Volatility Estimation of Multivariate Financial Time Series by ICA-GARCH Models

by

Yingwei Zhai

An Essay
presented to the University of Waterloo
in fulfillment of the
partial requirement for the degree of
Master of Mathematics
in
Statistics

Waterloo, Ontario, Canada, 2016

Abstract

In this essay, we propose an application of Independent Component Analysis (ICA) to decompose multivariate time series under the framework of a GARCH model, in order to estimate volatilities of financial return series. In the empirical analysis, we compare the ICA-GARCH model with other existing methods, including the PCA-GARCH, DCC and EWMA models. In addition, the Value at Risk (VAR) is computed by using the proposed methods, and the results of backtesting verify the effectiveness and stability of ICA-GARCH model in capturing the time-varying nature of volatilities.

Keywords: Multivariate Financial Time Series; Independent Component Analysis; Principle Component Analysis; GARCH Model; Value at Risk.

Acknowledgements

I would like to express my sincere gratitude to my advisor and mentor, Professor Tony Wirjanto, for his continuous support of the study and research throughout my master's program. It is his patience, inspiration, enthusiasm, and immense knowledge that help me present this paper.

Table of Contents

Chapter 1

Introduction

Modeling and forecasting multivariate time series has been one of the most popular subject in econometrics, as it can be applied broadly in many fields, such as risk management, asset pricing, and investment strategy. The multivariate volatility of financial time series, however, is difficult to capture, since it is complicated to estimate the parameters and its correlation structure in a given model has typical time-varying features. Hence there have been enormous efforts in developing the methodologies of volatility modeling and estimation.

One of the most important and famous volatility models for financial time series is the autoregressive conditional heteroscedasticity (ARCH) model. Firstly introduced by Engle in 1982 [9], it was then developed into the generalized ARCH (GARCH) model by Bollerslev (1986 [3]). Even though the univariate GARCH model has shown excellent performance of volatility estimation in financial time series, estimation of volatilities in high-dimensional settings, such as ones involved in a portfolio optimization, remains a challenging task for researchers. The main reason for this is that when the number of time series grows, the number of parameters to estimate in the multivariate GARCH models grows rapidly. Therefore, more efficient algorithms are demanded in order to manage this computationally prohibitive situation.

Wu, and Yu (2005 [17]) shows that Independent Component Analysis (ICA) is an effective way to reconstruct multivariate time series into statistically independent time series. Following this procedure, we firstly use vector autoregressive (VAR) model to find the residuals of return series. Then the ICA is applied to transform the residual series into statistically independent time series. Hence the volatility of each independent time series could be estimated by the univariate GARCH model. In other words, we can get the volatility of the initial multivariate time series by making linear transformations of the ICA procedure, with a relatively low cost of computation.

Additionally, Wu and Yu (2005[17]) also compares the PCA-GARCH model and the ICA-GARCH model. Both of these models are trying to find linear combinations of the residual series $\bar{r}$ unconditionally. However, PCA is only able to make the data uncorrelated while ICA renders the data independent. In addition, when we apply these methods to forecast the volatility of financial time series, ICA often performs better, since many financial time series follow a non-Gaussian distribution. To further evaluate the performance of ICA-GARCH, we also use other competing models to estimate the volatilities, such as Dynamic Conditional Correlation (DCC) and Exponentially Weighted Moving Average (EWMA) models.

Moreover, Value at Risk (VaR) is calculated based on the four models we proposed above. The reason of choosing VaR for model selection is its extensive use in risk management application. Both regulatory and financial institutions view VaR as a sufficient measurement of downside risk. Many studies are conducted based on the demand for a valid VaR estimation and prediction approach, such as the empirical research of VaR estimation using different GARCH models, presented by So and Yu (2005 [16]).

This essay is structured as follows: Chapter 2 briefly introduces several volatility models, which are the AR, ARCH, GARCH, DCC and EWMA models. Chapter 3 reviews the ICA algorithm, and extend this method into ICA-GARCH and PCA-GARCH for multivariate volatility modeling. Chapter 4 discusses the estimation of Value at Risk. Criteria of backtesting the VaRs are also constructed to justify the effectiveness of our proposed models. Eight stocks selected from the New York Stock Exchange market are used to compare the performance of different models by constructing the VaRs in Chapter 5. Chapter 6 concludes the essay.

Chapter 2

Volatility Models

Volatility is a statistical measure for the dispersion of returns in a given security or market index, which is commonly assessed by either standard deviation or variance of financial asset returns. Specifically, volatility can be viewed as the amount of uncertainty or risk in the change of a security's value. Generally speaking, a higher volatility means that a security's value can change dramatically over a short time period, either higher or lower. A lower volatility means that a security's value does not have great fluctuations, but may change at a stable pace over a certain period of time. In financial engineering, volatility estimation is a critical component in risk management, derivative pricing and hedging, market making and timing, portfolio selection and many other financial activities.

2.1 AR Model

Serial autocorrelations often characterize observed time series. As a result we can use past observations to predict current observations of a given time series. An autoregressive (AR) process models a time series value by regressing it on previous values from that same time series. An AR process depending on the p past observations of time series is called an AR model of degree p, which is denoted by AR(p). The AR(p) model can be defined as

$$r_t = \phi_0 + \phi_1 r_{t-1} + ... + \phi_p r_{t-p} + a_t \tag{2.1}$$

where a_t is assumed to be a uncorrelated innovation process, which is called a white noise series, with mean zero and variance σ^2. ϕ_0 is a constant that models the trend in this time series. When $\phi_0 > 0$, the trend is upwards; when $\phi_0 < 0$, the trend is downwards.

$\phi_1, \phi_2, ..., \phi_P$ are the lag coefficients, which determine how stable the process is. We usually restrict an autoregressive model to stationary data, and we have some constraints on the values of the parameters. For AR(1) model, the process is said to be second-order stationary when $|\phi_1| < 1$.

A vector autoregressive (VAR) model generalizes the univariate AR model for multivariate time series $\{\bar{r}_t\}$. A VAR model of degree p, denoted VAR(p), can be defined as

$$\bar{r}_t = \Phi_0 + \Phi_1 \bar{r}_{t-1} + ... + \Phi_p \bar{r}_{t-p} + \bar{a}_t \tag{2.2}$$

where $\{\bar{a}_t\}$ is a serially uncorrelated random vector with mean zero and covariance Σ. Φ_0, ... , Φ_p are n-dimensional vectors and n is the number of elements in the time series. In this model, we generally use the first p lags of each variable as the regression predictors for each variable.

2.2 ARCH Model

Motivated by volatility clustering, an Autoregressive Conditional Heteroskedasticity model of degree p, ARCH (p), was introduced by Engle (1982 [9]) to model volatility of financial asset return series as a second-order stationary solution of the equation

$$Z_t = \sqrt{h_t} e_t, \tag{2.3}$$

where $\{e_t\} \sim$ i.i.d. N(0,1) and h_t, as a positive function of $Z_s (s < t)$, is defined as

$$h_t = \alpha_0 + \Sigma_{j=1}^{p} \alpha_j Z_{t-j}^2, \tag{2.4}$$

with $\alpha_0 > 0$ and $\alpha_j \geqslant 0$, j = 1, 2, ..., p. Given $\{Z_s, s < t\}$, h_t is the conditional variance of Z_t.

2.3 GARCH Model

The time-dependent market information is able to influence the volatility of financial returns considerably. For such time series, the generalized autoregressive conditional heteroscedasticity (GARCH) process is a useful way to model and forecast the returns of financial indexes. The GARCH(p,q) model process (Bollerslev, 1986 [3]) is a generalization of the ARCH model with a replaced variance equation. In the GARCH models, we assume financial time series to be generated by a stochastic process with volatility fluctuations being characterized by squared past innovations and squared past volatilities. More specifically, a GARCH (1,1) model for a return series $\{y_t\}$ is defined as

$$y_t = \mu_t + \epsilon_t = \mu_t + \sqrt{h_t} z_t, \tag{2.5}$$

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{2.6}$$

where μ_t denotes the conditional mean of y_t based on information at time t-1, and can be described by a regression model; ϵ_t denotes the innovation; h_t is the conditional variance of y_t conditioning on the information up to time t-1; and the z_t's are independent random variables with zero mean and unit variance, i.e., $z_t \sim N(0,1)$. In order to ensure a positive h_t, we impose the restrictions of $\alpha_0 > 0$, $\alpha_1 > 0$, and $\beta_1 > 0$ (The restrictions of parameters may be weaker according to Nelson and Cao, 1992 [14]).

For univariate financial time series, the GARCH(1,1) model is capable to model the volatility for most of the times. In this paper, the GARCH model we use is able to fit the time series with different distributions of z_t, such as Gaussian and Students' t distribution. In Engle's original ARCH model [9], z_t is assumed to follow the Gaussian distribution. The maximum likelihood estimation is used as the main approach to estimate parameters in GARCH model. Bollerslev and Wooldridge (1992 [6]) indicates that the normality of z_t is often assumed and produces the quasi-maximum likelihood estimate, even though the standardized residual z_t does not follow the Gaussian distribution in many empirical studies.

Since the financial return series are often correlated, it is crucial not only to model and forecast the time-varying volatility of each series, but also the time-varying correlation among them. Therefore a development is needed to evolve the univariate GARCH model into a multivariate one.

Assume that we have a return series $\mathbf{r}_t = (r_{1t}, r_{2t}, ..., r_{nt})^T$. Let $\boldsymbol{\epsilon}_t$ be the mean-centered vector of $\mathbf{r}_t$, which is acquired by subtracting the mean from $\mathbf{r}_t$, and let $\mathbf{H}_t$ denote the conditional covariance matrix of $\mathbf{r}_t$, which is positive definite and measurable with respect to the information set $\boldsymbol{\Phi}_{t-1}$. To model the multivariate conditional covariance matrix, vech model is usually applied (Bollerslev et al., 1988 [4]). Let vech denote the vector-half operator, which stacks the lower triangular elements of an $N \times N$ matrix into a $[N(N+1)/2] \times 1$ column vector. Since the conditional covariance matrix $\mathbf{H}_t$ is symmetric, vech$(\mathbf{H}_t)$ contains all the unique elements in $\mathbf{H}_t$. Hence it models $\mathbf{H}_t$ as a linear combination of the lagged squared errors, the cross-products of errors, and the lagged elements in $\mathbf{H}_t$.

However, as the dimension of data increases, the number of parameters in this model grows rapidly. Consequently there exists a great computational cost in parameterizing the multivariate GARCH for data with high dimension (Bauwens et al. 2006 [2]). Nevertheless, in reality we often need to use a high-dimensional volatility model. For instance, in portfolio optimization, the portfolio may contain hundreds of capital assets. In order to handle this situation, new methods are needed to estimate the parameters more efficiently.

2.4 DCC Model

One development of the multivariate GARCH model is the Dynamic Conditional Correlation (DCC) model, which reveals important time-varying features in financial time series (Engle, 2002 [8]). It is a generalization of Constant Conditional Correlation (CCC) model (Bollerslev 1990 [5]). The CCC model considers the time variation in conditional covariance to be caused by the time variation in conditional variance of each return series, which can be denoted by

$$\boldsymbol{H}_t = \boldsymbol{D}_t \boldsymbol{R} \boldsymbol{D}_t \tag{2.7}$$

where $\boldsymbol{D}_t = diag\{\sqrt{h_{it}}\}$ and h_{it} stands for the conditional variance of the ith return.

In the DCC model, the correlation matrix $\mathbf{R}$ is allowed to be time-varying. For N return series $\boldsymbol{y_t} = (y_{1t}, y_{2t}, ..., y_{Nt})^T$, the DCC model is defined as

$$\boldsymbol{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t = \boldsymbol{\mu}_t + \boldsymbol{H}_t^{\frac{1}{2}} \boldsymbol{z}_t \tag{2.8}$$

$$\boldsymbol{H}_t = \boldsymbol{D}_t \boldsymbol{R}_t \boldsymbol{D}_t \tag{2.9}$$

where the conditional correlation matrix $\boldsymbol{R}_t$ is determined by another matrix $\boldsymbol{Q}_t$:

$$[\boldsymbol{R}_t]_{i,j} = \rho_{i,j,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t} q_{jj,t}}} \tag{2.10}$$

$$\boldsymbol{Q}_t = [q_{ij,t}] = \bar{\boldsymbol{Q}}(1 - \alpha - \beta) + \alpha(\boldsymbol{u}_{t-1} \boldsymbol{u}'_{t-1}) + \beta \boldsymbol{Q}_{t-1} \tag{2.11}$$

Here $\boldsymbol{z}_t$ is an independent and identically distributed random vector with mean zero and identity covariance matrix; $\boldsymbol{D}_t = diag(\boldsymbol{H}_t^{\frac{1}{2}})$, where $h_{ii,t}$, the conditional variance of y_{it} at time t, is modeled by univariate GARCH(1,1); $\boldsymbol{Q}_t$ is the covariance matrix of $\boldsymbol{u}_t$ with each $u_{i,t} = \epsilon_{i,t}/\sqrt{h_{ii,t}}$, which is used to model the dynamic structure of the standardized residual by univariate GARCH(1,1); and $\bar{\boldsymbol{Q}}_t$ denotes the unconditional covariance matrix of $\boldsymbol{z}_t$. The DCC model is degenerated to the CCC model when $\alpha = \beta = 0$.

2.5 EWMA Model

Another common approach for multivariate volatility estimation is to employ the Exponentially Weighted Moving Average (EWMA) model. It assigns weights to the periodic returns and hence improves on simple variance, which is defined as the average of the squared returns. In fact, the forecast of volatility at time t is a weighted average of both the volatility forecast and the corresponding squared error at time t-1. The EWMA model enables us to use a large sample size while give greater weight to more recent returns at the same time.

For univariate time series, the time-varying volatility is modeled by

$$h_t = (1-\lambda)\epsilon_{t-1}^2 + \lambda h_{t-1} \tag{2.12}$$

where ϵ_t denotes the residual series and λ is called the decaying factor or smoothing parameter ($\lambda < 1$). In this case, each weight is a constant multiplier (i.e., λ) of the prior day's weight, which explains the meaning of "exponential" in EWMA. Therefore the variance is weighted toward the most recent data. In practice, $\lambda = 0.94$ for the daily financial time series.

In a multivariate case, the covariance $h_{ij,t}$ of $y_{i,t}$ and $y_{j,t}$ can be estimated by the exponential weighting

$$h_{ij,t} = (1-\lambda)\epsilon_{i,t-1}\epsilon_{j,t-1} + \lambda h_{ij,t-1}. \tag{2.13}$$

If i=j, (2.13) is degenerated to (2.12).

The EWMA model is a particular case of the GARCH model, with $\alpha_0 = 0$ and $\alpha_1 + \beta_1 = 1$. One salient difference is that the GARCH model has a feature of mean reversion, which means that prices or returns will tend to move to the average over time, but the EWMA model does not. In addition, the EWMA model assumes the return series to follow the Gaussian distribution, while the GARCH model is able to support return series with other distribution, such as the Students' t distribution.

Chapter 3

The ICA-GARCH Models

In this chapter, we first introduce the ICA approach, explain the process of applying ICA into GARCH models, and then extend our method with a similar procedure, the PCA-GARCH model.

3.1 The ICA Model

Multivariate time series are often viewed as obscure measurements arising from certain underlying sources, which cannot be directly observed. Factor analysis is a classical statistical procedure that aims at identifying these latent sources. Independent component analysis (ICA) is a particular case of factor analysis, which can express the observed data in terms of a combination of latent variables. In this essay, we only consider a linear combination of these variables. Assume that the underlying latent variables are non-Gaussian distributed and statistically independent with each other. Our goal is to identify the latent variables and the corresponding mixing process. The ICA model, therefore, can be defined as

$$\boldsymbol{y}_t = \boldsymbol{A}\boldsymbol{s}_t \tag{3.1}$$

where $\boldsymbol{y}_t$ is the vector of the observed random variables, $\boldsymbol{s}_t$ is a vector of mutually independent latent variables, which are called independent components, and $\boldsymbol{A}$ is a mixing matrix with unknown constants. In other words, the number of observed time series is equal to the number of sources. Here we only discuss the situation when $\boldsymbol{A}$ is a square matrix.

To find the independent component $\boldsymbol{s}_t$, we need to calculate a matrix $\boldsymbol{W}$ such that $\boldsymbol{s}_t = \boldsymbol{W}\boldsymbol{y}_t$ (i.e., $\boldsymbol{W} = \boldsymbol{A}^{-1}$) up to some indeterminacies. One popular algorithm is the

FastICA proposed by Hyvarinen and Oja (1997, 2000 [11] [12]), which is a fast fixed point approach to separate independent sources signals. The FastICA algorithm is a robust fixed-point type of algorithm for an independent component analysis, which is efficient in computation. It seeks the extreme of $E\{G(\boldsymbol{w'y})\}$, where G stands for a non-quadratic function, such as $G(x) = log(cosh(x))$.

First of all, we transform the observed vector linearly to get a new vector $\bar{y}$, whose components are uncorrelated and variances equal unity, i.e., $\bar{y}$ is white. To be specific, the covariance matrix of $\bar{y}$ is equivalent to the identity matrix. It is a common way to use the Principle Component Analysis (PCA) to whiten the data. Then we apply the iterative fixed-point algorithm to find each unit, which can be defined as

$$\tilde{\boldsymbol{w}}_{n+1} = E[\bar{\boldsymbol{y}}g(\boldsymbol{w}'_n\bar{\boldsymbol{y}})] - E[g'(\boldsymbol{w}'_n\bar{\boldsymbol{y}})]\boldsymbol{w}_n \tag{3.2}$$

where g is the derivative of the non-quadratic function G and $\boldsymbol{w}_{n+1} = \frac{\tilde{\boldsymbol{w}}_{n+1}}{||\tilde{\boldsymbol{w}}_{n+1}||}$. After the estimation of $\boldsymbol{w}$, we can get an Independent Component (IC) by s = $\boldsymbol{w'y}$. Extending the algorithm above, we acquire the entire ICA transformation as $\boldsymbol{s}_t = \boldsymbol{W}\boldsymbol{y}_t$. In addition, the outputs $\boldsymbol{w}'_1\boldsymbol{y}, \boldsymbol{w}'_2\boldsymbol{y}, ..., \boldsymbol{w}'_n\boldsymbol{y}$ are decorrelated after each iteration in order to avert convergence to the same ICs. After estimating n ICs, i.e., n vectors $\boldsymbol{w}_1, \boldsymbol{w}_2, ..., \boldsymbol{w}_n$, we run the one-unit fixed-point algorithm for $\boldsymbol{w}_{n+1}$. We also subtract the projections of the previously estimated n vectors from $\boldsymbol{w}_{n+1}$ after each iteration procedure, and then $\boldsymbol{w}_{n+1}$ will be renormalized as follows:

$$\tilde{\boldsymbol{w}}_{n+1} \leftarrow \tilde{\boldsymbol{w}}_{n+1} - \sum_{j=1}^{n} \boldsymbol{w}_j\boldsymbol{w}'_j\tilde{\boldsymbol{w}}_{n+1} \tag{3.3}$$

The decorrelation scheme stated above is applicable to deflationary separation of the ICs.

3.2 ICA-GARCH Modeling Procedure

The ICA-GARCH model works as follows: first of all, we employ the AR(p) model, which is defined as (2.1), to obtain the residuals of the return series. Then the Bayesian Information Criterion (BIC), or the Schwarz Criterion (Schwarz, 1978 [15]), is applied to determine the unknown degree p.

BIC is an essential statistical criterion for model selection among a finite set of models. It is intended to select a model from a group of candidates by maximizing the posterior probability. In other words, BIC deals with models estimated by the maximum likelihood method, and the model with the lowest BIC value is selected. The BIC statistic is defined as follows:

$$BIC = -2LLF + N_{para} ln(N_{obs}) \tag{3.4}$$

where LLF stands for a vector of optimized log-likelihood objective function values, which are associated with the parameter estimates of the models to be justified, N_{para} is the number of estimated parameters corresponding to each value of LLF, and N_{obs} is the sample size of observed return series that are correlated with each LLF value. The second term of (3.4), $N_{para} ln(N_{obs})$, is a penalty function which penalizes a candidate model by the number of parameters used.

After choosing the appropriate degree p of the AR models for each return series, we employ the ICA procedure to decompose $\boldsymbol{a}_t$, the residual vector derived from the adapted AR model, into Independent Components $\{s_{i,t}\}$, $i = 1, ..., N$. In other words, $\boldsymbol{a}_t = \boldsymbol{A}\boldsymbol{s}_t$, while $\boldsymbol{s}_t = (s_{1,t}, s_{2,t}, ..., s_{N,t})'$. Then the univariate GARCH models are applied to determine the volatilities of each independent component series. Accordingly, we obtain the conditional covariance matrix of the return series at time t, by using the estimated mixing matrix $\boldsymbol{A}$, which can be expressed as

$$\boldsymbol{H}_t = \boldsymbol{A}\boldsymbol{V}_t\boldsymbol{A}' \tag{3.5}$$

where $\boldsymbol{V}_t$ represents a diagonal matrix and its diagonal elements can be viewed as the volatilities of independent components $\boldsymbol{s}_t$.

This algorithm does not significantly raise the computational complexity, but still holds a great precision at the same time, since the N components of $\boldsymbol{s}_t$ are independent. Hence the multivariate volatilities of N returns series can be generated from N univariate GARCH models under the ICA procedure. The other multivariate GARCH models, however, may encounter computational laboriousness in parameter estimation as the number of time series increases. Furthermore, there is no restrictions of data dimension for the ICA-GARCH model, since we only use the univariate GARCH models to estimate the volatilities.

3.3 PCA-GARCH Modeling Procedure

One comparable method of ICA-GARCH model is called the Principal Component GARCH (PCA-GARCH) model or orthogonal GARCH model, which was firstly presented by Alexander (2001 [1]). The Principal Component Analysis is used for extracting the key factors from the orthogonal GARCH model, in order to capture as much of the information as possible. To be specific, we first find the appropriate AR(p) model to obtain the residuals of the observed return series, $\boldsymbol{a}_t = (a_{1,t}, a_{2,t}, ..., a_{N,t})'$. Next we apply the PCA-GARCH procedure to find a matrix $\boldsymbol{B}$ such that

$$\boldsymbol{a}_t = \boldsymbol{B}\bar{\boldsymbol{s}}_t \tag{3.6}$$

and

$$E[\bar{\boldsymbol{s}}_t\bar{\boldsymbol{s}}_t'] \equiv \boldsymbol{V} \tag{3.7}$$

where $\boldsymbol{V}$ is a diagonal matrix. Furthermore, the univariate GARCH processes are employed to model the volatility of each principal components, using the method of maximum likelihood.

Both the PCA-GARCH and ICA-GARCH models aim at composing the uncorrelated linear combinations for $\boldsymbol{a}_t$. A major difference between the models is that ICA-GARCH makes the source signals independent while PCA-GARCH makes the source signals only uncorrelated. Consequently, we expect the ICA-GARCH model to have better performance than the PCA-GARCH model in the empirical study, since financial return series often exhibit non-Gaussian distributions.

Chapter 4

The ICA-GARCH Models in VaR: Volatility Estimation and Forecasting

In this chapter, we first briefly introduce the concept of Value at Risk (VaR), then present the estimation procedure of VaR by applying the ICA-GARCH model, which is for the estimation of the multivariate volatilities using a set of univariate GARCH models.

4.1 VaR

Value at Risk (VaR) is a broadly recognized measure in risk management. It is used to quantify and manage the level of financial risk within a company or investment portfolio over a particular time frame. In other words, VaR is measured in three variables: the maximal amount of possible loss, the probability of that amount of loss, and the period of time. Under a probabilistic structure, it can be defined as

$$p = Pr(\Delta V_{t+1} \geqslant VaR) \tag{4.1}$$

where p is the pre-determined probability of interest, such as p = 1 % or p = 5%, and ΔV_{t+1} is the profit or loss of the investment portfolio during the next time unit. For an specific financial asset, $\Delta V_{t+1} = Q_0(P_{t+1} - P_t)$ where Q_0 stands for the amount of underlying asset and P_t represents the market price of this asset at time t. In practice, ΔV_{t+1} can also represent the log-returns of the asset.

Our goal is to first find estimate the volatilities of multivariate time series by using the methods proposed in Chapter 2 and Chapter 3, and compute the time-varying VaRs. As a matter of fact, the performance of VaR estimation highly depends on the accuracy of volatility estimation, hence we mainly focus on the efficiency of our proposed models in VaR estimation.

In evaluating the efficiency of our proposed models with the respect of VaR estimation, the time series data is partitioned into in-sample data and out-of-sample data. The in-sample data is utilized to train the volatility model, while the out-of-sample data is reserved for the forecasting performance of VaR.

4.2 Estimation of VaR

The computation of portfolio VaR proceeds as follows: To begin with, we fit the ICA-GARCH, PCA-GARCH, EWMA and DCC models to the multivariate return series, one at a time. After obtaining the estimated parameters for each model, we use them to forecast the volatilities within the next time frame. Hence we predict the VaR performance by the volatility forecast.

For the most common cases, we hold the assumption that the return series follows the Gaussian distribution. Then the VaRs can be calculated by a relatively straightforward equation as

$$VaR_{t+1} = \alpha_p \sqrt{\hat{h}_{t+1} \Delta t} \tag{4.2}$$

where α_p represents the value of Z corresponding to a pre-specified significant level, such as $\alpha_{0.05} = 1.645$, $\hat{h}_{t+1}$ represents the forecasted conditional variance of the time t+1, and Δt denotes the period of time for VaR calculation. In this essay, we assume $\Delta t = 1$, i.e., the daily VaRs are analyzed.

In order to forecast the out-of-sample VaR, we need to obtain the Independent Components of out-of-sample data to fit the ICA-GARCH models. The ICs can be computed as

$$\boldsymbol{s}_o = \boldsymbol{A}^{-1} \boldsymbol{y}_o \tag{4.3}$$

where $\boldsymbol{A}$ is the mixing matrix, $\boldsymbol{y}_o$ is the out-of-sample return series, and $\boldsymbol{s}_o$ is the out-of-sample ICs we desired. Furthermore, the Principal Components of out-of-sample data can also be determined in a similar way. Therefore we can apply (4.2) to derive the VaRs of ICA-GARCH, PCA-GARCH, EWMA and DCC models and develop the procedure of model selection.

4.3 Backtesting VaR

Backtesting is a technique that simulates a model based on the past data to assess its accuracy and efficiency. Backtesting in VaR is a statistical algorithm for comparing the forecasting losses estimated from the calculated VaR with the actual losses received at the end of the specified period. This comparison identifies the time frame where the VaR is underestimated or where the portfolio losses are greater than the initial expected VaR. If the backtesting results are not satisfactory, the forecasting VaR can be recalculated to reduce the risk of unexpected losses.

Computing the failure rate is one commonly used method to substantiate the precision of a model. The failure or violation rate stands for the proportion of times when VaRs are greater than the actual loss in a given sample, i.e., the violation means $X_t < VaR_t$ where X_t is the actual profit or loss between the end time t and t + 1. Hence the counting function of violation can be defined as

$$I_t = \begin{cases} 1, \text{ if } X_t < VaR_t; \\ 0, \text{ otherwise} \end{cases} \tag{4.4}$$

and we define the number of violations (NoV) to be

$$E[I_t] = N(1-c) \tag{4.5}$$

where c is the confidence level, e.g., c = 99%. Furthermore, a likelihood-ratio (LR) test statistic, firstly introduced by Christoffersen (1998 [7]), is defined to develop the backtesting in VaR:

$$LR = 2log\left[\left(\frac{\gamma^*}{\gamma}\right)^n \left(\frac{1-\gamma^*}{1-\gamma}\right)^{N-n}\right] \sim \chi^2_{(1)} \tag{4.6}$$

where n denotes the NoV, N stands for the sample size, γ is the confidence level and γ^* represents the ratio between NoV and the sample size, which is also called the frequency of the occurrence of violations. At the significant levels of 1% and 5%, the critical values of this test statistic are accordingly 6.635 and 3.841. Based on the value of γ^*, we can also obtain the actual violation rate(ART), with $E[ART] = 1 - c$.

In addition, Lopez (1998 [13]) suggests constructing of a general loss function, $L(VaR_t, X_t, t+1)$, to measure the performance of the predicted VaR through the data of profits and losses in the past and the reported VaR series:

$$S_t = \begin{cases} 1 + (X_t - VaR_t)^2, \text{ if } X_t < VaR_t; \\ 0, \text{ otherwise} \end{cases} \tag{4.7}$$

Note that this quadratic function measures the extent of violation, or the Sum Square of Violations (SSV). Furthermore, we use the Average Size of Violations (ASV) as another approach to evaluate the calculated VaR, which can be defined as follows:

$$A_t = \begin{cases} \frac{X_t - VaR_t}{VaR_t}, \text{ if } X_t < VaR_t; \\ 0, \text{ otherwise} \end{cases} \tag{4.8}$$

The ASV intends to combine the frequency with the size of the exceptions, by focusing on the average size of the exception.

Chapter 5

Empirical Analysis

In this chapter, the empirical analysis are divided into two components: the first component focuses on testing and evaluating the performance of ICA-GARCH model as well as other competing models in modeling the volatility of multivariate time series. The second component is concerned with backtesting the forecasted VaR and checking the adequacy of the ICA-GARCH model as well as other competing models.

5.1 Data Description

We use eight stocks from the New York Stock Exchange (NYSE) market as our portfolio in this experiment, which are (1) American Telephone & Telegraph (T), (2) General Electric (GE), (3) Tiffany (TIF), (4) Bank of America (BAC), (5) McDonald's (MCD), (6) CVS Health (CVS), (7) Pepsico (PEP), (8) NIKE (NKE). The sample period is from January 2, 2002 to December 31, 2007, with 1510 daily observations in total.

The dataset are split into two sections in order to evaluate different models: the first section is the in-sample data for model training, containing the first 1310 observations from our sample; the second section is the out-of-sample data reserved for forecasting assessment, including the additional 200 observations.

In addition, the daily log-returns $y_i(t)$ are calculated by

$$y_i(t) = log(P_i(t)) - log(P_i(t-1)) \tag{5.1}$$

where $P_i(t)$ stands for the closing price of stock i on the trading day t. Table 1 shows a brief summary of our portfolio:

Table 1. Summary Statistics of Daily Log-return (In-sample Data)

STOCKS	MEAN	VARIANCE	SKEWNESS	KURTOSIS	MINIMUM	MAXIMUM
T	0.0000	0.0002960088	-0.08059955	4.237184	-0.1076626	0.08704414
GE	0.0000	0.0002336867	0.2276687	5.486546	-0.0973012	0.09107721
TIF	0.0000	0.0004702416	0.1529996	4.891599	-0.1602029	0.1160773
BAC	0.0000	0.0001501546	-0.7292968	8.547178	-0.1072804	0.07924406
MCD	0.0000	0.0002775743	0.290029	6.931586	-0.1376094	0.08785828
CVS	0.0000	0.0002931091	0.206911	4.930442	-0.08788058	0.0858646
PEP	0.0000	0.0001514232	0.7832987	21.40683	-0.1075421	0.1383693
NKE	0.0000	0.0002208809	0.3217195	6.189952	-0.07053208	0.08195612

Then the VAR(p) model is used to filter the autocorrelation within the multivariate return series where p is determined by use of the BIC. Figure 5.1 is the plot of the eight residual series after fitting the proper vector autoregressive (VAR) model.

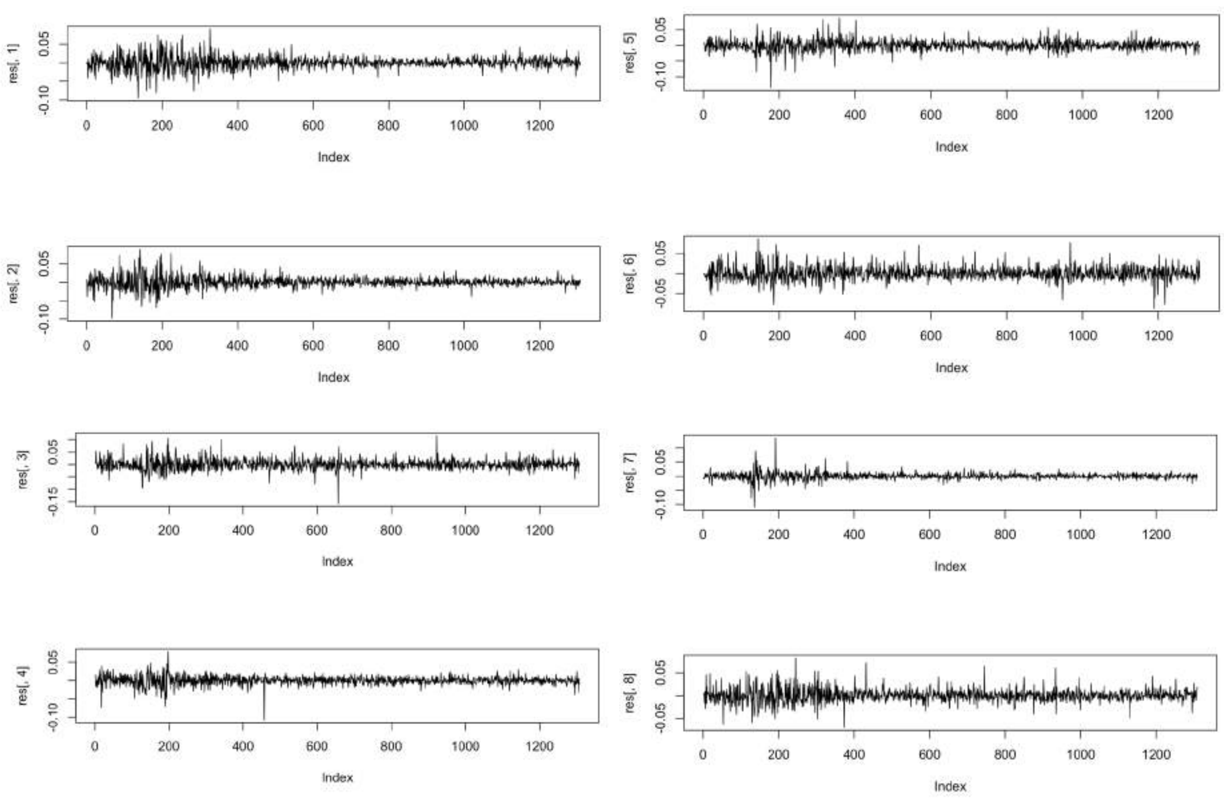

Figure 5.1: Residual Return Series Obtained by the VAR Model

Briefly, the plotted residuals in Figure 5.1 suggest that they are approximately serially uncorrelated in the return series, although they are obvious persistence in the squared return series exhibited by highly uneven widths with which each of the return series fluctuates over time.

5.2 Multivariate volatilities modeling

Before applying the ICA algorithm, we need to make sure that there is no more than one non-kurtic component in our portfolio (Hyvarinen, Karhunen, and Oja, 2001 [10]). According to Table 1, there exists significant excess kurtosis value of each return series, which demonstrates that the data we used follows a non-Gaussian distribution. Then we apply the ICA procedure stated in Section 3.1 and obtained eight independent components. The GARCH (1,1) model is employed to fit the ICs. Table 2 presents the estimates of the GARCH coefficients, where we find the significant GARCH effects through most ICs. Most of β_1's (the GARCH coefficients) are of high values, and the sum of α_1 and β_1, i.e. the ARCH coefficient and the GARCH coefficient, is close to one.

Table 2. Estimates of Independent Components by GARCH (1,1)

INDEPENDENT COMPONENTS	α_0	α_1	β_1
IC1	0.5215 (0.06493)	0.3402 (0.06485)	0.2036 (0.06817)
IC2	0.001518 (0.0008491)	0.02755 (0.00533)	0.9701 (0.005186)
IC3	0.004313 (0.002311)	0.0412 (0.007894)	0.9552 (0.007791)
IC4	0.20945 (0.1276)	0.081241 (0.03079)	0.70966 (0.1522)
IC5	0.00199 (0.000933)	0.02663 (0.005268)	0.9698 (0.005262)
IC6	0.0078755 (0.003189)	0.028429 (0.005799)	0.96412 (0.007032)
IC7	0.00189 (0.000817)	0.0082315 (0.00136)	0.98912 (0.001432)
IC8	0.0024306 (0.001435)	0.035649 (0.006216)	0.96193 (0.005922)

Note: Standard errors are shown in the brackets.

In order to evaluate the performance of ICA-GARCH model, we also consider the PCA-GARCH, DCC, and EWMA models for volatility estimation. The conditional volatilities of the four models are presented in Figure 5.2 to Figure 5.5 respectively.

There are several findings through the model comparison:

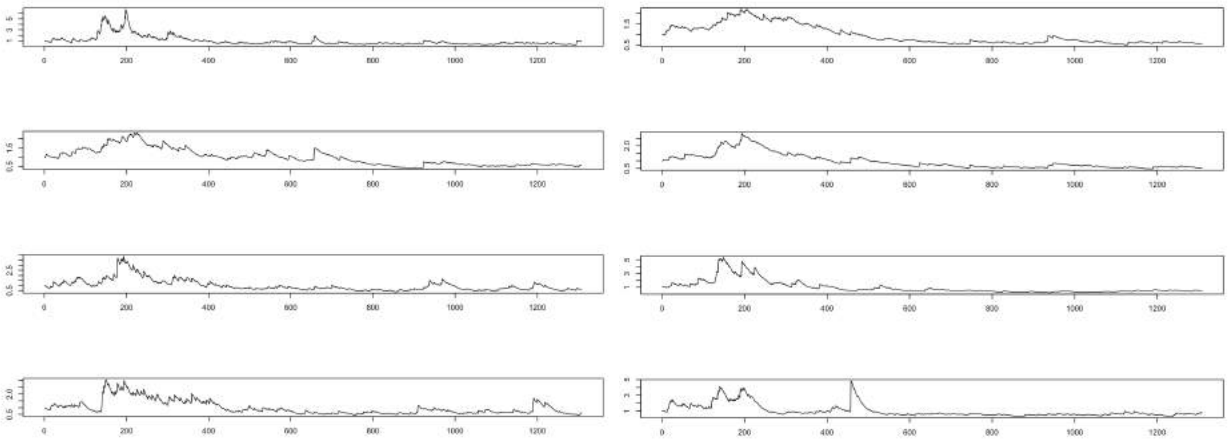

Figure 5.2: Conditional volatilities by ICA-GARCH Model

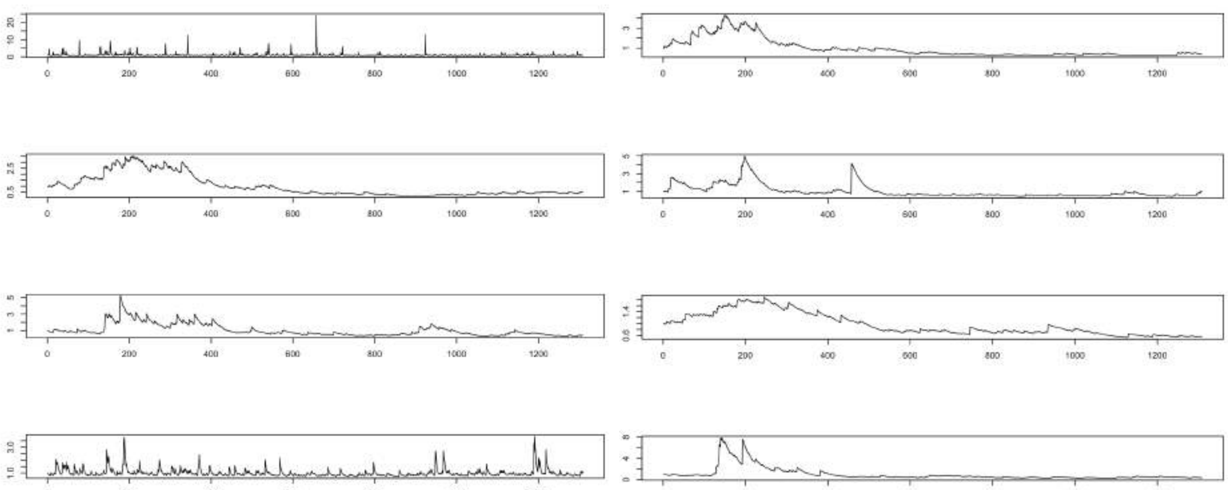

Figure 5.3: Conditional volatilities by PCA-GARCH Model

1. volatilities derived by PCA-GARCH fluctuate more than the result of ICA-GARCH, when the residuals have a great change, especially for the first stock. PCA-GARCH also makes the volatilities of different series oscillate together, even when the correlation coefficients of these time series are comparatively small. For instance, most of the volatility series have spikes around the 200th and the 400th observation. As a matter of fact, PCA-GARCH may give us the incorrect information in volatility, such as the 650th observation in the first plot and the 480th observation in the second plot

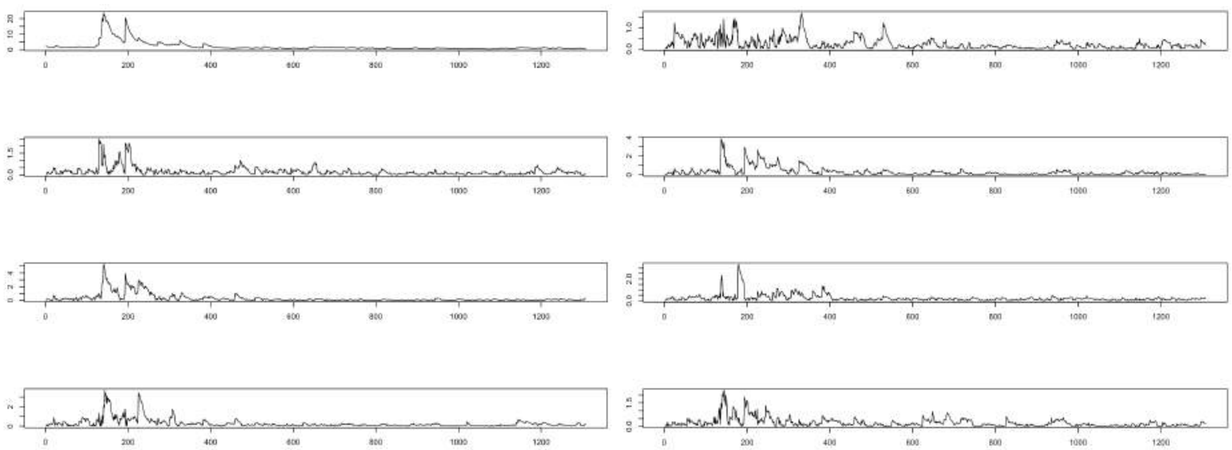

Figure 5.4: Conditional volatilities by EWMA Model

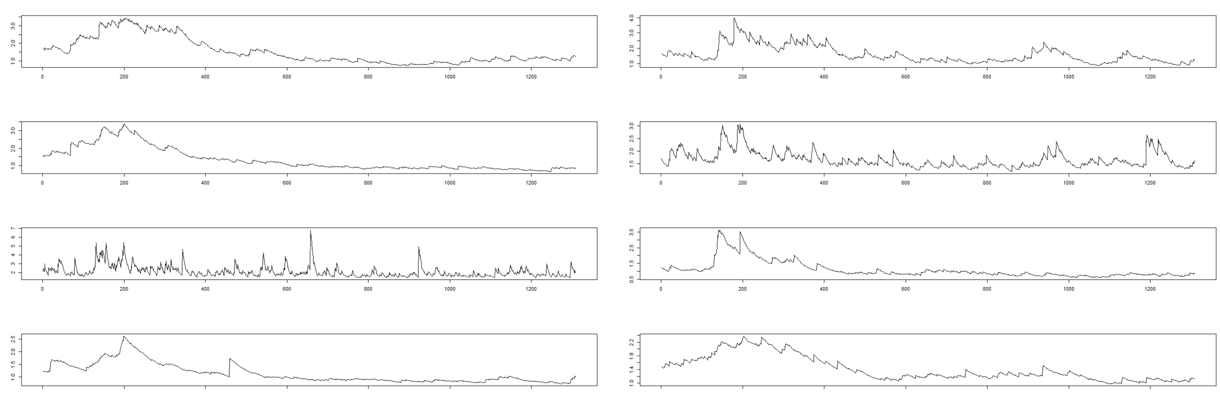

Figure 5.5: Conditional volatilities by DCC Model

of the second row. On the other hand, ICA-GARCH has a stronger power to predict the volatility. For example, ICA-GARCH is better at modeling the magnitude of outliers in the sample.

2. The patterns of volatilities modeled by ICA-GARCH models are the smoothest, while the other three models provides more pointed volatility series.

3. The DCC also seems to have a satisfactory volatility estimation for each return series. However, we can see that the spikes of each volatility is slightly left-skewed. This does not happen when we use other models.

5.3 Empirical Results of Backtesting VaRs

In this section, we use the models presented above to calculate the VaR at the 99 percent confidence level. For a better comparison, the empirical VaRs by each model and the original return series are plotted respectively in Figure 5.6. We can see that the ICA-GARCH, PCA-GARCH, and DCC models tend to provide similar VaR outcomes. The ICA-GARCH and PCA-GARCH, however, tend to be provide more pointed patterns than the DCC and EWMA models. In fact, the shape of VaRs by ICA-GARCH is even more sharpened than the shape of VaRs by PCA-GARCH. This result suggests that the ICA-GARCH model has much better performance in fitting a portfolio with a comparatively large number of assets where the linearity and normality assumptions are not satisfied.

Furthermore, the backtesting outcomes are shown in Table 3a and 3b. The in-sample data is used in the first part to evaluate the effectiveness of model estimation, while the out-of-sample data is used in the second part to assess the prediction abilities of different models. NoV, ASV and SSV are the measurement metrics we use to evaluate the VaR estimation. Our findings are listed as follows:

1. In general, the ICA-GARCH and the DCC model provides the best results for both in-sample data and out-sample data, as these two models produce the smallest test results. In fact, VaR of a portfolio with optimized weight of the stocks may provide a more significant evidence of the superior performance of the ICA-GARCH model.

2. EWMA tends to have the lowest accuracy in VaR estimation, both at the confidence level of 95% and 99%. The reason is that the EWMA model assumes the return to follow the Gaussian distribution. In practice, however, the distributions of most financial time series appear to have fat-tails.

3. There exists a similar performance trend between ICA-GARCH and PCA-GARCH, in both in-sample and out-of-sample data. Furthermore, the backtesting results of PCA-GARCH sometimes have fewer violations than the results of ICA-GARCH, such as the VaRs for CVS, PEP and NKE. However, figure 5.3 indicates that the volatilities estimated by PCA-GARCH tend to fluctuate simultaneously, which weakens the accuracy of this model for VaR estimation. Overall, we believe the ICA-GARCH model is superior than the PCA-GARCH model, and suggest to apply the DCC model to verify the VaR calculation.

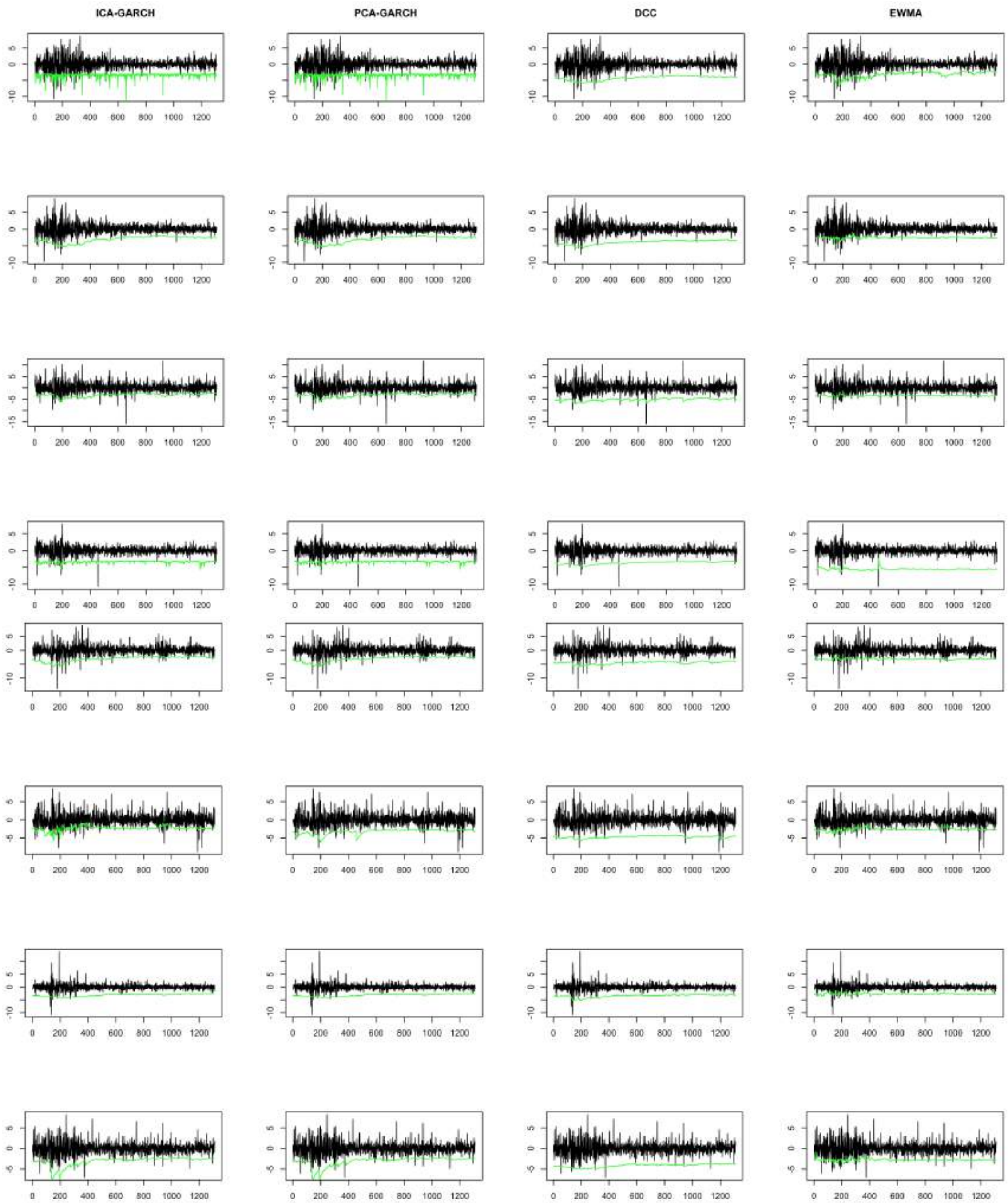

Figure 5.6: Return and VaR at 99% Confidence Level

Table 3a. Backtesting VaR at 95% Confidence Level with In-sample Data

NUMBER OF VIOLATION	ICA-GARCH	PCA-GARCH	DCC	EWMA
T	138	142	56	122
GE	64	59	45	384
TIF	550	586	89	507
BAC	97	103	72	135
MCD	102	97	102	459
CVS	255	82	76	510
PEP	85	77	65	263
NKE	92	90	70	784

Expected number of violations: 65

AVERAGE SIZE OF VIOLATION	ICA-GARCH	PCA-GARCH	DCC	EWMA
T	0.01264982	0.013110728	0.009345574	0.01212362
GE	0.00824394	0.007480471	0.005704960	0.02491300
TIF	0.02538828	0.025661306	0.010616538	0.02291913
BAC	0.00980939	0.018804781	0.007853329	0.01377355
MCD	0.01249970	0.010419087	0.013400668	0.02192829
CVS	0.01459184	0.009002206	0.008947975	0.02334484
PEP	0.01059988	0.009386000	0.008676303	0.01767862
NKE	0.00992793	0.009402197	0.007540244	0.02871208

SUM SQUARE SIZE OF VIOLATION	ICA-GARCH	PCA-GARCH	DCC	EWMA
T	0.095825344	0.103797312	0.072916054	0.092981019
GE	0.052483553	0.049771486	0.037407798	0.160575907
TIF	0.174146308	0.183959832	0.075004108	0.159164082
BAC	0.067730952	0.135699022	0.055849606	0.096284641
MCD	0.091945585	0.080098468	0.101526375	0.163305949
CVS	0.107227768	0.069136822	0.067724079	0.173681698
PEP	0.077510501	0.070773672	0.064473875	0.129134483
NKE	0.083366915	0.083537268	0.06602284	0.247125904

Table 3b. Backtesting VaR at 95% Confidence Level with Out-of-sample Data

NUMBER OF VIOLATION	ICA-GARCH	PCA-GARCH	DCC	EWMA
T	30	27	12	32
GE	9	11	8	17
TIF	75	52	15	60
BAC	18	20	11	20
MCD	17	14	6	41
CVS	55	39	14	72
PEP	10	9	12	14
NKE	12	15	11	23

Expected number of violations: 10

AVERAGE SIZE OF VIOLATION	ICA-GARCH	PCA-GARCH	DCC	EWMA
T	0.016516861	0.01583784	0.013723898	0.01733647
GE	0.008545917	0.00983799	0.005977205	0.01162991
TIF	0.020627330	0.01972388	0.010908252	0.02413744
BAC	0.011529278	0.013983800	0.010316807	0.01304954
MCD	0.012540873	0.01007942	0.005913215	0.02144786
CVS	0.015542787	0.012906076	0.010672109	0.02639356
PEP	0.009468771	0.008258671	0.009699557	0.01044710
NKE	0.010885938	0.011025852	0.010428991	0.01493653

SUM SQUARE OF VIOLATION	ICA-GARCH	PCA-GARCH	DCC	EWMA
T	0.06658266	0.063745395	0.056323686	0.069786662
GE	0.030947948	0.035625513	0.021648691	0.042002626
TIF	0.072697613	0.069617933	0.038491448	0.085951398
BAC	0.050447842	0.062187919	0.045143519	0.057099944
MCD	0.047839919	0.038450165	0.022557261	0.08181758
CVS	0.060390037	0.051975702	0.042152103	0.103247817
PEP	0.038015654	0.033544509	0.039244661	0.042153826
NKE	0.037922127	0.038308245	0.036434510	0.051795513

Chapter 6

Conclusion

In this essay, we proposed a computationally efficient algorithm, the ICA-GARCH, to estimate and forecast the volatilities of multivariate financial time series. In addition, we apply the results derived by our proposed models to Value at Risk estimation, which is widely used in risk management. The empirical study of this essay suggested that the ICA-GARCH procedure is reliable in capturing the time-varying volatilities and have superior performance in the VaR estimation comparing to the PCA-GARCH and EWMA models.

One noteworthy caveat in this paper is that we assume the ICs derived by an ICA algorithm to be statistically independent with each other. In order to verify the accuracy of an ICA algorithm, it is necessary to conduct tests for independence. One approach of IC selection is the smoothed bootstrap test established on mutual information (Wu, Yu and Li [18]). With an efficient validation device, we can develop our proposed method for further financial applications.

References

[1] Carol Alexander. Orthogonal garch. *Mastering risk*, 2:21–38, 2001.

[2] Luc Bauwens, Sébastien Laurent, and Jeroen VK Rombouts. Multivariate garch models: a survey. *Journal of applied econometrics*, 21(1):79–109, 2006.

[3] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

[4] Tim Bollerslev. On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis*, 9(2):121–131, 1988.

[5] Tim Bollerslev. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The review of economics and statistics*, pages 498–505, 1990.

[6] Tim Bollerslev and Jeffrey M Wooldridge. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric reviews*, 11(2):143–172, 1992.

[7] Peter F Christoffersen. Evaluating interval forecasts. *International economic review*, pages 841–862, 1998.

[8] Robert Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

[9] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

[10] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja. Independent component analysis, adaptive and learning systems for signal processing, communications, and control. *Thomson Learning, Inc*, 2001.

[11] Aapo Hyvärinen and Erkki Oja. A fast fixed-point algorithm for independent component analysis. *Neural computation*, 9(7):1483–1492, 1997.

[12] Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430, 2000.

[13] Jose A Lopez. Methods for evaluating value-at-risk estimates. *Economic Policy Review*, 4(3), 1998.

[14] Daniel B Nelson and Charles Q Cao. Inequality constraints in the univariate garch model. *Journal of Business & Economic Statistics*, 10(2):229–235, 1992.

[15] Gideon Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.

[16] Mike KP So and LH Philip. Empirical analysis of garch models in value at risk estimation. *Journal of International Financial Markets, Institutions and Money*, 16(2):180–197, 2006.

[17] Edmond HC Wu and LH Philip. Volatility modelling of multivariate financial time series by using ica-garch models. In *International Conference on Intelligent Data Engineering and Automated Learning*, pages 571–579. Springer, 2005.

[18] Edmond HC Wu, LH Philip, and Wai Keung Li. A smoothed bootstrap test for independence based on mutual information. *Computational Statistics & Data Analysis*, 53(7):2524–2536, 2009.