Lund Institute of Technology
Master Thesis in Risk Management and Safety Engineering

VaR methods for linear instruments

Birgir Viðarsson

LUND INSTITUTE OF TECHNOLOGY
Lund University

Title: VaR methods for linear instruments
Author: Birgir Viðarsson

Instructor: Dr. Freyr Harðarson. Risk Management, Landsbanki
Examiner: Dr. Henrik Johansson. Department of Fire Safety Engineering and Systems Safety, Lund University

Report nr.: 5268
ISSN: 1402-3504
ISRN: LUTVDG/TVBB–5268–SE

Keywords
VaR, Volatility, Covariance, GARCH, EWMA, Univariate, Multivariate, Market risk, Financial time series

Abstract
In this thesis various Value-at-Risk models are compared and evaluated towards finding the optimal model for the bank's trading book. The focus is on linear instruments (stocks, indicies and currency) and real market data, both domestic and foreign, is used for the calculations. I find that the GARCH(1,1) based model outperforms other models in volatility estimation and should thus be a wise choice when volatility estimation is needed, but GARCH based methods become rather complex for a multivariate covariance estimation. Therefore a mixture of simpler models such as EWMA is needed for making rational estimates.

Brandteknik och Riskhantering
Lunds tekniska högskola
Lunds universitet
Box 118
221 00 Lund

brand@brand.lth.se
http://www.brand.lth.se
Telefon: 046 - 222 73 60
Telefax: 046 - 222 46 12

Department of Fire Safety
Engineering and Systems Safety
Lund University
P.O. Box 118
SE-221 00 Lund
Sweden

brand@brand.lth.se
http://www.brand.lth.se/english
Telephone: +46 46 222 73 60
Fax: +46 46 222 46 12

Thanks:

Dr. Freyr Harðarson, Dr. Henrik Johansson and the employees at Risk Management department in Landsbanki, Reykjavík

Summary

In the last decades the growth in financial markets has been enormous. Global trading is easier then ever and business with all kinds of financial instrument such as stocks and bonds and their derivatives is common. Usually the aim with all trades is to save or gain money one way or another, but with all trades there exist as well the risk that the trade will not be profitable and result in a financial loss. This risk that leads to financial loss can be regarded as financial risk. One of the main category of financial risk is market risk which is the risk that market fluctuations will lead to financial loss. For financial institutions who have large amount of their assets in financial instrument (traded on markets), market risk can have great impact on their performance and therefore essential to quantify.

In this thesis the goal is to quantify market risk and for that cause use the term Value-at-Risk (VaR), which is commonly used among financial institutions. The term VaR is defined as the amount X that you are $\alpha\%$ certain of not losing more than the following N days. More general, VaR gives a kind of worst case scenario at preferred level of probability (α) and time period (N-days). There are various ways and techniques for calculating VaR, all with their pros and cons, and generally depended on presumptions. In this thesis the focus is on linear instruments, such as equity and currency, and for that sake the main categorize of calculating VaR are parametric and non-parametric approaches. VaR is also used in regulatory terms. The Basel Committee, which is a international banking supervisor, uses VaR to stipulate the minimum amount of regulatory capital that financial institution must have available at all times. This is done to prevent financial crisis and possible bankruptcy due to unforeseen market movements, credit defaults or any other risk faced by the financial institution. Both parametric and non-parametric methods are used to calculate VaR for real market data, with the goal of finding the method that suits the bank's trading book the best while fulfilling regulations set by the Basel Committee.

Criteria's are set up for judging the best model where the main conclusions are that a mixture of both parametric and non-parametric methods should be used. GARCH volatility estimate is found to be the best to describe market volatility conditions, but they become very complex as the dimensionality increases. Therefore I recommend that a mixture of both parametric and non-parametric methods with GARCH and/or EWMA volatility estimates should be used. I also find that a student's t-distribution fits the data better then the commonly used normal distribution. But student's t-distribution with GARCH volatility estimates are sensitive and therefore my recommendation is that normal distribution should be used as well as some alternative distribution's.

Contents

Chapter 1

Introduction

As pointed out by Kaplan and Garrick (1981) quantitative risk can be defined as combination of scenario, probability and a consequence. It is necessary to evaluate all these parts to get a relative and rational estimate of risk. More general in risk assessment and risk management we want to answer the question; What can go wrong, how likely is it and what will be the consequences. The main fields of risk assessment can be separated into 5 categories; *Safety risk*, *Health risk*, *Ecological and Environmental risk*, *Public welfare risk* and *Financial risk* (Kolluru, 1995).

Financial risk can be thought of as any risk concerned with financial loss due to some random changes in underlying risk factors (stock, currency, derivatives, interest rate etc.) and can furthermore be categorized into three main level of concerns; *Market risk* (due to movements in market factors), *Credit risk* (the risk that a person or an organization will not fulfill his/her obligations) and *Operational risk* (risk of loss because of systematic failures) (Dowd, 1998). Methods for measuring and evaluating financial risk are many and depend on what is of interest to examine. Just for market risk there are numerous of ways. For options and derivatives examining 'the greeks' might by a good choice, stress testing might by good for worst case scenario analysis while Value-at-Risk (VaR) could give a universal risk measure for the exposed market risk. In this thesis I will concentrate on VaR for measuring market risk[1].

In the last 50 years there has been enormous growth in trading worldwide, for example the New York Stock Exchange has grown from \$4 million in 1961 to \$1.6 trillion[2] in 2005. This growth and the massive increase of new instruments (all kinds of derivatives, swaps, CDO's (collateralized debt obligation) and so on) has invoked more need for good risk management in

[1] This was proposed by Landsbanki bank

[2] One trillion equals million millions (10^{12})

the financial sector. Managing risk improves the value of the company and can avoid major financial disaster such as Orange County (1994), Barings Bank (1995), Enron (2001), WorldCom (2002) and Socièté Générale (2008) which have all been related to poor risk management.

1.1 Aim and purpose

The aim of the thesis is to evaluate different methods for calculating VaR for bank's trading book, where the goal is to minimize VaR while fulfilling regulatory requirements. The focus will be on linear instruments (such as equity and currency) and actual market data will be used for the evaluation, both domestic and foreign (Icelandic and Swedish).

VaR estimation can be categorized into three main categories; *parametric methods*, *non-parametric methods* and *Monte Carlo simulation methods*. Both parametric and non-parametric approaches will be used to calculate VaR for the market data and compared with critical judgment towards obtaining the optimal VaR method for the bank's trading book. Monte Carlo methods are mostly used on non-linear financial instruments and will therefore not be the topic of this thesis.

1.2 What is VaR?

> You are responsible for managing your company's foreign exchange positions. Your boss, or your boss's boss, has been reading about derivatives losses suffered by other companies, and wants to know just how much market risk the company is taking. What do you say?
>
> You could start by listing and describing the company's positions, but this isn't likely to be helpful unless there are only a handful. Even then, it helps only if your superiors understand all of the positions and instruments, and the risks inherent in each. Or you could talk about portfolio's sensitivities, i.e., how much the value of the portfolio changes when various underlying market rate or prices change, or perhaps options delta and gammas. Even if you are confident of your ability to explain these in English, you still have no natural way to net the risk of your short position in Deutsche marks against the long positions in Dutch guilders. (It makes sense to do this because gains or losses on the short position in marks will be offset almost perfectly by gains or losses on the long position in guilders.) You could simply assure your superiors that you never speculate but rather use derivatives only

> to hedge, but they understand that this statement is vacuous. They know that the word 'hedge' is so ill-defined and flexible that virtually any transaction can be characterized as a hedge. So what do you say? (Linsmeier and Pearson, 1996)

Value at Risk (VaR) is an attempt to give a relatively simple measure of financial risk (not only market risk) with a single number answering the question 'how bad can things get?' (Dowd, 2005). VaR could thus be a fair attempt to answer the question in the example above. The statement we wan't to make with VaR is:

> We are α percent certain to lose not more than X much money in the following N days.

The amount X is a loss due to market movements and could be for a single asset or a portfolio (see section 2.2). The amount X is function of two variables, the confidence level α and time period, usually given in N days. The calculation of VaR is thus based on the probability of changes in asset or the portfolio value over the next N days. To be more mathematical VaR is the quantile corresponding to the (1 - α) of the return distribution, so if we set $p = 1 - \alpha$ and call q_p the p-th quantile of α, then VaR can be defined as; $\text{VaR}_{\alpha\%} = -q_p$ for specified confidence level and period. Meaning that we can be α certain to not lose more than VaR in that period (see figure 1.1)

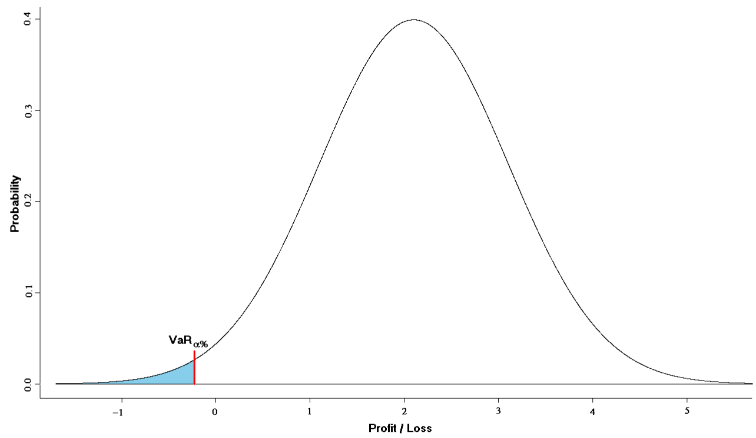


Figure 1.1: VaR at α% confidence level for some imaginary profit/loss series, assuming normality

Let's imagine that the x-scale represents the profit/loss distribution of some imaginary company (and the scale is perhaps in million euros). Say that a company is interested in knowing their market risk status and especially what the $1-\alpha$ worst case would be. Then $\text{VaR}_{\alpha\%}$ would give a estimate of the loss that would not be exceeded α percent of the time. Likewise, losses larger then $\text{VaR}_{\alpha\%}$ happen only (1 - α)% of the time. This $\text{VaR}_{\alpha\%}$ is equal to the boundaries between the blue and white area on figure 1.1.

The reason for VaR's popularity, as a definition of risk, is its simplicity in interpretation. It is relatively easy to understand, has the unit of the measure (i.e. euro, SEK, ISK)[3], it is probabilistic (concerned with probability), it can be used for any type of positions (bond, stock, currency, derivatives etc.) and portfolios (meaning that it will aggregate many sub-positions into one measure) and it is holistic (meaning that takes into account all underlying risk factors) (Dowd, 2005). VaR is also used in regulatory terms. Financial institutions are required to have some minimum regulatory capital available at all times for safety reasons and with the Basel accord[4], financial institution were allowed to base this minimum capital partly on their own VaR estimates. This is known as the internal approach, see section 5.1.1.

1.3 Data

In the thesis I will examine four stocks, two indices, two portfolios and one currency pair. Portfolio 1 will be made of equal shares (25%) in all four stocks, 25% LAIS, 25% MARL, 25% NDA and 25% ERIC and Portfolio 2 will consist stocks in the following ratios; 50% LAIS, 30% MARL, 10% NDA and 10% ERIC, see 1.1.

Name	Ticker	Description
Landsbankinn	LAIS	Icelandic bank
Marel	MARL	Icelandic food processing company
OMX Iceland 15	ISXI15	Index consisting of 15 Icelandic companies
Nordea	NDA	Swedish bank
Ericsson	ERIC	Swedish tele & datacommunication company
OMX	OMX	Index consisting of 30 Swedish companies
Portfolio 1	-	Equal share portfolio
Portfolio 2	-	50% LAIS, 30% MARL, 10% NDA and 10% ERIC
USD/ISK	USDISK	USA dollars to Icelandic króna

Table 1.1: The time series used for calculation

[3]In the thesis I will though use percentage for the sake of comparison, see section 5.1

[4]Banking supervision accords (recommendations on banking laws and regulations), Basel I and Basel II issued by the Basel Committee on Banking Supervision (BCBS), see Basel Committee (2006).

Various time periods will be used for modeling to try to capture the most efficient model. Time periods are given in table 1.2.

Name	From date	To date
Short	1. July 2006	30. June 2008
Long	28. June 2004	30. June 2008
Extra Long	1. July 1998	30. June 2008

Table 1.2: The time series used for calculation

where the following time periods are assigned to the time series, table 1.3.

Name	Ticker	Description
Landsbankinn	LAIS	Short and Long
Marel	MARL	Short and Long
OMX Iceland 15	ICEXI15	Short and Long
Nordea	NDA	Short and Long
Ericsson	ERIC	Short and Long
OMX	OMX	Short and Long
Portfolio 1	-	Long
Portfolio 2	-	Long
USD/ISK	USDISK	Extra Long

Table 1.3: The time periods used for each time series

All data is provided through the Bloomberg Terminal[5]. It should be noted that all data used are 'hypothetical' outcomes, meaning that it is assumed that no dividends are paid out, no trades are done and weights in portfolios are assumed to be fixed at all times. The 'window size' (sample size) used to estimate each days VaR estimate is 500 day's (meaning that each days VaR estimate is calculated from the previous 500 days) which are the requirement set by the Basel Committee (2006).

1.4 Overview

In chapter 2 I will give an overview of the main topics of financial time series modeling, introduce how data is treated, present various methods for estimating volatility and covariances. Chapter 3 describes parametric approaches for calculating VaR and chapter 4 decribes non-parametric approaches for calculating VaR. Analysis, criteria's and results are presented in chapter 5 and finally chapter 6 gives conclusions, discussions and ideas for further analysis.

[5]A computer system by Bloomberg that provides information on financial market data.

Chapter 2

Financial time series modeling

Financial time series modeling is the task of building a model used to predict, evaluate and/or forecast the performance of financial instruments. The models often use a mixture of theories from economics, engineering, statistics and business administration to obtain information. The main characteristics with financial data is uncertainty and sudden movements and therefore the central aim of the modeling is to be able to explain and predict these characteristics. Two known facts with financial time series is that low and high fluctuations tend to come in periods (resulting in periods of high and low returns) and the fact that their probability distributions usually have fatter tails than normal time series, meaning that extreme cases (high and low returns) are more likely than normal distribution describes.

2.1 Return analysis

For modeling financial data it is common to work with profit/loss data (P/L) and return series r_t. Profit/Loss data tells how much you have gained (profit) or lost (loss) over a time period between time $t-1$ and t and can be written as:

$$P/L_t = P_t - P_{t-1}$$

where P_t represents the value of the asset at time t and P_{t-1} the value at time $t-1$. Return series tells you how much in percentage you gained or lost between time $t-1$ and t. There are two ways of representing the return series, arithmetic (2.1) or geometric (2.2)

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{2.1}$$

$$r_t = ln(\frac{P_t}{P_{t-1}}) \tag{2.2}$$

The former is more common when working with non-parametric methods and the latter with parametric methods. It is also common to work with the deviation from the mean of the series at each time t, called error terms or residuals ϵ_t and represented as:

$$\epsilon_t = r_t - \bar{r}$$

where $\bar{r}$ is the average return of the series (often when dealing with daily returns the mean will be low and therefore the approximation $\bar{r} \to 0$ is often made, which leads to $\epsilon_t \approx r_t$).

2.2 Portfolio

In finance a portfolio is a mix or a collection of assets, $k \geq 2$. The idea with a portfolio is often to build up a more stable ownership and spreading the risk ('not putting all the eggs in the same basket'). This is called to diversify. Each asset in the portfolio is given weight depending on it's market-value MV, i.e.

$$w_i = \frac{MV_i}{\sum_{i=1}^{k} MV_i} \tag{2.3}$$

The total return of the portfolio r_p can then be calculated as:

$$r_p = \mathbf{w}^\mathsf{T}\mathbf{r} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix}^\mathsf{T} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix} \tag{2.4}$$

Variance of the portfolio can be obtained as:

$$\sigma_p^2 = \mathbf{w}^\mathsf{T}\Sigma\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix}^\mathsf{T} \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 & \cdots & \sigma_1\sigma_k \\ \sigma_2\sigma_1 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & \\ \sigma_k\sigma_1 & \dots & & \sigma_k^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix} \tag{2.5}$$

where Σ is variance-covariance matrix.

2.3 Volatility

The volatility σ of a variable is defined as the standard deviation of the returns r_t per unit of time t, when the returns are expressed using continuous compounding[1]. Usually the unit of time is one day so that the volatility is expressed as a standard deviations of the continuous compounded return per day (Hull, 2007).

By examining a long return series such as in figure 2.1 it is clear that variance, and therefore volatility, varies with time. Volatility represents risk and since volatility is changing that implies that market risk is changing. (**?**)

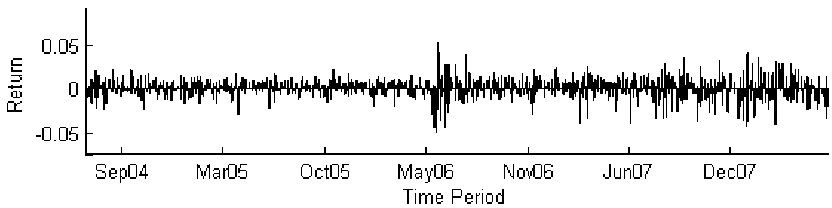


Figure 2.1: Return series for OMX

This is one of the main characteristics with financial volatility, called heteroscedasticity, which is the behavior of having time varying periods of low and high volatilities. We therefore want to take this characteristics into account when estimating volatility.

2.4 Estimating volatility

2.4.1 Historical volatility

The most obvious choice of estimating volatility is the historical (equal weighted) volatility defined as (here I talk about volatility although equation gives the variance σ^2, the volatility is obtained of course by taking the square-root of the variance, $\sigma = \sqrt{\sigma^2}$, this also applies for other equation in this chapter):

$$\sigma_t^2 = \frac{1}{N-1} \sum_{i=1}^{N} (r_{t-i+1} - \bar{r})^2 \qquad (2.6)$$

[1]Continuous compounding means that the growth and loss of the variable (asset) is expressed continuously, not just the profit or loss from the time when bought till the time when sold.

where N is the number of days used in the estimate. As can be seen in figure 2.2 the size of N has an effect on how the volatility estimate will look like. For a small N the estimate is usually more responsive and jumps a lot, but becomes more stable as N gets larger.

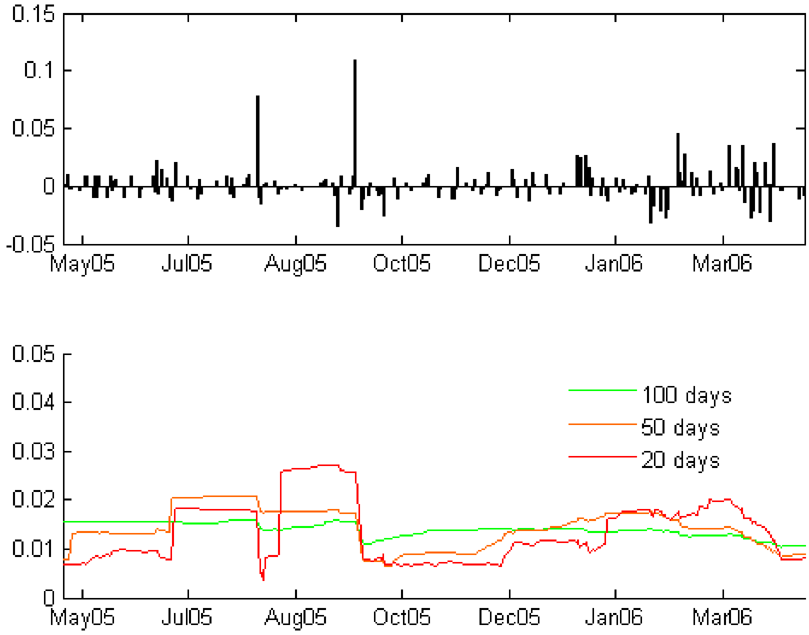


Figure 2.2: Volatility Estimates for MARL, shows how different time periods effect historical volatility estimates

By making the approximation that average of daily return is close to zero and it makes insignificant difference to the estimate, and that it makes insignificant difference to divide with N instead of $N-1$ when dealing with long time series, equation 2.6 can be modified as:

$$\sigma_t^2 = \frac{1}{N}\sum_{i=1}^{N} r_{t-i+1}^2 \tag{2.7}$$

The problem with this method is that it treats all observations equally, meaning that all observations in the estimate will have the same weight. That is old and new (in time) observations are treated equally and have as much contribution to the estimate. This means that if a shock on the market had occurred in the past it will have as much impact on the estimate as any other day until it falls out of the sample space. This could cause a overestimate while the shock is still in the sample space and a sudden jump when it falls out of the sample. This overestimate is known as 'ghost effects'. This problem and the fact that volatility tends to vary with time in financial time series has led to development of weighted volatility estimates in the form:

$$\sigma_t^2 = \sum_{i=1}^{N} \alpha_i r_{t-i+1}^2 \tag{2.8}$$

where α_i are the weights which are assigned to each return r_t. The weights, $0 < \alpha_i < 1$, decline as i gets larger and sum up to 1.

2.4.2 EWMA model

One of the most known weighting models is the EWMA model (exponentially weighting moving average) where the weights decrease exponentially as we move back in time;

$$\sigma_t^2 \approx (1-\lambda) \sum_{i=1}^{N} \lambda^{i-1} r_{t-i+1}^2 \tag{2.9}$$

Here λ is defined as $\lambda = \alpha_{i+1}/\alpha_i$ and thus a constant between 0 and 1. The choice of λ is depending of the behavior desired. Low λ values gives more responsive and rougher behavior in the volatility estimate and high λ values gives less responsive and smoother volatility estimates, see figure 2.3. RiskMetrics[2] suggests that $\lambda = 0,94$ should be used for equity and $\lambda = 0.97$ for foreign exchange (FX), such as currency trade. The parameter could also be optimized using traditional maximum log-likelihood estimate. With little modification equation (2.9) can be written as:

$$\sigma_t^2 \approx \lambda \sigma_{t-1}^2 + (1-\lambda) r_t^2 \tag{2.10}$$

which is a simple updating formula, where the only variable needed for estimating the volatility at time t is the most recent return r_t (the return after the market closes) and volatility σ_{t-1} (the estimated volatility from the day before).

The advantages with the EWMA model is that it only relies on one parameter, λ, tends to produce much less ghost effects than the historical equal weighted model and a very little data needs to be stored. The main disadvantages with the EWMA model is that it takes λ to be constant and can therefore be unresponsive to market conditions (Dowd, 2005).

[2] J.P. Morgan Guaranty Trust Company, see Morgan (1996)

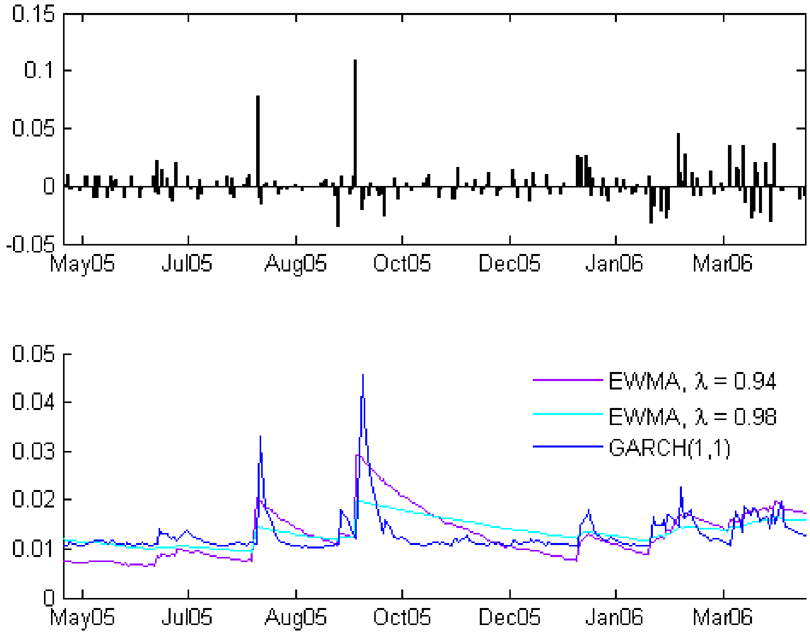


Figure 2.3: Volatility Estimates for MARL, shows how different volatility estimates react to a shock on market

2.4.3 GARCH models

GARCH (generalized autoregressive conditional heteroscedasticity) models proposed by Bollerslev (1986), which was an extension of Engle's (1982) ARCH models, give a solution to this kind of problem. GARCH models can show volatility clustering and leptokurtosis (fatter tails than normal tails) which are two of the most important facts with financial time series. The GARCH(p, q) model depends on q past volatilities and p last error terms and has the following representation:

$$\begin{cases} \epsilon_t = \sigma_t z_t \\ \sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \end{cases} \tag{2.11}$$

where the residuals ϵ_t are defined as before as $\epsilon_t = r_t - \bar{r}$. The parameters must be non-negative and fulfill;

$$\sum_{i}^{q} \alpha_i + \sum_{i}^{p} \beta_i < 1$$

The GARCH(p, q) model differs with the choice off distribution governing the residuals. The most common choice is normal distribution, but could also be for example t-distributed (which produces even fatter tails). A maximum likelihood is ideal for obtaining parameter values (α, β and ω).

To obtain the optimal number of parameters k and p tests such as deviance statistics test can be used:

$$D = 2\{\ell_2(M_2) - \ell_1(M_1)\} > c_\alpha \tag{2.12}$$

where $\ell_i(M_i)$ is the maximum log-likelihood parameter for model i. The maximum log-likelihood of the model with fewer parameters M_1 should be subtracted from the higher number of parameters M_2 and compared with c_α which is the $(1-\alpha)$ quantile of the χ^2 distribution. Model M_1 is rejected if $D > c_\alpha$ in favor of M_2.

GARCH(1,1)

The GARCH$(1,1)$ model is the most popular GARCH model. The reason is because of its simplicity (depends only on the 'last' volatility and return), seems to fit most financial data fairly well (higher values of k and p usually don't give significantly better result) and from the principle of parsimony (to choose as simple model as possible to fit the data). The model is given as:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{2.13}$$

where the parameters must as before be non-negative and fulfill $\alpha + \beta < 1$. If the governing distribution is assumed to be normal (most common) then the maximum log-likelihood function is obtained as:

$$\begin{aligned} \ell &= -\frac{n}{2}ln(2\pi) - \sum_{i=1}^{n} ln(\sigma_i^2) - \sum_{i=1}^{n} \frac{\epsilon_i^2}{2\sigma_i^2} \qquad (2.14) \\ &\simeq -\sum_{i=1}^{n} ln(\sigma_i^2) - \sum_{i=1}^{n} \frac{\epsilon_i^2}{2\sigma_i^2} \end{aligned}$$

since the constant doesn't matter in the maximization. The GARCH(1,1) model depends on the same variables as the EWMA model (volatility and return) but now there are three parameters instead of one. Figure 2.3 shows a comparison between GARCH and EWMA model. From the figure it is clear that the GARCH model adopts the market condition better than the EWMA model. The EWMA model can be regarded as a special case of GARCH(1,1) with $\omega = 0$, $\alpha = 1 - \lambda$, and $\beta = \lambda$. The main advantages with GARCH models is how they can accommodate heteroscedasticity and fat tails and how responsive they are to market condition (see figure 2.3).

The main disadvantages is that the GARCH models are more complex then the other volatility estimates.

Many modification have been developed to the GARCH(1, 1) model such as A-GARCH, E-GARCH, GJR-GARCH, I-GARCH, V-GARCH and many more[3]. Some of those models can explain 'leverage effect' which is the behavior of not responding the same way to negative and positive returns (good and bad news on market), i.e. are asymmetric not symmetric as the GARCH model. Here I will introduce two of them.

GJR-GARCH: Glosten, Jagannathan & Runkle
The GJR-GARCH model (also known as Threshold-GARCH or T-GARCH) was proposed by Glosten, Jagannathan and Runkle (1993) is similar to the GARCH(1,1) model but also exhibits the term S_{t-1} to capture the leverage effect.

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + (\alpha + \gamma S_{t-1})\epsilon_{t-1}^2 \tag{2.15}$$

where $S_{t-1} = 0$ if $\epsilon_{t-1} \geq 0$ and $S_{t-1} = 1$ if $\epsilon_{t-1} < 0$, so it doesn't react the same to positive and negative returns. The GJR-GARCH model has the same parameter restriction as GARCH(1,1), that is α and β must be non-negative (γ can be negative) and fulfill:

$$\alpha + \beta < 1$$

E-GARCH: Exponential GARCH
The E-GARCH model by Nelson (1991) is a little bit different from the GARCH(1,1) and GJR-GARCH, but also captures the leverage effect by letting the volatility estimate depend on the sign of the lagged residual.

$$\ln(\sigma_t^2) = \omega + \beta\ln(\sigma_{t-1}^2) + \alpha\left[\frac{|\epsilon_{t-1}|}{\sigma_{t-1}} - E\{z_{t-1}\}\right] + \gamma\frac{\epsilon_{t-1}}{\sigma_{t-1}} \tag{2.16}$$

One of the main advantages with the E-GARCH model is that it has no parameter restrictions as the GARCH model. Figure 2.4 shows a comparison between the GARCH models introduced, GARCH(1,1), E-GARCH and GJR-GARCH. By looking at figure 2.4 the leverage effect can be examined, GARCH(1,1) rises more than the other two when large positive shocks occur and rises less when a short period of negative returns occurs.

[3] see for example Hansen and Lunde (2005) who make a comparison of volatility forecasting with many GARCH based methods

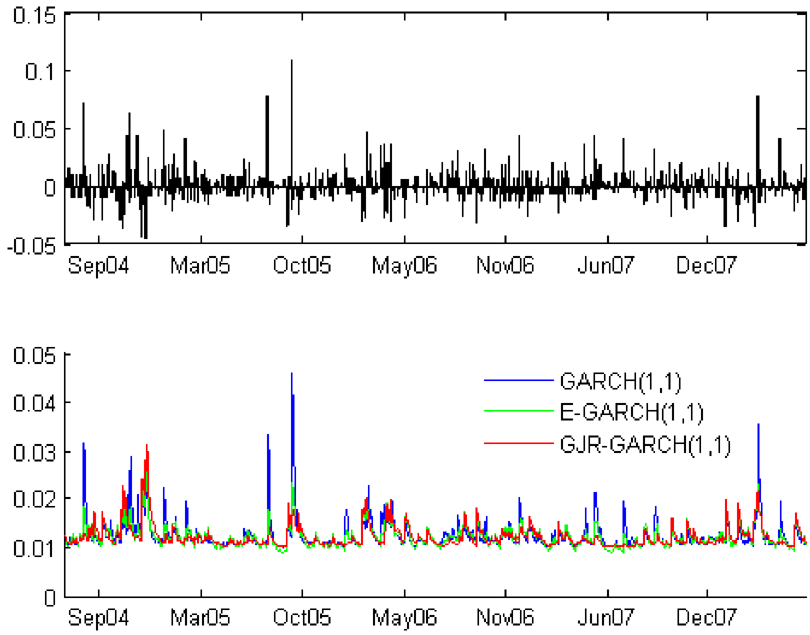


Figure 2.4: Volatility Estimates for MARL, shows how differently the GARCH models behave.

2.5 Covariance and correlation

In statistics the covariance gives an estimate of how much two variables change together and is only used when two or more variables are of concern (bivariate or multivariate). For two variables x and y the covariance can be defined as:

$$cov(x, y) = E[xy] - E[x]E[y] \tag{2.17}$$

A strictly related term is the correlation between two variables which gives a measure of how the two variables move together. The correlation is always between -1 and $+1$ and can be defined as

$$corr(x, y) = \frac{cov(x, y)}{\sigma_x \sigma_y} \tag{2.18}$$

where -1 means that instruments x and y move totally against each other, 0 means that nothing can be said about their movement together and 1 means that they move totally the same.

2.6 Estimating covariance and correlation

2.6.1 Historical covariance and correlation

Estimation of covariance and correlation is parallel to estimation of volatility. The historical correlation estimate can be written as:

$$corr(x,y)_t = \frac{\sum_{i=1}^{n} x_{t-i}y_{t-i}}{\sqrt{\sum_{i=1}^{n} x_{t-i}^2 \sum_{i=1}^{n} y_{t-i}^2}} \tag{2.19}$$

and then the covariance could be obtained with:

$$cov(x,y) = \sigma_x \sigma_y corr(x,y) \tag{2.20}$$

2.6.2 EWMA covariance

EWMA covariance can be estimated as:

$$cov(x,y)_t = \lambda cov(x,y)_{t-1} + (1-\lambda)x_{t-1}y_{t-1} \tag{2.21}$$

and the correlation can be obtained with equation 2.18. As before RiskMetrics suggests that $\lambda = 0,94$ should be used for equity and $\lambda = 0.97$ for FX. To ensure that the matrix is positive definite or semi-positive definite[4] it is important to use the same λ value for all parameters in the matrix.

2.6.3 GARCH covariance

Since the GARCH based volatility model seem to be powerful tool for the estimation of volatility, it seem to be obvious idea to generate a multivariate version of GARCH estimation. The problem is that multivariate GARCH models are computationally complex and the number of parameters to be estimated grow rapidly as the number of assets in the portfolio grow. Multivariate GARCH models, often called VECH can be written as:

$$\mathbf{H}_t = \mathbf{W} + \mathbf{A}(\epsilon_{t-1}\epsilon_{t-1}^{\mathsf{T}}) + \mathbf{B}(\mathbf{H}_{t-1}) \tag{2.22}$$

where $\mathbf{H}_t$ is the vector of volatilities, $\mathbf{W}$ is a vector of ω coefficients and $\mathbf{A}$ and $\mathbf{B}$ are matrix for α and β coefficients, respectively. For a portfolio consisting of only two assets the multivariate VECH becomes:

[4]Positive definite or semi-positive definite means in this case that the portfolio volatility σ_p has to be larger or equal to zero, i.e. $\sigma_p^2 = w^{\mathsf{T}}\Sigma w \geq 0$ for all$w$.

$$\begin{bmatrix} \sigma_{1,t}^2 \\ \sigma_{1,t}\sigma_{2,t} \\ \sigma_{2,t}^2 \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} \epsilon_{1,t-1}^2 \\ \epsilon_{1,t-1}\epsilon_{2,t-1} \\ \epsilon_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{bmatrix} \begin{bmatrix} \sigma_{1,t-1}^2 \\ \sigma_{1,t-1}\sigma_{2,t-1} \\ \sigma_{2,t-1}^2 \end{bmatrix} \tag{2.23}$$

Here the total number of parameters to be estimated are 21 and with a 3 asset portfolio the number of parameters needed to be estimated becomes 78. Furthermore this formulation doesn't ensure $\mathbf{H}_t$ to be positive definite (?).

Because of this complexity there has been developed several simplified GARCH based covariance models such as diagonal VECH (DVECH) proposed by Bollerslev, Engle and Woolridge (1988) where the $\mathbf{A}$ and $\mathbf{B}$ are assumed to be diagonal and the total number of parameters becomes $3(k(k+1)/2)$ where k is the number of asset in the portfolio, the Constant Conditional Correlation model (CCC) proposed by Bollerslev (1990) where total number of parameters becomes $k(k+5)/2$ and Dynamic Conditional Correlation model (DCC) proposed by Engle (2002) where total number of parameters becomes $(k+1)(k+4)/2$. Table 2.1 gives an comparison of parameters needed to be estimated in multivariate GARCH models.

Number of assets, k	VEC	DVEC	CCC	DCC
2	21	9	7	9
3	78	18	12	14
4	210	30	18	20

Table 2.1: Number of parameters to be estimated in multivariate GARCH models

Recent researches such as Sheppard (2003) and Sigurdarson (2007) have shown the advantages with the CCC and DCC model, both because of their 'simplicity' (compared to the other) and behavior. I will therefore focus on those two and here I will give a short introduction to them.

Constant conditional correlation, CCC

The model assumes k assets which all are conditionally normal distributed. The covariance matrix, Σ_t, is defined as:

$$\Sigma_t = \mathbf{D}_t\mathbf{R}\mathbf{D}_t \tag{2.24}$$

where $\mathbf{D}_t$ is the $k \times k$ diagonal volatility matrix, estimated from univariate GARCH(1,1) process (one at a time as in section 2.4.3)

$$\mathbf{D}_t = \begin{bmatrix} \sigma^2_{1,t} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma^2_{k,t} \end{bmatrix} \tag{2.25}$$

and $\mathbf{R}$ is the correlation matrix. The maximum log-likelihood function for the multivariate case when assuming normality can be written as:

$$\begin{aligned} \ell &= -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + \log(|\Sigma_t|) + \mathbf{r}_t^{'}\Sigma_t^{-1}\mathbf{r}_t) & (2.26) \\ &= -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + 2\log(|\mathbf{D}_t|) + \log(|\mathbf{R}|) + \epsilon_t^{'}\Sigma_t^{-1}\epsilon_t) \\ &\simeq -\sum_{t=1}^{T}(2\log(|\mathbf{D}_t|) + z_t^{'}\mathbf{R}^{-1}z_t) & (2.27) \end{aligned}$$

since the constants don't matter in the maximization. Where $z_t \sim N(0, R)$ when $\epsilon_t \sim N(0, \Sigma_t)$ are univariate GARCH standardized residuals. The univariate volatility process can be any kind of GARCH process (E-GARCH, GJR-GARCH or some other) and doesn't have to be the same for all assets in the portfolio. The correlation is estimated as a constant historical correlation.

$$\mathbf{R} = (\rho_{ij}) \tag{2.28}$$

The advantages with the CCC method is that it is much simpler then the full multivariate GARCH model (VECH), a univariate GARCH process can be used for the estimation and the formulation ensures positive definiteness of $\mathbf{H}_t$. The disadvantages is that it assumes the correlation to be constant, which is unrealistic (**?**).

Dynamic conditional covariance, DCC

Here the idea is the same in all steps as in CCC except to let the correlation be time varying. A structure for estimating the dynamic correlation was introduced by Engle (2002) as:

$$\mathbf{Q}_t = (1 - \alpha - \beta)\bar{\mathbf{Q}} + \alpha(\epsilon_{t-1}\epsilon_{t-1}^{'} + \beta\mathbf{Q}_{t-1}) \tag{2.29}$$

where $\bar{\mathbf{Q}}$ is the $k \times k$ unconditional covariance matrix of ϵ_t and α and β are parameters > 0, that have to satisfy; $\alpha + \beta < 0$. The correlation matrix $\mathbf{R}_t$ can now be obtained as:

$$\mathbf{R}_t = \mathbf{Q}_t^{\star-1}\mathbf{Q}_t\mathbf{Q}_t^{\star-1} \tag{2.30}$$

where $\mathbf{Q}_t^{\star-1}$ is obtained as:

$$\mathbf{Q}_t^{\star-1} = \begin{bmatrix} \frac{1}{\sqrt{q_{11}}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{\sqrt{q_{kk}}} \end{bmatrix} \tag{2.31}$$

where q_{ii} are the i-th diagonal element of the matrix Q_t, where $i \in [1, k]$.

Some of the main advantages of the DCC model is that it doesn't rely on a constant correlation matrix, can be calculated in steps, only few extra parameters are needed and univariate estimate can be used for obtaining a large part of the parameters.

Chapter 3

Parametric methods

3.1 Univariate parametric methods

In the parametric approach a distribution is fitted to the data and the VaR is estimated from the fitted distribution. The parametric approach is more appealing mathematically than the non-parametric, since it has a distribution (and density) function, which can give a relatively straight forward way of calculating VaR. For example if the normal distribution fits the data well, the VaR at α confidence level can be calculated as

$$\mathrm{VaR}_{\alpha\%} = \mu + \sigma \cdot z_\alpha \tag{3.1}$$

where z_α comes from the standard normal distribution table ($\simeq 2.326$ for 99% confidence level), see figure 3.1.

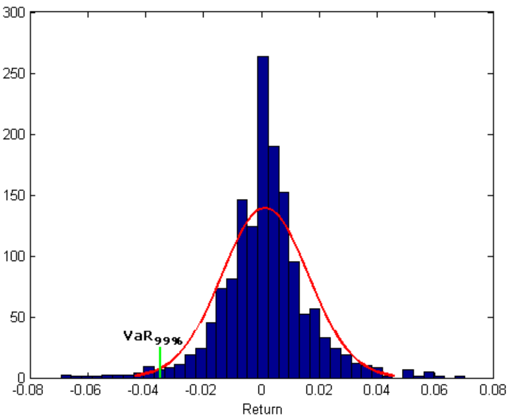


Figure 3.1: VaR assuming normal distribution, LAIS long period data

If the governing distribution is assumed be student's t-distribution, then the VaR can be calculated as:

$$\text{VaR}_{\alpha\%} = \mu + \sqrt{\frac{\upsilon - 2}{\upsilon}}\sigma \cdot t_\alpha \tag{3.2}$$

with υ degrees of freedom and t_α comes from standard t-distribution table. If for example the interest would be to find $\text{VaR}_{99\%}$ assuming that t-distribution with 4 degrees of freedom, $\upsilon = 4$, fits the data well while $\mu = 0$ and $\sigma = 0.02$, then it would be:

$$\text{VaR}_{99\%} = \sqrt{\frac{4-2}{4}} \cdot 0.02 \cdot 3.747 = 0.053$$

while normal distribution would have given $\text{VaR}_{99\%} = 0.047$. Here it is important to understand that volatility tends to be time varying, see section 2.3, therefore the distribution of the residuals is expressed as for example:

$$r_t \mid \Theta_t \sim N(0, \sigma_t^2)$$

where Θ_t is information set know at time t, for example past returns $\{r_0, \ldots, r_{t-1}\}$ and/or past volatilities $\{\sigma_0, \ldots, \sigma_{t-1}\}$. Therefore it is said that the returns r_t are conditionally normal distributed, meaning that the returns at time t are normal distributed conditional on the information set.

The choice of distribution can differ a lot. The most common is to assume that normality (that is a normal distribution) is sufficient to fit the data well, although this has been debated[1]. Other common choice of distributions are for example log-normal distribution and extreme value distribution. As stated before, financial data tend to be clustered, have fat tails and are possibly skewed and thus we would like to fit a distribution to the data that can show these characteristics.

3.2 Multivariate parametric methods

Multivariate parametric methods are analogous to the univariate case where assumptions are made about the portfolio rather then a single asset, although making the assumption that each asset in the portfolio is normal is the same as assuming that the portfolio is normal distributed (holds only for normal distribution). With k assets in a portfolio, assuming normal distribution, the VaR can be obtained by:

[1] see for example Hull (2007)

$$\text{VaR}_{\alpha\%} = \mathbf{w}^\mathsf{T}\mathbf{r} + \sqrt{\mathbf{w}^\mathsf{T}\Sigma\mathbf{w}}z_\alpha \tag{3.3}$$

see section 2.2 for further details. As before z_α is obtained from the standard normal distribution. Likewise the for student's t-distributed data portfolio $\text{VaR}_{\alpha\%}$ can be obtained as:

$$\text{VaR}_{\alpha\%} = \mathbf{w}^\mathsf{T}\mathbf{r} + \sqrt{\frac{\upsilon - 2}{\upsilon}}\sqrt{\mathbf{w}^\mathsf{T}\Sigma\mathbf{w}}z_\alpha \tag{3.4}$$

Correlation can have much effect on VaR estimate which can be shown with a simple example. Suppose we have two equal weighted ($w_1 = w_2 = 0.5$) assets, A_1 and A_2, in a portfolio which both are; $A_1, A_2 \sim \text{N}(0,1)$. The VaR of each asset is obtained by equation 3.1. The portfolio volatility could be calculated as:

$$\sigma_P = \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2\rho w_1 w_2 \sigma_1 \sigma_2} = \sqrt{\frac{1+\rho}{2}}$$

where ρ is the correlation between the two assets. Since $\mu_1 = \mu_2 = 0$ and $\sigma_1 = \sigma_2 = 1$ the VaR estimate can then be written as:

$$\text{VaR}_{\alpha\%} = r_P + \sigma_P z_\alpha = \sqrt{\frac{1+\rho}{2}}z_\alpha \tag{3.5}$$

The VaR estimate will be less then the individual VaR estimate for all values of ρ except for $\rho = 1$ (when short-selling[2] is not allowed). So generally it could be stated that:

$$\text{VaR}_{portfolio} < \sum_{i=1}^{N} \text{VaR}_i$$

where VaR_i is the Value-at-Risk for asset i in the portfolio. This is one of the fundamentals with portfolios, called to diversify (see section 2.2). Let's take a simple example. Say that we have two equal weighted assets, A_1 and A_2, with are found to have following characteristics; $A_1, A_2 \sim \text{N}(0, 0.02^2)$ and the correlation is found to be 0.6. We are interested in finding $\text{VaR}_{99\%}$. Start by finding the portfolio volatility as:

$$\sigma_P = \sqrt{0.5^2 \cdot 0.02^2 + 0.5^2 \cdot 0.02^2 + 2 \cdot 0.6 \cdot 0.5 \cdot 0.5 \cdot 0.02 \cdot 0.02} = 0.0179$$

and therefore the $\text{VaR}_{99\%}$ can be obtained as:

$$\text{VaR}_{\alpha\%} = 0 + 0.0179 \cdot 2.326 = 0.0416$$

[2]In finance short-selling is the act of getting of selling a asset you do not own, in hope repurchasing it back for a lower price.

while student's t-distribution would have given $\text{VaR}_{99\%} = 0.0474$. Both give lower VaR estimate then in the individual case, found for the univariate section.

It is easy to see the advantages with the parametric approach, but as strong as they can be they can be equally weak if assumptions about the fitted distributions are bad. Therefore obtaining the parametric assumption right is the most important part of the parametric approaches (Dowd, 2005). The main disadvantages with the parametric approaches is dealing with non-linear instruments, such as options. In that case a linear approximation is needed which cannot capture the behavior of the instrument and therefore lead to large error.

Chapter 4

Non-parametric methods

4.1 Basic historical simulation

The attempt with the non-parametric models is to let the data (profit/loss or return series) speak for themselves as much as possible, rather than some fitted distribution. The main assumption with non-parametric models is that the recent past can be used to model the near future, meaning that some past returns, say two years, are used to model tomorrow's VaR. This way the data (returns) can accommodate any behavior, such as fat tails and skewness, without having to make any distributional assumptions, if the past returns showed that kind of behavior.

The most popular and known non-parametric model is the basic historical simulation (HS). For the basic HS the general idea is to sort the historical returns and estimate the VaR from the sorted historical returns at preferred confidence level α. Suppose for example we have 1000 observation of historical returns and would want to estimate VaR for tomorrow at a 99% confidence level. We would start by sorting the data, then we would know that 10 returns would lie in 'left' of the VaR estimate ($1\% \cdot 1000$) and therefore a rational estimate of tomorrow's VaR would be the 11^{th} one (or some interpolation between the 10^{th} and the 11^{th} one). Meaning that 99% of the time the loss is not more than the $\text{VaR}_{99\%}$.

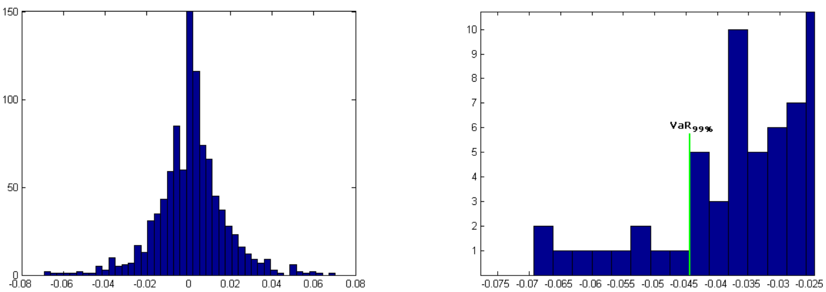
If for example the 15 worst returns the last 4 years have been the following (LAIS data):

{-0.0693, -0.0685, -0.0645, -0.0604, -0.0569, -0.0548, -0.0537, -0.0509, -0.0497, -0.0459, **-0.0443**, -0.0441, -0.0435, -0.0434, -0.0420, ... }

and the interest is to examine the 1% quantile for the period (i.e. the $\text{VaR}_{99\%}$) a reasonable estimate would be the 11^{th} one (since 4 years are equals to 1000 days[1] and 10 observation are allowed to lie 'left' of the estimate)

$$\text{VaR}_{99\%} = 0.0443$$

therefore it could be said that we are 99% sure of not getting worse return than 4.43% for tomorrow. Graphically this can be done by plotting a histogram and examine the tail as is shown in figure 4.1



(a) Histogram LAIS data (b) The tail

Figure 4.1: Shows how VaR is obtained graphically with the basic HS

The advantages with the basic HS is that it is really simple. The main disadvantages is that in the basic HS approach all observations are treated the same, that is all observations have the same weight (called equally-weighted). If all observation are treated the same a shock on the market today could be 'averaged' out if the sample size is large enough and not noticed at all except at high confidence levels. In other words risk grows without VaR showing it. Another example could be a major financial crisis in the past. This shock could produce high VaR estimate while it is in the sample space, called 'ghost effects', and then produce a jump in the estimate when it falls out of the sample space, see section 2.4.1.

One of the most attractive facts with the non-parametric approaches is that they can be applied as well for a multivariate case as well as univariate case and there is no need for estimating a variance-covariance matrix Σ, which is often the 'difficult' part of a multivariate estimation.

There are several implementations that can be added to the basic HS, such as bootstrapping (re-sampling the data over and over) and combination of non-parametric density function (for being able to treat the data as continuous, not discrete). One of the most popular implementation to the basic HS is to weight the data certain way so that not all observations are treated

[1] There are roughly 250 trading days each year

equally. These method's are called 'Weighted Historical Simulation' and can be thought as 'semi-parametric method' since they combine features of both non-parametric and parametric methods (Dowd, 2005).

4.2 Weighted historical simulation

There are various ways to adjust the data to overcome problems such as 'ghost effects'. Here I will introduce few of them.

4.2.1 Age weighted historical simulation (Age-WHS)

Observations are given weight according to their age as them name implies, so that recent (in time) observation will have more weight than older ones. Boudoukh, Richardson and Whitelaw (1998) introduced a formula for calculating observations weight as function of the decay factor λ

$$w(i) = \lambda^{i-1} \frac{1-\lambda}{1-\lambda^n} \tag{4.1}$$

where $w(i)$ is the weight to i days old observation (i.e $w(1)$ is the weight for the newest observation) and λ is the rate of decay, $0 < \lambda < 1$. High λ (close to 1) gives slow rate of decay and low λ gives high rate of decay, Boudoukh et al. (1998) recommend using $\lambda = 0.98$. As said returns are given weight according to their age, then the returns are sorted. Their weight's are then summed up, until preferred confidence level is achieved and corresponding return will give the VaR estimate.

Boudoukh et al. (1998) age-weighting formula is a nice generalization of basic equal weighted HS (the same as $\lambda \to 1$) and gives a more responsive VaR estimates with a well chosen decay factor, λ. The method is also helpful in reducing ghost effects, since old observation will have had weight close to zero and a large jump is thus less likely to be observed in the sample space.

4.2.2 Volatility weighted historical simulation (VWHS)

Proposed by Hull and White (1998) to update the return with volatility changes. As pointed out by them if for example the volatility on market today is 1.5% per day on average and two months ago it was 1% on average then the 'old' volatility will give an underestimate for changes in near future and vice versa. They therefore introduced a formula for updating return with volatility as:

$$r_t^\star = \sigma_T \cdot \frac{r_t}{\sigma_t} \tag{4.2}$$

where σ_t is the estimated daily volatility at time t, r_t is the historical return and σ_T is the most recent estimate of volatility made at the end of date T Hull and White (1998). More generally, daily returns are standardized with their volatility and then scaled with current volatility. This approach is a straight forward extension of the basic HS where volatility fluctuations has been taken into account in estimating VaR. Advanced techniques as EWMA or GARCH process could be used to estimate the volatility process to explain for example volatility clustering. This method has proved a higher estimate of VaR then basic HS (Dowd, 2005).

4.2.3 Filtered historical simulation (FHS)

Approach developed by Barone-Adesi, Bourgoin and Giannopoulos (1998)[2] has becoming more and more popular among risk analyzers. First a volatility process is fitted to the return (EWMA, GARCH or any other), then the returns (or the residuals, ϵ_t) are standardized with the volatility estimate as $z_t = r_t/\sigma_t$. Here heteroscedasticity should be removed (can be checked by for example looking at autocorrelation plot, see section 5.2.1). The standardized returns z_t are then bootstrapped. Bootstrapping involves drawing observations randomly from the sample, until the original sample size is reached. This is done N times (typically 500, 1000 or 5000 times). It should be mentioned that in the bootstrapping procedure the same observation can be drawn more often than once. Finally the new samples are scaled with current volatility σ_T, and then each sample can give an estimate of tomorrow's return and the VaR can be obtained at preferred confidence level. The procedure is shown graphically on figure 4.2.

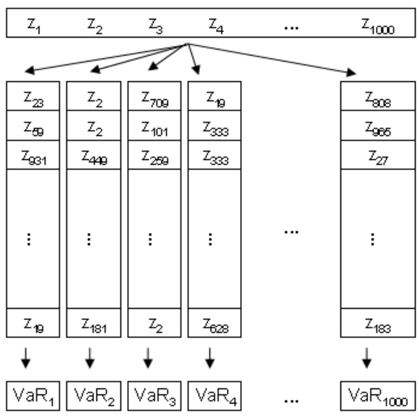


Figure 4.2: Bootstrap procedure

[2]and other papers by the same authors

Chapter 5

Analysis

5.1 Methodology

All methods described in chapters 3 and 4 are now compared with the goal of finding the method that gives fewest exceptions and lowest VaR estimate while fulfilling regulations stipulated by the Basel Committee (2006). By stating as low as possible, the aim is to minimize the regulatory capital, as that is depended on the VaR figure if the internal approach as set forth by the Basel Committee is used. If VaR limit is set very high or overestimated for some reasons, more capital will have to be kept in reserve.

All calculation were done in Matlab. For the multivariate GARCH calculations the UCSD Matlab toolbox by Kevin Sheppard was used[1]. As I mentioned in section 1.2, VaR has the unit of money, although in the analysis I will calculate VaR as a percentage of return for the sake of comparison between different instruments. Before going further I will give a short introduction to the regulations set by the Basel Committee (see Basel Committee (2006)).

5.1.1 Backtesting

Backtesting is a test performed to check the accuracy of internal VaR model, historically, meaning that over some period (at least 1 year) estimated daily VaR and actual P/L series[2] are compared and exceptions, when $-\mathrm{VaR}_\alpha > \mathrm{R}_t$, are counted. The internal model is given as:

$$\max\{\mathrm{VaR}_t, \frac{k}{60}\sum_{i=1}^{60}\mathrm{VaR}_{t-i+1}\} \tag{5.1}$$

[1] see http://www.kevinsheppard.com/wiki/UCSD_GARCH

[2] In my case return series, since VaR is in percentage

where VaR_t is the 10-day VaR estimate for day t, and k is known as the hysteria factor which is determined by the bank's backtesting result (somewhere between 3 and 4, see section 5.1.2). The reason for using the 10-day VaR, or 10-day holding period, is that it may take that long time to liquidate a position[3]. For interpolating the 1-day VaR to 10-day VaR the Basel Committee allows that the infamous 'square-root of time rule' should be used. The rule is given as:

$$\text{N-day VaR} = \sqrt{\text{N}} \times \text{1-day VaR} \tag{5.2}$$

The origin of this rule comes from that if you have 2 independent and identical normal distributed (normal iid) variables x_t and x_{t+1} with variance σ^2, the sum of their variance will be:

$$Var(x_t + x_{t+1}) = Var(x_t) + Var(x_{t+1}) = 2\sigma^2$$

Which implies that their volatility is scaled by $\sqrt{2}$. However this holds only if all observations are assumed to be normal iid, else it is a approximation (**?**).

Stylized facts such as heteroscedasticity violates the normal iid assumptions and therefore many have debated Basel Committee's recommendation using the square-root of time rule, see for example Daníelsson and Zigrand (2005). The most straight forward way of calculating a 10-day VaR is to use 10 times more data and divide it into 10-day intervals instead of 1-day. This means that for the short period 20 years of data would be needed and for the long period 40 years of data would be needed. The problem is that there are not that many stocks, indices or other financial time series with such a long history, and those who exist are likely to have changed drastically over last decades (it could thus be debated to use the same model for such a long period). Because of these debates and approximations I will concentrate on calculating 1-day VaR and skip any scaling or interpolations to other time intervals.

5.1.2 Basel zones

The Basel Committee requires that models are at least 99% accurate and uses a general hypothesis test, in order to balance two types of errors; (I) the possibility that an accurate risk model would be classified as inaccurate on the basis of its backtesting result, and (II) the possibility that an inaccurate model would not be classified that way based on its backtesting result. The Basel Committee categorizes backtesting result into 3 zones to minimize the type I and type II errors. Green zone, indicating that the model is probably good, minimal probabilities of type I error, yellow zone indicating uncertainty

[3]Meaning that it could take 10 days to sell the position.

and possibilities of both types of error and red zone indicating a probably bad model with a minimal chance of type II error. If the model ends in yellow zone it is up to the financial institution to prove it's goodness (Basel Committee, 2006). Table (1.2) shows the categorization for 1 year of data.

Zone	Number of exceptions	Increase in hysteria factor	Cumulative probability
Green Zone	0	0	8,11 %
	1	0	28,58%
	2	0	54,32%
	3	0	75,81%
	4	0	89,22%
Yellow Zone	5	0,40	95,88%
	6	0,50	98,63%
	7	0,65	99,60%
	8	0,75	99,89%
	9	0,85	99,97%
Red zone	10 or more	1	99,99%

Table 5.1: Backtesting zones in Basel accord based on 250 observations

Increase in hysteria factor is what adds to the default, $k = 3$, in equation 5.1. The cumulative probability is the binomial probability of getting the number of exceptions or fewer, see equation 5.3. For example the probability of getting 5 exceptions or less is equal to 95,88%, when dealing with 1 year of data and 99% confidence.

To interpolate the table to other time intervals the boundaries between green and yellow zone is when the cumulative binomial distribution is equal to/or exceeds 95% and the boundaries between yellow and red zone 99,99%. Limits for the short period are therefore, up to 8 exceptions is green zone, up to 14 is yellow zone and 15 or more will be red zone. For the long period the limits will be, up to 14 is green zone, up to 22 is yellow zone and 23 or more will be red zone.

5.1.3 The basic frequency test

The basic frequency test was proposed by Kupiec (1995) where the idea was to check with simple hypothesis testing whether to accept or reject a model. When checking number of exceptions the cumulative binomial distribution can be used:

$$P(K \leq x) = \sum_{i=0}^{K} \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \tag{5.3}$$

where n is the number of observations (days), K is the number of exceptions and p is the probability ($p = 1 - \alpha$). In hypothesis testing the idea is to propose a 'null-hypothesis' H_0 which is assumed to be true (in this case expected number of exceptions) and a 'alternative-hypothesis' H_1 which is the actual outcome from the model. If for example the total number of observation $n = 1000$, the actual number of exceptions $K = 20$, the confidence level $\alpha = 0.99$ ($\Rightarrow p = 0.01$) therefore the expected number of exceptions would be 10 ($1000 \times 0{,}01$) which is less then 20, the hypothesis test proposed could be:

$$\begin{cases} H_0 : p = 0.01 \\ H_1 : p > 0.01 \end{cases}$$

More generally, the idea is to check whether the model used for obtaining K exceptions is ok, when the expected number of exceptions is $n \cdot p$. This is done by putting the values of n, K and p are put into equation 5.3 and the results checked. $P(K \geq 20) = 1 - P(K \leq 19) = 0.0033$. Normally 5% confidence is used for validating the statistical test (Hull, 2007), therefore the 'alternative-hypothesis' would be rejected if $P(K = x)$ is less than the confidence level of the test. In this case $P(K \geq 20) = 0.0033 < 0.05$, therefore the null-hypothesis is rejected, which leads to that the model used for calculating this number of exceptions is rejected. If the number of exceptions had been less than the expected number of exceptions, say $K = 7$, the hypothesis had looked like:

$$\begin{cases} H_0 : p = 0.01 \\ H_1 : p < 0.01 \end{cases}$$

and the result had been $P(K \leq 7) = 0.2189 > 0.05$, and thus the 'alternative-hypothesis' is not rejected and therefore the model used is not rejected (found to be ok).

By using a combination of the two theories (Basel zones and frequency testing), we can sort the data into zones and also reject models who have too few exceptions according to 5% confidence interval, for preventing that the VaR limit is set to high. This will be used as a main criteria between methods. If models seem to be giving similar results, further analyzing criteria can be achieved by looking at means, standard deviations, minimum values and the models complexity.

5.2 Univariate case

First lets look at univariate analysis. Univariate means that there is only one asset (stock, currency pair, index) underlying. First I will present the results for the parametric approach, then non-parametric and finally compare them together.

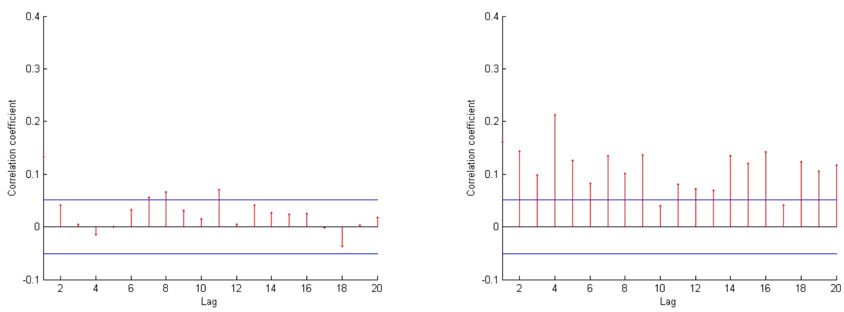
5.2.1 Parametric methods

Before modeling the data using parametric approach it is good to try to get as much information from the data as possible, to make as rational decisions as possible. One of those things is to examine the autocorrelation of the data which can give information about repeating patterns in the data. Autocorrelation is given as:

$$R(k) = \frac{E[(X_i - \mu)(X_{i+k} - \mu)]}{\sigma^2} \tag{5.4}$$

where k is the lag between observations. By calculation autocorrelation for the first two moments, i.e. r_t and r_t^2 (or ϵ_t and ϵ_t^2), repeating patterns in the mean and variance can be examined. As has been said before, one of the most stylized facts about financial time series is that they tend to have autocorrelation in the second moment (the variance, called heteroscedasticity). This fact cannot be dismissed, and therefore fitting a volatility process that can explain these characteristics should be rational. As was explained in section 2.4.3 a GARCH process is good in simulating heteroscedasticity and could thus be a wise choice.

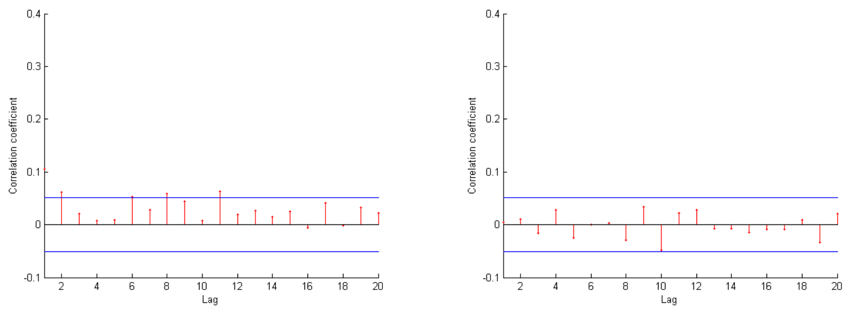
Here I shown the autocorrelation for the first two moments for LAIS data long period, see section 1.3 for information on time series.



(a) Autocorrelation in ϵ_t (b) Autocorrelation in ϵ_t^2

Figure 5.1: Examining autocorrelation in first two moments of the residuals, LAIS long period. The red bars show the autocorrelation for respective lag and the blue lines are the 95% confidence interval.

From figure 5.1 there is no indicator of autocorrelation in the first moment, the mean (figure (a)) while there is a strong indicator of autocorrelation in the second moment, the variance (figure (b)), almost all lags have autocorrelation higher then the 95% confidence interval. A GARCH process was fitted to the data and then the residuals standardized ($z_t = \epsilon_t/\sigma_t$). By examining autocorrelation in the standardized residuals (figure 5.2), it is possible to see whether the fitted volatility process did a good job in removing the autocorrelation or not.

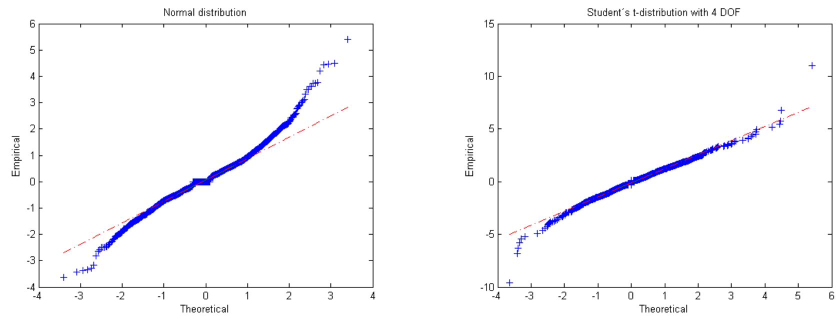


(a) Autocorrelation in z_t (b) Autocorrelation in z_t^2

Figure 5.2: Here the autocorrelation is shown for the standardized residuals, LAIS long period.

Figure 5.2 shows that there is no longer any clear autocorrelation in the standardized residuals and the next step would be to make some assumptions about the distribution. Similar results were achieved for other assets (see appendix A.1.1).

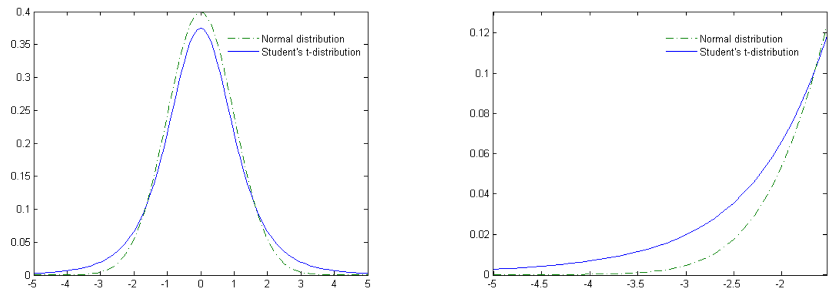
By looking at qq-plots, which are plots where empirical quantiles are plotted against theoretical quantiles, it is possible to check wheter a certain distribution fits the data. If a chosen distribution fits the data well the qq-plot should form a straight line. As before I show the results for LAIS long period.



(a) Normal distribution (b) Student's t-distribution

Figure 5.3: Estimating distribution, LAIS long period

From the qq-plots it's clear that Student's t-distribution (with 4 degrees of freedom) fits the LAIS data better then the normal distribution and should therefore be the choice (between those two), although I try both. Figure 5.4 shows the difference between a normal distribution and Student-t distribution with 4 degrees of freedom.



(a) Normal vs. t distribution (b) Tail behavior

Figure 5.4: The difference between normal distribution and Student's t-distribution with 4 degrees of freedom

Figure 5.4 shows the characteristic difference between normal and student's t-distribution (with 4 degrees of freedom).The figure shows that student's t-distribution has fatter tails then the normal distribution, which leads to higher VaR estimate, since the estimate is equal to the $\alpha\%$ area under the curve.

Now simulation can take place. Methods described in chapter 3 are calculated and results for LAIS are plotted in the following page. First figures for the long period, figures 5.5 and 5.6 then figures 5.7 and 5.8 shows how the methods react differently to shock (zoomed in on a shock). Figures for other stocks are shown in appendix A.1.1. For the EWMA model decay factor $\lambda = 0.94$ is used for stock and $\lambda = 0.97$ is used for the currency pair.

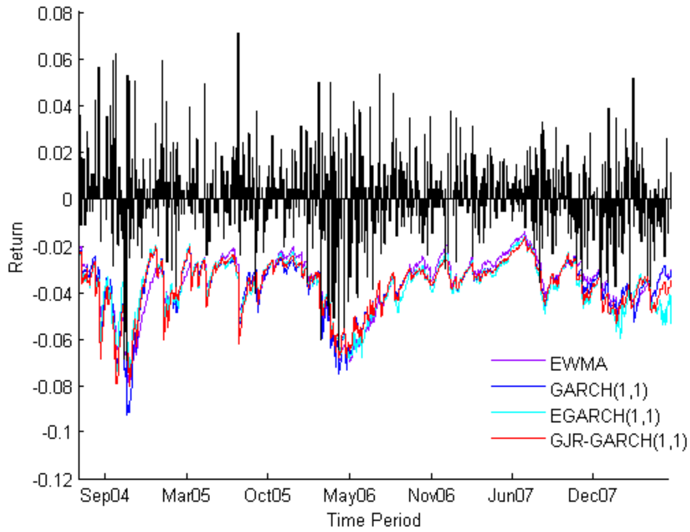


Figure 5.5: LAIS long period, assuming normal distribution

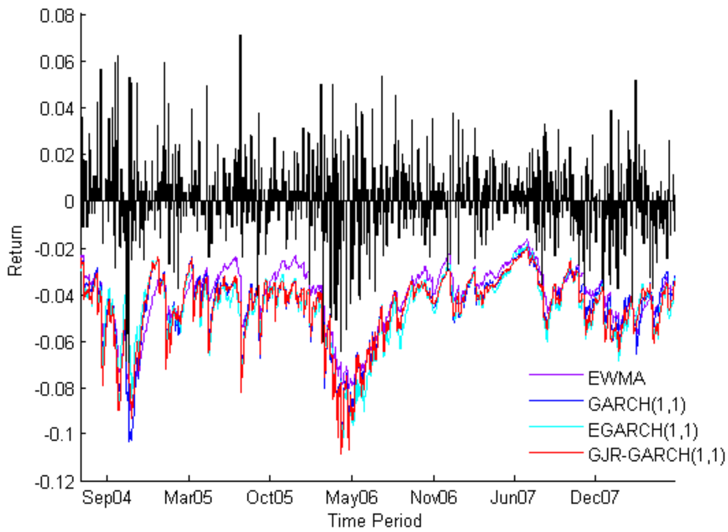


Figure 5.6: LAIS long period, assuming student's t-distribution

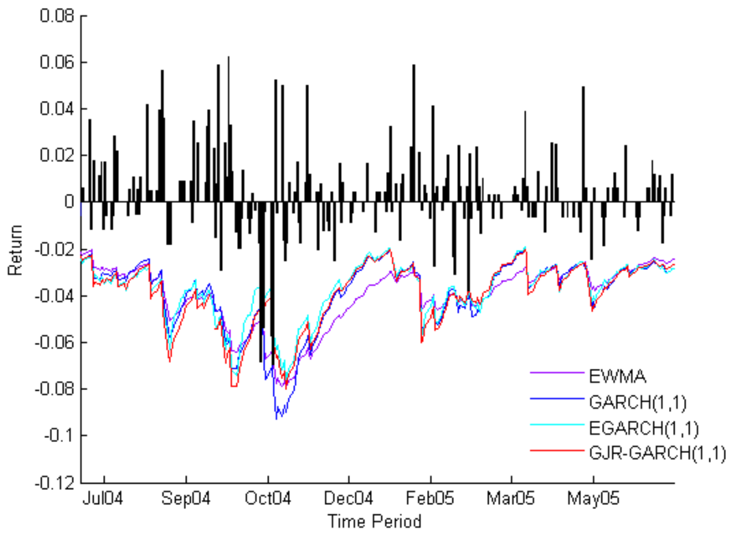


Figure 5.7: LAIS shock behavior, assuming normal distribution

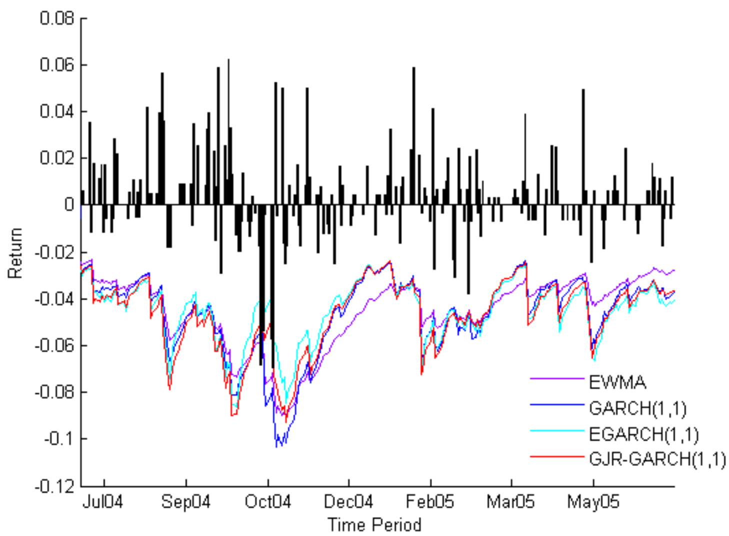


Figure 5.8: LAIS shock behavior, assuming student's t-distribution

Exceptions were counted and are presented in tables 5.2 to 5.5. In the tables the green color stands for green zone, no color for yellow zone and red color for red zone. Outlined numbers are those where a model has been rejected due to hypothesis testing. The number in brackets is observed exceptions divided by expected number of exceptions.

Method	LAIS	MARL	ERIC	NDA	ISXI15	OMX
EWMA	8 (1,6)	9 (1,8)	11 (2,2)	9 (1,8)	11 (2,2)	14 (2,8)
GARCH(1,1)	8 (1,6)	8 (1,6)	13 (2,6)	9 (1,8)	13 (2,6)	16 (3,2)
E-GARCH	6 (1,2)	8 (1,6)	17 (3,4)	8 (1,6)	12 (2,4)	14 (2,8)
GJR-GARCH	6 (1,2)	8 (1,6)	12 (2,4)	9 (1,8)	12 (1,4)	14 (2,8)

Table 5.2: Exceptions for the short period, assuming normal distribution

Method	LAIS	MARL	ERIC	NDA	ISXI15	OMX
EWMA	6 (1,2)	5 (1,0)	9 (1,8)	8 (1,6)	6 (1,2)	10 (2,0)
GARCH(1,1)	4 (0,8)	3 (0,6)	10 (2,0)	7 (1,4)	7 (1,4)	9 (1,8)
E-GARCH	2 (0,4)	2 (0,4)	7 (1,4)	7 (1,4)	9 (1,8)	9 (1,8)
GJR-GARCH	2 (0,4)	3 (0,6)	9 (1,8)	6 (1,2)	7 (1,4)	10 (2,0)

Table 5.3: Exceptions for the short period, assuming student's t-distribution

Method	LAIS	MARL	ERIC	NDA	ISXI15	OMX
EWMA	16 (1,6)	13 (1,3)	23 (2,3)	18 (1,8)	20 (2,1)	28 (2,8)
GARCH(1,1)	15 (1,5)	11 (1,1)	20 (2,0)	16 (1,6)	26 (2,6)	25 (2,5)
E-GARCH	16 (1,6)	13 (1,3)	24 (2,4)	16 (1,6)	25 (2,5)	23 (2,3)
GJR-GARCH	16 (1,6)	13 (1,3)	18 (1,8)	15 (1,5)	25 (2,5)	23 (2,3)

Table 5.4: Exceptions for the long period, assuming normal distribution

Method	LAIS	MARL	ERIC	NDA	ISXI15	OMX
EWMA	10 (1,0)	8 (0,8)	13 (1,3)	13 (1,3)	13 (1,3)	18 (1,8)
GARCH(1,1)	7 (0,7)	~~3~~ (0,3)	15 (1,5)	10 (1,0)	13 (1,3)	16 (1,6)
E-GARCH	8 (0,8)	~~2~~ (0,2)	12 (1,2)	11 (1,1)	17 (1,7)	16 (1,6)
GJR-GARCH	8 (0,8)	~~3~~ (0,3)	13 (1,3)	10 (1,0)	14 (1,4)	16 (1,6)

Table 5.5: Exceptions for the long period, assuming student's t-distribution

All models show improvements when t-distribution is assumed to fit the data rather then a normal distribution, except for GARCH methods on MARL data. The reason for that can be instability and that the model fails on finding a local maxima in the log-likelihood function. A further examination of means, standard deviation and minimum values of the VaR estimate are given in tables 5.6 - 5.9.

Method	LAIS Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	MARL Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	8	0.0324	0.0077	0.0529	9	0.0244	0.0060	0.0558
GARCH(1,1)	8	0.0338	0.0068	0.0526	8	0.0268	0.0065	0.1203
E-GARCH	6	0.0359	0.0087	0.0599	8	0.0268	0.0055	0.0795
GJR-GARCH	6	0.0346	0.0073	0.0532	8	0.0262	0.0063	0.1094

Method	ERIC Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	NDA Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	11	0.0513	0.0248	0.1587	9	0.0367	0.0108	0.0667
GARCH(1,1)	13	-	-	0.1242	9	0.0345	0.0094	0.0734
E-GARCH	17	0.0428	0.0098	0.0728	8	0.0346	0.0086	0.0732
GJR-GARCH	12	-	-	0.1315	9	0.0344	0.0098	0.0809

Method	ISXI15 Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	OMX Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	11	0.0272	0.0107	0.0557	14	0.0307	0.0083	0.0548
GARCH(1,1)	13	0.0274	0.0123	0.0755	16	0.0285	0.0085	0.0636
E-GARCH	12	0.0270	0.0110	0.0636	14	0.0277	0.0094	0.0592
GJR-GARCH	12	0.0273	0.0120	0.0678	14	0.0280	0.0100	0.0656

Table 5.6: Detailed information assuming normal distribution, short period

Method	LAIS Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	MARL Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	6	0.0369	0.0088	0.0602	5	0.0278	0.0068	0.0635
GARCH(1,)	4	0.0399	0.0091	0.0657	3	0.0375	0.0132	0.1771
E-GARCH	2	0.0418	0.0105	0.0685	2	0.0662	0.0384	0.2276
GJR-GARCH	2	0.0408	0.0095	0.0681	3	0.0363	0.0118	0.1445

Method	ERIC Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	NDA Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	9	0.0585	0.0282	0.1807	8	0.0418	0.0123	0.0760
GARCH(1,1)	10	0.0556	0.0208	0.1641	7	0.0395	0.0116	0.0922
E-GARCH	7	0.0551	0.0203	0.1466	7	0.0399	0.0112	0.0923
GJR-GARCH	9	0.0566	0.0263	0.2058	6	0.0395	0.0122	0.1001

Method	ISXI15 Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	OMX Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	6	0.0310	0.0122	0.0634	10	0.0349	0.0094	0.0625
GARCH(1,1)	7	0.0316	0.0146	0.0891	9	0.0330	0.0103	0.0762
E-GARCH	9	0.0316	0.0135	0.0772	9	0.0320	0.0112	0.0689
GJR-GARCH	7	0.0316	0.0144	0.0804	10	0.0324	0.0119	0.0774

Table 5.7: Detailed information assuming student's t-distribution, short period

Method	LAIS				MARL			
	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	16	0.0363	0.0121	0.0796	13	0.0277	0.0093	0.0703
GARCH(1,1)	15	0.0370	0.0117	0.0928	11	0.0310	0.0078	0.1203
E-GARCH	16	0.0375	0.0111	0.0775	13	0.0305	0.0074	0.0822
GJR-GARCH	16	-	-	0.0799	13	0.0307	0.0080	0.1094

Method	ERIC				NDA			
	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	23	0.0464	0.0207	0.1587	18	0.0317	0.0109	0.0667
GARCH(1,1)	20	-	-	0.1242	16	0.0319	0.0087	0.0734
E-GARCH	24	0.0443	0.0136	0.0885	16	0.0315	0.0083	0.0732
GJR-GARCH	18	-	-	0.1315	15	0.0318	0.0089	0.0809

Method	ISXI15				OMX			
	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	20	0.0257	0.0102	0.0557	28	0.0254	0.0097	0.0556
GARCH(1,1)	26	0.0254	0.0105	0.0755	25	0.0252	0.0084	0.0658
E-GARCH	25	0.0254	0.0093	0.0636	23	0.0247	0.0084	0.0592
GJR-GARCH	25	0.0253	0.0102	0.0678	23	0.0250	0.0092	0.0686

Table 5.8: Detailed information assuming normal distribution, long period

Method	LAIS				MARL			
	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	10	0.0413	0.0138	0.0906	8	0.0316	0.0106	0.0801
GARCH(1,1)	7	0.0449	0.0149	0.1033	~~3~~	0.0476	0.0179	0.2257
E-GARCH	8	0.0460	0.0145	0.1012	~~2~~	0.1192	0.1659	1.9051
GJR-GARCH	8	0.0457	0.0147	0.1082	~~3~~	-	-	0.2439

Method	ERIC				NDA			
	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	13	0.0529	0.0236	0.1807	13	0.0362	0.0124	0.0760
GARCH(1,1)	15	0.0526	0.0188	0.1641	10	0.0366	0.0101	0.0922
E-GARCH	12	-	-	0.1466	11	0.0366	0.0100	0.0923
GJR-GARCH	13	-	-	0.2058	10	0.0366	0.0107	0.1001

Method	ISXI15				OMX			
	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA	13	0.0293	0.0116	0.0634	18	0.0289	0.0110	0.0633
GARCH(1,1)	13	0.0300	0.0132	0.0891	16	0.0288	0.0098	0.0762
E-GARCH	17	0.0300	0.0125	0.0772	16	0.0283	0.0099	0.0689
GJR-GARCH	14	-	-	0.0823	16	0.0288	0.0108	0.0774

Table 5.9: Detailed information assuming student's t distribution, long period

The models give fairly similar results. EWMA, E-GARCH and GJR-GARCH give 11 green zones out of 24 and GARCH(1,1) 10 out of 24. When means, standard deviations and minimum values are examined, it comes clear that EWMA usually has the lowest mean and minimum values while GARCH based methods have the lowest standard deviations. This supports descriptions given in section 2.3, i.e. GARCH models are quicker to simulate market condition and spike higher, while EWMA is slower to follow market fluctuations (see figures 5.7 and 5.8). The GARCH models are more sensitive than the EWMA model, especially GJR-GARCH, and fails on getting results when high jumps occur in time series (MARL and ERIC). Finally I compare the parametric results with the extra long currency pair time series (10 years). Exceptions, means, standard deviation and minimum values are given in table 5.10.

	Normal distribution				Student's t-distribution			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	min(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	min(VaR)
EWMA	44	0.0179	0.0076	0.0591	22	0.0204	0.0086	0.0673
GARCH(1,1)	31	0.0179	0.0073	0.1036	14	0.0206	0.0081	0.0855
E-GARCH	42	0.0172	0.0081	0.1537	18	0.0199	0.0079	0.0853
GJR-GARCH	32	-	-	0.1389	18	-	-	0.1046

Table 5.10: Detailed information, currency pair

For the currency pair GARCH and GJR-GARCH has green zones for both distribution, while EWMA and E-GARCH has only green zone when t-distribution is assumed to be the governing distribution of the residuals. Means, standard deviation and maximum values are pretty similar between the methods. Here it is worth mention that although a frequency test would suggest to reject models who have 17 or fewer exceptions they will not be rejected because of central limit theorem[4].

Altogether none of the method is showing any superior characteristics. E-GARCH and GJR-GARCH do not improve the regular GARCH significantly and seem to be more sensitive than the regular one. E-GARCH has though the tendency to lower the $\overline{\text{VaR}}$, which is appealing. Since GARCH is known to be better in explaining fat tails and heteroscedasticity it is recommended as a first choice, but since EWMA isn't giving any fewer green zones it is recommended as second choice. It could thus be wise to model both for the sake of comparison and to minimize model and implementation risk.

[4]Central limit theorem indicates that if you for example toss a fair dice 10 times the probability of getting say 7 heads isn't that unlikely, but as you toss the dice more often a fair dice would give equal probabilities of getting head and tails. In other words the probability of getting 7 heads in 10 tosses isn't the same as getting 700 heads out of 1000 tosses. The limit gets narrower.

5.2.2 Non-parametric methods

For the non-parametric case any distributional assumption aren't necessary and therefore a pre analysis on the data isn't needed. Calculation for all the non-parametric method's described in the chapter 4 were done for the data. For the FHS simulation a GARCH(1,1) process was used to standardize returns and 1000 bootstraps were used. The mean of those 1000 were taken as the final solution. Solutions for all calculation are in tables. Here I present only LAIS figures (see figures 5.9 and 5.10) , the rest is in appendix.

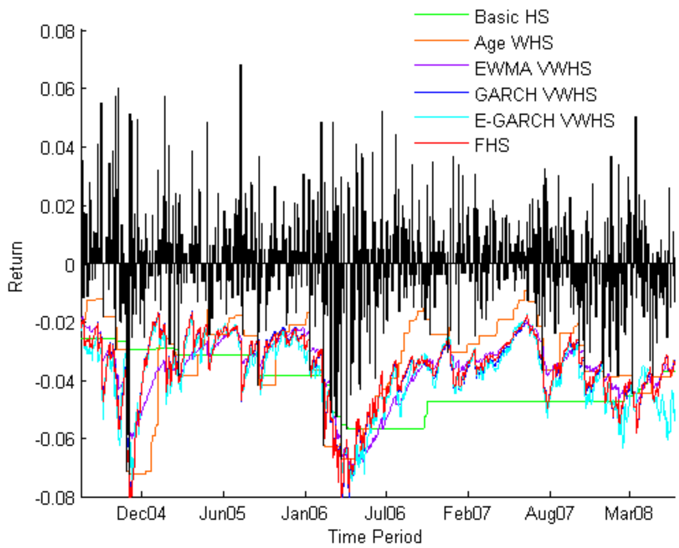


Figure 5.9: LAIS long period

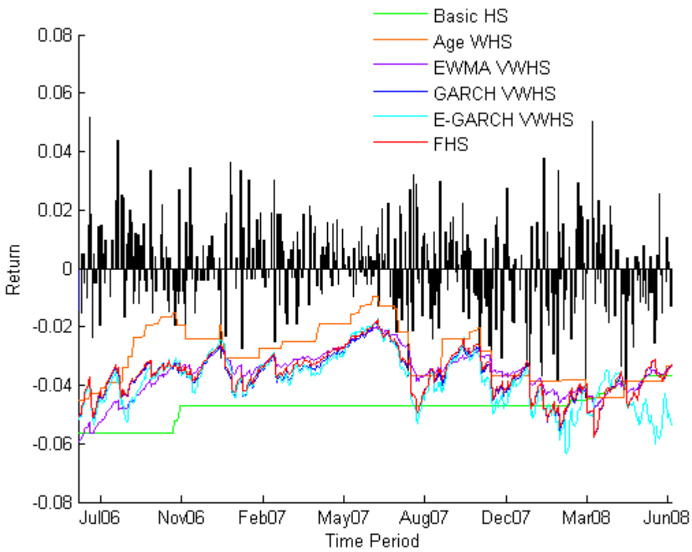


Figure 5.10: LAIS short period

Exceptions where counted and are displayed in tables 5.11 and 5.12. Overlined numbers indicates that the model is rejected due to hypothesis testing (too few exceptions). Color indicates which zone the model ends in (green means green zone, no color means yellow zone and red means red zone).

Method	LAIS	MARL	ERIC	NDA	ISXI15	OMX
Basic HS	~~1~~ (0.2)	4 (0.8)	8 (1.6)	6 (1.2)	7 (1.4)	7 (1.4)
Age WHS	13 (2.6)	13 (2.6)	12 (2.4)	12 (2.4)	15 (3.0)	14 (2.8)
EWMA VWHS	11 (2.2)	13 (2.6)	11 (2.2)	7 (1.4)	11 (2.2)	8 (1.6)
GARCH(1,1) VWHS	7 (1.4)	6 (1.2)	10 (2.0)	8 (1.6)	8 (1.6)	8 (1.6)
E-GARCH VWHS	5 (1.0)	8 (1.6)	11 (2.2)	6 (1.2)	9 (1.8)	10 (2.0)
FHS	7 (1.4)	9 (1.8)	9 (1.8)	8 (1.6)	8 (1.6)	7 (1.4)

Table 5.11: Exceptions for the short period

Method	LAIS	MARL	ERIC	NDA	ISXI15	OMX
Basic HS	15 (1,5)	~~8~~ (0,8)	13 (1,3)	13 (1,3)	19 (1,9)	14 (1,4)
Age WHS	27 (2,7)	30 (3,0)	26 (2,6)	27 (2,7)	31 (3,1)	31 (3,1)
EWMA VWHS	22 (2,2)	21 (2,1)	23 (2,3)	18 (1,8)	27 (2,7)	22 (2,2)
GARCH(1,1) VWHS	17 (1,7)	11 (1,1)	17 (1,7)	13 (1,3)	24 (2,4)	15 (1,5)
E-GARCH VWHS	16 (1,6)	13 (1,3)	17 (1,7)	11 (1,1)	23 (2,3)	17 (1,7)
FHS	17 (1,7)	11 (1,1)	15 (1,5)	13 (1,3)	22 (2,2)	15 (1,5)

Table 5.12: Exceptions for the long period

Means, standard deviations and maximum values are given in table 5.13 and 5.14.

	LAIS				MARL			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
Basic HS	1	0.0475	0.0048	0.0563	4	0.0278	0.0018	0.0316
Age WHS	13	0.0296	0.0096	0.0450	13	0.0218	0.0078	0.0352
EWMA VWHS	11	0.0305	0.0076	0.0490	13	0.0223	0.0061	0.0545
GARCH(1,1) VWHS	7	0.0365	0.0074	0.0578	6	0.0273	0.0071	0.1285
E-GARCH VWHS	5	0.0383	0.0090	0.0632	8	0.0264	0.0056	0.0862
FHS	7	0.0358	0.0075	0.0581	7	0.0270	0.0069	0.1259

	ERIC				NDA			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
Basic HS	8	0.0525	0.0078	0.0641	6	0.0351	0.0042	0.0453
Age WHS	12	0.0654	0.0669	0.3124	12	0.0337	0.0119	0.0572
EWMA VWHS	11	0.0539	0.0311	0.1940	7	0.0387	0.0130	0.0750
GARCH(1,1) VWHS	10	0.0518	0.0157	0.1356	8	0.0388	0.0118	0.0888
E-GARCH VWHS	11	0.0504	0.0135	0.0943	6	0.0387	0.0097	0.0833
FHS	9	0.0522	0.0169	0.1376	8	0.0381	0.0113	0.0861

	ISXI15				OMX			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
Basic HS	7	0.0338	0.0026	0.0395	7	0.0344	0.0039	0.0394
Age WHS	15	0.0263	0.0112	0.0470	14	0.0316	0.0098	0.0500
EWMA VWHS	11	0.0274	0.0113	0.0583	8	0.0353	0.0100	0.0647
GARCH(1,1) VWHS	8	0.0302	0.0139	0.0849	8	0.0358	0.0105	0.0773
E-GARCH VWHS	9	0.0304	0.0125	0.0733	10	0.0323	0.0102	0.0658
FHS	8	0.0301	0.0138	0.0838	7	0.0361	0.0115	0.0819

Table 5.13: Detailed information, short period

	LAIS				MARL			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
Basic HS	15	0.0414	0.0096	0.0563	8	0.0336	0.0075	0.0514
Age WHS	27	0.0331	0.0154	0.0717	30	0.0239	0.0098	0.0455
EWMA VWHS	22	0.0332	0.0114	0.0724	21	0.0247	0.0084	0.0573
GARCH(1,1) VWHS	17	0.0361	0.0118	0.0821	11	0.0308	0.0080	0.1285
E-GARCH VWHS	16	0.0378	0.0119	0.0751	13	0.0305	0.0082	0.0862
FHS	18	0.0358	0.0118	0.0817	11	0.0309	0.0084	0.1257

	ERIC				NDA			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
Basic HS	13	0.0634	0.0191	0.1429	13	0.0370	0.0091	0.0714
Age WHS	26	0.0551	0.0511	0.3124	27	0.0288	0.0119	0.0572
EWMA VWHS	23	0.0464	0.0251	0.1940	18	0.0326	0.0124	0.0750
GARCH(1,1) VWHS	17	0.0511	0.0165	0.1356	13	0.0358	0.0104	0.0888
E-GARCH VWHS	17	0.0498	0.0163	0.0989	11	0.0354	0.0092	0.0833
FHS	15	0.0516	0.0170	0.1372	13	0.0355	0.0099	0.0860

	ISXI15				OMX			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
Basic HS	19	0.0290	0.0061	0.0395	14	0.0310	0.0058	0.0433
Age WHS	31	0.0246	0.0118	0.0487	31	0.0261	0.0108	0.0500
EWMA VWHS	27	0.0243	0.0111	0.0583	22	0.0274	0.0124	0.0647
GARCH(1,1) VWHS	24	0.0266	0.0124	0.0849	15	0.0302	0.0111	0.0816
E-GARCH VWHS	23	0.0265	0.0114	0.0733	17	0.0282	0.0098	0.0692
FHS	22	0.0266	0.0123	0.0831	15	0.0304	0.0116	0.0818

Table 5.14: Detailed information, long period

The basic HS has the most green zone results 8 out of 12, then GARCH VWHS with 7 out of 12 and then FHS with 6 out of 12. The only model rejected due to hypothesis testing is the basic HS, rejected twice, which also has the highest means in all cases for the long period and for half of the cases in the short period (indicating that basic HS is overestimating VaR, can be checked by looking at plots 5.9 and 5.10). As for the parametric case

GARCH VWHS has the highest maximum values in most of the times, although the difference between the VWHS isn't that much most of the time. The Age WHS gives the poorest result, a red zone for all the cases in the long period.

Results for the extra long currency pair modeling are given in table 5.15.

	USDISK			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
Basic HS	44	0.0177	0.0036	0.0277
Age WHS	57	0.0172	0.0077	0.0523
EWMA VWHS	38	0.0178	0.0057	0.0408
GARCH(1,1) VWHS	29	0.0188	0.0067	0.0992
E-GARCH VWHS	37	0.0180	0.0077	0.1514
FHS	32	0.0186	0.0067	0.0976

Table 5.15: Detailed information, currency pair

Result for the currency pair supports the result from the stock and index analysis. GARCH VWHS and FHS have green zones, while AGE WHS gives red zone and the rest yellow zone.

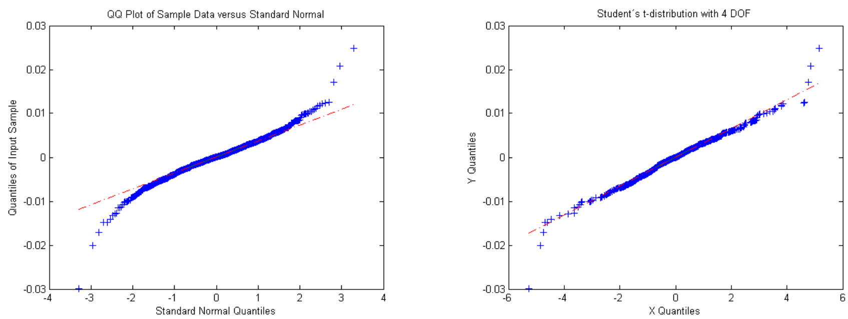
Since the basic HS is rejected twice in the stock and index analysis (and has the highest means), GARCH VWHS is regarded as the best of the univariate non-parametric models, since it gives more green zones then the other models and is slightly simpler then the FHS.

5.3 Multivariate case

As for the univariate case I first present the results for the parametric approach and then the non-parametric approach. For the parametric approach methods described in section 3.2 are calculated with covariance estimated as described in section 2.6.1.

5.3.1 Parametric methods

As in the univariate case plotting qq-plots can be helpful to see what distribution fits the data well. As before I check how the residuals fit to normal distribution and students t-distribution.



(a) Normal distribution (b) Student's t-distribution

Figure 5.11: Estimating distribution for Portfolio 1

As for the univariate case, student's t-distribution (with 4 degrees of freedom) fits the residuals better then the normal distribution, although as before I will try both. Portfolios described in section 1.3 are analyzed and VaR estimate is obtained both with multivariate EWMA and multivariate GARCH models. For the multivariate GARCH models, CCC and DCC, the univariate volatility estimates in matrix $\mathbf{D}_t$ (see equation 2.25) are all obtained by GARCH(1,1) process, since it showed the best result for the univariate case.

Figure 5.12 shows a comparison between multivariate EWMA, CCC and DCC model assuming normal distribution of the residuals and figure 5.13 shows a comparison between multivariate EWMA, CCC and DCC model assuming student's t-distribution of the residuals. Here I present only figures for Portfolio 1 (figures for Portfolio 2 is in appendix).

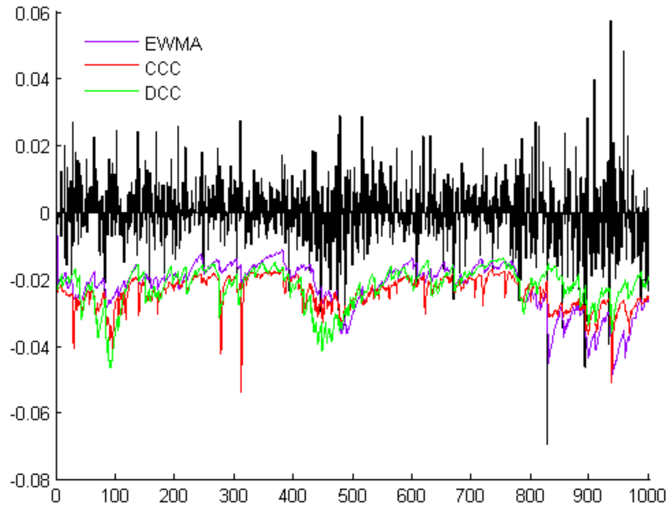


Figure 5.12: Multivariate case, assuming normal distribution

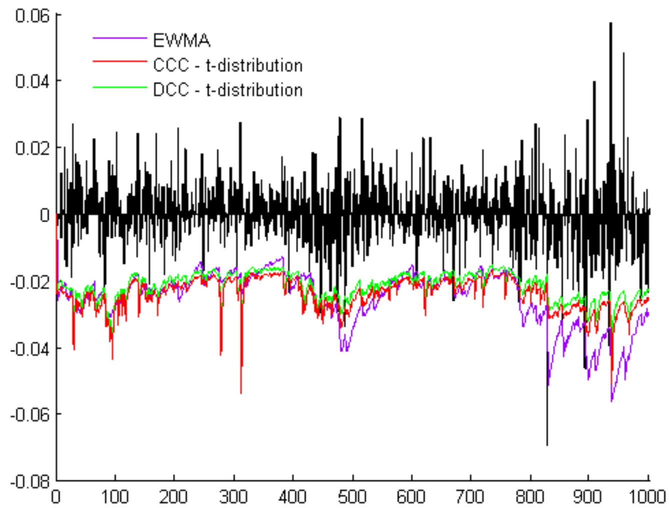


Figure 5.13: Multivariate case, assuming Student's t-distribution

Then I examine if taking covariances into account really matters. This can be done by estimating the VaR of each asset in the portfolio separately, then summing them together (depended on their weight) and comparing to regular multivariate case.

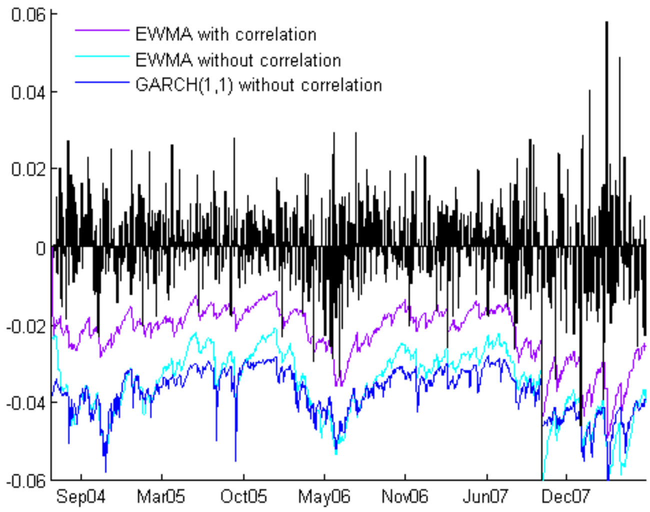


Figure 5.14: Multivariate case, assuming normal distribution

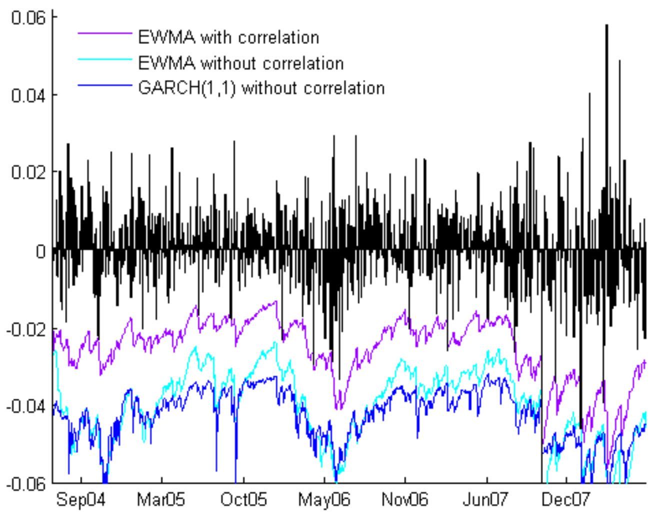


Figure 5.15: Multivariate case, assuming Student's t-distribution

Finally exceptions are counted and presented in tables 5.14 and 5.15. EWMA† and GARCH(1,1)† stands for EWMA and GARCH(1,1) without taking covariance into account.

Method	Portfolio 1	Portfolio 2
EWMA†	~~1~~ (0.1)	~~0~~ (0.0)
GARCH(1,1)†	~~2~~ (0.2)	~~4~~ (0.4)
EWMA	22 (2.2)	14 (1.4)
CCC	18 (1.8)	15 (1.5)
DCC	17 (1.7)	16 (1.6)

Table 5.16: Exceptions for the multivariate case, assuming normality

Method	Portfolio 1	Portfolio 2
EWMA†	~~0~~ (0.0)	~~0~~ (0.0)
GARCH(1,1)†	~~1~~ (0.1)	~~0~~ (0.0)
EWMA	13 (1.3)	12 (1.2)
CCC	13 (1.3)	12 (1.2)
DCC	19 (1.9)	13 (1.3)

Table 5.17: Exceptions for the multivariate case, assuming t-distribution

As can be seen from the figures 5.14 and 5.15 and tables 5.16 and 5.17 not taking covariances into account raises the VaR estimate a lot, as was expected (see section 2.5), and with hypothesis testing all of the cases when correlation are not taken into account are rejected due to too few exceptions. Further analysis of means, standard deviations and maximum values are presented in tables 5.18 and 5.19.

	Portfolio 1				Portfolio 2			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA†	~~1~~	0.0353	0.0087	0.0615	~~0~~	0.0341	0.0085	0.0632
GARCH(1,1)†	~~2~~	0.0373	0.0060	0.0668	~~4~~	0.0355	0.0066	0.0673
EWMA	22	0.0223	0.0073	0.0492	14	0.0235	0.0070	0.0461
CCC	18	0.0240	0.0047	0.0538	15	0.0249	0.0067	0.0535
DCC	17	0.0223	0.0059	0.0464	16	0.0247	0.0067	0.0530

Table 5.18: Detailed information assuming normal distribution

	Portfolio 1				Portfolio 2			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	max(VaR)
EWMA†	~~0~~	0.0403	0.0098	0.0702	~~0~~	0.0389	0.0096	0.0720
GARCH(1,1)†	~~1~~	0.0424	0.0068	0.0761	~~0~~	0.0404	0.0075	0.0766
EWMA	13	0.0255	0.0082	0.0560	12	0.0267	0.0080	0.0526
CCC	13	0.0240	0.0047	0.0538	12	0.0263	0.0064	0.0658
DCC	19	0.0219	0.0044	0.0532	13	0.0258	0.0064	0.0657

Table 5.19: Detailed information assuming Student's t-distribution

Although the multivariate GARCH models are theoretically more 'fancy' then multivariate EWMA, they do not provide more green zones then multivariate EWMA. The multivariate EWMA model has 3 green zones out of 4, while CCC has 2 out of 4 and finally DCC 1 out of 4. Detailed information shows that multivariate EWMA has the lowest maximum VaR value in 3 out of 4, but the highest standard deviation in all cases. Mean values are pretty similar for all the multivariate methods.

Since multivariate EWMA has the most green zones and multivariate GARCH doesn't show any superior skills multivariate EWMA is valued as the best method for the multivariate parametric case. An really important factor is also that the multivariate EWMA is much more simpler to calculate than the very complex multivariate GARCH cases.

5.3.2 Non-parametric methods

For the Non-parametric case the methods described in chapter 4 were compared. The portfolio return and the portfolio volatility was calculated with equation 2.4 and 2.5 for the VWHS case, and then proceeded as in univariate case. For the FHS a GARCH(1,1) volatility process was used to standardize the portfolio returns, 1000 bootstraps were made and the mean of those gave the final VaR estimate. Here I only present the figure for portfolio 1 (portfolio 2 is in appendix).

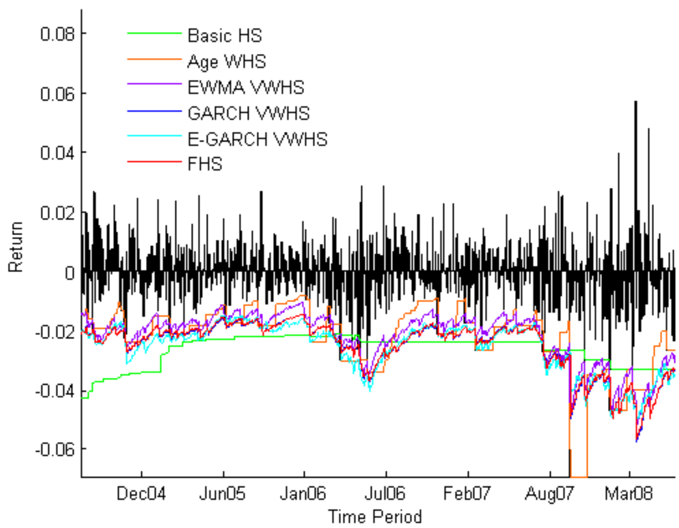


Figure 5.16: Multivariate case, assuming Student's t-distribution

Exceptions are presented in table 5.20

Method	Portfolio 1	Portfolio 2
Basic HS	14 (1.4)	17 (1.7)
Age-WHS	25 (2.5)	24 (2.4)
EWMA VWHS	23 (2.3)	15 (1.5)
GARCH(1,1) VWHS	16 (1.6)	13 (1.3)
E-GARCH VWHS	13 (1.3)	12 (1.2)
FHS	15 (1.5)	14 (1.4)

Table 5.20: Exceptions for the Non-parametric multivariate case

The only method that has green zones for both portfolios is the E-GARCH. Age-WHS gives red zones for both portfolios, which was also the case in the univariate solutions, see 5.12. For further examination means, standard deviations and maximum values are given in table 5.21.

	Portfolio 1				Portfolio 2			
Method	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	min(VaR)	Exceptions	$\overline{\text{VaR}}$	σ_{VaR}	min(VaR)
Basic HS	14	0.0264	0.0052	0.0427	17	0.0262	0.0034	0.0302
Age WHS	25	0.0226	0.0119	0.0693	24	0.0222	0.0098	0.0446
EWMA VWHS	23	0.0219	0.0085	0.0534	15	0.0215	0.0081	0.0479
GARCH(1,1) VWHS	16	0.0251	0.0086	0.0573	13	0.0247	0.0080	0.0561
E-GARCH VWHS	13	0.0253	0.0083	0.0507	12	0.0240	0.0077	0.0478
FHS	15	0.0250	0.0086	0.0563	14	0.0242	0.0074	0.0541

Table 5.21: Detailed information

By examining the two tables (5.20 and 5.21) it is clear that GARCH based methods give fewer exceptions then the other, and of the GARCH based methods, E-GARCH seems to be the best, since it has green zones for both portfolios and lower minimum value than the other two. The basic HS has pretty good results for the portfolios. It doesn't have that many exceptions and has the lowest standard deviation and minimum value for both portfolios. But as was shown in univariate case they can overestimate the VaR value drastically. Therefore E-GARCH is deemed the best method for the multivariate non-parametric case.

Chapter 6

Conclusion

As was said in section 1.1 the aim of the thesis was to evaluate different methods for calculating VaR in order to obtain the optimal method for banks trading book. Many calculations have been carried out with various results. Of course it is difficult to point out one method that is 'definitely' superior than the others and results are dependent on assumptions and prerequisites. Factors such as whether the data can be regarded as univariate or multivariate, how long the time period is, which distribution is assumed, what complexity is allowed and how long does it take to perform calculation are all things that can have great impact on the valuation.

If the conclusion could be separated into to two main fields, the univariate and the multivariate case, which both have parametric and non-parametric approaches the best models would be as in table 6.1

	Univariate	Multivariate
Parametric	GARCH(1,1) (EWMA)	EWMA
Non-Parametric	GARCH(1,1) VWHS	E-GARCH VWHS

Table 6.1: Separated conclusion

Of the parametric approaches the GARCH(1,1) and EWMA are found to give the best result in the univariate case, but despite GARCH's 'fancy' behavior, it's complexity affects the multivariate result and EWMA model is regarded as the best. In all cases student's t-distribution lowered the number of exceptions, although it made minimum values larger in most cases. Student's t-distribution is more sensitive to finding a local maxima and for example failed for the MARL data. I would therefore recommend that normal distribution is always used as well as a alternative distribution to avoid such incidents. In the non-parametric case volatility weighted historical simulations (VWHS) seem to be the best. They are relatively easy to implement and can be used for both univariate and multivariate case without complications. They manage to minimize ghost effects and react to volatility fluctuations. Filtered historical simulation also gave satisfactory results and are quite interesting with their bootstrapping procedure, which could developed further by for example enlarging the number of bootstraps, and/or obtaining the final estimate by some other way then taking the average. Age weighted historical simulation didn't give good results in any case and the basic historical simulation showed tendency to overestimate VaR drastically.

GARCH based models are more responsive and react to market conditions better than other models. Their main failure is that they become intolerably complex as the dimensionality grows, while EWMA is always quite simple to implement. Bank's trading book is usually very large and holds all sorts of instruments. It can therefore regarded as strictly multivariate and complex. When evaluating VaR for bank's trading book, I would recommend using a mixture of parametric and non-parametric models both to get a comparison as well as to prevent failures (such as model and implementation failures). My proposition for multivariate book would be parametric EWMA model and a non-parametric GARCH based VWHS. But as said, bank's trading book is usually very large and complex and therefore it is important to know the underlying risk factors when choosing a model.

It is also worth mentioning that the last 2 years have been really unique in financial markets. Stock and indices have never been as high, and they have been fallen really quickly past 2 years therefore the results of the this thesis are colored by that. Not meaning that the results are any less important, but just that it is important to keep in mind this extreme fluctuations on financial markets over the past years when results are examined. Figure 6.1 shows the development of the OMX index from January 1993.

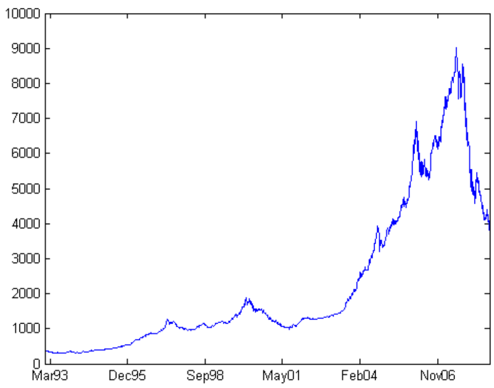


Figure 6.1: OMX index since January 1993

Finally I will list the main advantages and disadvantages with the parametric and the nonparametric approaches.

The main advantages with the parametric approach are;

- It's mathematically more advanced to use the parametric approach. You will get a density and distribution function so there will be a straight forward way of calculating VaR.
- They will give a better results than the non-parametric approach, if assumptions about distributions are correct (or to be more specific good) so the most important step is to make rational decisions about the distribution.

while the main disadvantages are

- That you have to make an assumption about the distribution, and since financial data tend to be clustered, skewed and/or fat tailed picking a distribution can be hard.
- Computationally more complex than non-parametric approach
- For non-linear assets they are an approximation and very complicated.

For the non-parametric methods, the main advantages could be summed up as;

- They don't have any distribution and can therefore accommodate features as fat tails, skew and any non-normal features, which are harder to describe with parametric methods.
- They are relatively easy and simple
- Can be used for any type of risk, linear and non-linear (derivatives)
- They don't have any high dimensional problems
- Several implementations (or refinements) have been developed for the Basic HS, which can reduce 'ghost effects'.

while their main disadvantages are;

- They will only simulated from the past history, meaning that results will never be worse then what is in the sample (if the period was quiet the VaR estimate will be low and vice versa).
- If the past (sample space) has some extreme losses, the VaR estimate will be reflected by that loss (unless some refinement are used).

- Ghost effects are likely to corrupt the estimate.
- Can be computationally heavy, especially for more complicated assets, as the assets has to be revalued.

6.1 Criticism on VaR

Although VaR is a very popular measure of risk in the financial sector there are many who have criticized that it isn't a very sophisticated risk measure. Artzner, Delbaen, Eber and Heath (1999) debated that VaR isn't sub-additive, meaning that the VaR of a portfolio may be larger than the sum of individual instruments in the portfolio. This is a fundamental rule, since it means that diversification isn't guaranteed to reduce risk. Let's take an example.

Imagine two identical bonds A and B both with the default probability of 4%. Now if default occurs the loss will be the value of the bond, say equal to 100. Therefore the $\text{VaR}_{95\%}$ of each bond is 0 (higher than the 4% chance of default) and the $\text{VaR}_{95\%}(\text{A}) = \text{VaR}_{95\%}(\text{B}) = 0$. Now imagine a portfolio of A and B. The probability of a loss equal to 200 is P(A will default) $\times$ P(B will default) $= 0.04^2 = 0.0016$, and likewise the probability of no loss is $0.96^2 = 0.9216$. Therefore the loss equal to 100 (one of the bond will default) is P(A or B will default) = 1 - P(no default) - P(both will default) $= 1 - 0.9216 - 0.0016 = 0.0768$. Therefore the $\text{VaR}_{95\%}(A+B) = 100 > \text{VaR}_{95\%}(A)+\text{VaR}_{95\%}(B)$. Diversification has failed.

Generally speaking VaR doesn't say anything about the potential loss when the loss occurs. VaR only gives the amount we are α percent sure of losing not more than. Say that the $\text{VaR}_{99\%} = 1$ million euros. Now if the unlikely occurs and we suffer from a potential loss greater than $\text{VaR}_{99\%}$ we have no idea if it will be 1.1 million euros or 100 million euros. (Daníelsson, 2002) The loss is only depending on the tail of the distribution (in left of the VaR value).

Other criticism proposed is for example by Taleb (1997) and Hoppe (1998) who argued the statistical assumptions of VaR could lead to major errors, Beder (1995) argued that different VaR models can give different VaR estimates which makes the estimate impercise, furthermore Marshall and Siegel (1997) argued that a similar techniques could give different estimates due to implementations of the models. All this uncertainty concerning VaR can lead to that experts and traders do not fully trust the VaR proposed and take on larger risk than suggested by the VaR number, making the VaR biased downwards (Ju and Pearson, 1999). Daníelsson (2002) argued that regulatory constraints might discourage good risk management.

6.2 Further analysis

There are various ways to extend the research of this thesis. In the parametric approach different distribution could be tested, for example log-normal, extreme value distribution or perhaps some skewed distribution, and in the non-parametric approach fitting a distribution to the histogram could be quite interesting and would combine the features of both parametric and non-parametric approach (called Kernel's). Comparing other types of stocks, portfolios and currency pairs is straight forward and a whole new landscape would be obtained by trying modeling non-linear instruments, such as options and futures, where Monte Carlo method's could come handy. Alternative volatility estimates such as stochastic and implied volatility are interesting as well as Copula theory for estimating covariance. There would also be interesting to research model/methods for interpolating VaR to other time intervals, both longer, for example 10-day VaR which Basel Committee demands, as well as shorter intra-day VaR (10 or 20 minute VaR) which are common among traders.

Appendix A

Pictures

A.1 Univariate case

A.1.1 Parametric methods

First I show how GARCH(1,1) removes the autocorrelation in the 2^{nd} moment by plotting autocorrelation plot of the squared residuals, ϵ_t^2, and the squared standardized residuals, z_t^2.

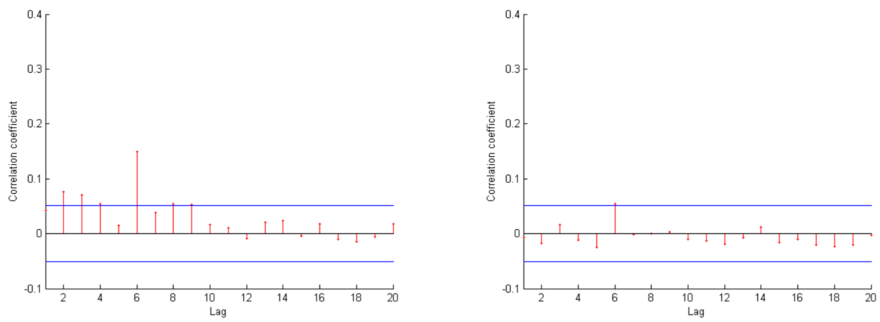


Figure A.1: Autocorrelation for ϵ_t^2 and z_t^2, MARL

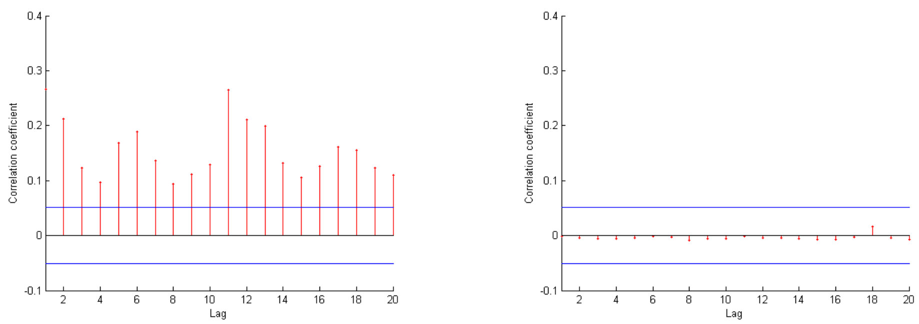


Figure A.2: Autocorrelation for ϵ_t^2 and z_t^2, ERIC

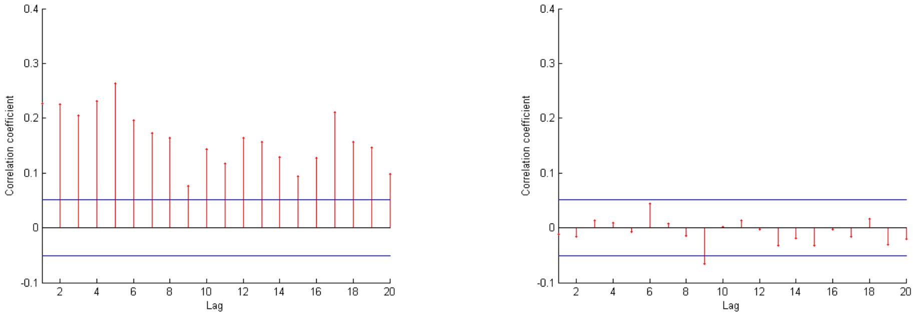


Figure A.3: Autocorrelation for ϵ_t^2 and z_t^2, NDA

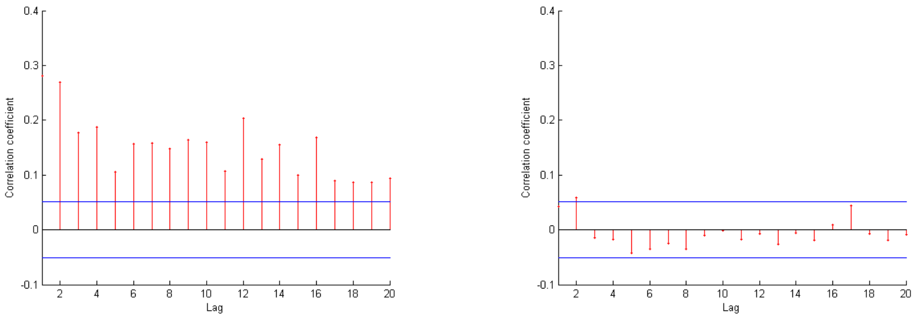


Figure A.4: Autocorrelation for ϵ_t^2 and z_t^2, ISXI15

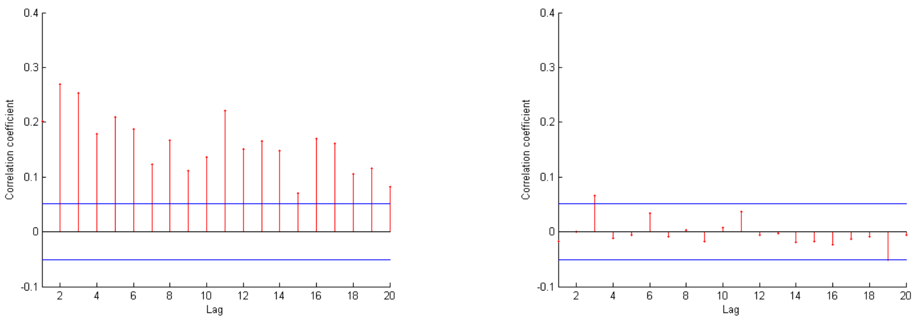


Figure A.5: Autocorrelation for ϵ_t^2 and z_t^2, OMX

As can be seen in figures GARCH(1,1) succeeds well in removing heteroscedasticity. Following are rest of the figures for the Parametric approach. Only the long period is showed (the short period is the latter half of the long period).

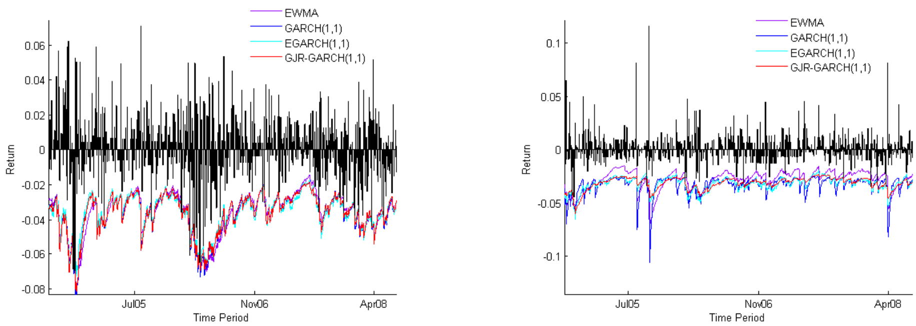


Figure A.6: LAIS and MARL

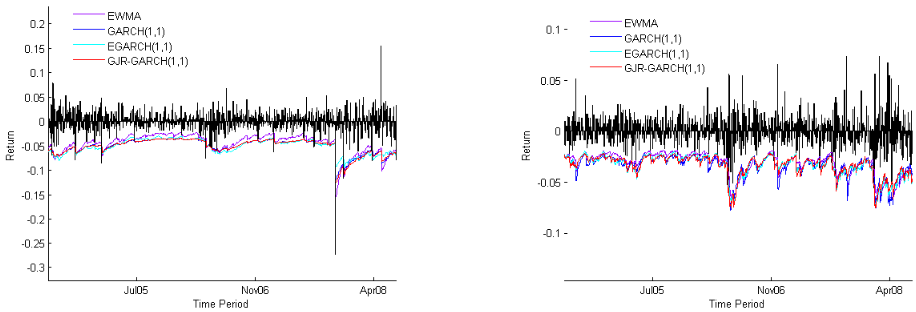


Figure A.7: ERIC and NDA

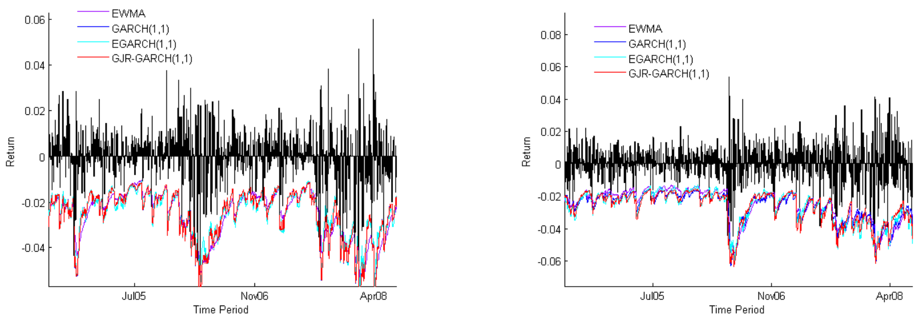


Figure A.8: ISXI15 and OMX

Then figures for the long period assuming student's t-distribution

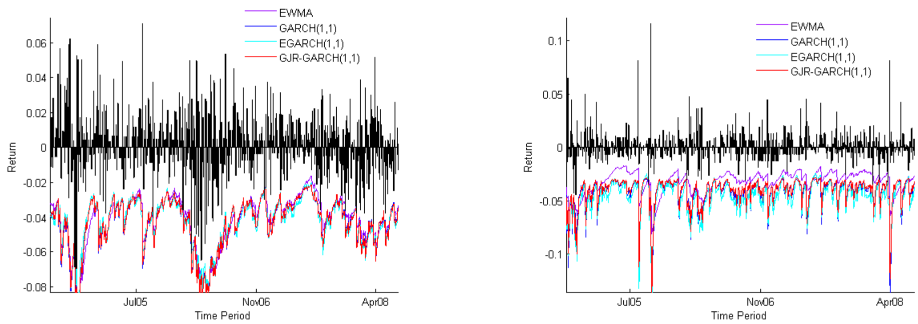


Figure A.9: LAIS and MARL

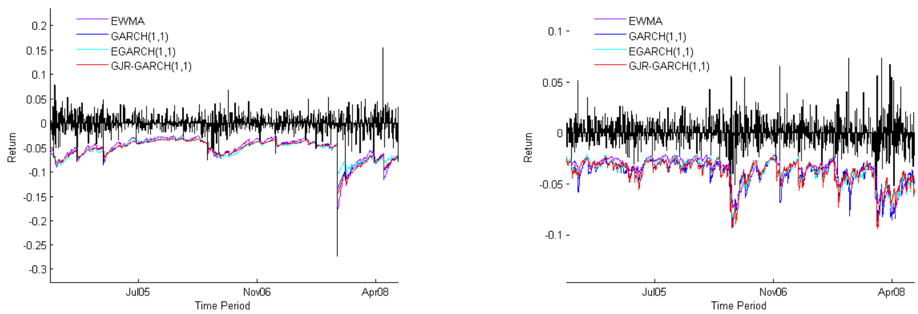


Figure A.10: ERIC and NDA

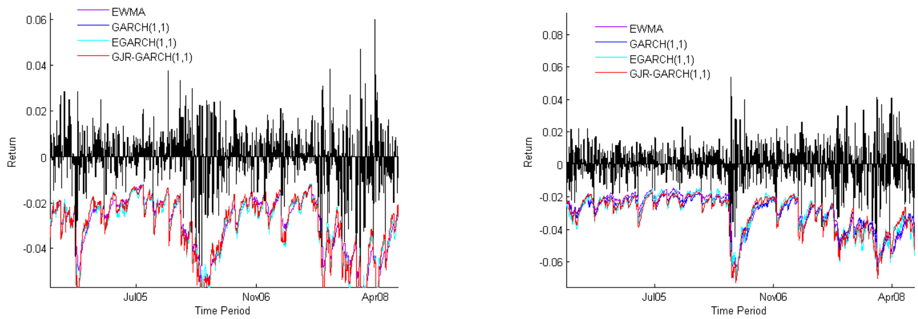


Figure A.11: ISXI15 and OMX

A.1.2 Non-parametric methods

Following are the rest of the figures for the Parametric approach. Only the long period is showed (the short period is the latter half of the long period).

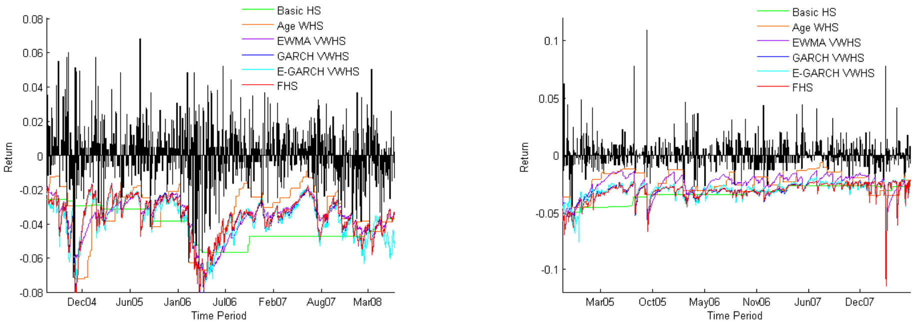


Figure A.12: LAIS and MARL

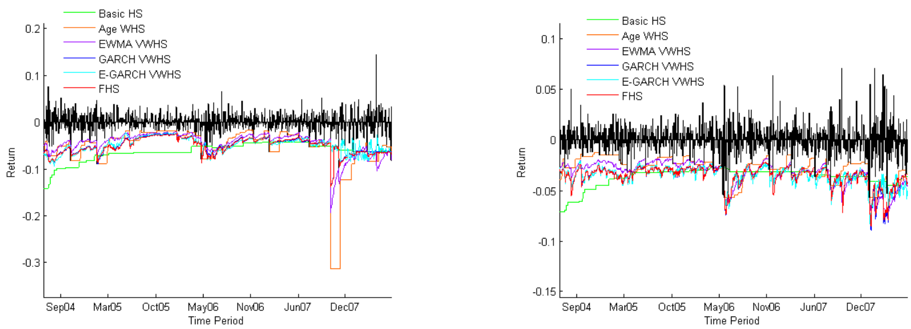


Figure A.13: ERIC and NDA

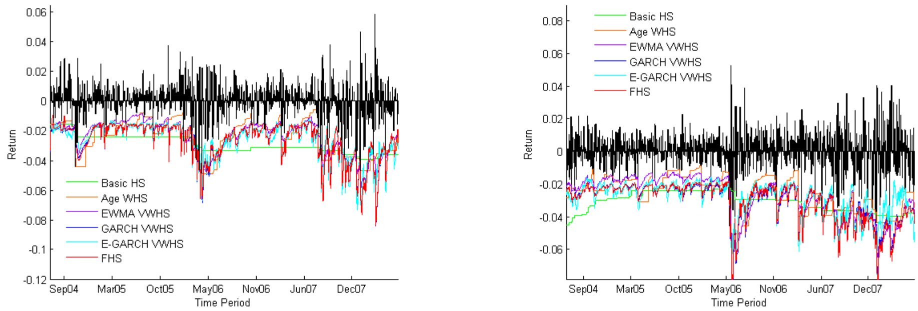


Figure A.14: ISXI15 and OMX

A.2 Multivariate case

A.2.1 Parametric methods

Figures for Portfolio 2 are presented here below.

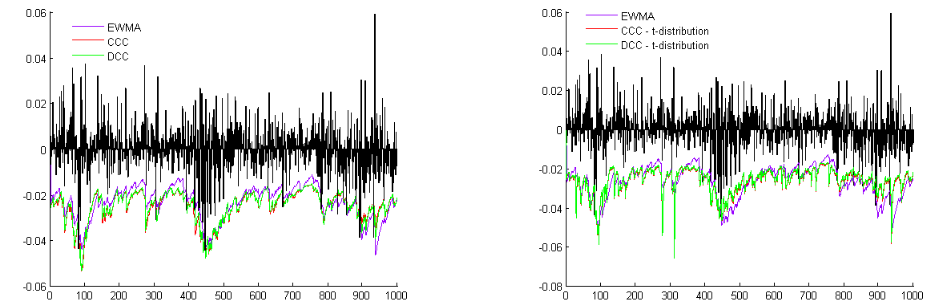


Figure A.15: Normal distribution and student's t-distribution

Then pictures showing that not taking covariances into account overestimates the value of VaR for portfolio 2.

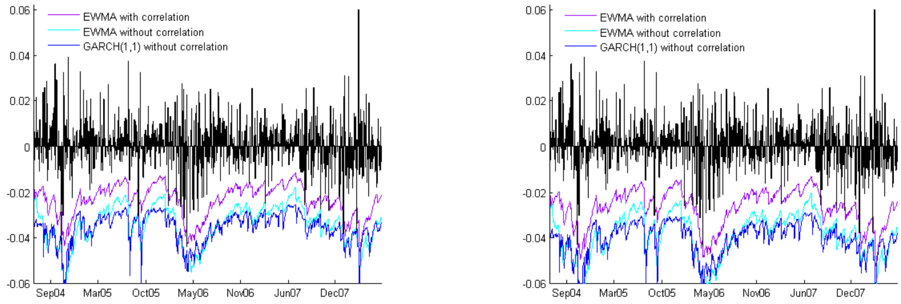


Figure A.16: Normal distribution and student's t-distribution

A.2.2 Non-parametric methods

Finally the non-parametric result for portfolio 2

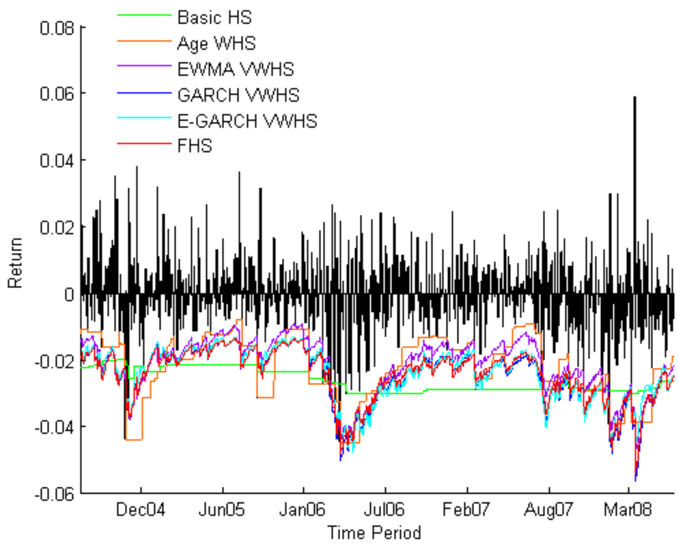


Figure A.17: Portfolio 2, non-parametric

Bibliography

Artzner, P., Delbaen, J., Eber, M. and Heath, D. (1999). Coherent measure of risk. *Mathematical Finance*. 9. 203–228.

Barone-Adesi, G., Bourgoin, F. and Giannopoulos, K. (1998). Don't look back. *Risk*. 11. 100–104.

Basel Committee on Banking Supervision. (2006). *International convergence of capital measurement and capital standards*. Basel: Bank for International Settlements.

Beder, T. (1995). VaR: Seductive but dangerous. *Financial Analysts Journal*. 51. 12–24.

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

Bollerslev, T., Engle, R. and J. Wooldridge. (1988). A capital asset pricing model with time varying covariances, *Journal of Political Economy*. 96. 116–131.

Bollerslev, T. (1990) Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model, *The Review of Economics and Statistics*, 72, 498–505.

Boudoukh, J., Richardson, M. and Whitelaw, R. (1998). The best of both worlds: A hybrid approach to calculate value at risk. *Risk*. 11. 64–67.

Daníelsson, J. (2002). The emperor has no clothes: Limits to risk modelling. *Journal of Banking and Finance*. 26. 1273–1296.

Daníelsson, J. and Zigrand, J. P. (2005). On time-scaling of risk and the square-root-of-time-rule. EFA 2004 Maastricht Meetings. No. 5339.

Dowd, K. (1998). *Beyond value at risk: The new science of risk management*. New Yotk: John Wiley and sons Ltd.

Dowd, K. (2005). *Measuring market risk* (2nd edition). New York: John Wiley and sons Ltd.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*. 50. 987–1008.

Engle, R. (2002). Dynamic conditional correlation - A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics*. 20. 339–350.

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

Hansen, P. R. and Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*. 20. 873–889.

Hoppe, R. (1998). VaR and the unreal world. *Risk*. 11. 45–50.

Hull, J. C. and White, A. (1998). Incorporating volatility updating into the historical simulation method for value at risk. *Journal of Risk*. 1. 5–19.

Hull, J. C. (2007). *Risk management and financial institutions*. New Jersey: Pearson Prentice Hall.

Ju, X. and Pearsson, N. D. (1999). Using value-at-risk to control taking: How wrong can you be?. *Journal of Risk*. 1. 5–36

Kaplan, S. and Garrick, B.J. (1981). On the quantitative definition of risk. *Risk Analysis*, 1, 11–27.

Kolluru, R. V. (1995). *Risk assessment and management: Handbook for environmental*. New York: McGraw-Hill Education

Kupiec, P. H. (1995). Techniques for verifying the accurancy of risk management models. *Journal of Derivatives*. 3. 73–84.

Linsmeier, T. J. and Pearson, N. D. (1996). *Risk measurement: An introduction to value at risk*. Mimeo, University of Illinois at Urbana-Champaign.

Marshall, C. and Siegel, M. (1997). Value at risk: Implementing a risk measurement standard. *Journal of Derivatives*. 4. 91–110.

Morgan, J. P. (1996). *RiskMetrics technical document*. New York, J.P. Morgan Bank.

Nelson D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*. 59. 347–370.

Sheppard, K. (2003). *Multi-step estimation of multivariate GARCH models*. Proceedings of the international ICSC symposium: Advanced computing in financial markets.

Sigurdarson, H. A. (2007). Margvid GARCH líkön og verdlagning afleidna. Unpublished M.Sc. thesis. Háskóli Íslands.

Taleb, N. (1997). Against VaR. *Derivatives Strategy*. 2. 21–26.