❑   1087

Modelling volatility of Kuala Lumpur composite index (KLCI) using SV and garch models

Ezatul Akma Abdullah, Siti Meriam Zahari, S.Sarifah Radiah Shariff,
Muhammad Asmu'i Abdul Rahim
Centre for Statistics and Decision Science Studies, Faculty of Computer & Mathematical Sciences,
Universiti Teknologi MARA Shah Alam, Malaysia

Article Info	ABSTRACT
	It is well-known that financial time series exhibits changing variance and this can have important consequences in formulating economic or financial decisions. In much recent evidence shows that volatility of financial assets is not constant, but rather that relatively volatile periods alternate with more tranquil ones. Thus, there are many opportunities to obtain forecasts of this time-varying risk. The paper presents the modelling volatility of the Kuala Lumpur Composite Index (KLCI) using SV and GARCH models. Thus, the aim of this study is to model the KLCI stock market using two models; Stochastic Volatility (SV) and Generalized Auto-Regressive Conditional Heteroscedasticity (GARCH). This study employs an SV model with Bayesian approach and Markov Chain Monte Carlo (MCMC) sampler; and GARCH model with MLE estimator. The best model will be used to forecast the future volatility of stock returns. The study involves 971 daily observations of KLCI Closing price index, from 2 January 2008 to 10 November 2016, excluding public holidays. SV model is found to be the best based on the lowest RMSE and MAE values.

Corresponding Author:

Siti Meriam Zahari,
Centre for Statistics and Decision Science Studies,
Faculty of Computer & Mathematical Sciences,
Universiti Teknologi MARA Shah Alam, Malaysia
Email: mariam@tmsk.uitm.edu.my

1. INTRODUCTION

Many financial time series exhibits changing variance and this can have important consequences in formulating economic or financial decisions. In much recent evidence shows that volatility of financial assets is not constant, but rather that relatively volatile periods alternate with more tranquil ones. Thus, there are many opportunities to obtain forecasts of this time-varying risk. Modeling of financial time series has focused on estimating the time-varying volatility. The latter key is for measuring risk, pricing asset derivatives, and hedging strategies [1]. When we talk about volatility, it is actually a statistical measure of the dispersion of returns for a given security or market index. Volatility can either be measured by using the standard deviation or variance between returns from the same security or market index. Commonly, riskier security has higher volatility. There are many models that can be used in financial time series data. Those models can be simple model such as random walk, smoothing models, simple regression, moving average and exponential smoothing, or complex model such as Autoregressive Conditional Heteroscedasticity (ARCH), Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and Stochastic Volatility (SV) models. One of the obvious interests in empirical finance is forecasting future returns of assets such as stocks and currencies' exchange rates. If one were able to forecast tomorrow's return with some degree of precision, one could use this information in an investment decision today. It is true that it is not easy to generate a very accurate prediction for asset returns, since we are aware that with all forecasting methods, success is not guaranteed,

but we can at least forecast and get a sign of tomorrow's return. However, one is not only interested in obtaining accurate forecasts of returns on financial assets, but also in forecasts in associated volatility. This is because; volatility is often regarded as a measure of the risk of return.

A problem in the analysis of financial time series data is to forecast the volatility of future returns since it can reflect the risk of the returns. The first attempt to solve this problem is using the Black - Scholes model. This model was first introduced by Scholes and Black in 1973. The advantages of the Black-Scholes model have been well known for years and this model is simple to implement. According to the Black Scholes formula, the model assumes that continuously compounded stock returns are normally distributed with constant mean and variance. This means that measure of how much a stock can be expected to move is constant over time. However, in the real financial world, volatility of financial data is dynamic. The assumption of constant volatility is very restrictive since a significant number of empirical studies show that volatility in asset prices is time-varying. Therefore, Black-Scholes is not suitable to use. To overcome this problem, time varying volatility model named Autoregressive Conditionally Heteroscedastic (ARCH) model was expressed. Then, the model was extended to Generalized Auto-Regressive Conditional Heteroscedasticity (GARCH) models. GARCH model is the most appropriate model to use in measuring the volatility [2]. This model can be estimated using maximum likelihood estimator (MLE) or other robust estimators. However, a basic GARCH model can only allow for a single error process while Stochastic Volatility (SV) model assumes two error processes. This shows that, SV will provide a better in-sample fit and it is believed to be a better model to use [3]. Financial econometrics literatures give lots of attention in the Stochastic volatility (SV) models [4] reviewed the univariate and multivariate of Stochastic Volatility (SV) model and stated that SV model provides more flexibility in describing stylized facts. One of the models that can be used to describe the stochastic dynamic is definitely the Heston Stochastic Volatility model. Similarly, [5] also adopted the Heston Stochastic Volatility model to model the implied volatility index in continuous time setting and price options on it [6] showed that the basic SV model gives a more accurate result of prices for European call options on currencies compared to the Black–Scholes model [7], applied stochastic volatility factors in the model and found that it provides good result for short term [8], proposed a simple extension of a standard GARCH (1,1) model, which can capture SV-like properties of data that called Stochastic GARCH (SGARCH). It shows that SGARCH model is better to use since it captures the typical SV model property which has two errors, but with only single parameters that makes it easy to estimate [9], found that SV model outperformed the GARCH model. This findings is also supported by [10]. Stochastic Volatility (SV) model is found to be the best model to predict the NZ stock market, [11]. Other advantages of Stochastic Volatility (SV) model can be found in [12].

Nevertheless, the estimation of SV models is the main challenge in applying the model. Depending on the model, some moments may or may be unknown as in closed form resulting the unknown of the transition density of the state vector [13]. However, this issue can be solved by performing the efficient and fast Bayesian Markov chain Monte Carlo (MCMC) estimation algorithm [14]. Therefore, the aim of this study is to model and compare between both the SV and GARCH models by applying the MCMC method in the estimation process for SV model. The best model is used for forecasting purpose.

2. RESEARCH METHOD

The data of this study consists of 971 daily closing indexes from Kuala Lumpur Composite Index (KLCI), from 2nd January 2008 to 10th November 2016. The data is extracted from Thomson DataStream. Kuala Lumpur composite index or KLCI is the main index and market indicator in Malaysia. This index is the representative of Malaysia's stock market. KLCI is now enhanced and known as FTSE Bursa Malaysia KLCI. The FTSE Bursa Malaysia KLCI was introduced on 4th April 1986 as the Kuala Lumpur Composite Index (KLCI). This index is derived from 100 companies. The companies have been chosen by Bursa Malaysia from a cross section of the total listed companies in Malaysia.

2.1. Stochastic Volatility (SV) Model

For the SV model, Bayesian parameter estimation via Markov Chain Monte Carlo (MCMC) method is used. Let $y = (y_1, y_2, \ldots, y_n)^T$ be a vector of KLCI returns. The feature of SV model is that each observation of y_t is assumed to have its "own" contemporaneous variance of e^{h_t}. Therefore, it will relax the usual assumption of homoscedasticity. The variance is not allowed to vary unrestrictedly over time in order to make the estimation of a model is feasible. The logarithm of the model is assumed to follow an autoregressive process (AR) of order one.

The SV model can be expressed in the form (1).

$$y_t = \exp(\tfrac{h_t}{2})\varepsilon_t \tag{1}$$

$$h_t = \mu + \phi h_{t-1} + \sigma_\eta \eta_t \qquad\qquad (2)$$

The above equation is the equation for SV model. SV model is simply interested in market news and in modeling their effects on volatility. From the equation, it shows that the volatility in SV model is not affected by past returns. SV is supposed to describe financial time series better since it essentially involves two error processes, which are ε_t and η_t. The added noise process in the equation makes the model more flexible. h_t in (1) and (2) is equal to log volatility of KLCI at time t. The parameter vector is $\theta = (\mu, \phi, \sigma_\eta)$. ε_t and η_t are the error processes. The parameter involved level of log-variance, μ, the persistence of log-variance ϕ, and the volatility of log-variance σ_η. The value of ϕ,(-1<ϕ<1), measures the autocorrelation present in the logged squared data. Thus ϕ can be interpreted as the persistence in the volatility. High ϕ indicating volatility clustering. ϕ is assumed to follow a stationary process ($|\phi|$<1). Parameter σ_η in the equation is the volatility of the log-volatility. Log-variance process in the equation is $h = (h_0, h_1, \ldots, h_n)$.. It is unobserved.

2.2. Bayesian Approach: Prior Distribution

A prior distribution needs to be specified first to complete the estimation setup for SV model. Each parameter is assumed to be prior independent. The Bayesian statistics is a mathematical process that applies probabilities to statistical problems by providing the tools to update the beliefs in the evidence of new data. The most important of Bayesian is the establishment of parameters and models. The prior belief distribution is used to represent the strengths on beliefs about the parameters based on previous experience. However, there might be no previous experience. Therefore, mathematicians have formulated methods to overcome this problem known as uninformative priors. This study used prior parameters from the past study.

2.3. Markov Chain Monte Carlo Methods (MCMC)

Stochastic Volatility (SV) model can be described in three conditional distributions that is f(θ),y|h, and h|θ. The function of f(θ) is the prior distribution of θ. The joint distribution of (θ,h) is derived from Bayes theorem. Therefore, f(θ,h|y) is proportional to the product of the f(y│h)f(h│θ)f(θ). To construct the Markov chain with invariant distribution, f(θ,h|y), algorithm called Metropolis-Hashtings algorithm is proposed as ilustrated in Figure 1.
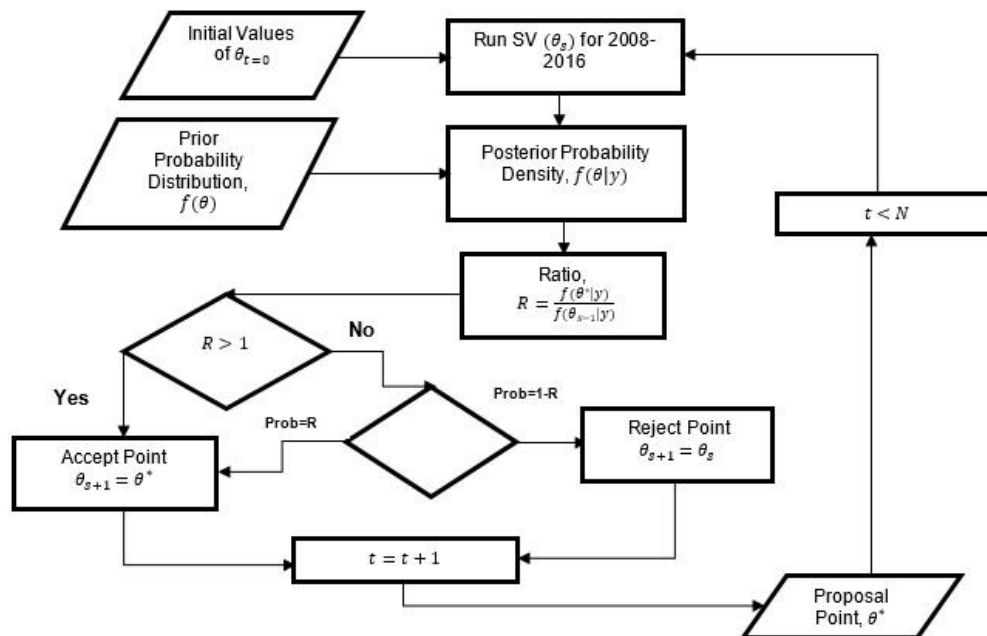

Figure 1. Process flow diagram for Metropolis-Hasting MCMC

The process of this method generates the proposal point by using the last sample and adding a random noise which generated from a random noise distribution. Next, the height of the posterior parameter of the new proposal is then compared with the height of parameters of the recent proposal and calculated as R, which represent the ratio. If the new proposal has higher posterior than the new proposal, then the new proposal will

be accepted. Otherwise, MCMC will randomly accept or rejected the new proposal. If the new proposal is accepted, then it becomes the next sample of MCMC chain, $t + 1$. If the new proposal is rejected, then it will take the most recent proposal as the new sample. This step will continue and start again as the first step. When the sample, N is enough, the iteration will stop.

MCMC is one good approach to estimate the Bayesian. MCMC shows how the posterior parameter is affected by the distribution. MCMC is not representing the certainty but just the probability. Note that, having more information about the sample will help to make a better plan for the future. Analysis is carried out using Stochvol package.

2.4. Garch Model

Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model is the extension of the ARCH model. The model for GARCH is:

$$y_t = \sigma_t \varepsilon_t \tag{3}$$

$$\sigma_t^2 = \mu + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{4}$$

From (3), y_t or the returns of KLCI is influenced by standard deviation σ_t and error ε_t. (4) shows that the volatility, σ_t^2 is determined by squared past returns, y_{t-1}^2 and squared past volatility, σ_{t-1}^2. The GARCH model captures volatility better than the ARCH model. The steps in modeling the volatility using GARCH are as follows:

3. RESULTS AND ANALYSIS

From the graph, it is clear that the growth of the economy has slowed down after 2008. Drop in stock price during 2008-2009 not only affect other countries, but also Malaysian stock market because of the global financial crisis. The global financial crisis gives an impact towards financial, trade and real economy [15]. Therefore, it significantly affects the Malaysia's stock market, KLCI. Figure 2 shows the closing price index that is not stationary.
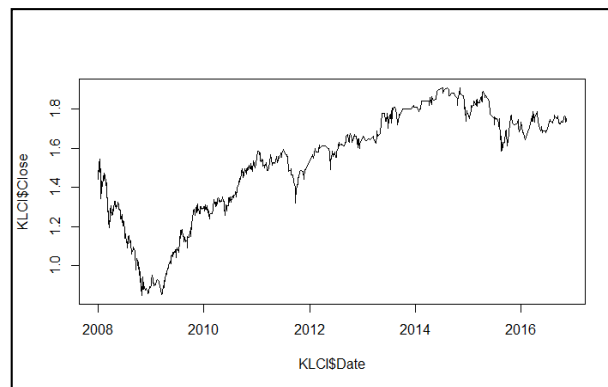

Figure 2. Daily Plot of Kuala Lumpur Composite Index (KLCI)

Since the index graph as shown in Figure 2 is not stationary, therefore the log return is used. Figure 3 presents the KLCI returns from 2nd January 2008 until 10th November 2016. The x-axis represented time while the y-axis is log returns.

Figure 3 shows that the graph is stationary. Log returns represent the volatility. Based on the graph, volatility changes over time, indeed there exists volatility clustering because there are several large points. The plot also shows that negative returns occur during the observed period. The largest negative returns are observed in the year 2008.
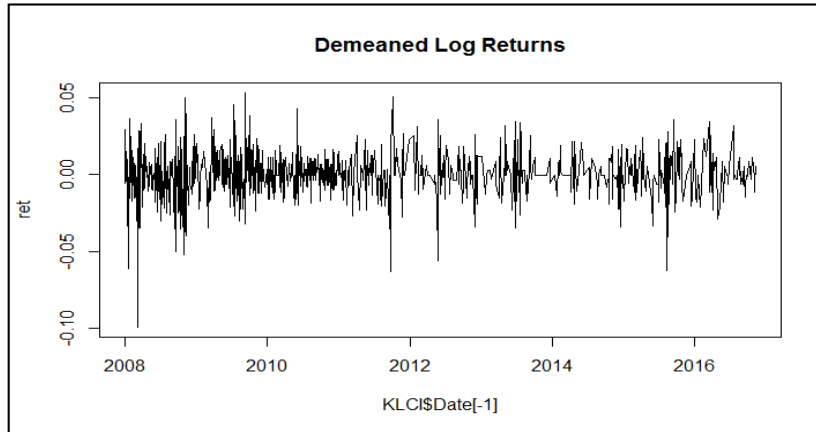
Figure 3. Demeaned Log Returns of KLCI

3.1. Simulation using MCMC

Simulation using Markov Chain Monte Carlo (MCMC) is carried out. Since MCMC is computer-based calculation, therefore the analysis is done by MCMC using the algorithm, discussed in sub section 2.3. As stated before, the SV model involved parameter $\theta = (\mu, \phi, \sigma_\eta)$. Since the value for each parameter is unknown and must be specified first before running the simulation, this study referred to a past study in identifying the parameter values. This study followed [12] who is the author of the stochvol package. The author used $\mu = -9$, $\phi = 0.99$, and $\sigma_\eta = 0.1$ to run the simulation. However, with less information about the prior parameters, it will cause the posterior parameter being far from the true values. The length of simulated time series is 1000. After running the function with the prior parameter, result will draw the initial log volatility, h_0 from the stationary distribution of the AR (1) process. Then, using h_0 value, the value for $h_1, h_2, \ldots, h_n$ is generated iteratively. MCMC calculate the new proposal according to the most recent sample. Lastly, using normal distribution with mean 0 and standard deviation, $\exp(\frac{h}{2})$, the log-returns are simulated. From the simulated conditional volatility, the initial value will be generated. From the result of simulation, the value for initial log-volatility, $h_0 = 0.009107516$. Using this initial value, MCMC will iterate the next value until the 1000th value. The algorithm used in MCMC is called Metropolis algorithm. The algorithm starts with possible initial value, in this case $h_0 = 0.009107516$. Next, it generated new proposal by using the last sample and adding with some noise. The result also gives the descriptive statistics of simulated conditional volatility. The minimum value for simulated conditional volatility is 0.08357, first quadrant equal to 0.41850, median value is 0.56650, 0.64760 for the mean and 0.83020 and 2.14000 for third quadrant and maximum value. Results for simulated conditional volatility and simulated data are visualized in Figure 4.
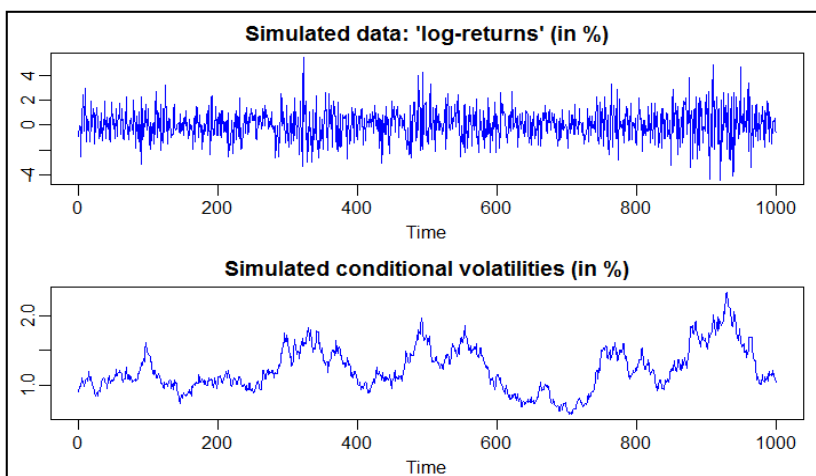

Figure 4. (a) Simulated Data (log-returns) (in%), (b) Simulated Conditional Volatilities (in %)

From Figure 4, there exists clustering pattern, where periods of high volatility tend to be followed by periods of even higher volatility. Similarly, periods with low volatility tends to be followed by low volatility. The patterns of Figure 4(a) shows that it tends to fluctuate at the mean close to 0. Figure 4(b) shows that periods of high volatility cause large changes in the returns, whereas periods of low volatility cause small changes in the returns. The simulated distribution exhibits numbers of outliers which reflect frequent sizeable changes in volatility.

Next step is to simulate from the joint posterior distribution of the SV parameter $\theta(\mu, \phi, \sigma_\eta)$ using the function "svsample". Value of latent log volatilities $h_0, h_{1,...} h_n$ and returns draw in MCMC is used in this simulation. The result is summarized in Table 1.

Table 1. Posterior Parameter

	Mean	Standard Error	5%	50%	95%	ESS
Mu	-8.944	0.14724	-9.184	-8.945	-8.703	4295
Phi	0.902	0.03743	0.834	0.907	0.953	117
Sigma	0.389	0.08065	0.273	0.381	0.537	111
exp(mu/2)	0.011	0.00085	0.010	0.011	0.013	4295
sigma^2	0.158	0.06700	0.075	0.145	0.288	111

Table 1 shows the value of posterior parameters given by MCMC sampler. The value of mu is equal to -8.944 while phi is equal to 0.902. Since value of phi is high, thus, it shows high persistence in the volatility. Therefore, there exists the volatility clustering. The sigma or volatility of the log-volatility value is 0.389. The values for exp (mu/2) and sigma^2 are 0.011 and 0.158 respectively. The percentages given are, 5%, 50%, and 95%. It represents the belief of the possible values will fall within it. The value for the mean, mu is guaranteed that 95% of its possible values will fall at -8.703. ESS in the last column is the effective sample size. It is the number of effectively independent draws from the posterior distribution that the MCMC is equivalent to.

Trace plots in Figure 5 indicate whether convergence can be safely diagnosed and how well the output samples perform. Trace plot shows the values that parameter took during MCMC runtime. All three trace plots show that the MCMC algorithm has converged. From the trace plot, it shows that the parameters do not exhibit significant bias.
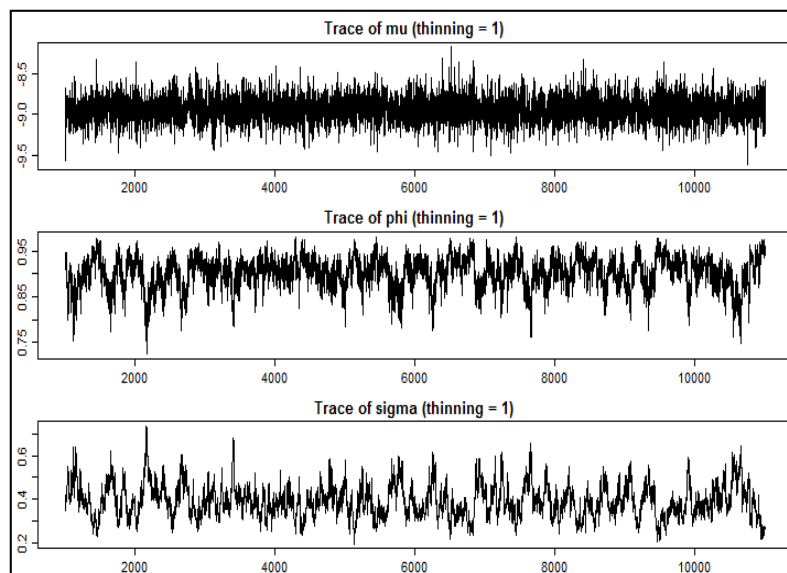

Figure 5. Trace Plot for MCMC Estimate of θ of the Obtained Posterior Sample

Stocks that maintain a relatively stable price is said to have low volatility. However, from Figure 6, it shows that the volatility is not relatively stable. Therefore, the KLCI is said to have high volatility, hence indicating high risk in the Malaysian Stock Market for the forecasting period.
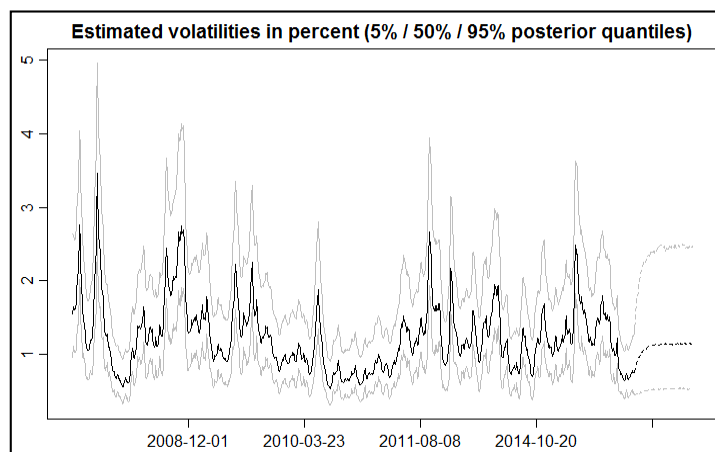
Figure 6. Estimated Volatilities in Percent (%)

3.2 Comparison Between GARCH and SV models

The value of RMSE and MAE for both models is compared to find the best model. The comparison between these two models is shown in Table 2.

Table 2. Comparison between GARCH and SV model

	GARCH	SV
RMSE	0.014218	0.014213
MAE	0.0098740	0.0098702

The value for RMSE and MAE for SV are slightly lower compared to GARCH model. This indicates that the SV model performed slightly better than GARCH model. Therefore, SV model is the best model to be employed for forecasting the KLCI returns.

4. CONCLUSION

The paper presents the modelling volatility of the Kuala Lumpur Composite Index (KLCI) using SV and GARCH models. The findings show that both models are superior to each other, with slight differences in terms of RMSE and MAE values. For forecasting purposes, it is suggested to employ SV model for managing the risks since it provides more information about the data. It also focuses on the market news and also the impact of the volatility towards the returns. For future studies, it is recommended to use different prior parameter values since the choice for parameter may produce different results, especially in terms of persistence of volatility.

REFERENCES

[1]   Hull J., White A., "The Pricing of Options on Assets with Stochastic Volatilities," *The Journal of Finance*, vol.42(2), pp. 281-300, 1987.
[2]   Matei M., "Assessing Volatility Forecasting Models: Why GARCH Models Take the Lead." *Romanian Journal of Economic Forecasting*. vol. 4(4), 42-65, 2009.
[3]   Franses P.H., van der Leij, M.J., Paap R., "A Simple Test for GARCH against a Stochastic Volatility Model," *Econometric Institute Research Papers EI 2005-41*. 2005.
[4]   Platanioti E. J., "A Review of Stochastic Volatility: Univariate and Multivariate Models," *Imperial College London, London. Technical Paper*, 2005.

[5] Grunbichler, A, Longstaff, F.A. "Valuing Futures and Options on Volatility," Journal of Banking & Finance, vol. 20(6), pp. 985-1001, 1996.

[6] Melino A. T., "Pricing foreign currency options with stochastic volatility," *Journal of Econometrics,* vol. 45, pp. 239–265, 1990

[7] Anderson T.G., Lund J., "Estimating continuous-time stochastic volatility models of the short-term interest rate," *Journal of econometrics*, vol. 77(2), pp. 343-377, 1997.

[8] Kim S., Shephard N., Chib S., "Stochastic volatility: Likelihood Inference and Comparison with ARCH Models." *Review of Economic Studies,* vol. 65, pp. 361–393, 1998.

[9] Nakajima J., "Bayesian analysis of GARCH and stochastic volatility: modeling everage, jumps and heavy-tails for financial time series," *Discussion Paper No. 2008-E-23*, 2009.

[10] Jawed S. M., "Application of Stochastic Volatility Model to KSE-100," *Pakistan Journal of Engineering, Technology & Science*, vol. 4(1), pp. 28-40, 2014

[11] Yu J., "Forecasting Volatility in the New Zealand Stock Market." *Applied Financial Economics*, vol. 12(3). Pp. 193-202, 2012.

[12] Kastner G., "Dealing with Stochastic Volatility in Time Series Using the R Package Stochvol." *Journal of Statistical Software*, vol. 69, pp. 1-30, 2016.

[13] Aı Y., & Kimmel R., "Maximum likelihood estimation of stochastic volatility models," *Journal of financial economics*, vol. 83(2), pp. 413-452, 2007.

[14] Chib, S., Nardari, F., & Shephard, N. "Markov chain Monte Carlo methods for stochastic volatility models," *Journal of Econometrics*, vol. 108(2), pp. 281-316, 2002.

[15] Goh Soo Khoon, M. L.-H., "The Impact of the Global Financial Crisis: The Case of Malaysia," *TWN Global Economy Series*, pp. 1-38, 2010.