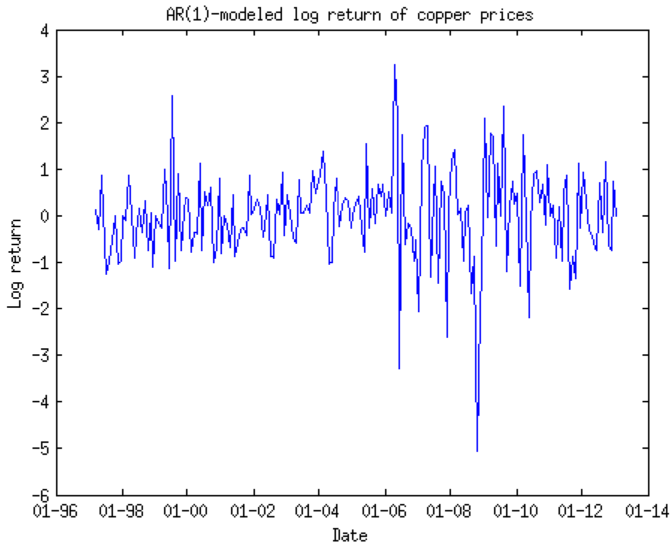

Figure 6.34: Log-return of AR(1)-modeled data. The varying volatility that was so clear for daily data is not as visible for monthly average copper prices.

6.4 Value at Risk three months predictions

After concluding which GARCH model to use, it is time to start estimating the value at risk for the monthly average copper price. As mentioned earlier, the $\mathrm{VaR}_\gamma(M_{k+3}) = M_k e^{\mathrm{VaR}_\gamma(\hat{r}_{k+1}+\hat{r}_{k+2}+\hat{r}_{k+3})}$. The values of $\hat{r}_{k+1}, \hat{r}_{k+2}$ and $\hat{r}_{k+3}$ will be estimated using simulations to get data points from the distribution of the normalized residuals to create one-, two- and three-month predictions of the residuals. It is done according to the algorithm below which is iterated N times (i goes from 1 to N). N is chosen sufficiently large to be able to estimate VaR for 95%, 99% and 99.5%. At least 2000 iterations is needed for VaR at 99.5%. $N = 3000$ was chosen in this thesis.

$$
\begin{aligned}
\sigma^2_{k+1,i} &= \alpha_0 + \hat{\alpha_1}\hat{\epsilon}^2_k + \beta_1\sigma^2_k \\
\hat{\epsilon}_{k+1,i} &= \sigma_{k+1,i}z_{k+1,i},\, z_{k+1,i} \in N(0,1) \\
\sigma^2_{k+2,i} &= \alpha_0 + \hat{\alpha_1}\hat{\epsilon}^2_{k+1} + \beta_1\sigma^2_{k+1} \\
\hat{\epsilon}_{k+2,i} &= \sigma_{k+2,i}z_{k+2,i},\, z_{k+2,i} \in N(0,1) \\
\sigma^2_{k+3,i} &= \alpha_0 + \hat{\alpha_1}\hat{\epsilon}^2_{k+1} + \beta_1\sigma^2_{k+2} \\
\hat{\epsilon}_{k+3,i} &= \sigma_{k+3,i}z_{k+3,i},\, z_{k+3,i} \in N(0,1)
\end{aligned} \tag{6.20}
$$

Now $\hat{r}_{k+1,i}, \hat{r}_{k+2,i}$ and $\hat{r}_{k+3,i}$ are given recursivly from $r_k = \phi r_{k-1} + \epsilon_t$ by

$$
\begin{aligned}
\hat{r}_{k+1,i} &= \hat{\phi} r_k + \hat{\epsilon}_{k+1,i} \\
\hat{r}_{k+2,i} &= \hat{\phi}^2 r_k + \hat{\phi}\hat{\epsilon}_{k+1,i} + \hat{\epsilon}_{k+2,i} \\
\hat{r}_{k+3,i} &= \hat{\phi}^3 r_k + \hat{\phi}^2\hat{\epsilon}_{k+1,i} + \hat{\phi}\hat{\epsilon}_{k+2,i} + \hat{\epsilon}_{k+3,i}.
\end{aligned}
\tag{6.21}
$$

VaR is estimated from the sample of N $\hat{r}_{k+1,i} + \hat{r}_{k+2,i} + \hat{r}_{k+3,i}$

$$
\text{VaR}_\gamma(M_{k+3}) = M_k e^{\text{VaR}_\gamma(\hat{r}_{k+1}+\hat{r}_{k+2}+\hat{r}_{k+3})} \tag{6.22}
$$

	$\text{VaR}_{0.05}$	$\text{VaR}_{0.01}$	$\text{VaR}_{0.001}$
Number of breaches	9	4	2
Breach ratio	0.0479	0.0213	0.0106

When doing the unconditional coverage test proposed by Kupiec (1995) which is used in *A review of backtesting and backtesting procedures* by Sean D. Campbell in 2005, the following results were found:

	$\text{VaR}_{0.05}$	$\text{VaR}_{0.01}$	$\text{VaR}_{0.001}$
POF	0.0158	1.8234	5.8509

The POF-test in this case indicates that the null hypothesis can be be rejected in the $\text{VaR}_{0.001}$ case on the 5% significance level. But on the 1.5% significance level it can no longer be rejected.

In the case of 95%-VaR it is expected that 5% of the residuals break through the confidence band. Looking in the table it is seen that breach ratio is 0.0479 which is fairly close to 0.05. If the model would be perfect, then the breach ratio would be exactly 0.05, but since no model is perfect 0.0479 is close enough for saying that the model works. Since the breach ratio is a bit less then 0.05 the model has made the the confidence bands a bit to wide and thus overestimating the losses. Since only having 187 data points in this set, every breach will affect the breach ratio. 5% out of 187 is 9.35 and since 9 breaches were received and only a whole number of breaches are possible, this is as close as can be expected to get.

In the case of 99%-VaR the breach ratio is 0.0209 and the number of breaches 4. Here, if the model works, one would expect a breach ratio of 0.01 and a breach ratio two times that was received. But again, due to the small number of data points in the set, the expected number of breaches is 1.87 and the number of breaches received was 4. Here the model has made the the confidence bands a bit too narrow and thus underestimating the losses.

When looking at the 99.9%-VaR the breach ratio is actually quite high and is exceeding the 0.1% level by almost a factor of 16. From the low

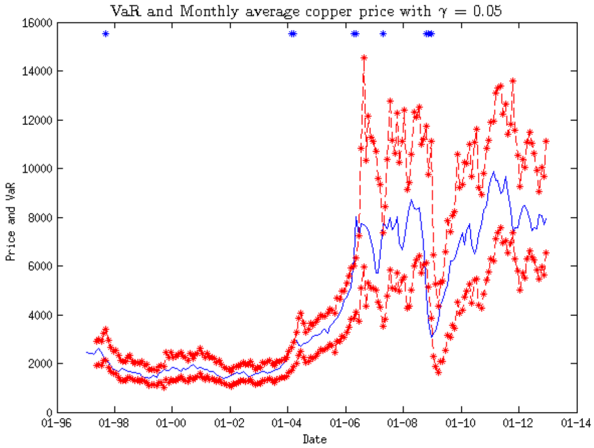

Figure 6.35: Confidence bands for the estimated residuals of the log-returns of the monthly average copper price for GARCH(1,1) with Normally distributed innovations, of the daily copper prices from LME from January 4th 1997 to December 31st 2012, for $\text{VaR}_{0.05}$ and asterisks for every time the residuals broke the interval barriers (in the figure only 5 asterisks are seen. The reason is that on two occasions two breaches occurred very close to each other, thus the asterisks are to close together to be visible).

number of data points in the set, getting any breach of the 0.1% confidence band, is unlikely. In this dataset 3 breaches were received and thus making the breach ratio very high.

Again using the Markov test suggested by Christoffersen(1998) these results were received:

	$\text{VaR}_{0.05}$	$\text{VaR}_{0.01}$	$\text{VaR}_{0.001}$
N_{00}	173	181	181
N_{01}	5	2	2
N_{10}	5	2	2
N_{11}	4	2	0
N	187	187	187

which gives us these quotients:

	$\text{VaR}_{0.05}$	$\text{VaR}_{0.01}$	$\text{VaR}_{0.001}$
$\frac{N_{00}}{N_{00}+N_{01}}$	0.9719	0.9891	0.9892
$\frac{N_{10}}{N_{10}+N_{11}}$	0.5556	0.5000	1.0000
$\frac{N_{00}+N_{10}}{N}$	0.9519	0.9786	0.9893

According to p.10 in *A Review of Backtesting and Backtesting Procedures* unconditional coverage and independence can be jointly tested and for VaR to be accurate then $\frac{N_{00}}{N_{00}+N_{01}} = \frac{N_{10}}{N_{10}+N_{11}} = \frac{N_{00}+N_{10}}{N} = 1 - \gamma$.

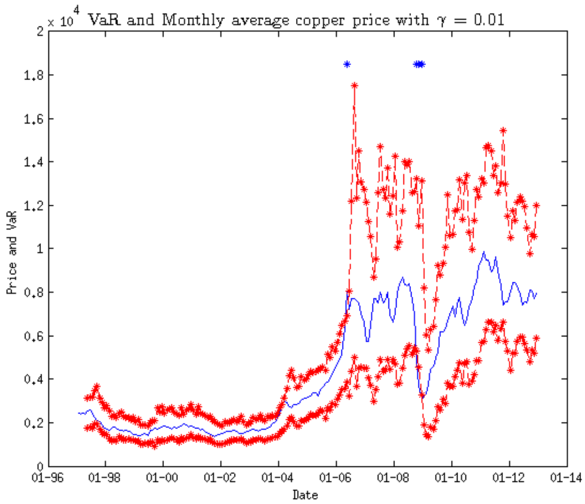

Figure 6.36: Confidence bands for the estimated residuals of the log-returns of the monthly average copper price for GARCH(1,1) with Normally distributed innovations, of the daily copper prices from LME from January 4th 1997 to December 31st 2012, for $\mathrm{VaR}_{0.01}$ and asterisks for every time the residuals broke the interval barriers.

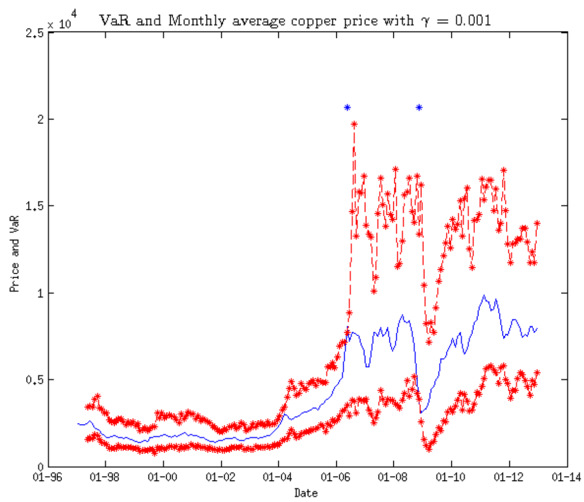

Figure 6.37: Confidence bands for the estimated residuals of the log-returns of the monthly average copper price for GARCH(1,1) with Normally distributed innovations, of the daily copper prices from LME from January 4th 1997 to December 31st 2012, for $\mathrm{VaR}_{0.001}$ and asterisks for every time the residuals broke the interval barriers.

As can be seen in the table above the quotients in each column are not equal and hence the monthly VaR-measure can not be seen as independent.

Chapter 7

Conclusion and summary

The aim with this thesis was to get a model that describes the copper price movements and to get a sense on big a risk one takes from dealing in copper. As for Cavotec Sweden, who among other things buys and sells copper cables, the risk is two sided in that they both buy and sell copper cables. Thus regardless of what the copper price is, the movements is the more important.

The GARCH model for the one day modeling of the copper price was estimated to be a GARCH(1,1) with innovations from a GED with $\hat{\nu} = 1.33$. The other distributions for which the innovations was chosen from, Normal and student-T, gave too thick tails (student-T) and thin tails to fit the data. This was confirmed by the estimated tail-thickness parameter $\hat{\nu}$ who was estimated to a value between 1 and 2.

Looking at the estimated parameters of the GARCH model, it is clear that previous values of the logreturns are not as important as the previous volatility values. The volatility terms play a large role in the model. The sum of the parameters are very close to one. If the sum of the parameters are larger than one, the model is no longer convergent and should not be used. Getting this close to the limit, this might indicate that improvements could be made using a different model. Since several other GARCH models and EGARCH models were tested in this thesis and discarded in favour of GARCH(1,1), perhaps some other class of models should be used to model the volatility.

The VaR of the three month predictions gives much wider confidence bands compared to the one day VaR.

Given that the number of data points for the average monthly copper price is only 192, the data set might not be large enough to do a VaR-measure. This is especially true for VaR-99% and VaR-99.9%, since they constructed to get 2 respectivly 0 events that cross the barriers.

Even though the monthly average copper price is, for Cavotec and from a trading point of view, more relevant compared to the one day copper price, the number of data points available makes the measurement poor.

In both cases, the risk metric captures the changes in the copper price fairly good, though in the case of the three month predictions, the measure was neither independent nor was the number of data points were perhaps not large enough to say that it was not by chance that the model worked as well as it did for the VaR-95%. However, it is still useful as a tool to use for evaluting if it is worth the effort to take positions on the financial markets in order to hedge against the price of copper cable being to high.

However, in order to catch the high volatility in the copper price, the risk metric produced very wide bands for the three month predictions. Since the part of the large volatility happened during a time period of economic crises in the western world, there are for sure other variables which influence the copper price and thus as an improvement on the model, one suggestion is to add some kind of explanatory variable to assist the model when extreme events such as the finacial crisis of 2008-2009 hit.

As for Cavotec, with the code developed in MATLAB® for this thesis and by continuing to get the historical copper prices, they can keep updating the model to see whether it predicts to much risk for them to act and go on the financial markets to hedge.

Since the VaR of the three month predictions are based on the current copper price and predictions on the coming logreturns, it is clear from the plots that the model lags a bit, which is to be expected as the further into the future one tries to predict, the more uncertain the predictions are. But, since lead time to delivery of the copper cable is quite long, there might still be time for Cavotec to go on the finacial markets and hedge against price changes that will affect the price of the copper cable upon delivery. Since there are several different products available at LME with almost endless variations, the type of scheme optimal for Cavotec to use is not clear and has not been addressed in this thesis and is thus yet another avenue to continue the research of the topic of this thesis.

Bibliography

[1] Black, Fischer; Scholes, Myron: *The Pricing of Options and Corporate Liabilities*. Journal of Political Economy Vol. **81** , No. 3, pp. 637-654, (1973)

[2] Engle, Robert F.: *Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation*, Econometrica vol. **50**, pp. 987-1008, (1982)

[3] Nelson, Daniel B.: *Conditional Heteroskedasticity In Asset Returns: A New Approach* Econometrica, Vol. **59** No. 2, pp. 347-370, (1991)

[4] Jorion, P. *Value at Risk: The new benchmark for managing financial risk*, second edn, McGraw-Hill (2001)

[5] Campbell, Sean D.: *A Review of Backtesting and Backtesting Procedures*, Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C., (2005)