LUND UNIVERSITY
School of Economics and Management

Measuring Portfolio Value at Risk

Chao Xu[1] , Huigeng Chen[2]

Supervisor: Birger Nilsson

Department of Economics

School of Economics and Management, Lund University

May 2012

[1]saintlyjinn@hotmail.com
[2]chenhuigeng@gmail.com

Abstract

On estimating portfolio Value at Risk, the application of traditional univariate VaR models is limited. Under specific circumstance, the VaR estimation could be inadequate. Facing the financial crises and increasing uncertainty in financial markets, effective multivariate VaR models have become crucial. This paper gives an overview of various multivariate VaR models. The main aim is to compare the one day out-of-sample predictive performances of different models, including basic multivariate VaR models, volatility weighted multivariate VaR models and copula-based multivariate VaR models. Performance is evaluated in terms of Christoffersen test, quadratic probability score and root mean squared error. The findings show that basic multivariate VaR models such as multivariate normal VaR model and multivariate t VaR model behave poorly and fail to generate reliable VaR estimations. By contrast, volatility weighted multivariate VaR models and copula-based multivariate VaR models show notable improvements in the predictive performance.

Keywords: Multivariate Value at Risk, portfolio risk measures, Copula, Monte Carlo simulation, DCC-GARCH, multivariate EWMA, Christoffersen test, quadratic probability score, root mean squared error, R software.

Acknowledgements

We wish to thank our supervisor, Professor Mr. *Birger Nilsson* for his useful comments, suggestions and valuable guidance.

Contents

1 Introduction

Value at Risk (VaR) is a widely used measurement of financial risk and plays a decisive role in risk management. In recent years, globalization of financial markets, financial integration and more complex derivatives have caused a more volatile environment. Firms and investors are exposed to more financial risks than before. A better and more liable risk management is demanded as the enlargement of financial risks. Although VaR is a simple measurement and easy to be interpreted, it is not easy to be estimated. The estimation of VaR is sensitive to the model assumption. Any deviations from the assumption would lead to an inadequate estimation. Facing the financial crises and increasing uncertainty in financial markets, effective measures of market risks have become crucial.

Traditional studies of VaR focus their attention on the univariate approaches. Univariate VaR models are easily constructed, but ignore the time varying covariance or correlation between financial assets. Assuming constant time-varying volatility may lead to an inadequate estimation of VaR in the long-run if changes in the dependence structure are not taken into account. Moreover, in some circumstances, univariate approaches are inapplicable as some of the portfolio returns are not observable. Furthermore, estimating portfolio VaR simply by aggregating the VaR of each portfolio component can be problematic. On one hand, due to the diversification effects, the portfolio VaR can be smaller than the sum of the portfolio components' VaR; on the other hand, the portfolio VaR can be larger than the sum of the portfolio components' VaR, as a result of non-subadditive property of VaR (In specific cases, $VaR_{A+B} > VaR_A + VaR_B$, the diversification effects are ignored). Both of them can lead to an inadequate result from the regulatory purposes or users' perspective.

Compared with the univariate approaches of VaR, the multivariate approaches of VaR are far from well-developed. Up to now, there are several multivariate approaches for estimating VaR, such as the variance-covariance approach, historical simulation and the Monte Carlo method. But most of them are developed directly from the univariate approaches and work with unrealistic and inadequate assumptions. In addition, newly developed statistical tools such as the advanced volatility model and the advanced kernel density estimation method are seldom applied to the estimation of VaR. The theory of multivariate VaR models is still not mature and faces many problem when they are applied. For example, the variance-covariance approach or analytical approach assumes a multivariate normal distribution of portfolio returns and estimation is made based on the expected return and sample standard deviation: It is widely used after the publishing of Riskmetrics^{TM} technology. However, the multivariate normality is rarely an adequate assumption in finance. Sheikh and Qiao (2010) found evidence that in many cases, financial returns were not independent and not normally distributed. If one financial model incorporates non-normality, standard deviation would become an ineffective measurement of risk. In this case, the portfolio could be riskier than desired.

In this paper, we discuss various approaches of estimating multivariate VaR and propose a copula-based Monte Carlo approach. In order to model VaR adequately, some recent advanced techniques are employed. The performance of both multivariate VaR models are evaluated by an application on the portfolio that consists of S&P 500 stock index and Hang Seng stock index (HSI). This paper contributes to the literature on multivariate approach VaR models, giving a detailed summary of various multivariate models and offering several ways on computational realizations. In addition, it expands the application of statistical software R to the area of multivariate VaR models.

In summary, our research on the multivariate VaR models is trying to answer four research questions,

- What is the performance of basic multivariate VaR models?
  (Including historical simulation, parametric VaR model based on multivariate normal/t distribution, age-weighted historical simulation)
- How to construct multivariate volatility models? How is the accuracy?
- Do volatility adjusted multivariate VaR models have better predictive performances than the basic multivariate VaR models?
- Do VaR models based on copula theory and Monte Carlo simulation method have better predictive performances than the basic multivariate VaR models?

The conclusive evidence of this study indicates that the basic multivariate VaR models do not perform well in estimating future losses. Most of them estimates VaR inadequately, which leads to an unacceptable number of violations in the test period and a failure in passing the christoffersen test. By contrast, both volatility adjusted multivariate VaR models and copula-based multivariate VaR models perform well in VaR estimation. Both of them show notable improvements on the predictive performance than the basic multivariate VaR models.

The paper is organized as follow. The first section introduces the theoretical background of various multivariate VaR models which will be followed by a copula-based Monte Carlo VaR model. Section 3 gives a description of the data as well as basic analysis of the data. The methodology is presented in section 4. The time-varying volatilities are modelled in section 5. In section 6, the empirical results of both models are presented. Finally, section 7 concludes our study.

The time frame is limited for this study, and quite understandably, it is difficult to cover all aspects of multivariate VaR models. Detailed analysis on this topic would require extensive research; therefore, several aspects of this paper have to be delimited.

- The methodology discussed in the theory part can be applied to the multivariate case when the portfolio is consist of more than two financial assets. But for simplicity, we only focus on the bivariate case and choose a portfolio that consists of two assets as an illustration.
- There are a considerable number of VaR models or assumptions on distribution that are available; however we are limited to the 22 models we are using.

- Multivariate DCC-GARCH with leverage effects and conditional copula methods are not employed. We believe they can significantly improve the estimation results, but they are rather time-consuming and computationally intensive. Due to the restriction on the time-horizon of this study, we have to abandon them. However, they are available for future study and they can be easily realized by extending the models discussed in this paper(original DCC-GARCH model and unconditional copula theory).

2 Theory

In this section, the theoretical background of multivariate VaR models is presented. It starts with the definition of univariate and multivariate Value at Risk. In addition, the advantages and the shortcomings of VaR models are discussed. Next, four different types of multivariate VaR models are introduced, including non-parametric VaR models, VaR model under multivariate normal distribution, VaR model under multivariate t distribution and copula-based multivariate VaR models. Furthermore, the multivariate volatility models are introduced as an improvement on the basic multivariate VaR models. In the end, the backtesting and evaluation methodologies are presented.

2.1 Value at Risk

2.1.1 Definition

In 1994, J.P morgan published a risk control methodology known as Riskmetrics^{TM}, which was mainly based on a newly developed financial risk measurement named Value at Risk. It was regarded as a masterpiece in financial risk management, and soon became popular. Over the last few years, VaR has become a key component in the management of market risk for many financial institutions. It is used as an internal risk management tool, as well as chosen by the Basel Committee as the international standard for regulatory purposes.

Given confidence level $\alpha \in (0, 1)$ and holding period (H), the Value at Risk of a portfolio is defined as the smallest number l, such that the probability of a future portfolio loss L exceeds l is no larger than $1-\alpha$. It measures the risk of future losses from a specific financial assets for a certain holding period. In probabilistic terms, VaR is simply a quantile of the loss distribution (McNeil *et al*, 2002). Formally,

$$VaR_\alpha(L) = \inf\{l \in \mathbb{R} : Pr(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\}$$

In the equation, inf is short for *infimum* and $\inf(S)$ represents the greatest lower bound of a subset S, i.e. the biggest number that is smaller than or equal to every number in S.

2.1.2 Parameters

VaR involves two arbitrarily chosen parameters, confidence level (α) and holding period (H). The confidence level indicates the probability that we will get a future outcome no worse than estimated VaR. Holding period determines the length of interval within which the loss is calculated.

Dowd (2005) shows VaR is contingent on the choice of confidence level and is non-decreasing with the confidence level. VaR cannot fall when the confidence level rises. In choosing confidence levels, investors or managers should consider "the worst-case loss amounts that are large enough to be material" (Laubsch 1999, p.10). On the contrary, for a capital adequacy purpose, a relatively high confidence level is recommended. Basel Committee recommends the 99% confidence level (Basel Committee, 1996). Higher confidence level would benefit when faced with a unexpected high market risk. However, choosing an unnecessary high level of confidence, such as 99.9%, would lead to a false sense of risk management as the losses will rarely exceed that level. Moreover, due to fat-tailed distribution of market returns, it is difficult to select a proper theoretical probability distribution and a high confidence level VaR is both time consuming and costly to be correctly modelled (Laubsch, 1999). As a result, lower confidence levels are often used for internal management purpose. For example, J.P Morgan uses a 95% confidence level, Citibank uses a level of 95.4% (Dowd, 1998). Furthermore, confidence level also varies with the different risk attitudes of managers. A risk averse and conservative manager would prefer a higher confidence level.

In practice, holding periods (H) are usually defined as one day or one month. But VaR can also operate on other length of holding period, depend on investment horizons of the investors or managers. For model validation or backtesting purposes, a short holding period is preferable. Reliable validation requires a large dataset and thus requires a short holding period.

2.1.3 Attractions and criticism of Value at Risk

The reasons behind the popularity of VaR can be concluded into three main attractions. The primary reason is, it provides a common consistent measurement of risk across different positions and risk factors. As a result, VaR can be applied to all asset classes (stocks, bonds, derivatives etc.). In addition, VaR makes it possible to measure the risk in both portfolio components level and overall level, which enables managers to take a detailed measurement of portfolio risks. Finally, VaR is conceptual simplicity and its results are easy to be interpreted.

From among the critics, Einhorn and Brown (2008) argue that VaR focus on the manageable risks near the center of the distribution, but ignore the tails. Taleb (1997) claims that VaR is impossible to estimate the risks of rare events. As a result, VaR could be destabilizing during a crisis. Another criticism of VaR is its non-coherence due to its non-subadditive property. In specific conditions, VaR increases when financial assets are aggregated into portfolio. VaR does not always encourage diversification. It is seen as the most serious drawback of VaR as a risk measurement.

2.2 Multivariate Value at Risk

The portfolio Value at Risk can be seen as a combination of the multivariate Value at Risk of portfolio components. In this part, we discuss the definition and features of multivariate Value at Risk, as well as its implication on the portfolio Value at Risk. From the definition in the univariate VaR model, we know the VaR is provided by a quantile function $Q_X(\alpha)$ which accumulates a probability α

to the left tail or $1-\alpha$ to the right tail. The definition of multivariate VaR is similar. Embrechts and Puccetti (2006), Nappo and Spizichino (2009) propose to define an intuitive and immediate generalization of the VaR models in the case of a d-dimensional loss distribution. According to their researches, multivariate VaR is denoted as the α quantile curves of the d-dimensional loss distribution.

$$VaR^i_\alpha(\mathbf{X}) = \mathbb{E}[X_i|F(\mathbf{X}) = \alpha]$$

Cousin and Bernardino (2011) point out some characters of multivariate VaR. Before presenting their results, the definition of regularity condition has to be introduced: A random vector satisfies regularity conditions, when the vector is non-negative absolutely-continuous and with partially increasing multivariate distribution function F.

With the definition, considering a random vector $\mathbf{X}$ satisfying the regularity conditions and assuming its multivariate distribution function F is a quasi concave (the upper level sets of function F are convex sets), for all $\alpha \in (0, 1)$, the estimation of multivariate VaR is always greater than or equal to the estimation of univariate VaR,

$$VaR^i_\alpha(\mathbf{X}) \geq VaR_\alpha(X_i)$$

According to the results, multivariate $VaR^i_\alpha(\mathbf{X})$ is a more conservative measurement than the vector consists of the univariate VaR ($VaR_\alpha(X_i)$). As a result, the portfolio VaR estimated with multivariate VaR model is more conservative than the VaR estimations from univariate VaR models. From an empirical point of view, multivariate VaR takes the correlation between asset returns into account. Compared with univariate VaR, more information and more risk factors are considered in the estimation.

2.3 Approaches of multivariate Value at Risk

Traditional univariate VaR models focus on a financial asset or portfolio individually. Portfolio losses are assumed to be observable. However, we can not always observe portfolio return directly in the practice. In order to study a generalized portfolio VaR, we have to use multivariate approaches of VaR which explicitly model the correlation structure or covariance structure between portfolio components. Similar with univariate VaR models, there exists a vast number of ways of multivariate VaR calculation which differ in their assumptions and have their own advantages and disadvantages. In this paper, we review major approaches of multivariate VaR estimation and we believe that addressing the problem of comparison of various VaR would offer useful information for VaR users.

2.3.1 Multivariate historical simulation

The most widely used non-parametric approach of multivariate VaR models is the multivariate historical simulation (multivariate HS). Under this approach, the estimation of VaR is based on the empirical loss distribution. All informations about the distribution of future returns are assumed to be reflected by the empirical loss distribution. This assumption enables forecasting future VaR directly from the historical observation of portfolio returns, instead of estimating the loss distribution under some explicit statistical models. The multivariate version of historical simulation is similar with the univariate basic historical simulation. But before doing the procedures of historical simulation, the assets returns are transformed into portfolio returns.

$$R_p = wR_a$$

where R_p denotes the composed portfolio returns, w denotes the weights of financial assets in the portfolio, R_a denotes the vector of historical returns of portfolio components.

Afterwards, the VaR of the next day (VaR_{t+1}) is estimated by the $1-\alpha$ quantile ($Q_{1-\alpha}$) of historical distribution of portfolio returns R_p, multiplied by the current value of the portfolio ($\bar{P}$).

$$VaR_{t+1} = -Q_{1-\alpha}(R_p(t), R_p(t-1), \cdots, R_p(1))\bar{P} \tag{1}$$

Taking a sliding windows of 1000 observations as the illustration, $VaR_{0.99}$ at (t+1) is simply the negative of the 10^{th} (1000×0.01) lowest portfolio return in the sorted observations multiply by current value of the portfolio $\bar{P}$.

The historical simulation method has obvious attractions: it is easy to implement and does not depend on certain assumptions of loss distribution. It is an appealing feature among the risk measurements on portfolio level. As in some circumstances, it is not possible to model the dependence structure between portfolio components and the joint probability distribution is hard to be constructed. In that case, multivariate historical simulation method is the only choice of risk measurement.

However the success of this approach is highly dependent on the user's ability to collect sufficient quantities of relevant, synchronized data for all risk factors (McNeil, 2005). An insufficient dataset would lead to the destabilizing of the empirical loss distribution. Furthermore, historical simulation approaches of VaR models suffer from the so-called ghost effect. Namely, when a large loss observation falling out of the sample, there would be a jump in the estimated VaR. Hence, multivariate historical simulation could perform well only if there are no gaps in the volatility of portfolio returns overtime.

2.3.2 Age-weighted multivariate historical simulation

In order to reduce the ghost effects of basic historical simulation approach, Boudoukh *et al* (1998) suggested weighting the observations according to their age, instead of giving equal weights $1/N$ for all historical observations. Accordingly, observations farther away from today are given lower weights, while latest observations are given higher weights. In practice, the weights are often defined as exponentially decreasing, with the form:

$$\begin{aligned} w_1 &= \frac{1-\lambda}{1-\lambda^n} \\ w_2 &= \lambda w_1 \\ &\cdots \\ w_N &= \lambda^{N-1} w_1 \\ \sum_{i=1}^{N} w_i &= 1 \end{aligned}$$

where w_i represents the adjusted weights according to the 'age' of the observed returns. Constant λ lies between 0 and 1, a λ close to zero will make older observations irrelevant quickly and a λ close to one will transform the age weighted simulation into the equally weighted basic historical simulation. In our research, λ is set to 0.94, which is consistent with major previous researches on AWHS.

Dowd (2005) gives a summary of improvement of age weighted historical simulation against basic historical simulation,

- It provides a generalisation of basic historical simulation models. Basic historical simulation can be regarded as a special case with zero decay ($\lambda = 1$).
- A suitable choice of λ can make the VaR estimates more responsive to large loss observations. It also makes this approach better at handling clusters of large losses (Volatility clustering).
- Age-weighting reduces the so-called ghost effects.

2.3.3 VaR under multivariate normal distribution

VaR under multivariate normal distribution is the most widely used parametric approach of multivariate VaR models. This approach assumes the returns of portfolio components are multivariate normally distributed with mean vector μ and covariance matrix Σ,

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} \quad and \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix}$$

The mean vector μ captures the average level of returns, while the covariance matrix Σ captures the interactions between the returns to different assets. Additionally, the current value of the portfolio is defined as $\bar{P}$. Given the weights of the portfolio components $w = (w_1, w_2, \cdots, w_n)$, the portfolio expected return (μ_p)

and the portfolio return variance σ_p^2 are given by,

$$\mu_p = w\mu$$

$$\sigma_p^2 = w\Sigma w'$$

Then, VaR under the assumption of multivariate normal distribution returns can be estimated by equation (2).

$$VaR_\alpha(L) = \bar{P}(-\mu_p - \sigma_p z_{1-\alpha}) \tag{2}$$

In the equation, the mean vector and the covariance matrix are usually unknown and we have to explicitly model them based on the actual observations.

The simplest way to estimate the mean vector and the covariance matrix is using the sample mean vector $\hat{\mu}$ and sample covariance matrix $\hat{\Sigma}$ directly. Denotes N vectors of portfolio component's return as $r_1, \cdots, r_N$,

$$\hat{\mu} = \frac{1}{N-1}\sum_{i=1}^{N} r_i$$

$$\hat{\Sigma} = \frac{1}{N-1}(r_i - \hat{\mu})(r_i - \hat{\mu})'$$

An alternative method for estimating parameters μ and Σ of multivariate normal distribution is the well-known maximum likelihood estimation (MLE). The method of maximum likelihood is widely used in statistical inference to estimate parameters. Maximum likelihood estimation begins with the mathematical expression known as a likelihood function of the sample data. Recall the probability density function of a d-dimensional multivariate normal distribution $N(\mu, \Sigma)$,

$$p(x|\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} e^{-\frac{(x-\mu)'\Sigma(x-\mu)}{2}}$$

Given observed returns $(r_1, r_2, \cdots, r_n)$, the log-likelihood function is defined as,

$$l(\mu, \Sigma|(r_1, r_2, \cdots, r_n)) = -\frac{Nd}{2}log(2\pi) - \frac{N}{2}log(\det\Sigma) - \frac{1}{2}\sum_{i=1}^{N}(r_i - \mu)'\Sigma^{-1}(r_i - \mu)$$

Parameters of the multivariate normal distribution can be estimated by maximizing the log-likelihood function $l(\mu, \Sigma|(r_1, r_2, \cdots, r_n))$. The maximizing process with multiple variables is a bit complex. One widely used numerical optimization algorithm is L-BFGS-B algorithm. It is a class of hill-climbing numerical optimization techniques that seeks a stationary point of a function. As the aim of the L-BFGS-B is to minimize the objective function, the log-likelihood function should be multiplied by (-1) to make the algorithm applicable, when applied to parameters estimation of multivariate normal distribution. Hence, the target function can be defined as,

$$\begin{aligned} f(\mu, \Sigma) &= -l(\mu, \Sigma|(r_1, r_2, \cdots, r_n)) \\ &= \frac{Nd}{2} log(2\pi) + \frac{N}{2} log(\det \Sigma) + \frac{1}{2}\sum_{i=1}^{N}(r_i - \mu)'\Sigma^{-1}(r_i - \mu) \end{aligned}$$

The L-BFGS-B algorithm proceeds roughly as follow. Before the approximation, a starting point is chosen. At each iteration, the Cauchy point is first computed by algorithm CP. Then a search direction is computed by either the direct primal method, or the conjugate gradient method. Afterwards, a line search is performed along the search direction, subject to the bounds on the problem. The optimum point is find after several repeating of the process above. (For a detailed algorithm, see Byrd *et al*, 1995, p.17)

2.3.4 VaR under multivariate t-distribution

Empirical studies show that financial returns do not follow the normal distribution. An estimation under the multivariate normality can be inadequate. As a result, multivariate student's t-distribution is introduced into VaR modelling to dealing with fat-tailed and leptokurtic features of portfolio returns.

Similar with the multivariate normal distribution approach, denote $\bar{P}$ as the current price of the portfolio. VaR under the multivariate t-distribution is given by the equation (3).

$$VaR_\alpha(L) = \bar{P}[-\mu_p - \sqrt{\frac{v-2}{v}}\sigma_p t_{1-\alpha,v}] \tag{3}$$

where

$$\mu_p = w\mu$$

$$\sigma_p = \sqrt{w\Sigma w'}$$

In the equation, portfolio return matrix (μ), portfolio covariance matrix (Σ) and degree of freedom (v) is unknown and needed to be estimated. Exactly as the approaches of multivariate normal distribution, we have to estimate parameters of the multivariate t distribution.

Aeschliman *et al* (2010) developed a Batch approximation algorithm for estimating parameters of multivariate t distribution. At the expense of a slightly decreased accuracy, the proposed algorithm is significantly faster and easier to implement. The algorithm starts with the estimation of sample mean vector μ, simply by taking median of the portfolio returns.

$$\hat{\mu} = median(r_i)$$

With the estimated $\hat{\mu}$, we can get the degree of freedom $\hat{v}$ afterwards.

$$b = \frac{1}{n}\sum_{i=1}^{n}(log\|x_i - \hat{\mu}\|^2 - \frac{1}{n}\sum_{i=1}^{n} log\|x_i - \hat{\mu}\|^2)^2 - \psi_1(\frac{p}{2})$$

$$\hat{v} = \frac{1 + \sqrt{1 + 4b}}{b}$$

where $\psi_1(x)$ is the trigamma function ($\psi_1(x) = \frac{d^2}{dx^2} ln\Gamma(x) = \frac{d}{dx}\psi(x)$) and p is the number of portfolio components.

Afterwards, the covariance matrix Σ can be derived by

$$\Sigma = \frac{exp\{\frac{1}{n}\sum_{i=1}^{n} log\|x_i - \hat{\mu}\|^2 - log\hat{v} + \psi_0(\frac{\hat{v}}{2}) - \psi_0(\frac{p}{2})\}}{tr(\frac{1}{n}\sum_{i=1}^{n}\frac{(x_i-\hat{\mu})(x_i-\hat{\mu})'}{\|x_i-\hat{\mu}\|^{2log_2 p/(\hat{v}^2+2log_2 p)}})}\sum_{i=1}^{n}\frac{(x_i - \hat{\mu})(x_i - \hat{\mu})'}{\|x_i - \hat{\mu}\|^{2log_2 p/(\hat{v}^2+2log_2 p)}}$$

where $\psi_0(x)$ is the digamma function ($\psi_0(x) = \frac{d}{dx} ln\Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$)

An alternative for estimating the parameters of multivariate t-distribution is Maximum likelihood estimation (MLE). Recall the probability density function for a multivariate t-distribution with mean vector μ, covariance matrix Σ and degrees of freedom parameter v is,

$$f(x, \mu, \Sigma, v) = \frac{\Gamma[(v + n)/2]}{\sqrt{\det\Sigma[(v - 2)\pi]^{n/2}\Gamma(v/2)}}[1 + \frac{1}{v - 2}(x - \mu)'\Sigma^{-1}(x - \mu)]^{-(v+n)/2}$$

The corresponding target log-likelihood function can be derived as,

$$l(\mu, \Sigma, v|(x_1, x_2, \cdots, x_n)) = \sum_{i=1}^{n} log(f(x_i, \mu, \Sigma, v))$$

The vector μ, covariance matrix Σ and degrees of freedom v can be estimated by maximizing the log-likelihood function. The procedure of MLE is similar with the multivariate normal approach.

2.3.5 Monte Carlo simulation method

Monte Carlo simulation methods are by far the most flexible and powerful tools for estimating Value at Risk. They are able to take into account all non-linearities of the portfolio value with respect to its underlying risk factors. However, This method still has one potential weakness. Specific stochastic processes need to be selected before the simulation. As a result, this method is very sensitive to the selection of stochastic processes.

The basic idea of this approach is to simulate repeatedly from a stochastic processes which governing the returns of the financial assets. Dowd (2005) gives a general simulation process for Monte Carlo simulation.

1. Select a model for the stochastic variables of interest.
2. Construct fictitious or simulated paths for the stochastic variables
3. Repeat these simulations enough times to be confident that the simulation distribution is sufficiently close to the 'true' distribution of actual portfolio values to be a reliable proxy for it.
4. Infer the VaR from this proxy distribution

Consider a simple case that we have two financial assets. The return vector μ, covariance matrix Σ and portfolio weights w vector are assumed to be,

$$\mu = \begin{pmatrix} 0.1 \\ 1 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 1 & 0.4 \\ 0.4 & 1 \end{pmatrix} \qquad w = \begin{pmatrix} 0.5 & 0.5 \end{pmatrix}$$

The simulation procedure starts with defining the path to generate possible scenarios of portfolio return. For simplicity, the path is specified as the random number generated from the multivariate distribution with mean (μ) and covariance matrix (Σ). In each iteration, we get a simulated portfolio return. And after repeating the iteration 100000 times, we get the probability distribution of simulated portfolio returns (Figure 10, see Appendix A). $VaR_{0.99}$ can be inferred from the figure as the 99% quantile of the loss distribution.

2.4 Modelling Volatility

Both approaches discussed in section 2.3 are under the assumption of constant volatility overtime. Hence, recent changes in the volatility of financial assets are not taken into account. However, under constant volatility assumption, estimated VaR would not incorporate the observed volatility clustering of financial returns. And the model may fail in generating the adequate VaR estimations.

Hull and White (1998) suggest one possible solution to the historical simulation approach. The basic idea is to adjust the return to take account of recent changes in volatility. For example, in forecasting VaR for day T+1, we transform the historical return (r_t) into volatility weighted return (r_t^*) before performing historical simulation approach.

$$r_t^* = \frac{\sigma_{T+1}}{\sigma_t} r_t \qquad t = 1, 2, 3, \cdots, T$$

where σ_t denotes the volatility associated with the observed losses and σ_{T+1} denotes the forecast volatility (conditional volatility) based on the historical changes in volatility.

For parametric approaches (multivariate Normal/t VaR models), forecast volatility σ_{T+1} enters the VaR formula directly and replace the portfolio volatility σ_p associated with the observed losses. The task is thus to forecast conditional volatility σ_{T+1} for each day.

In the univariate volatility weighted VaR models, volatility σ_{T+1} is estimated by univariate GARCH model or univariate exponentially weighted moving average (EWMA) model. Similar with the univariate case, we use multivariate GARCH models or multivariate EWMA model to take the historical changes in volatility into account. In practice, there are numerous multivariate GARCH models can be chosen from, such as VEC model (Bollerslev *et al*, 1988) and BEKK model (Engle *et al*, 1995). In this paper, we just focus on (extended) dynamic conditional correlation GARCH model and Multivariate EWMA model.

2.4.1 Multivariate EWMA model

EWMA is developed on the basis of equally weighted moving average and captures the dynamic features of volatility. But different with the equally weighted estimator of volatility, the most recent observations of returns are assigned with higher weights. As a result, the volatility reacts faster to shocks in the market.

In practice, it is more reasonable to use EWMA and assume today's volatility is more affected by the most recent events. Previous research based on the EWMA volatility model shows its reliable performance in VaR estimation.

In the univariate EWMA volatility model, the estimator of conditional variance defines variance of next period σ_{t+1} as a weighted average of the current period's variance σ_t^2 and squared current deviations from the average loss ε_t^2.

$$\sigma_{t+1}^2 = \lambda\sigma_t^2 + (1-\lambda)\varepsilon_t^2$$

The equation can be expanded to the multivariate EWMA, with the definition of covariance matrix (Σ_t). The future covariance of portfolio components Σ_t can be estimated by today's changes in returns ϵ_t and covariance of portfolio components at t-1, Σ_{t-1}.

$$\Sigma_t = \lambda\Sigma_{t-1} + (1-\lambda)\epsilon_t\epsilon_t'$$

where λ is a fixed constant and with the range from 0 to 1. A lower λ makes older changes in volatility irrelevant quickly and vice versa. In this paper, we prefer to use $\lambda = 0.94$ which is consistent with the choice of Riskmetrics$^{\text{TM}}$.

2.4.2 Conditional Correlation GARCH Model

An alternative for multivariate EWMA model is conditional correlation GARCH model. It can be viewed as a non-linear combination of univariate GARCH models. And the model can be separated into two parts, GARCH models (conditional variance) and correlation matrices. In the model, any individual conditional variance can be specified separately and a conditional correlation matrix can be constructed to describe the dependence structure between the individual series.

Bollerslev (1990) proposes a constant conditional correlation GARCH (CCC-GARCH) model in which the conditional correlations are constant. The conditional covariances are proportional to the product of the corresponding conditional standard deviations. The CCC-GARCH model is defined as,

$$H_t = D_t R D_t$$

where D_t is the $k \times k$ diagonal matrix of time varying standard deviations from univariate GARCH models with $\sqrt{h_{ii}}$ on the ith diagonal, and R_t is the time varying correlation matrix.

$$D_t = diag(h_{11t}^{1/2} \cdots h_{NNt}^{1/2})$$

$$R_t = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{pmatrix}$$

h_{iit} in matrix D_t is the conditional variances and can be defined as any univariate GARCH model, taking GARCH(1,1) as example,

$$h_{iit} = w_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{ii,t-1}$$

However, in practice, the assumption that conditional correlations are constant overtime, is unrealistic. As an improvement on the CCC-GARCH model, Engel (2002), Christodoulakis and Satchell (2002) and Tse and Tsui (2002) propose a generalization of the CCC-GARCH model called dynamic conditional correlation GARCH (DCC-GARCH) model.

In this paper, we only focus on the Engel's DCC-GARCH model which define the conditional correlation matrix as time-varying. Mathematically,

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1}$$

where Q_t is the unconditional covariance of standardized residuals resulting from the univariate GARCH models and Q_t^* is a diagonal matrix consists of the square root of the diagonal elements of Q_t

$$Q_t^* = Diag(\sqrt{q_{11}}, \sqrt{q_{22}}, \cdots, \sqrt{q_{nn}})$$

Engel (2002) performed a comparison of several conditional covariance and showed that DCC-GARCH model was overall best in estimation. Despite its accurate estimation of future covariances, potential weaknesses still exist. One potential drawback of the DCC-GARCH model is that all conditional correlations follow the same dynamic structure is unrealistic in practice.

2.5 Copula-based Monte Carlo approach

2.5.1 Introduction

As discussed in section 2.3.5, Monte Carlo approaches of multivariate VaR estimation require the joint distributions of portfolio component returns to be known. In addition, the accuracy of Monte Carlo method is very sensitive to the assumption of joint distribution. A deviation from the actual distribution may lead to inadequate VaR estimations. Thus, the feasibility of the approach highly depends on an accurate modelling of joint distribution.

The copula theory was first developed in Sklar (1959). It is a very powerful tool for modelling joint distribution because it does not require any assumptions on the selection of distribution function and allows us to decompose any n-dimensional joint distribution into n marginal distributions and a copula function.

In this section, we take the advantage of copula theory and develop a copula-based Monte carlo approach. In consistent with the other parts of our research, only bivariate copula is introduced in this paper.

The study starts with a definition of the bivariate copula functions.

Definition 1 *A 2-dimensional copula is a function $C(u, v)$ defined in the domain $[0, 1] \times [0, 1]$ and with the range of $[0, 1]$, i.e. $[0, 1]^2 \to [0, 1]$. The copula function satisfied following properties,*

(1) Boundary condition

For all $u, v \in [0, 1]$,

$$C(u, 0) = C(0, v) = 0$$

$$C(u, 1) = u, C(1, v) = v$$

(2) Monotonic condition

For all $u_1, u_2, v_1, v_2 \in [0, 1]$, when $u_1 \leq u_2, v_1 \leq v_2$

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

With the definition of copula function, Sklar (1959) proposes the sklar's theorem that shows the importance and usefulness of copula function.

Theorem 1 *(Sklar's Theorem) Let $H(x, y) = P[X \leq x, Y \leq y]$ be a joint distribution function with marginal distribution $F(x) = P(X \leq x)$ and $G(y) = P(Y \leq y)$. Then there exists a copula function C: $[0, 1]^2 \to [0, 1]$ such that,*

$$H(x, y) = C(F(x), G(y))$$

If $F(x)$ and $G(y)$ are continuous, then copula function C is unique. Namely, if C is a copula and $F(x)$ and $G(y)$ are distribution functions, then the function $H(x,y)$ is a joint distribution function with margins $F(x)$ and $G(y)$.

The main implication of Sklar's theorem is that a joint distribution can be decomposed into two univariate marginal distributions $F(x)$ and $G(y)$. Conversely, we can link any group of two univariate distributions with a copula function and construct a valid joint distribution for the two variables. This implication offers an effective way for modelling joint distributions.

Despite its convenience of constructing joint probability distribution, bivariate copula function is also a measurement of dependence structure between two random variables. Each bivariate copula functions has its specific features of describing the dependence structure. Some of them focus on the linear correlations, while the others focus on the tail dependence/independence. As a result, VaR models with different assumptions on the copula functions are expected to have different results.

2.5.2 Some families of Copula

Five families of copula functions are introduced in this paper: Gaussian copula, Student's t-copula, Gumbel copula, Clayton copula and Frank copula. In addition, Gumbel copula, Clayton copula and Frank copula are also known as Archimedean class copulas. In this part, both the definitions of copula functions and their features are discussed.

Bivariate Gaussian Copula

The bivariate Gaussian copula is a dependence function associated with bivariate normality and is given by,

$$\begin{aligned} C^{Ga}(u,v) &= \Phi_\rho(\Phi^{-1}(u),\Phi^{-1}(v)) \\ &= \int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} exp(\frac{2\rho st - s^2 - t^2}{2(1-\rho^2)})dsdt \end{aligned}$$

where Φ^{-1} is the quantile function of the corresponding standard normal cumulative distribution function and $\Phi_\rho(x,y)$ is the standard bivariate normal distribution with correlation parameter ρ. Since it is parametrized by the correlation coefficient ρ, we can also write the bivariate Gaussian copula function as C_ρ^{Ga}.

In the bivariate Gaussian copula function, the dependence structure is described by the linear correlation coefficient ρ. As a result, the bivariate Gaussian copula gives an overall description of the dependence structure between the stochastic variables.

Figure 1 illustrates the joint density function constructed with bivariate normal copula and standard normal marginal distributions. The correlation coefficient $\rho = 0.5$

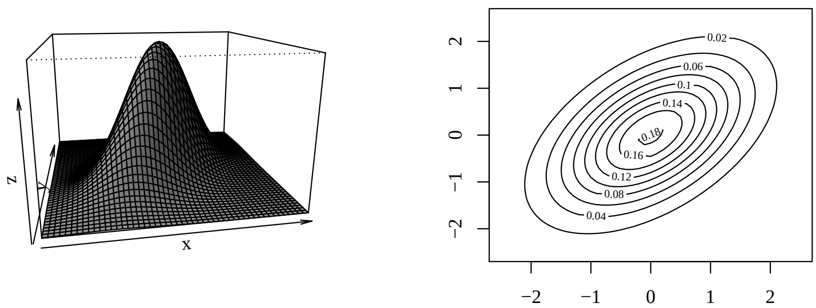

Figure 1: Density and level curves of the Gaussian Copula with $\rho = 0.5$

Bivariate Student's t-Copula

The student's t-copula function is defined as,

$$\begin{aligned} T_{\rho,v}(u,z) &= t_{\rho,v}(t_v^{-1}(u), t_v^{-1}(z)) \\ &= \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(z)} \frac{1}{2\pi\sqrt{1-\rho^2}} (1 + \frac{s^2 + t^2 - 2\rho st}{v(1-\rho^2)})^{-\frac{v+2}{2}} ds dt \end{aligned}$$

where ρ and v are the parameters of the copula, $t_v^{-1}(v)$ is the inverse of the standard student t-distribution with degrees of freedom v. The stronger correlation ρ and the lower the degree of freedom v, the stronger is the tail dependence. As a result, the student's t copula consider both the tail dependence and overall dependence in composing joint distributions.

Figure 2 shows the joint density function constructed with bivariate student's t-copula and standard normal marginal distributions. The correlation coefficient $\rho = 0.5$ and degree of freedom $df = 3$.

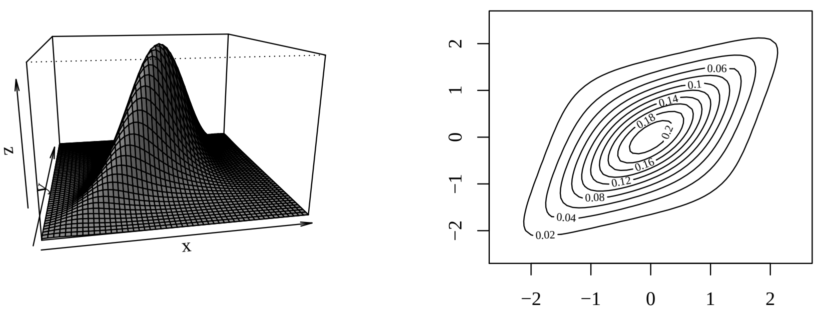

Figure 2: Density and level curves of the Student's t-Copula with $\rho = 0.5$ and $df = 3$

Archimedean Copulas

Archimedean copulas is an important class of copula functions that are easy to construct and have good analytical properties. Before introducing it, two important concepts have to be defined: *generator function* ϕ and *pseudo-inverse of generator function* $\phi^{[-1]}$.

Definition 2 *Function ϕ can be a generator function, if it satisfies,*

- $\phi : [0, \infty) \to [0, 1]$, $\phi(0) = 1$, $\lim_{x\to\infty} \phi(x) = 0$
- ϕ is continuous
- ϕ is strictly decreasing on $[0, \phi^{-1}(0)]$
- ϕ^{-1} is given by $\phi^{-1}(x) = \inf\{u : \phi(u) \leq x\}$

Definition 3 *The pseudo-inverse of generator function ϕ is defined as,*

$$\phi^{[-1]}(v) = \begin{cases} \phi^{-1}(v) & 0 \leq v \leq \phi(0) \\ 0 & \phi(0) \leq v \leq +\infty \end{cases}$$

With the definitions above, an bivariate Archimedean copula function can be 'generated' by the generator function:

$$C^A(u, v) = \phi^{[-1]}(\phi(u) + \phi(v))$$

Numerous Archimedean copula functions can be generated, with different assumptions of generator functions. In this paper, we will present three widely used Archimedean family copula functions. Figure 3 shows the level curves of the probability density function of them with standard normal margins and $\alpha = 2$.

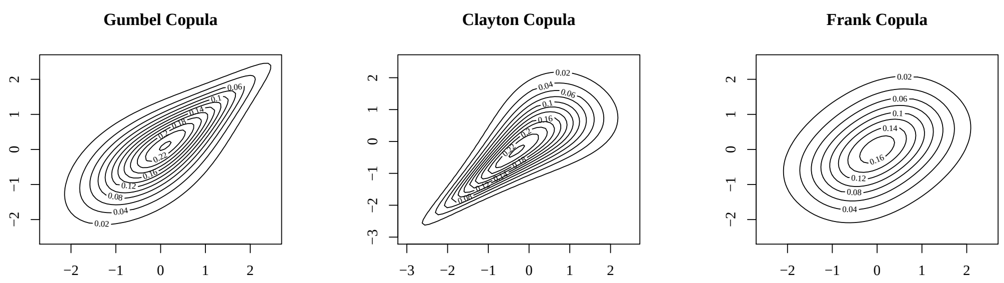

Figure 3: Level curves of the Archimedean Copula density with $\alpha = 2$

The first Archimedean copula employed is the Gumbel copula. It was first proposed by Gumbel (1960). The generator function is in the form of $\phi_\alpha(t) = (-ln(t))^\alpha$. The Gumbel copula is an asymmetric copula but exhibits greater

greater tail dependence in the upper tail (Figure 3, left). This copula function is give by,

$$C(u, v) = exp\{-[(-ln(u))^{\alpha} + (-ln(v))^{\alpha}]^{1/\alpha}\}$$

The parameter α determines the degree of dependency. Independence is obtained when $\alpha = 1$, while perfect dependence is obtained as $\alpha \to \infty$.

The second Archimedean copula used is the Clayton copula. It is also asymmetric but exhibits greater tail dependence in the lower tails (Figure 3, middle). It was first proposed by Clayton (1978). The generator function of Clayton copula $\phi_{\alpha}(t) = \frac{1}{\alpha}(t^{-\alpha} - 1)$. And the copula function is,

$$C(u, v) = max[(u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha}), 0]$$

When $\alpha \to \infty$, perfect tail dependence is obtained. When $\alpha \to 0$ implies tail independence.

The third Archimedean copula is Frank copula which is first introduced in Frank (1979). The generator function is $\phi_{\alpha}(t) = -ln\frac{exp(-\alpha t)-1}{exp(-\alpha)-1}$. Different with Gumbel/-Clayton copula, Frank copula exhibits tail independence (Figure 3, right). The copula function,

$$C(u, v) = -\frac{1}{\alpha}ln(1 + \frac{(exp(-\alpha u) - 1)(exp(-\alpha v) - 1)}{exp(-\alpha) - 1})$$

2.5.3 Marginal distribution of copula function

Marginal distribution plays an important role in copula theory. As the bivariate copula functions are defined in the space $[0, 1] \times [0, 1]$, the real observations cannot be substituted into the copula function directly. The marginal distributions can work as a proxy between copula function and the real observations. In a portfolio case, the marginal distribution is simply the probability distribution (CDF) of the portfolio components.

Theoretically, copula method do not restrict the choice of marginal distribution and it works with any assumption of marginal distribution. In previous researches, normal distribution, t distribution and generalized pareto distribution (GPD) are frequently used. In this paper, for simplicity and illustration purposes, we select the normal distribution as the marginal distribution.

2.5.4 Pseudo observations: An alternative for marginal distribution

In section 2.5.1-2.5.3, we discuss the definition of a traditional copula. In the traditional copula framework, a marginal distribution should be defined and the parameters of the marginal distribution have to be estimated before modelling copula function. However, the estimation procedure copula is computationally intensive and time consuming. The application of traditional copula on the Monte Carlo simulation would be limited, as a full the estimations of margins and copula function have to be performed in each iteration.

Yan (2007) proposed an alternative approach for constructing copula function. This approach uses the empirical cumulative probability distribution instead of marginal distribution. The original datasets $(X_{i1}, X_{i2}, \cdots, X_{in})$ are transformed into pseudo-observations $(U_{i1}, \cdots, U_{in})$.

$$U_{ij} = \frac{rank(X_{ij})}{n+1}$$

Thus, the copula function can be estimated based on the pseudo-observations instead of real data. There is no need to specify and estimate the marginal distribution of the copula function.

2.5.5 Estimation: Inference functions for margins method

In general, there are two approaches can be used for estimating copula parameters, including *one step maximum likelihood estimation* and *inference functions for margins (IFM) method.* In this paper, we choose the IFM method (Joe and Xu, 1996). It is less efficient than one-step maximum likelihood method, but it is computationally more attractive and allows larger flexibility in choosing the estimation techniques for the marginal distribution. The procedures of IFM method is presented in this part.

Suppose that we observe n independent observations $X_t = (x_{t1}, x_{t2}, \cdots, x_{tp})$ from an multivariate distribution, which can be constructed with p marginal distributions and a copula function $C(F_1(x), \cdots, F_n(x); \alpha)$ with parameter α. Furthermore, the probability density function (PDF) of the marginal distributions is defined as $f_i(x; \theta_i)$ and the corresponding cumulative density distribution (CDF) is denoted as $F_i(x; \theta_i)$, where θ_i is the parameter of marginal distributions. The IFM method estimates the parameters of marginal distribution in the first step. The log-likelihood function of the first step could be written as,

$$Logl(\theta) = \sum_{i=1}^{n} \sum_{j=1}^{p} log f_i(x_{ij}; \theta_i).$$

The estimation of the parameter $\theta = (\theta_1, \cdots, \theta_n)$ of marginal distributions can be made through maximizing the log-likelihood function.

$$\hat{\theta}_i = \arg\max \sum_{i=1}^{n} \sum_{j=1}^{p} log f_i(x_{ij}; \theta_i)$$

Then the parameter α of the copula function is estimated in the second step of

IFM, with the parameter $\hat{\theta}$ of the p marginal distributions.

$$\hat{\alpha} = \arg\max \sum_{t=1}^{n} logC(F_1(x_{i1}; \hat{\theta}_1), \cdots, F_p(x_{ip}; \hat{\theta}_p); \alpha)$$

2.5.6 Copula-based Monte Carlo approach

Based on the Monte Carlo simulation method and the theory of copula discussed in this section, we propose a detailed procedure of copula-based Monte Carlo approach of estimating portfolio VaR,

1. Select a class of Copula model (Gaussian/Student's t/Archimedean etc.) according to their different features.
2. Select a marginal distribution for each portfolio component and estimate the parameters of the marginal distribution.
3. Transform the original data into the domain of copula function by using each margin's distribution function $F_x(x)$.
4. Fit the copula model to the stochastic variables and estimate the parameters of the copula function.
5. Use the estimated copula function to generate random variables from the estimated joint probability density.
6. Invert the generated random variables by using the quantile function of the marginal probability function.
7. Calculate the portfolio loss/profit based on the simulated variables.
8. Repeat these simulation enough times to be confident that the simulation distribution is sufficiently close to the 'true' distribution.
9. Infer the VaR from the distribution of the simulated portfolio returns

2.6 Evaluation Methods

2.6.1 Christoffersen frequency test

Christoffersen frequency test is a standard tool that evaluates the performance of VaR models individually. It aims at examining whether the observed frequency of violations satisfy the unconditional coverage property and the independent property (Christoffersen, 1998). If a VaR model is adequate, the frequency of violations of the estimated VaR should be consistence with the expected frequency of tail losses and violations are independent and identical distributed.

The Christoffersen frequency test is constructed following its aims. It consists of two individual tests and an overall conditional coverage test.

- Unconditional coverage test (or Kupiec test)
- Independence test
- Conditional coverage (overall) test

Unconditional coverage test

The unconditional coverage test examines unconditional coverage property of VaR estimates. The null hypothesis for this test is,

$$H_0\text{: The probability of occurrence of a violation is } p$$

Denote the number of observations in the test period by N, the expected frequency of violations by the p and the observed frequency of losses exceeds VaR by $\pi = x/N$. The test statistics,

$$LR_{uc} = -2[ln(p^x(1-p^{N-x})) - ln(\pi^x(1-\pi)^{N-x})] \sim \chi^2(1)$$

Under the 95% confidence level, when $LR_{uc} > LR_{critical} = 3.841$, the null hypothesis is rejected. It indicates the VaR model fails to generate the adequate VaR estimations.

Independence test

The independence of frequency test was first proposed in Christoffersen (1998). It examines if the probability of a violation at time t given a violation occurred at time $t-1$ is equal to the probability of a violation at time t given no violation occurred at time $t-1$. The null hypothesis and alternative hypothesis of this test,

$$H_0\text{: VaR non-violations and violations are independent over time}$$

$$H_1\text{: VaR non-violations and violations follow a two state Markov chain.}$$

Assume that the violations and non-violations follows a Markov chain with transition matrix,

$$\Pi = \begin{pmatrix} \pi_{00} & \pi_{10} \\ \pi_{01} & \pi_{11} \end{pmatrix}$$

Where state 1 represents violation, state 0 represents violation. Denote $n_0, n_1, n_{00}, n_{01}, n_{10}, n_{11}$ as the number of the states or transitions of Markov stochastic process. Then,

$$\pi_{00} = \frac{n_{00}}{n_{00}+n_{01}}, \quad \pi_{01} = \frac{n_{01}}{n_{00}+n_{01}}$$

$$\pi_{10} = \frac{n_{10}}{n_{10}+n_{11}}, \quad \pi_{11} = \frac{n_{11}}{n_{10}+n_{11}}$$

And Define $\pi_0 = n_0/N$, $\pi_1 = n_1/N$. The log-likelihood ratio test statistic,

$$LR_{ind} = -2[ln(\pi_0^{n_0}\pi_1^{n_1}) - ln(\pi_{00}^{n_{00}}\pi_{01}^{n_{01}}\pi_{10}^{n_{10}}\pi_{11}^{n_{11}})] \sim \chi^2(1)$$

Similarly, under the 95% confidence level, if $LR_{ind} > LR_{critical} = 3.841$, the null hypothesis is rejected and indicates non-violation and violation is not independent over time. Hence, the model does not pass the independence test.

Conditional coverage test

It is an overall test of the unconditional coverage and independence test. The test statistic is the sum of the test statistic for unconditional coverage and independence test.

$$LR_{cc} = LR_{uc} + LR_{ind} \sim \chi^2(2)$$

Under the 95% confidence level, the $LR_{critical}$ for the conditional coverage test is 5.991. Namely, when $LR_{cc} \leq 5.991$, the VaR model passes the test.

2.6.2 Ranking alternative VaR models

It is often the case that management and investors are not only interested in the performance of an individual VaR model, but also in the comparison of different VaR models. Previous researches on the evaluation of VaR models already develops several effective ranking methods such as quadratic probability score function (Lopez, 1998), quadratic score function (Blanco and Ihle, 1999). Both of them offers possible measurements of relative performance of VaR models.

In evaluating the relative performance of different VaR models, two conflicting objectives are often taken into account. On one hand, we expect the estimated VaR to be high and as a result, the difference between VaR and actual loss would be low at violation days. It is because if a violation occurs and the reserved capital is too small to cover the losses, the firm would face financial distress or even go bankruptcy. On the other hand, we expect the estimated VaR to be low. It is because a high VaR means high capital reserves for the potential loss. But as the capital is costly, a firm or an investor want a low amount of reserve.

In this part, we discuss the quadratic probability score (QPS) function as a measurement of the first objective of evaluating relative performance and root mean squared error (RMSE) as a measurement of the second objective.

Quadratic probability score function

Lopez (1998) introduces the quadratic probability score function as a measurement of relative performance of VaR models. It is defined as,

$$QPS = \frac{2}{n}\sum_{t=1}^{n}(C_t - p)^2$$

where n is the number of observations, p is the expected probability of violation, i.e. the actual loss is larger than estimated VaR. C_t is a predetermined loss function which reflects the interest of users. In this paper, we use the binary loss function proposed by Lopez (1998). This loss function is intended for the user who is concerned with the frequency of violations.

$$C_t = \begin{cases} 1 & L_t > VaR_t \\ 0 & L_t \leq VaR_t \end{cases}$$

The QPS takes a value in the range $[0, 2]$, and under general conditions, accurate VaR estimates will generate the lowest possible numeric score (Lopez, 1998). Namely, smaller QPS indicates better performance in the violation-days.

Root mean squared error

Root mean squared error, or RMSE is a common measurement of the difference between the estimated value and the true value. Denote the estimated VaR as VaR_t and the actual losses as L_t, the definition of RMSE is,

$$RMSE = \sqrt{E[(VaR_t - L_t)^2]} = \sqrt{\frac{1}{n}\sum(VaR_t - L_t)^2}$$

In this paper, we employ root mean square error as a measurement of excess reserved capital during non-violation days. Hence, the t in the above equation represents the non-violation days in the test period. If estimated VaR_t has a smaller RMSE, the corresponding VaR model is considered as the better one.

3 Data

3.1 Data description

The theories presented in last section are applied to a portfolio composed by S&P 500 index and Hang Seng Index (HSI). The dataset contains 2974 daily closing prices from January 3rd, 2000 to March 29th, 2012. The daily closing prices are presented in Figure 11 (see Appendix A). In order to apply the multivariate VaR models, the original indexes are transformed into log-returns. We denote the log-returns of S&P 500 index as variable 1, the log-returns of Hang Seng Index (HSI) as variable 2. Figure 4 presents the daily log-returns of both series. In the figure, we can observe the evidence of stylized fact known as volatility clustering. Large returns follow with large returns, and similar for small returns.

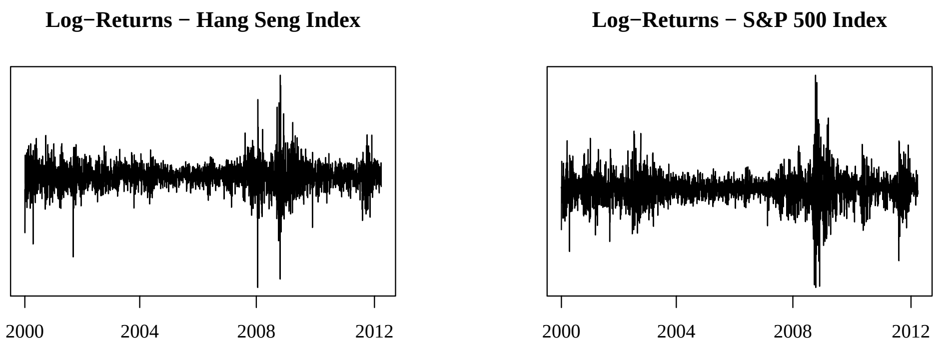

Figure 4: Daily log-returns of HSI and S&P 500 Index

3.2 Setting of rolling window periodic sampling

In order to analyse the performance of various multivariate VaR models, we employ the method of rolling window with sample size 2600, i.e. for each VaR estimation, we use the 2600 observations ahead of it. Figure 5 illustrates the rolling window vividly.

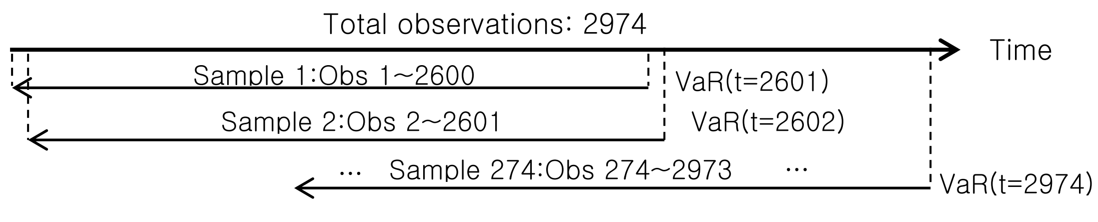

Figure 5: Rolling window periodic sampling

The whole dataset is divided into two parts: in-sample period and test period. The in-sample period starts on January 3rd, 2000 and ends with September 20th,

2010. It consists of 2600 daily return of each stock index and offers the historical information needed for estimating VaR. The test-period starts on September 21th, 2010 and ends on March 29th, 2012. It is used for testing the performance of VaR models. The size of the test period is 374. VaR is estimated for each day in the test period, with the information offered by the 2600 observations ahead of it. The accuracy of different VaR model can be assessed by comparing the estimated VaR and the actual loss. Table 1 summaries the sample division.

Table 1: In-sample period and test period

Period	In-sample period	test period	Total
Date	3/1/2000-20/9/2010	21/9/2010-29/3/2012	
Number of observations	$N_1 = 2600$	$N_2 = 374$	N=2974

3.3 Static and dynamic analysis on probability distributions

In this part, we discuss the static and dynamic statistical features of both indexes' log-returns. The study begins with a descriptive statistics on the log-return series, which is shown in Table 2.

Table 2: Descriptive statistics of daily log-return of HSI and S&P 500 indexes

Statistics	Hang Seng Index	S&P 500 Index
Mean	5.753×10^{-5}	-1.222×10^{-5}
Min	-0.147	-0.095
Max	0.134	0.110
Kurtosis	11.749	9.981
Skewness	-0.253	-0.150
Jarque-Bera Test	9513.012	6048.956

The table shows that Hang Seng Index has a positive average daily return, while S&P 500 Index has a negative average daily return. Both series are nearly symmetric, but fat-tailed ($kurtosis > 3$). In addition, the Jarque-Bera normality test rejects its normality null hypothesis (critical value for Jargue-Bera test is 5.991, at 95% significance level), i.e. the returns of both indexes are not normally distributed.

Furthermore, a comparison of descriptive statistics is made between the sample period and test period (Table 3). The results indicate that the distributions of both index returns have large differences between sample-period and test-period. Estimating their VaR with multivariate normal distribution or multivariate t distribution assumptions can be problematic and result in inadequate VaR estimations.

Table 3: Descriptive statistics of daily log-return of HSI and S&P 500 indexes

Statistics	Hang Seng Index		S&P 500 Index	
	In-sample period	Test period	In-sample period	Test period
Mean	9.053×10^{-5}	-1.718×10^{-4}	-9.302×10^{-5}	5.492×10^{-4}
Min	-0.147	-0.058	-0.095	-0.069
Max	0.134	0.055	0.110	0.046
Kurtosis	12.116	5.158	10.211	7.216
Skewness	-0.251	-0.285	-0.105	-0.544

By employing multivariate kernel smoothing and kernel density estimation (KDE) techniques (Duong, 2007), we present the estimated density of the joint probability distribution (Figure 6). The figure on the right shows the shape of empirical probability distribution is asymmetry and very sharp. It indicates no evidence of multivariate normality, but shows evidence of excess kurtosis.

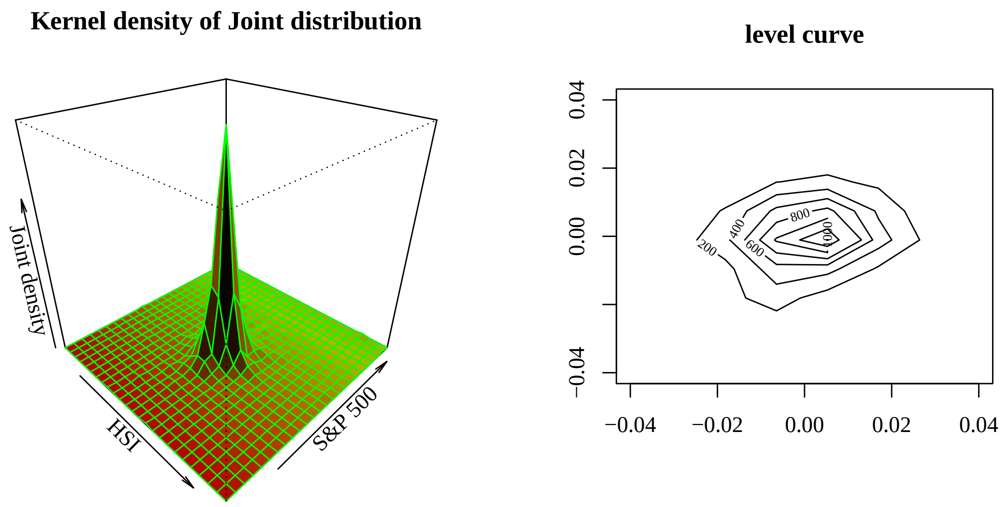

Figure 6: Joint kernel density and level curve of HSI and S&P 500 index

Afterwards, we analyse the dynamics of the sample distribution, by observing the rolling samples of both index. The results are illustrated in Figure 12 and Figure 13 (see Appendix A). Figure 12 indicates the sample distribution of Hang Seng index is unstable and varies with time. The volatility of Hang Seng index is decreasing along the test period, and the probability distribution tends to be more and more fat-tailed. Different with Hang Seng index, the continuously increasing average return in Figure 13 indicates the strong performance of S&P 500 index. Furthermore, the dynamics of sample volatility and kurtosis show no significant pattern. To conclude with, distributions of both index returns are unstable and the research based on them should pay more attention to their time-varying sample probability distribution.

4 Methodology: VaR models and notations

The accuracy of VaR models depends heavily on the model settings. For an adequate estimation of VaR, the characteristics of financial data must be taken into account. Brooks (2008) exhibit a number of interesting statistical property of financial time series which are common to a wide range of markets and time periods. In this paper, we focus on leptokurtosis and volatility clustering.

- *Leptokurtosis.* The distribution of financial returns displays a heavy tail with excess kurtosis ($kurtosis > 3$).
- *Volatility clustering* Large returns are expected to follow large returns, small returns to follow small returns.

Regarding the characteristics of financial data, some techniques have been developed. In previous studies on VaR estimation, the most common way to deal with leptokurtosis is by assuming a more proper probability distribution of financial returns. And volatility clustering effects are reduced by using time varying volatility instead of constant volatility.

In this paper, we discuss 22 different multivariate VaR models. According to different assumptions on the loss distributions, they can be separated into four groups: non-parametric approaches (based on empirical loss distribution), multivariate normal approaches (based on multivariate normal distribution), multivariate t approaches (based on multivariate t distribution) and copula approaches (based on joint distributions composed by copula). Table 4 summaries the models and gives their notations in our research.

There are four main highlights in this table:

1. The multivariate historical simulation approach (HS) is performed according to the theory discussed in section 2.3.1. The age-weighted multivariate historical simulation approach (AWHS) is performed with exponential decreasing age weighing assumption. Constant λ is set to be 0.94. As discussed in section 2.4, the volatility weighted historical simulation approaches are modelled with Hull and White transformation, $r_t^* = \frac{\sigma_{T+1}}{\sigma_t} r_t$ (Hull and White, 1998)
2. In multivariate Normal VaR model (mvn), VaR is estimated by the equation (2) in section 2.3.3. In multivariate t VaR model (mvt), VaR is estimated by equation (3) in section 2.3.4. The volatility adjusted models (DVW-mvn/EVW-mvn/DVW-mvt/EVW-mvt) is estimated by replacing the σ_p in equation (2) and (3) with the adjusted volatility (σ_{EWMA} or $\sigma_{DCC-GARCH}$).
3. The aim of Monte Carlo multivariate normal/t models is examining if there is a difference between the basic model (mvt or mvn) and Monte Carlo simulation model (MC-mvt or MC-mvn). Theoretically, there should be no difference between them. These models can be seen as benchmarks for assessing the effectiveness of copula theory.

Table 4: VaR models and notation

Model	Notation
Historical Simulation	HS
Age weighted Historical Simulation	AWHS
Volatility weighted (DCC-GARCH) Historical Simulation	DVWHS
Volatility weighted (EWMA) Historical Simulation	EVWHS
Multivariate Normal approach	mvn
Monte Carlo-multivariate normal	MC-mvn
Volatility adjusted (DCC-GARCH) Multivariate Normal	DVW-mvn
Volatility adjusted (EWMA) Multivariate Normal	EVW-mvn
Multivariate t approach	mvt
Monte Carlo-multivariate t	MC-mvt
Volatility adjusted (DCC-GARCH) Multivariate t	DVW-mvt
Volatility adjusted (EWMA) Multivariate t	EVW-mvt
Monte Carlo-Gaussian Copula(pseudo)	MC-GCp
Monte Carlo-Gaussian Copula(normal)	MC-GCn
Monte Carlo-Student's t-Copula(pseudo)	MC-tCp
Monte Carlo-Student's t-Copula(normal)	MC-tCn
Monte Carlo-Gumbel Copula(pseudo)	MC-GuCp
Monte Carlo-Gumbel Copula(normal)	MC-GuCn
Monte Carlo-Clayton Copula(pseudo)	MC-ClCp
Monte Carlo-Clayton Copula(normal)	MC-ClCn
Monte Carlo-Frank Copula(pseudo)	MC-FrCp
Monte Carlo-Frank Copula(normal)	MC-FrCn

4. Monte Carlo-Gaussian/t/Gumbel/Clayton/Frank copula represents the copula-based multivariate VaR model. The models are performed following the procedures proposed in section 2.5.6. Pseudo/normal in the parentheses shows the assumption of marginal distribution. 'Pseudo' denotes the copula is constructed on the pseudo observations and without specifying the marginal distribution. 'Normal' denotes the normal distribution is specified as the marginal distribution of copula function.

5 Modelling Volatility

The section starts with a focus on the time series properties of the S&P 500 Index and Hang Seng Index. Next, multivariate DCC-GARCH model and EWMA model are employed as an estimation of time-varying volatility. A discussion on the results of multivariate volatility model is presented. This section ends with a study on the dependence structure between S&P 500 Index and Hang Seng Index.

5.1 Time series properties

The construction of DCC-GARCH multivariate volatility model is based on a time-varying correlation matrix and conditional volatility of each stochastic variable. For an accurate estimation, we have to focus on the time series property of each stock index before the multivariate volatility modelling. Figure 7 and 8 show the autocorrelation function (ACF) and partial autocorrelation function (PACF) for the log-returns of Hang Seng Index and S&P 500 Index.

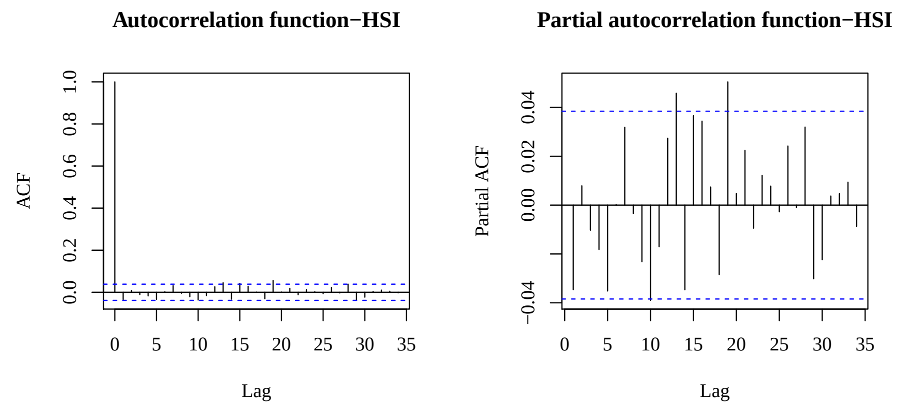

Figure 7: ACF and PACF of Hang Seng Index

The autocorrelation function is insignificant after lag 0, shows no discernible pattern at any order lags of moving average process. Together with the partial autocorrelation function, the time series of Hang Seng Index log-return should follow the ARMA(0,0) process, with the form,

$$r_t = \mu + \epsilon_t$$

Recall the GARCH(1,1) model, the conditional volatility model of Hang Seng index is conducted.

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

The ARMA(0,0)-GARCH(1,1) models can be estimated by maximum likelihood estimation (MLE) method. Table 6 presents the estimated parameters. The * indicates the corresponding estimated parameter is statistically significant at the 95% significance level. And both parameters in the table are significant and reliable. In addition, the results of LM ARCH test show 'ARCH-effects' presents in residuals of the ARMA(0,0) model and it makes sense to employ an ARCH/GARCH model.

Table 5: ARMA-GARCH model estimation results, Hang Seng Index

Parameters	Estimates	Standard Error	p-value
μ	5.597×10^{-4}	2.312×10^{-4}	0.01547*
α_0	1.292×10^{-6}	4.438×10^{-7}	0.00359*
α_1	6.870×10^{-2}	8.559×10^{-3}	0.00000*
β_1	9.279×10^{-1}	8.488×10^{-3}	0.00000*
Loglikelihood	7412.912		
AIC	-5.699		
BIC	-5.690		
LM Arch Test	p=0.07535		

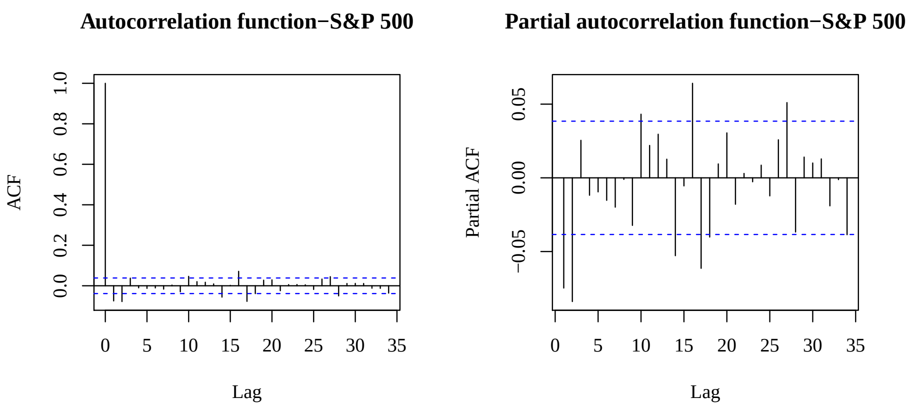

Figure 8: ACF and PACF of S&P 500 Index

The autocorrelation function of S&P 500 log-return shows a different pattern. The correlations at lag 1 and 2 are significant and negative. We can identify this series as it follows ARMA(0,2) process. Similar with Hang Seng Index, we construct the ARMA(0,2)-GARCH(1,1) model for S&P 500 index,

$$r_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

The results of ARMA(0,2)-GARCH(1,1) are presented in Table 6. In a similar manner, the results show the estimated parameters are significant (significance level, 95%) and reliable. The results of LM ARCH test show 'ARCH-effects' presents in residuals of the ARMA(0,2) model and it makes sense to use an ARCH/GARCH model.

Table 6: ARMA-GARCH model estimation results, S&P 500 Index

Parameters	Estimates	Standard Error	p-value
μ	3.652×10^{-4}	1.625×10^{-4}	0.02462*
θ_1	-6.050×10^{-2}	2.068×10^{-2}	0.00344*
θ_2	-4.252×10^{-2}	2.066×10^{-2}	0.03904*
α_0	1.302×10^{-6}	3.220×10^{-7}	0.00005*
α_1	8.193×10^{-2}	9.494×10^{-3}	0.00000*
β_1	9.110×10^{-1}	9.647×10^{-3}	0.00000*
Loglikelihood	8010.256		
AIC	-6.157		
BIC	-6.144		
LM Arch Test	p=0.077241		

5.2 DCC-GARCH model

With the results of univariate ARMA-GARCH model, we model the time-varying volatility of the portfolio by employing DCC-GARCH model. The form of DCC-GARCH model is,

$$H_t = D_t R D_t$$

The estimation starts with the diagonal matrices of conditional variances ($D_t = Diag(h_{1t}, \cdots, h_{nt})$).

$$h_t = a + A\epsilon_{t-1} + Bh_{t-1}$$

where a, A, B are coefficient matrices of the DCC-GARCH model, h_t is the matrices consists of each component's volatility (h_{1t}, h_{2t}). For the diagonal specification (original DCC-GARCH model, volatility spillover not allowed), the coefficient matrices,

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} \quad B = \begin{pmatrix} B_{11} & 0 \\ 0 & B_{22} \end{pmatrix}$$

Further, the dynamic conditional correlation matrix is defined as

$$R_t = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix}$$

The DCC-GARCH model is estimated by maximizing likelihood. The parameters of the DCC-GARCH model and the dynamic conditional correlation matrix at $t = 2600$ are presented in Table 7. Both estimates are reliable.

Finally, define $D_t = diag(h_{1t}, h_{2t})$. With the estimated dynamic conditional correlation matrix R_t and portfolio weights matrix w, the time-varying volatilities $\sigma_{p,t}$ are derived as,

$$\sigma_{p,t} = w' H_t w$$

Table 7: DCC-GARCH estimation results

Parameters	Estimates	Standard Error
a_1	1.292×10^{-6}	4.910×10^{-7}
a_2	1.302×10^{-6}	1.378×10^{-2}
A_{11}	6.870×10^{-2}	1.111×10^{-2}
A_{22}	8.193×10^{-2}	5.960×10^{-7}
B_{11}	9.279×10^{-1}	1.251×10^{-2}
B_{22}	9.110×10^{-1}	1.149×10^{-2}
Dynamic conditional correlation matrix at $t = 2600$		
$\rho_{11} = \rho_{22}$	1.000	
$\rho_{12} = \rho_{21}$	0.187	
Loglikelihood	38242.61	

5.3 Multivariate EWMA

Compared with modelling volatility with DCC-GARCH model, multivariate EWMA approach is easier to be realized. This approach starts with the unconditional covariance matrix of the first 2600 observations (In-sample period).

$$\Sigma_0(r_{HSI}, r_{S\&P500}) = \begin{pmatrix} 0.01713 & 0.00736 \\ 0.00736 & 0.01409 \end{pmatrix}$$

Then, the covariance matrix at time $t = 2601$ (denote as Σ_1),

$$\Sigma_1 = \lambda \Sigma_0(r_{HSI}, r_{S\&P500}) + (1 - \lambda)\epsilon_0 \epsilon_0'$$

Where $\epsilon_0 = r_0 - \mu$. And in this paper, we assume $\lambda = 0.94$. Then the covariance matrix for any time in the test period (Σ_t) can be derived by the following equation.

$$\Sigma_t = \lambda \Sigma_{t-1} + (1 - \lambda)\epsilon_{t-1}\epsilon_{t-1}'$$

With the calculated time-varying covariance matrix $(\Sigma_0, \Sigma_1, \cdots, \Sigma_n)$, the time-varying portfolio volatility $\sigma_{p,t}$ is derived by,

$$\sigma_{p,t} = w' \Sigma_t w$$

5.4 Dependence structure

In this part, we discuss the dependence structures between the portfolio components. It is an important concept to the estimation of portfolio VaR. It determines the covariance matrix and defines the risk level of the portfolio.

The discussion starts with the static measurement of the dependence structures. Dependence structure between Hang Seng Index and S&P 500 index is measured in terms of Pearson's ρ (linear dependence), Kendall's τ and Spearman's ρ (rank correlation coefficient). The result is shown in Table 8. Both the results show the two indexes are positive correlated, but their correlation is not strong. They are facing with different risk factors. It is worth consisting a portfolio that diversify the unsystematic risks.

Table 8: Static dependence structure measurement

Measurement of dependence structure	Value
Pearson's ρ	0.2247
Kendall's τ	0.1111
Spearman's ρ	0.1610

Afterwards, we focus on the dynamic conditional correlation matrices estimated by DCC-GARCH model. It gives the measurement of time-varying dependence structures between Hang Seng Index and S&P 500 Index (Figure 9).

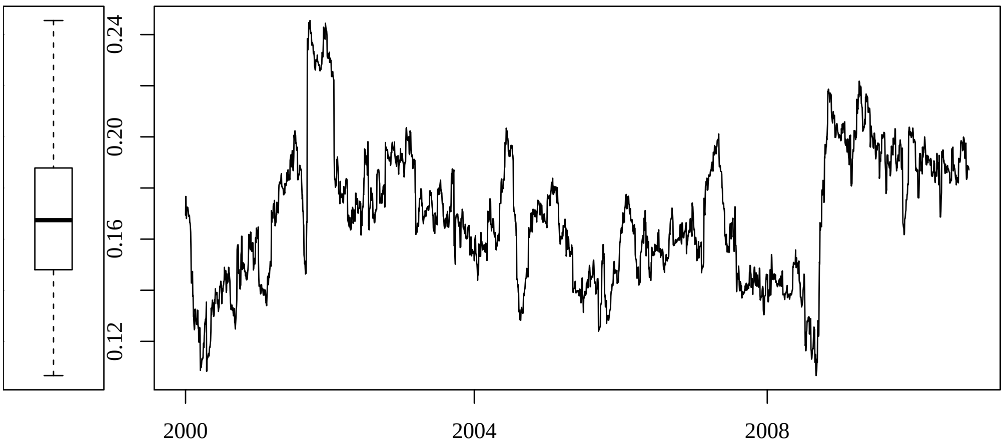

Figure 9: Dynamic of conditional correlation between HSI and S&P 500

The figure shows the dynamic correlations of HSI and S&P 500 are time-varying and volatiles at the range between 0.11 and 0.2455. In addition, by observing the probability distribution of the dependence structures, more than 70% observations of correlation lying in range between 0.15 and 0.2. It is highly concentrated and tends to continue fluctuating in the interval 0.15-0.2, which can be treated as relatively stable in a short period of time.

Thus, the correlation matrix of the HSI and S&P indexes could be assumed to be constant overtime in the test period. With the assumption, there is no need to estimate DCC-GARCH model for each day in the test period. The correlation between Hang Seng index and S&P 500 index is assumed to be consistent with the estimated correlation matrix of DCC-GARCH model (Table 6) at the end of the sample period ($t = 2600$), in the matrix form:

$$\rho = \begin{pmatrix} 1 & 0.187 \\ 0.187 & 1 \end{pmatrix}$$

6 Empirical Results

This section presents the empirical results of various multivariate VaR models that defined in methodology section. Both models are calculated on the 99% confidence level ($\alpha = 0.99$) and the holding period (H) is one-day. For each Monte Carlo simulations process, 10000 iterations are performed. The results for each model are presented (Table 9) in terms of three evaluation criteria: Christoffersen test, quadratic probability score (QPS) and root mean squared error (RMSE). Significant LR statistics are highlighted in bold, which indicate the VaR models fail to pass the corresponding test.

Table 9: Model evaluation, $VaR_\alpha : 1 - \alpha = 1\%$

		Christoffersen Test			Relative Performance	
VaR Model	Violations	LR_{uc}	LR_{ind}	LR_{cc}	QPS	RMSE
HS	1	2.862	2.666	5.528	0.005441	353.656
AWHS	13	**14.106**	2.666	**16.772**	0.068328	299.312
DVWHS	7	2.284	2.666	4.950	0.036884	277.616
EVWHS	4	0.018	2.666	2.684	0.021163	278.234
mvn	11	**9.357**	2.666	**12.023**	0.057847	314.098
MC-mvn	10	**7.256**	2.666	**9.922**	0.052606	313.526
DVW-mvn	4	0.018	2.666	2.684	0.021163	282.902
EVW-mvn	4	0.018	2.666	2.684	0.021163	281.837
mvt	11	**9.357**	2.666	**12.023**	0.057847	314.157
MC-mvt	11	**9.357**	2.666	**12.023**	0.057847	313.981
DVW-mvt	4	0.018	2.666	2.684	0.021163	282.951
EVW-mvt	4	0.018	2.666	2.684	0.021163	281.886
MC-GCp	4	0.018	2.666	2.684	0.021163	341.869
MC-GCn	1	2.862	2.666	5.528	0.005441	374.886
MC-tCp	3	0.159	2.666	2.825	0.015922	351.993
MC-tCn	1	2.862	2.666	5.528	0.005441	375.996
MC-GuCp	6	1.166	2.666	3.832	0.031643	337.798
MC-GuCn	2	0.984	2.666	3.650	0.010681	376.275
MC-ClCp	2	0.984	2.666	3.650	0.010681	362.967
MC-ClCn	1	2.862	2.666	5.528	0.005441	375.681
MC-FrCp	7	2.284	2.666	4.950	0.036884	333.797
MC-FrCn	1	2.862	2.666	5.528	0.005441	376.747
$LR_{critic}, \alpha^* = 95\%$		3.841	3.841	5.991		
Confidence Interval, violations (Obs=374, $\alpha^* = 99\%$)					0-9	
Confidence Interval, violations (Obs=374, $\alpha^* = 95\%$)					1-8	

The results can be summarized as follows:

1. Basic multivariate VaR models (mvn/mvt/HS/AWHS) do not perform well in predicting future losses. As indicated by the christoffersen test (LR_{uc})

(Table 9), three of them (mvn/mvt/AWHS) fail to generate adequate estimation of future losses. As a results, their QPS is larger than other models. In addition, the number of violations during the test period is relatively large, which indicates their underestimation of future losses.

However, one of them - the historical simulation VaR model (HS) passes the christoffersen test and have few violation during the test period. Consequently, its QPS is low which indicates its good performance in the violation days. The historical simulation VaR model could have a lower probability of occurring violations. Despite its good performance in violation days, some evidences of overestimating the future losses are found. Compared with the other VaR models, the RMSE of historical simulation VaR model is larger. It shows its relatively poor performance in non-violation days: users have to hold a higher level of capital reserves, which is costly. Figure 14 (see Appendix A) also shows some evidences of overestimating future losses. The estimated VaR is slightly higher than the other basic VaR models. Less violations are at the expense of higher reserves in non-violation days. For a regulatory purposes, the model is acceptable and conservative enough. However, from the users' perspective, it is costly to accept the historical simulation model.

2. Volatility adjusted multivariate VaR models (DVWHS/EVWHS/DVW-mvn/EVW-mvn/DVW-mvt/EVW-mvt) shows notable improvements in the performance of predicting future losses, compared with the basic multivariate VaR models. Both volatility adjusted multivariate VaR models pass the christoffersen test (Table 9, $LR_{uc}, LR_{ind}, LR_{cc}$). The relatively low QPS and the relatively low RMSE shows its good performance in the violation days as well as the non-violation days. Compared with the other VaR models, they have less probability of violations and at the same time, do not require a high-level capital reserve.

   The DCC-GARCH model and the multivariate EWMA model have similar performances in estimating the future volatilities. However, their features are slightly different. Among volatility adjusted parametric VaR models (DVW-mvn/EVW-mvn/DVW-mvt/EVW-mvt), the RMSE statistics in Table 9 indicate VaR models with multivariate EWMA volatility perform slightly better in non-violation days. In addition, figure 15 and 16 (see Appendix A) show DCC-GARCH model overestimates the volatility's sudden change in August, 2011. On other days of the test period, it is hard to figure out any difference between multivariate EWMA volatility model and multivariate DCC-GARCH volatility model. In non-parametric VaR models (DVWHS/EVWHS), the RMSE statistics indicate VaR models with DCC-GARCH have better performance in non-violation days. While the QPS statistics indicate VaR models with multivariate EWMA have lower violations in the test period. VaR models with multivariate EWMA are more conservative, compared with VaR models with multivariate DCC-GARCH.

3. Monte Carlo approaches of multivariate VaR (MC-mvn/MC-mvt) show consistent results with basic multivariate VaR models. Theoretically, Monte Carlo approach and its corresponding basic VaR model are under the same

assumptions and they should have the same results. The differences between QPS and RMSE statistics come from the standard errors of estimation. The bias can be decreased by increasing the simulation iterations.

4. Compared with basic multivariate VaR models, copula-based multivariate VaR models have better predictive power on future losses. Both copula-based multivariate VaR models pass the christoffersen test (Table 9), which indicate their adequate estimation of future losses. Their number of violations in the test period is low. Consequently, the QPS statistics is lower than the basic models (mvn/mvt/AWHS). From the regulatory perspective, copula-based multivariate VaR models performs well.

   In basic multivariate VaR models, we assume the probability distribution of loss as a specific statistical probability distribution. However, in practice, the actual losses do not always follow a certain probability distribution that can be specified by a simple equation. Sometimes, the probability distribution can be complex. In that case, the VaR estimation based on a specific probability distribution could be inadequate. By contrast, the copula theory shows its brilliant ability of describing complex multivariate probability distributions. The selection of probability distribution is not limited on the existing probability distribution. With the copula theory, we can construct an unknown distribution that fits the data best. It is the reason behind the better performance of copula-based multivariate VaR models.

   Despite its advantages over the basic models, they face the same problem with non-volatility adjusted VaR models: in order to lower the probability of violations, the estimated VaR values have to be more conservative than the volatility adjusted models. As a results, their RMSE statistics can be relatively high. From the users' perspective, it is costly to have a high level capital reserves.

   However, the problem can be solved by introducing volatility models into the copula theory, namely, conditional copula models. Due to the limited time frame, we do not employ the conditional copula models. But it is available for future study. We believe it can generate a VaR estimation with less violations and demanding less capital reserves.

   In the paper, five families of copula functions and two different assumptions on marginal distributions are employed. As indicated by the results, they have shown different features in estimating multivariate VaR. As indicated by the RMSE statistics and the figures (Figure 17-21, see Appendix), copula-based copula VaR models with normal marginal distribution are more conservative (RMSE is higher) than the copula-based copula VaR models based on pseudo observations. As a results, they have lower probability of violations.

   Among different copula functions (Gaussian/t/Gumbel/Clayton/Frank), Clayton copula and student's t copula perform the best in lowering violations. Gaussian copula performs on the average level and is less conservative than Clayton/student's t copula. Gumbel copula ranks at the fourth and Frank copula performs the least conservative. The results are consistent with some basic features of copula functions. The comparisons of different copula-based multivariate VaR models are presented in Table 9.

Table 10: Features and comparison of different copula-based VaR models

VaR Model	Violations	QPS	RMSE	Features of copula function
Gaussian(P)	4	0.021163	341.869	Reflecting overall dependence
Gaussian(N)	1	0.005441	374.886	structure.
t(P)	3	0.015922	351.993	Focus on the tail dependence
t(N)	1	0.005441	375.996	(Both upper and lower sides)
Gumbel(P)	6	0.031643	337.798	Focus on the tail dependence
Gumbel(N)	2	0.010681	376.275	(Upper side)
Clayton(P)	2	0.010681	362.967	Focus on the tail dependence
Clayton(N)	1	0.005441	375.681	(Lower side)
Frank(P)	7	0.036884	333.797	Tail independence
Frank(N)	1	0.005441	376.747	

In the Table, 'N' represents the copula model with normal margins and 'P' represents the copula model based on pseudo observations. The table indicates that the estimated VaR would be more conservative (larger RMSE and less violations) with the increasing focus on the lower tail dependence and tail losses. The model with the most focus on the lower tail dependence (Clayton copula) performs best in lowering number of violations. The model assumes the tail independence (Frank copula) performs the worst. Dependence structure is an important part of multivariate VaR models, especially the lower tail dependence (Actually, the VaR estimation locates here). More conservative results will be get, if the VaR model describes the characteristics of tail losses better.

7 Conclusion

The aim of this paper has been to examine the one-day predictive power of several multivariate VaR models, including basic multivariate VaR models, volatility-adjusted multivariate VaR models and copula-based multivariate VaR models. The comparison is made on a portfolio consisting of S&P 500 index and Hang Seng Index. Christoffersen test, quadratic probability score and root mean square error are used as standard tools to evaluate the performance. Following the research questions in the introduction section, the empirical results can be concluded as follows:

1. Based on the results of quadratic probability score and the root mean squared error, basic multivariate VaR models (mvn/mvt/HS/AWHS) show poor performances in estimating future losses. Additionally, three basic multivariate VaR models (mvn/mvt/AWHS) fail to pass the christoffersen test. Their VaR estimation can be treated as inadequate.

2. Multivariate EWMA and DCC-GARCH model are employed as the multivariate volatility model. Both models are easy to implement (section 5) and the results indicate the accurate estimation of the time-varying multivariate volatility.

3. Both volatility adjusted multivariate VaR models pass the christoffersen test. Compared with the basic multivariate VaR models, they have higher quadratic probability scores and lower root mean squared errors. The results indicate that volatility adjusted multivariate VaR models have better predictive performances than the basic VaR models.

4. Compared with basic multivariate VaR models, copula-based multivariate VaR models show notable improvements in lowering probability of violations. The copula theory constructs multivariate distributions with attentions on the tail losses and tail dependence. As a result, the copula-based multivariate VaR models have a better predictive power on the tail losses and show similar characters with extreme value theory (EVT) VaR models. All the copula-based multivariate VaR models pass the christoffersen test and have a lower quadratic probability score.

As a final point, the main implication of this research for practitioners is it offers a practical guidance for estimating portfolio VaR by multivariate VaR models. The comparison between different multivariate VaR models gives an overview of their performances and features. This paper can be effective as a reference when facing portfolio risk measurement problems or facing the problem of selecting adequate multivariate VaR models.

References

Aeschliman, C., Park, J. and Kak, A.C., (2010). A novel parameter estimation algorithm for the multivariate t-distribution and its application to computer vision. *11th European conference on computer vision (ECCV)*, Berlin Heidelberg. Springer, 594-607.

Basel Committee on Banking Supervision, (1996). Amendment to the capital accord to incorporate market risks. *Basel Committee Publications*, 24.

Belzunee, F., Castano, A., Olvera-Cervantes, A. and Suárez-Lorens, A., (2007). Quantile curves and dependence structure for bivariate distribution. *Computational statistics and data analysis*, 45(3), 5112-5129.

Blanco, C. and Ihle, G., (1999). How good is your VaR? Using back-testing to assess system performance. *Financial Engineering News*, 11(8), 1-2.

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

Bollerslev, T., Engle, R.F. and Wooldridge, J.M., (1988). A capital asset pricing model with time varying covariances. *Journal of political economy*, 96(1), 116-131.

Boudoukh, J., Richardson, M. and Whitelaw R., (1998). The best of both worlds: a hybrid approach to calculating value at risk. *Risk*, 11, 64-47.

Bouyè, E., Durrleman, V., Nikeghbali, A., Riboulet, G. and Roncalli T., (2000). Copulas for finance-a reading guide and some applications. Working Paper, Group de Recherche Opèrationelle, Crédit Lyonnais.

Brooks, C., (2008). *Introductory econometrics for finance*. 2nd ed. Cambridge: Cambridge University Press.

Byrd, R.H., Lu, P. and Nocedal, J., (1995). A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific and Statistical Computing*, 16(5), 1190-1208.

Cherubini, U., Luciano, E. and Vecchinato, W., (2004). *Copula methods in finance*. Chichester: John Wiley and Sons, Ltd.

Christodoulakis, G.A. and Satchell, S.E., (2002). Correlated ARCH: modelling the time-varying correlation between financial asset returns. *European journal of operations research*, 139, 351-370.

Christoffersen, P.F., (1998). Evaluating interval forecasts. *International economic review*, 39, 841-862.

Clayton, D.G., (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65, 141-151.

Cousin, A. and Di Bernardino, E., (2011). A multivariate extension of Value-at-Risk and Conditional-Tail-Expectation. *ArXiv e-prints*, 1111.1349.

Dowd, K., (1999). Backtesting Market Risk Models. *ICBI conference*, Geneva.

Dowd, K., (2005). *Mesuring market risk*. 2nd ed. Chichester: Wiley.

Duong, T., (2007). ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R. *Journal of statistical software*, 21(7).

Einhoen, D. and Brown, A., (2008). Private profits and socialized risk. *Global association of risk professionals*, 42, 10-26.

Embrechts, P. and Puccetti, G., (2006). Bounds for functions of multivariate risks. *Journal of multivariate analysis*, 97(2), 526-547.

Engel, R.F., (2002). Dynamic conditional correlation-a simple class of multivariate GARCH models. *Journal of business and economic statistics*, 20, 339-350.

Engel, R.F., Kroner, K. and Kraft, D., (1995). Multivariate simultaneous generalized ARCH. *Econometric theory*, 11, 122-150.

Engel, R.F. and Sheppard, K., (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. NBER Working Paper 8554, National Bureau of Economic Research.

Engelbrecht, R. and West, G., (2003). A comparison of value-at-risk methods for portfolios consisting of interest rate swap and FRAs. Economic Series Working Papers, University of Oxford.

Fernández-Ponce, J.M. and Suárez-Lloréns, A., (2002). Central regions for bivariate distributions. *Austrian journal of statistic*, 31(2-3), 141-156.

Frank, M.J., (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes math*, 19, 194-226.

Guermat, C. and Harris, R., (2002). Robust conditional variance estimation and value-at-risk. *Journal of Risk*, 4, 25-41.

Gumbel, E.J., (1960). Bivariate exponential distributions. *J. Amer. Statist. Assoc.*, 55, 698-707.

Hull, J. and White, A., (1988). Value at risk when daily changes in market variables are not normally distributed. *Journal of derivatives*, 5(Spring), 9-19.

Ivanov, S., Jordan, R., Panajotova, B., and Schoess, S., (2003). Dynamic Value-at-Risk with heavy tailed distribution. *9th Annual IFCI Risk Management Round Table*.

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

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

Laubsch, A.J., (1999). *Risk management:a practical guide.* 1st ed. United States: MSCI.

Lehmann, E., (1966). Some concepts of dependence. *Ann. Math. Statist.*, 37, 1137-1153.

Lopez, J.A., (1998). Regulatory evaluation of Value-at-Risk models. *Federal Reserve Bank of New York Economic Policy Review*, 4(3), 119-124.

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

Morgan Guaranty Trust Company, (1996). *RiskMetrics-Technical document.* 4th ed. New York: Morgan Guaranty Trust Company.

Nappo, G. and Spizzichino, F., (2009). Kendall distributions and level sets in bivariate exchangale survival models. *Information sciences*, 179, 2878-2890.

Nelsen, R.B., (1999). *An introduction to copulas.* Springer-Verlag Inc.

Palaro, H.P. and Hotta, L.K., (2006). Using conditional copula to estimate value at risk. *Journal of data science*, 4, 93-115.

Sheikh, A.Z. and Qiao, H.T., (2010). Non-normality of market returns: a framework for asset allocation decision making. *The journal of alternative investments*, 12(3), 8.

Sklar, A., (1959). Fonctions de repartition à n dimensions et leurs marges. *Pub. Inst. Statist. Univ. Paris*, 8, 229-231.

Taleb, N.N., (1997). *Dynamic hedging: managing vanilla and exotic options.* New York: Wiley.

Tibiletti, L., (1993). On a new notion of multidimensional quantile. *International Journal of statistics*, 51(3-4), 77-83.

Tse, Y.K. and Tsui, A.K.C, (2002). A multivariate GARCH model with time-varying correlations. *Journal of business and economic statistics*, 20, 351-362.

Yan, J., (2007). Enjoy the joy of copulas: with a package copula. *Journal of statistical software*, 21(4).

Zhu, C., Byrd, R.H., Lu, P. and Nocedal, J., (1994). L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization. Report NAM-11, EECS Department, Northwest university.

Appendix A

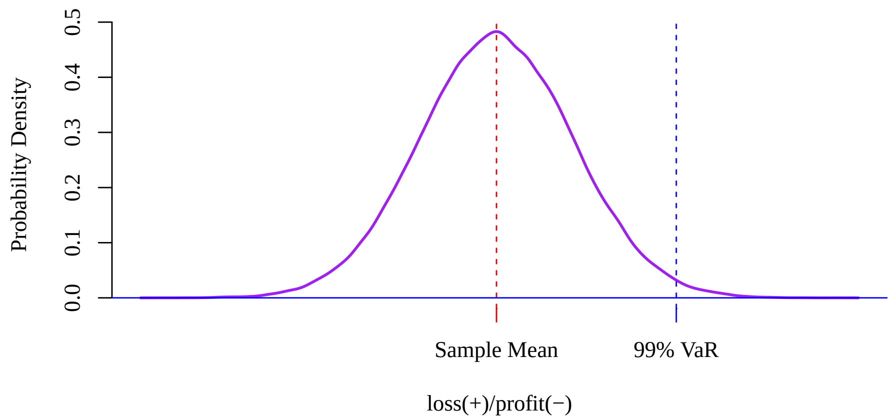

Figure 10: A simple illustration of Monte Carlo approach

This figure graphs a realization of 100000 simulated portfolio returns. The portfolio is consist of two assets with equal weights. The assets follow a multivariate normal distribution, with $\mu = (0.1, 1)$ and $\Sigma = \begin{pmatrix} 1 & 0.4 \\ 0.4 & 1 \end{pmatrix}$. Additionally, the 99% VaR can be point out as the 99% quantile of the loss distribution.

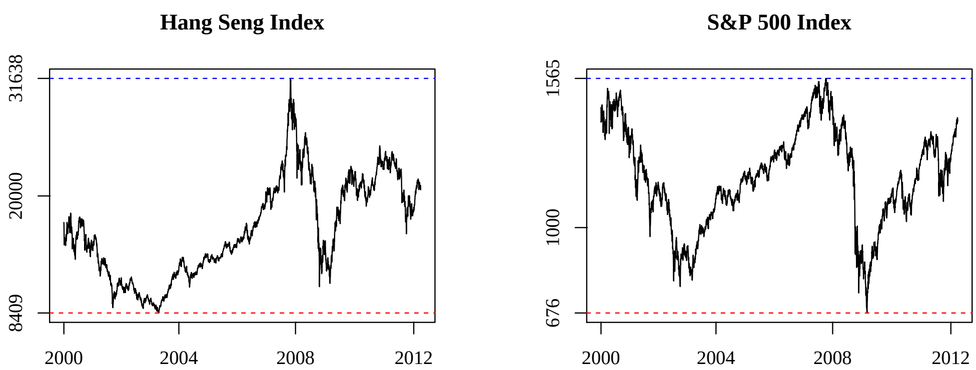

Figure 11: Daily closing price of Hang Seng Index and S&P 500 Index

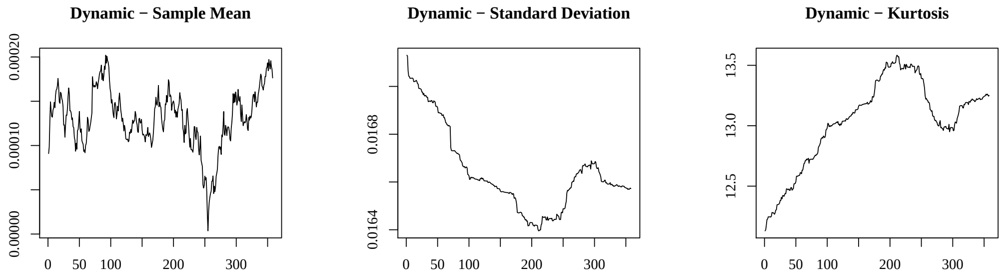

Figure 12: Dynamic of log-return distributions (Hang Seng Index)

The figure on the left shows the time-varying sample mean of Hang Seng Index. The average return seems to have a gap in $t = 240 \sim 260$. The figure in the middle shows the time-varying sample standard deviation of Hang Seng Index. It shows a trend that the volatility of the index returns is decreasing. The figure on the right shows the time-varying kurtosis of Hang Seng Index. It seems the probability distribution tends to be more and more fat-tailed overtime. It can be improper to assume the loss distributions follow normal distribution.

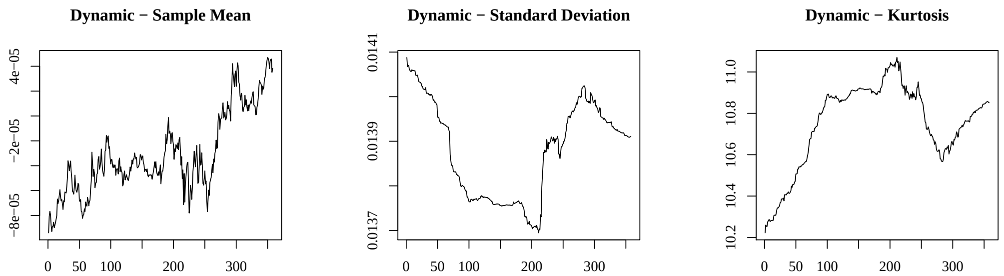

Figure 13: Dynamic of log-return distributions (S&P 500 Index)

The figure on the left shows the time-varying sample mean of S&P 500 Index. The average return is increasing overtime, and the performance of the index is strong. The figure in the middle shows the time-varying sample standard deviation of S&P 500 Index. The figure on the right shows the time-varying kurtosis of S&P 500 Index.

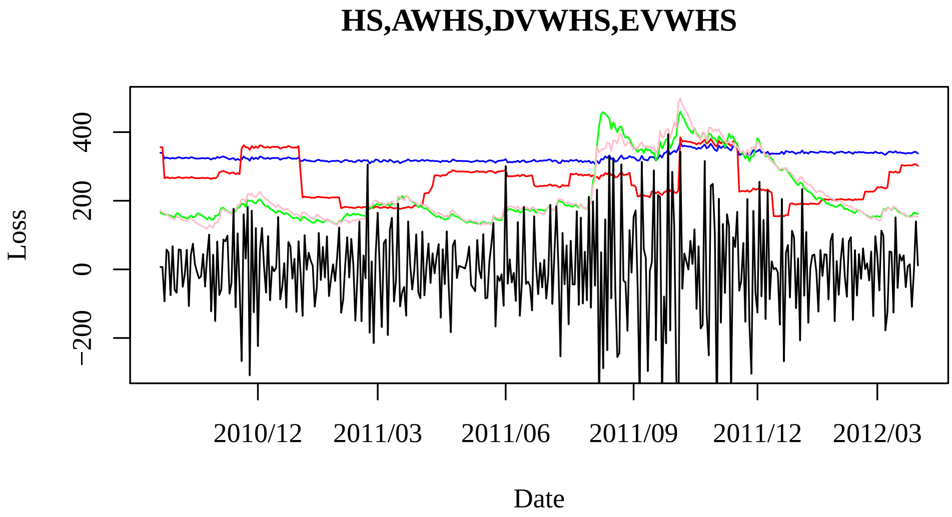

Figure 14: Estimated VaR: Multivariate historical simulation approach

The black fold line represents the actual losses of the portfolio. The blue line represents the estimated VaR by multivariate historical simulation VaR model (HS). The red line denotes the estimated VaR by multivariate age-weighted historical simulation VaR model (AWHS). The green line denotes the estimated VaR by volatility weighted (EWMA) historical simulation VaR model (EVWHS). The pink line represents the estimated VaR by volatility weighted (DCC-GARCH) historical simulation VaR model (DVWHS).

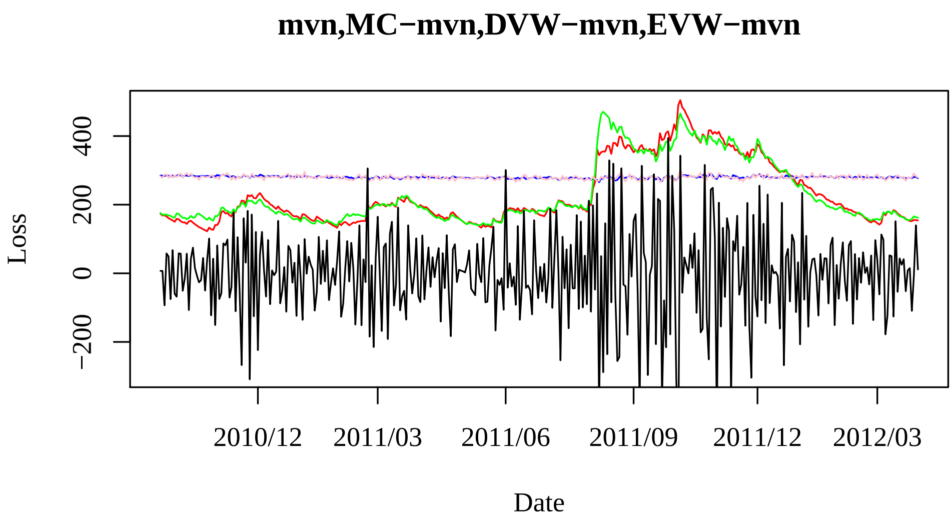

Figure 15: Estimated VaR: Multivariate normal approach

The black fold line represents the actual losses of the portfolio. The blue and pink line represents the estimated VaR by multivariate normal VaR model (mvn) and Monte Carlo multivariate normal VaR model (MC-mvn). The red line denotes the estimated VaR by volatility weighted (EWMA) multivariate normal VaR model (EVW-mvn). The green line denotes the estimated VaR by volatility weighted (DCC-GARCH) multivariate normal VaR model (DVW-mvn).

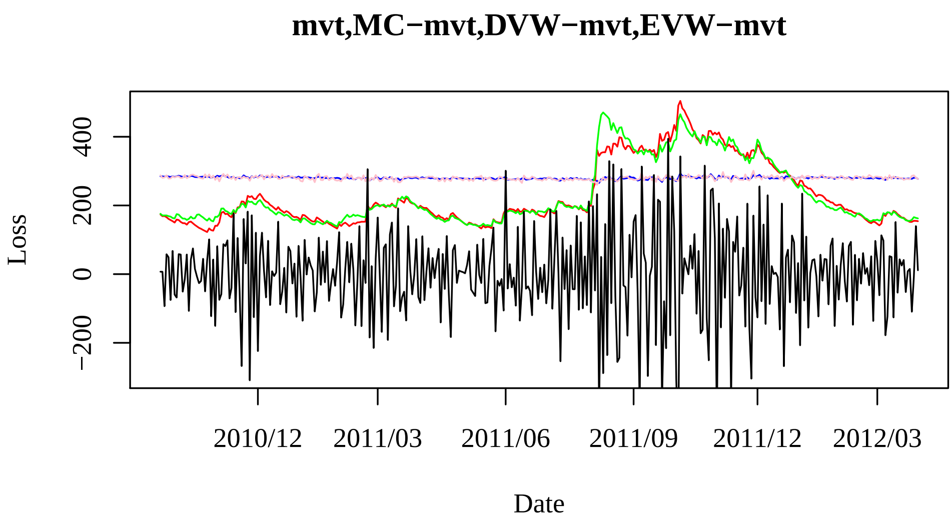

Figure 16: Estimated VaR: Multivariate t approach

The black fold line represents the actual losses of the portfolio. The blue and pink line represents the estimated VaR by multivariate t VaR model (mvt) and Monte Carlo multivariate t VaR model (MC-mvt). The red line denotes the estimated VaR by volatility weighted (EWMA) multivariate t VaR model (EVW-mvt). The green line denotes the estimated VaR by volatility weighted (DCC-GARCH) multivariate t VaR model (DVW-mvt).

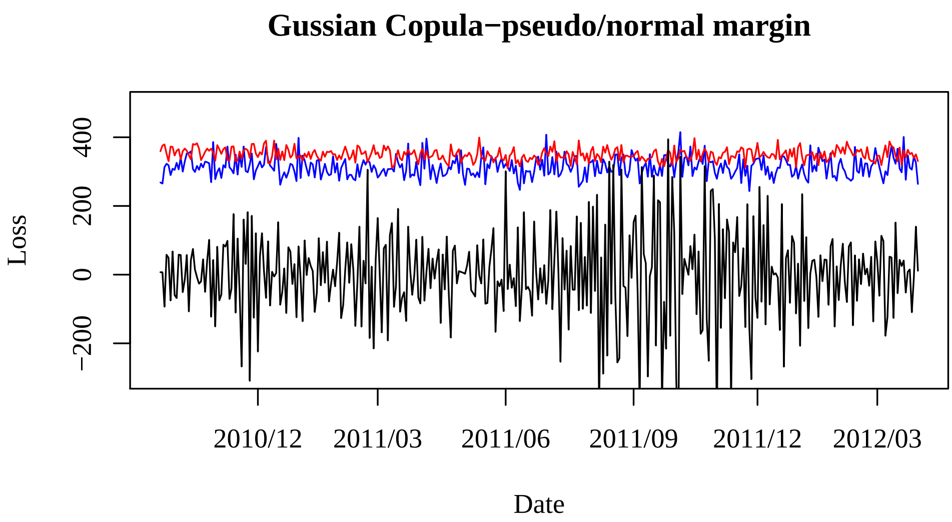

Figure 17: Estimated VaR: Gaussian Copula Monte Carlo approach

The black fold line represents the actual losses of the portfolio. The red line denotes the estimated VaR by Gaussian Copula (normal margin) multivariate VaR model (MC-GCn). The blue line denotes the estimated VaR by Gaussian Copula (pseudo observations) multivariate VaR model (MC-GCp).

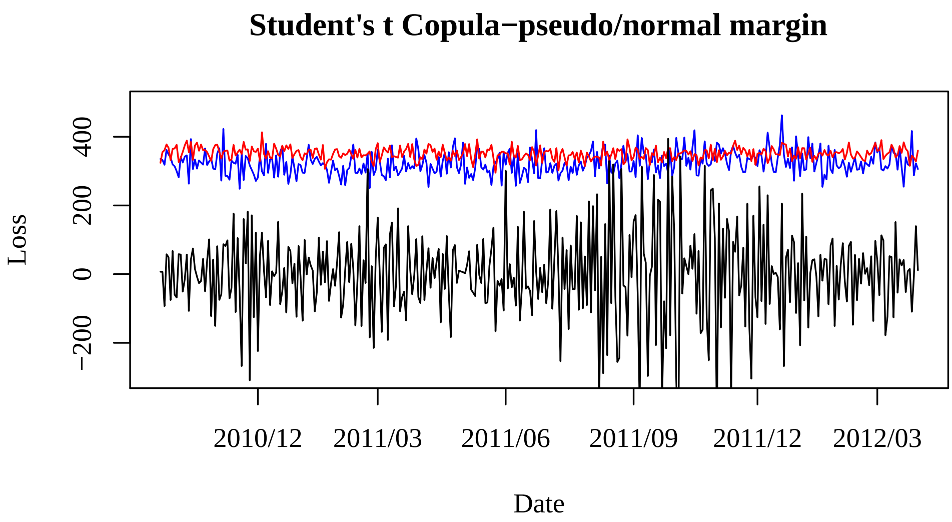

Figure 18: Estimated VaR: Student's t Copula Monte Carlo approach

The black fold line represents the actual losses of the portfolio. The red line denotes the estimated VaR by student's t Copula (normal margin) multivariate VaR model (MC-tCn). The blue line denotes the estimated VaR by student's t Copula (pseudo observations) multivariate VaR model (MC-tCp).

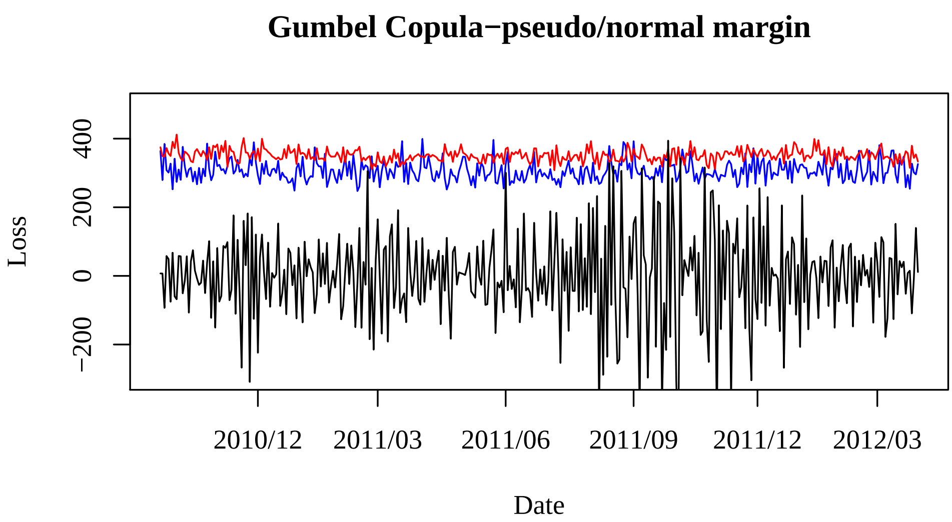

Figure 19: Estimated VaR: Gumbel Copula Monte Carlo approach

The black fold line represents the actual losses of the portfolio. The red line denotes the estimated VaR by Gumbel Copula (normal margin) multivariate VaR model (MC-GuCn). The blue line denotes the estimated VaR by Gumbel Copula (pseudo observations) multivariate VaR model (MC-GuCp).

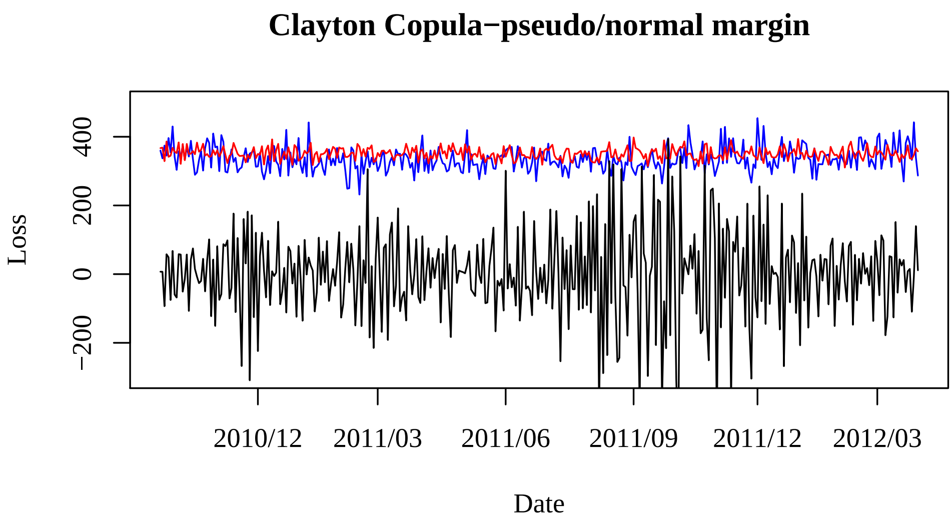

Figure 20: Estimated VaR: Clayton Copula Monte Carlo approach

The black fold line represents the actual losses of the portfolio. The red line denotes the estimated VaR by Clayton Copula (normal margin) multivariate VaR model (MC-ClCn). The blue line denotes the estimated VaR by Clayton Copula (pseudo observations) multivariate VaR model (MC-ClCp).

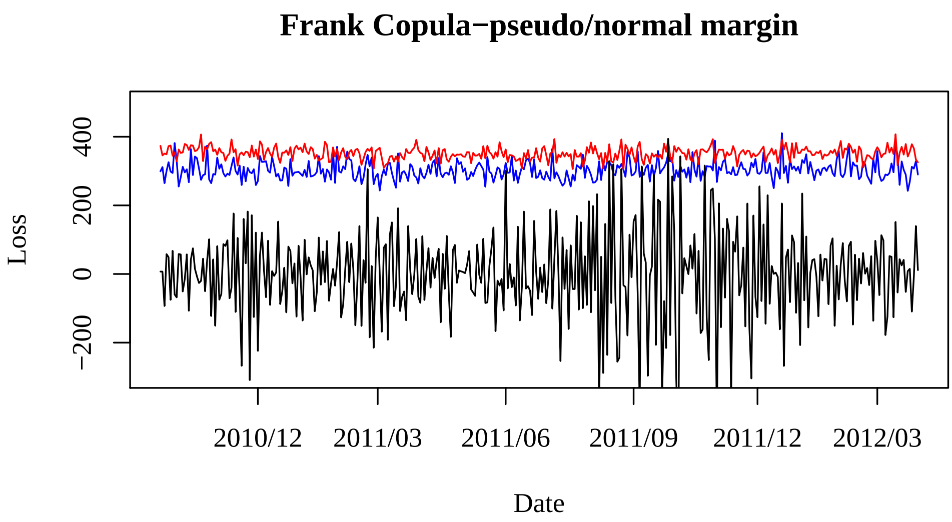

Figure 21: Estimated VaR: Frank Copula Monte Carlo approach

The black fold line represents the actual losses of the portfolio. The red line denotes the estimated VaR by Frank Copula (normal margin) multivariate VaR model (MC-FrCn). The blue line denotes the estimated VaR by Frank Copula (pseudo observations) multivariate VaR model (MC-FrCp).

Appendix B

R codes for this paper(304 lines)

R: A Language and Environment for Statistical Computing, Version 2.15.0 (2012-3-30). R Development Core Team. http://www.R-project.org

```
#Define Portfolio weights, sample size and alpha level
w=c(0.5,0.5)
tP=374
si=0.01

#Multivariate EWMA volatility
xdata=cbind(hsreturn,spreturn)
res=xdata-apply(xdata,2,mean)
s=var(sdata)
lambda=0.94
s=lambda*s
sigma=sqrt(t(w)%*%s%*%w)
EWMAsigma=rep(0,T+1)
EWMAsigma[1]=sigma
T=length(hsreturn)
for(i in 2:(T+1)){
s=lambda*s+(1-lambda)*res[(i-1),]%*%t(res[(i-1),])
sigma=sqrt(t(w)%*%s%*%w)
EWMAsigma[i]=sigma
}

#Dynamic conditional correlation GARCH volatility
library(ccgarch)
library(fGarch)
f1=garchFit(~garch(1,1),sdata[,1],trace=FALSE)
f1=f1@fit$coef
f2=garchFit(~arma(0,2)+garch(1,1),sdata[,2],trace=FALSE)
#f2=garchFit(~garch(1,1),sdata[,2],trace=FALSE)
f2=f2@fit$coef
inia=c(f1[2],f2[4])
iniA=diag(c(f1[3],f2[5]))
iniB=diag(c(f1[4],f2[6]))
dcc.para=c(0.01,0.97)   #intial Value
dcc.results=dcc.estimation(inia,iniA,iniB,dcc.para,sdata,model="diagonal")
#dcc.results=dcc.estimation(inia,iniA,iniB,dcc.para,sdata,model="extended")
DCCGarchsigma=rep(0,(length(hsreturn)+1))
for(i in 1:2600){
D=diag(sqrt(dcc.results$h[i,]))
R=matrix(dcc.results$DCC[i,],2,2)
H=D%*%R%*%D
DCCGarchsigma[i]=sqrt(t(w)%*%H%*%w)
}
T=length(sdata[,1])
```

```
h=matrix(0,(tP+1),2)
h[1,]=dcc.results$h[T,]
for(i in 2:(tP+1)){
h[i,]=c(dcc.results$out[1,1],dcc.results$out[1,2])+diag(c(dcc.results$out[1,3],
    dcc.results$out[1,4]))%*%c(res[(i+2598),1]^2,res[(i+2598),2]^2)+diag(c(dcc.
    results$out[1,5],dcc.results$out[1,6]))%*%h[i-1,]
}
R=matrix(dcc.results$DCC[T,],2,2)
for(i in 2601:2974){
D=diag(sqrt(h[(i-2599),]))
H=D%*%R%*%D
DCCGarchsigma[i]=sqrt(t(w)%*%H%*%w)}

#Multivariate t VaR model
VaRmt=rep(0,tP)
VaRmtEWMA=rep(0,tP)
VaRmtDCCG=rep(0,tP)
mupts=rep(0,tP)
sigmapts=rep(0,tP)
for(i in 1:tP){
data1=NULL
data1=cbind(hsi[i,],sp[i,])
kurt1=kurtosis(data1[,1],method="moment")
#kurt1=kurtosis(data1[,1])
kurt2=kurtosis(data1[,2],method="moment")
#kurt1=kurtosis(data1[,2])
kavg=(kurt1+kurt2)/2
df=(4*kavg-6)/(kavg-3)
mu=apply(data1,2,mean)
T=length(sdata[,1])
sigmat=(T-1)*var(data1)/T
cort=cor(data1)
params=c(mu,df)
out<-nlm(mlogl,params,cort,data1)
mue=c(out$estimate[1],out$estimate[2])
df=out$estimate[3]
mupts[i]=w%*%mue
sigmapts[i]=sqrt(t(w)%*%sigmat%*%w)
VaRmt[i]=cpprice[i]*(-mupts[i]-sigmapts[i]*sqrt((df-2)/df)*qt(si,df))
VaRmtEWMA[i]=cpprice[i]*(-mupts[i]-EWMAsigma[i+2600]*sqrt((df-2)/df)*qt(si,df))
VaRmtDCCG[i]=cpprice[i]*(-mupts[i]-DCCGarchsigma[i+2600]*sqrt((df-2)/df)*qt(si,df
    ))
}

#Multivariate normal VaR models
VaRmn=rep(0,tP)
VaRmnEWMA=rep(0,tP)
VaRmnDCCG=rep(0,tP)
mupns=rep(0,tP)
```

```
sigmapns=rep(0,tP)
library(mvnmle)
for(i in 1:tP){
data1=NULL
data1=cbind(hsi[i,],sp[i,])
fit=mlest(data1)
mun=fit$muhat
sigman=fit$sigmahat
mupns[i]=w%*%mun
sigmapns[i]=sqrt(t(w)%*%sigman%*%w)
VaRmn[i]=cpprice[i]*(-mupns[i]-sigmapns[i]*qnorm(si))
VaRmnEWMA[i]=cpprice[i]*(-mupns[i]-EWMAsigma[i+2600]*qnorm(si))
VaRmnDCCG[i]=cpprice[i]*(-mupns[i]-DCCGarchsigma[i+2600]*qnorm(si))
}

#Multivariate historical simulation VaR model
T=length(sdata[,1])
op=T*si
VaRHS=rep(0,tP)
for(i in 1:tP){
data1=NULL
data1=cbind(hsi[i,],sp[i,])
HSp=data1%*%w
sdata1=sort(HSp)
VaRHS[i]=-sdata1[op]*cpprice[i]
}
T=length(sdata[,1])
op=T*si
VaRVWHSD=rep(0,tP)
VaRVWHSE=rep(0,tP)
for(i in 1:tP){
data1=NULL
data1=cbind(hsi[i,],sp[i,])
HSp=data1%*%w
HSpWD=rep(0,length(data[,1]))
HSpWE=rep(0,length(data[,1]))
for(j in 1:length(data1[,1])){
HSpWD[j]=HSp[j]*DCCGarchsigma[i+2600]/DCCGarchsigma[i+j]
HSpWE[j]=HSp[j]*EWMAsigma[i+2600]/EWMAsigma[i+j]
}
sdata1=sort(HSpWD)
VaRVWHSD[i]=-sdata1[op]*cpprice[i]
sdata2=sort(HSpWE)
VaRVWHSE[i]=-sdata2[op]*cpprice[i]
}
lambda=0.94
T=length(sdata[,1])
w2=rep(0,T)
w2[1]=(1-lambda)/(1-lambda^T)
```

```
VaRAWHS=rep(0,tP)
for(i in 2:T){w2[i]=w2[i-1]*lambda}
for(i in 1:tP){
data1=NULL
data1=cbind(hsi[i,],sp[i,])
HSp=data1%*%w
AWHS=cbind(HSp,w2)
sAWHS=AWHS[order(AWHS[,1]),]
prob=0
j=1
while(prob<=si){
prob=prob+sAWHS[j,2]
j=j+1
}
op=j
VaRAWHS[i]=-sAWHS[op,1]*cpprice[i]
}

#Monte Carlo Method- mvt & mvn
library(MASS)
MCVaRn=rep(0,tP)
for(i in 1:tP){
data1=NULL
data1=cbind(hsi[i,],sp[i,])
fit=mlest(data1)
mun=fit$muhat
sigman=fit$sigmahat
MC=mvrnorm(10000,mun,sigman)
MCportfolio=MC%*%w
MCVaRn[i]=-cpprice[i]*quantile(MCportfolio,p=si)
}
library(mvtnorm)
MCVaRt=rep(0,tP)
for(i in 1:tP){
data1=NULL
data1=cbind(hsi[i,],sp[i,])
kurt1=kurtosis(data1[,1],method="moment")
kurt2=kurtosis(data1[,2],method="moment")
kavg=(kurt1+kurt2)/2
df=(4*kavg-6)/(kavg-3)
mu=apply(data1,2,mean)
T=length(sdata[,1])
cort=cor(data1)
params=c(mu,df)
sigmat=(T-1)*var(data1)/T
out<-nlm(mlogl,params,cort,data1)
mue=c(out$estimate[1],out$estimate[2])
df=out$estimate[3]
MC=rmvt(10000,sigmat,df,mue)
```

```
MCportfolio=MC%*%w
MCVaRt[i]=-cpprice[i]*quantile(MCportfolio,p=si)
}

#Copula-based multivariate VaR models
library(copula)
VaRGaussianC=rep(0,tP)
VaRGaussianC2=rep(0,tP)
VaRtC=rep(0,tP)
VaRtC2=rep(0,tP)
VaRGC=rep(0,tP)
VaRGC2=rep(0,tP)
VaRCC=rep(0,tP)
VaRCC2=rep(0,tP)
VaRFC=rep(0,tP)
VaRFC2=rep(0,tP)
for(i in 1:tP){
v=4
data1=NULL
data1=cbind(hsi[i,],sp[i,])
fit1=fitdistr(data1[,1],"normal")
fit2=fitdistr(data1[,2],"normal")
u1=pnorm(data1[,1],fit1$estimate[1],fit1$estimate[2])
u2=pnorm(data1[,2],fit2$estimate[1],fit2$estimate[2])
Un=cbind(u1,u2)
T=length(data1[,1])
Up=apply(data1,2,rank)/(T+1)  #pseudo-observations
fitGp=fitCopula(normalCopula(0.6,dim=2,dispstr="ex"),Up,method="mpl")
fitGn=fitCopula(normalCopula(0.6,dim=2,dispstr="ex"),Un,method="mpl")
fitTp=fitCopula(tCopula(0.6,dim=2,dispstr="ex",df=v,df.fixed=TRUE),Up,method="mpl")
fitTn=fitCopula(tCopula(0.6,dim=2,dispstr="ex",df=v,df.fixed=TRUE),Un,method="mpl")
fitGUp=fitCopula(gumbelCopula(2,dim=2),Up,method="mpl")
fitGUn=fitCopula(gumbelCopula(2,dim=2),Un,method="mpl")
fitCp=fitCopula(claytonCopula(2,dim=2),Up,method="mpl")
fitCn=fitCopula(claytonCopula(2,dim=2),Un,method="mpl")
fitFp=fitCopula(frankCopula(2,dim=2),Up,method="mpl")
fitFn=fitCopula(frankCopula(2,dim=2),Un,method="mpl")
xGp=rcopula(normalCopula(fitGp@estimate,dim=2,dispstr="ex"),10000)
xGn=rcopula(normalCopula(fitGn@estimate,dim=2,dispstr="ex"),10000)
xTp=rcopula(tCopula(fitTp@estimate,dim=2,dispstr="ex",df=v,df.fixed=TRUE),10000)
xTn=rcopula(tCopula(fitTn@estimate,dim=2,dispstr="ex",df=v,df.fixed=TRUE),10000)
xGUp=rcopula(gumbelCopula(fitGUp@estimate,dim=2),10000)
xGUn=rcopula(gumbelCopula(fitGUn@estimate,dim=2),10000)
xCp=rcopula(claytonCopula(fitCp@estimate,dim=2),10000)
xCn=rcopula(claytonCopula(fitCn@estimate,dim=2),10000)
xFp=rcopula(frankCopula(fitFp@estimate,dim=2),10000)
xFn=rcopula(frankCopula(fitFn@estimate,dim=2),10000)
```

```
yGp=w[1]*quantile(data1[,1],xGp[,1])+w[2]*quantile(data1[,2],xGp[,2])
yGn=w[1]*qnorm(xGn[,1],fit1$estimate[1],fit1$estimate[2])+w[2]*qnorm(xGn[,1],fit2$estimate[1],fit2$estimate[2])
yTp=w[1]*quantile(data1[,1],xTp[,1])+w[2]*quantile(data1[,2],xTp[,2])
yTn=w[1]*qnorm(xTn[,1],fit1$estimate[1],fit1$estimate[2])+w[2]*qnorm(xTn[,1],fit2$estimate[1],fit2$estimate[2])
yGUp=w[1]*quantile(data1[,1],xGUp[,1])+w[2]*quantile(data1[,2],xGUp[,2])
yGUn=w[1]*qnorm(xGUn[,1],fit1$estimate[1],fit1$estimate[2])+w[2]*qnorm(xGUn[,1],fit2$estimate[1],fit2$estimate[2])
yCp=w[1]*quantile(data1[,1],xCp[,1])+w[2]*quantile(data1[,2],xCp[,2])
yCn=w[1]*qnorm(xCn[,1],fit1$estimate[1],fit1$estimate[2])+w[2]*qnorm(xCn[,1],fit2$estimate[1],fit2$estimate[2])
yFp=w[1]*quantile(data1[,1],xFp[,1])+w[2]*quantile(data1[,2],xFp[,2])
yFn=w[1]*qnorm(xFn[,1],fit1$estimate[1],fit1$estimate[2])+w[2]*qnorm(xFn[,1],fit2$estimate[1],fit2$estimate[2])
VaRGaussianC[i]=-quantile(yGp,si)*cpprice[i]
VaRGaussianC2[i]=-quantile(yGn,si)*cpprice[i]
VaRtC[i]=-quantile(yTp,si)*cpprice[i]
VaRtC2[i]=-quantile(yTn,si)*cpprice[i]
VaRGC[i]=-quantile(yGUp,si)*cpprice[i]
VaRGC2[i]=-quantile(yGUn,si)*cpprice[i]
VaRCC[i]=-quantile(yCp,si)*cpprice[i]
VaRCC2[i]=-quantile(yCn,si)*cpprice[i]
VaRFC[i]=-quantile(yFp,si)*cpprice[i]
VaRFC2[i]=-quantile(yFn,si)*cpprice[i]
}

#Backtesting methdology
#Christoffersen Test
CC<-function(ploss,VaR,p){
n=length(VaR)
v=sum(ploss>VaR)
LRuc=0
LRuc=2*(log((v/n)^v*(1-v/n)^(n-v))-log(p^v*(1-p)^(n-v)))
LRind=0
I=rep(0,n)
n0=n1=n00=n01=n10=n11=0
pi0=pi1=pi00=pi01=pi10=pi11=0
for(i in 1:n){if(ploss[i]>VaRmt[i]){I[i]=1}}
for(i in 2:n){
if(I[i-1]==0 & I[i]==0){n00=n00+1}
if(I[i-1]==1 & I[i]==1){n11=n11+1}
if(I[i-1]==0 & I[i]==1){n01=n01+1}
if(I[i-1]==1 & I[i]==0){n10=n10+1}
}
n0=n00+n10
n1=n01+n11
pi0=n0/n
pi1=n1/n
```

```
pi00=n00/(n00+n01)
pi01=n01/(n00+n01)
pi10=n10/(n10+n11)
pi11=n11/(n10+n11)
LRind=2*(log(pi00^n00*pi01^n01*pi10^n10*pi11^n11)-log(pi0^n0*pi1^n1))
LRcc=LRuc+LRind
return(list(LRuc=LRuc,LRind=LRind,LRcc=LRcc,chisq1=qchisq(0.95,df=1),chisq2=
    qchisq(0.95,df=2)))
}
#QPS statistic
QPS<-function(ploss,VaR,p){
n=length(VaR)
QPS=0
C=rep(0,n)
for(i in 1:n){if(ploss[i]>VaR[i]){C[i]=1}}
QPS=2*sum((C-p)^2)/n
return(QPS)
}
#RMSE statistic
RMSE<-function(ploss,VaR){
temp1=ploss[ploss<=VaR]
temp2=VaR[ploss<=VaR]
n=length(temp1)
RMSE=0
RMSE=sqrt((sum((temp1-temp2)^2))/(n-1))
return(RMSE)
}
```