MATEMATIKA, 2014, Volume 30, Number 1a, 44-58

Probability Distribution of Returns In Heston Model And Hurst Exponent Estimaton for Index Prices Of FTSE Bursa Malaysia KLCI

[1]Kho Chia Chen, [2]Tey Seah Ying and [3]Arifah Bahar

[1, 2]Universiti Teknologi Malaysia
Department of Mathematical Sciences, Faculty of Science, UTM,81310, Skudai Johor
[3]UTM Centre for Industrial & Applied Mathematics (UTM-CIAM), UTM
Email: [1]cc_kho88@hotmail.com, [2]teyseahying@yahoo.com, [3]arifah@utm.my

Abstract In this study, the probability distribution of returns for the index prices of FTSE Bursa Malaysia KLCI based on the Heston Model with stochastic variance is constructed to analyse its capability in describing the returns. The values of parameters in Heston Model of the index prices of FTSE Bursa Malaysia KLCI have been estimated using Simulated Maximum Likelihood method. Euler-Maruyama method has been used to approximate the solutions of the stochastic differential equation. It is found that for complete set of the data, the probability distribution of log returns for closing prices of FTSE Bursa Malaysia KLCI fitted the theoretical curve better at time lag, t =1 and 20. In addition, we used Hurst exponent to measure the existing of long range dependence of time series towards the real data.

Keywords Financial mathematics; Heston Model; Simulated Maximum Likelihood; Stochastic Differential Equation; Hurst Exponent.

2010 Mathematics Subject Classification 62P05.

1 Introduction

In 1993, Steven Heston [1] proposed a model for the stock price dynamics using Geometric Brownian Motion with stochastic volatility. The model described by Heston is well known as Heston model. This model assumes that the volatility of the asset is not constant but follows a random process. Index prices behave randomly so it is not accurate to be represented by initial value problem with ordinary differential equation. The Heston model, which is governed by Geometric Brownian Motion with stochastic volatility, is always used to describe the evolution of the volatility of an underlying asset since the volatility of the underlying asset is not constant. Hence, the incorporation of stochasticity into index prices is important in determining correctly the probability distribution of log returns for stock indexes.

However, Heston model itself does not provide the probability distribution of log returns for stock prices. Dragulescu and Yakovenko [2] had derived the analytical formula for the probability distribution of log returns in the Heston model. According to them, their equations reproduce the probability distribution of returns for time lag 1 and 250 trading days very well by using only four parameters. But, the number of parameters of other similar models can easily go to a few tens. Moreover, Silva and Yakovenko [3] had tested that the probability distribution of log returns agreed very well with the analytical formula for NASDAQ, S&P 500 and Dow-Jones for the period of 1982-1999 via visual inspection.

Kuala Lumpur Stock Exchange started to have a downward trend in July 1997 as it fell below its psychological level of 1000 points. According to Hughes [4], the contagious effects of speculative activity on the Thai Baht had led to the crisis in 1997. Based on the economic report 1997/98, in September 1997, the KLCI touched its lowest level since 20 April 1993 of 675.15. In September 1998, the KLCI fell sharply to as low as 262.7 points from 1077.3 points in June 1997 following the implementation of exchange control on 1st September 1998 [4]. According to Goh and Lim [5], the Dot Com bubble that burst in March 2000 had adversely affected the growth of Malaysia's economy. The KLCI which was recovering from the 1997 crash took another hit and went down. From 2003 to 2007, the index point of KLCI started to pick up and eventually surpassed the 1500 points in the early of 2008. Unfortunately, in 2008, severe global recession crisis that began with the burst of the U.S. housing bubble was transmitted to Malaysia [5]. KLCI fell below 1000 points markedly in September 2008. According to The World Bank [6], the Malaysian economy completed the rebound from the downturn but the growth momentum is volatile. Thus, we have studied the real data of the closing prices of FTSE Bursa Malaysia KLCI from 1st September 1998 until 31st July 2008 for the probability distributions of log returns to analyse the trend of Malaysia economy during the period.

In this study, we will use the techniques given in Dragulescu and Yakovenko [2] to assess the suitability of the index prices for FTSE Bursa Malaysia KLCI (formerly known as KLCI) based on the analytical formula derived by Dragulescu and Yakovenko for the Heston model with stochastic volatility. Parameters of the model will be estimated based on Simulated Maximum Likelihood method. The stochastic differential equation will be solved by Euler-Maruyama method. With the analytical formula, the probability of particular log return and time lag for index prices can be calculated. Investors can then plan their investments accordingly based on the probability of getting certain return. Last but not least, the index prices were estimated with Hurst exponent [7, 8] to measure the existing of long range of dependence of time series.

2 Methodology

2.1 Heston Model

We first consider stock price S_t, as a time dependent function of *t*, which obeys the stochastic differential equation of a Geometric Brownian Motion

$$dS_t = \mu S_t dt + \sigma_t S_t dW_t^{(1)}. \tag{1}$$

Subscript *t* indicates time dependence, *μ* is the drift parameter, $W_t^{(1)}$ is a standard random Wiener process, and σ_t is the time dependent volatility. We are interested in log-returns, hence change the stock price S_t to log-return $r_t = \ln(S_t / S_0)$ and eliminate the drift parameter μ by introducing $x_t = r_t - \mu t$, we get

$$dx_t = -\frac{v_t}{2} dt + \sqrt{v_t} dW_t^{(1)} \tag{2}$$

where $v_t = \sigma_t^2$ is the variance.

Now, let us assume that the variance v_t obeys the CIR process

$$dv_t = -\gamma\left(v_t - \theta\right)dt + \kappa\sqrt{v_t}dW_t^{(2)} \tag{3}$$

.

Here θ is the long time mean of *v*, γ is the rate of relaxation to this mean, $W_t^{(2)}$ is a standard Wiener process, and κ is the variance noise. The Wiener process in (1) can be correlated with the Wiener process in (3)

$$dW_t^{(2)} = \rho dW_t^{(1)} + \sqrt{1-\rho^2} dZ_t$$

where Z_t is a Wiener process independent of $W_t^{(1)}$, and $\rho \in [-1,1]$ is the correlation coefficient.

Stochastic processes (2) and (3) constitute the Heston model. According to Gardiner [9], the Fokker-Planck equation can be derived for the transition probability $P_t(x, v \mid v_i)$ to have log-return *x* and variance *v* at time *t* given the initial log-return *x = 0* and variance v_i at $t = 0$,

$$\frac{\partial}{\partial t} P = \gamma \frac{\partial}{\partial v}\left[\left(v-\theta\right)P\right] + \frac{1}{2}\frac{\partial}{\partial x}\left(vP\right) + \rho\kappa \frac{\partial^2}{\partial x \partial y}\left(vP\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(vP\right) + \frac{\kappa^2}{2}\frac{\partial^2}{\partial v^2}\left(vP\right) \tag{4}$$

The initial condition for (4) is a product of two delta functions

$$P_{t=0}\left(x, v|v_i\right) = \delta\left(x\right)\delta\left(v - v_i\right).$$

The probability distribution for the variance itself, $\Pi_t\left(v\right) = \int dx \, P_t(x,v)$, satisfies the equation

$$\frac{\partial}{\partial t}\Pi_t(v)=\frac{\partial}{\partial v}\left[\gamma(v-\theta)\Pi_t(v)\right]+\frac{\kappa^2}{2}\frac{\partial^2}{\partial^2 v}\left[v\Pi_t(v)\right]$$

The Fokker-Planck equation above had been solved analytically [2] for the transition probability $P_t(x,v\,|\,v_i)$. Then $P_t(x,v\,|\,v_i)$ is integrated over the variance, *v* to give

$$P_t\left(x|v_i\right)=\int_0^{+\infty}dv\,P_t(x,v\,|\,v_i) \tag{5}$$

However, equation (5) cannot be directly compared with financial time series data, since it depends on an unknown variable v_i .

In order to solve the problem, Dragulescu and Yakovenko [2] assumed that v_i follows stationary probability distribution $\Pi_*(v_i)$ and averaging (5) over v_i with the weight $\Pi_*(v_i)$ to give the PDF $P_t(x)$:

$$P_t\left(x\right)=\int_0^{\infty}dv_i\Pi_*\left(v_i\right)P_t\left(x|v_i\right) \tag{6}$$

This is the PDF of log-returns *x* after the time lag *t* and it can be directly compared with the financial time series data.

We can find in Dragulescu and Yakovenko [2] that the fitted data are not very sensitive to the parameter ρ, hence we only consider the case $\rho=0$ for simplicity in this project. The final expression for (6) at $\rho=0$ has the form of a Fourier integral

$$P_t\left(x\right)=\frac{e^{-x/2}}{x_0}\int_{-\infty}^{+\infty}\frac{d\tilde{p}}{2\pi}e^{i\tilde{p}\tilde{x}+F_{\tilde{t}}(\tilde{p})}, \tag{7}$$

$$F_{\tilde{t}}\left(\tilde{p}\right)=\frac{\alpha\tilde{t}}{2}-\alpha\ln[cosh\frac{\tilde{\Omega}\tilde{t}}{2}+\frac{\tilde{\Omega}^2+1}{2\tilde{\Omega}}sinh\frac{\tilde{\Omega}\tilde{t}}{2}, \tag{8}$$

where $\tilde{\Omega}=\sqrt{1+\tilde{p}^2}$, $\quad \tilde{t}=\gamma t$, $\quad \tilde{x}=x/x_0$,

$$x_0=\kappa/\gamma, \qquad \alpha=2\gamma\theta/\kappa^2.$$

In the long-time limit $\tilde{t}\gg 2$, equation (7) and (8) exhibit scaling behaviour. $P_t\left(x\right)$ becomes a function of a single combination z of the two variable, x and t .

$$P_t\left(x\right)=N_t e^{-x/2}K_1(z)/z$$

where $z = \sqrt{\tilde{x}^2 + \bar{t}^{\,2}}$, $\bar{t} = t\theta / x_0^2$, $N_t = \bar{t}e^{\bar{t}} / \pi x_0$

and $K_1(z)$ is the first-order modified Bessel function.

2.2 Parameters Estimation of Heston Model

We use the stochastic differential equation (SDE) toolbox in Matlab to obtain the estimated parameters μ, γ, κ and θ. According to Picchini [10], the integration of the stochastic differential equations is based on the Euler-Maruyama method [11] while the implementation of simulation is based on Simulated Maximum Likelihood.

According to Picchini [10], the general framework of stochastic differential equations in Ito form

$$dS_t = f\left(t, S_t; \theta\right) dt + g\left(t, S_t; \theta\right) dW_t \ ,$$

with initial condition $S_0 = s_0$ and $t \geq 0$,where W is standard Wiener process is known function depending on parameter θ. Since S is Markovian, the maximum likelihood estimator of θ can be calculated if the transition densities of S are known. The log-likelihood function of θ is given by

$$l_n\left(\theta\right) = \sum_{i=1}^{n} \log p(s_i; s_{i-1}, \theta) \tag{9}$$

and the maximum likelihood estimator $\hat{\theta}$ can be found by maximising (9) with respect to θ.

However, the transition density function of the underlying diffusion process is often unknown. SDE toolbox computes an approximation to the transition density function numerically by simulating R times the process using Monte-Carlo Simulations.

2.3 Algorithm for Calculating Probability Distribution of Log-returns For KLCI

Time series $\left\{S_\tau\right\}$ is formed by the data set of the index prices of FTSE Bursa Malaysia KLCI, where τ is the number of trading days. We will follow the procedures mentioned in Dragulescu and Yakovenko [2] to extract the probability density of log-return r , $P_t^{(KLCI)}\left(r\right)$ for a given time lag t from the $\left\{S_\tau\right\}$.

Step 1

For fixed *t*, calculate the set of log-returns $\{r_\tau = \ln S_{\tau+t} / S_\tau\}$ for all possible times τ .

Step 2

Partition the *r*-axis into equally spaced bins of width Δr .

Step 3

Count the number of the log-returns belonging to each bin.

Step 4

Omit the bins with occupation numbers less than five.

Step 5

Divide the occupation number of each bin by Δr and by the total occupation number of all bins. This is the probability density $P_t^{(KLCI)}(r)$ for a given time lag *t*.

Step 6

To find $P_t^{(KLCI)}(x)$, replace $r \to x + \mu t$.

2.4 Hurst Exponent Estimation

The Hurst Exponent is used as a measure of long memory of time series [12, 13]. There are three distinct classifications for the Hurst exponent *H*.

For the $0.5 < H \leq 1$, it shows that the time series have long-term positive autocorrelation. This means that a high value in the series will probably be followed by another high value and that the values along time into the future will also tend to be high. This series is considered to be persistent. The strength of this persistent trend depends on how close *H* is to 1.

For the case $0 \leq H < 0.5$, it denotes a time series with long-term switching between high and low values in next pairs, which mean that a single high value will probably be followed by a low value and that the value after that will tend to be high, with this trend to switch between high and low values lasting a long time into the future. This type of series is called an anti-persistent. The strength of this anti-persistent trend depends on how close *H* is close to zero.

For the case *H* equals to 0.5 denotes that the series are random and uncorrelated. In fact, the value is applicable to series for which the autocorrelations at small time lags can be positive or negative but where the absolute values of the autocorrelations decay exponentially quickly to zero.

To estimate the Hurst exponent, we need to estimate the dependence of the rescaled range, the R/S statistic [12, 13]. The adjusted range *R* is defined as

$$R(n) = \max_{1\le k\le n}\left\{\sum_{i=1}^{k} X_i - k\bar{X}\right\} - \min_{1\le k\le n}\left\{\sum_{i=1}^{k} X_i - k\bar{X}\right\}$$

where $\bar{X}$ represents the sample mean. The standard deviation S is given by

$$S(n) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

The Hurst exponent is estimated by calculating the average rescaled range over multiple regions of the data. The expected value of R/S converges on the Hurst exponent power function are shown in (10) as below

$$E\left[\frac{R(n)}{S(n)}\right] = Cn^H \qquad \text{as} \quad n \to \infty \tag{10}$$

By plotting the logarithm of equation (10) as a function of $\log n$ with fitting a straight line, the slope of lines gives H.

3 Results

3.1 Data Description

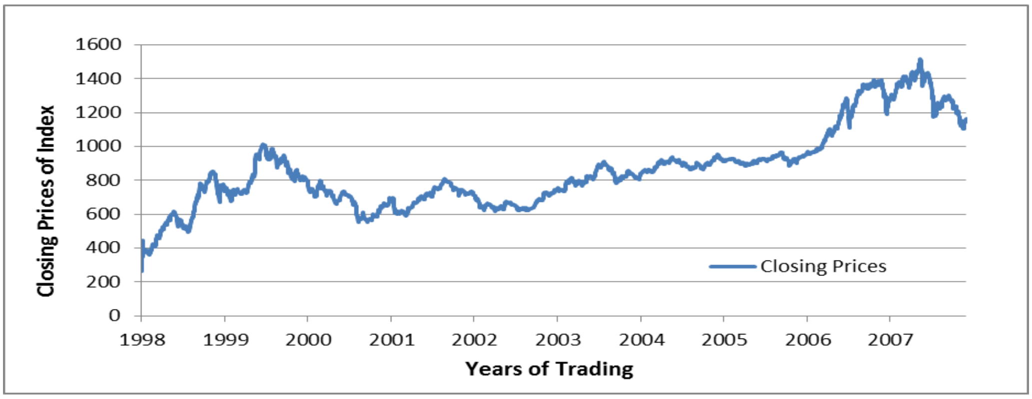

Figure 1: Closing Prices of FTSE Bursa Malaysia KLCI

The historical prices of FTSE Bursa Malaysia KLCI were described through graphical representation. The graphing representation of the time series data was done via scatter plot.

3.1.1 Historical Prices FTSE Bursa Malaysia KLCI (1/9/1998-31/7/2008)

The real data of the closing prices of FTSE Bursa Malaysia KLCI from 1st September 1998 until 31st July 2008 were plotted against its trading days.

From figure 1, it showed that the real prices of FTSE Bursa Malaysia KLCI during this period have an increase trend. The economy growth is quite stable along the period.

3.2 Estimation of the Heston Model Parameters

The values of the μ, γ, κ and θ parameters were estimated based on simulated maximum likelihood method using Matlab. Table 1 shows the values of the estimated μ, γ, κ and θ parameters of Heston Model for the closing prices of FTSE Bursa Malaysia KLCI.

Table 1: Estimated Values of Heston Model Parameters

Types of Parameter	**Values of Parameter**
Drift, μ	0.003750684
Long-time mean of variance, θ	0.0000713
Rate of relaxation to θ , γ	0.07797
Variance noise, κ	0.00225

3.3 Comparison with the FTSE Bursa Malaysia KLCI Time Series

Figure 2, 3, 4 and 5 compare the theoretical curve that was derived by Dragulescu and Yakovenko [4] with the probability distribution of log returns for historical prices of FTSE Bursa Malaysia KLCI at delay of time, t = 1, 20,40 and 250 respectively.

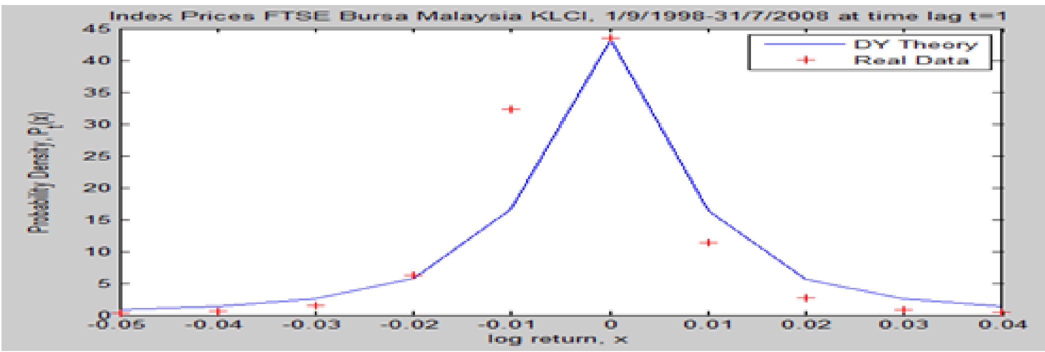

Figure 2: Theoretical curve and the probability distribution of log returns for historical prices of FTSE Bursa Malaysia KLCI at a delay of time, $t = 1$

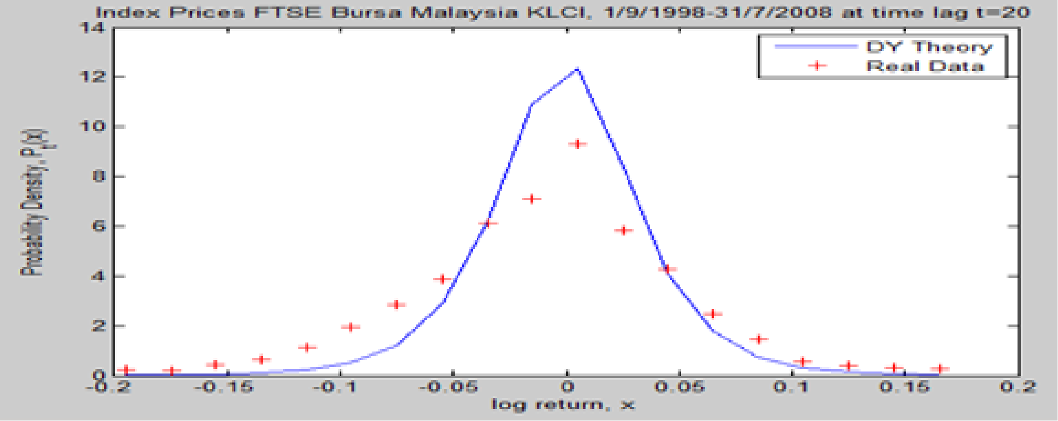

Figure 3: Theoretical curve and the probability distribution of log returns for historical prices of FTSE Bursa Malaysia KLCI at a delay of time, $t = 20$

From Figure 2, we are able to notice that the real data points fitted the theoretical curve quite well. On the other hand, Figure 3 shows that the real data points have flatter shape than curve of the theoretical curve. The data points that have log return x nearer to value zero tend to fall below the curve. However, the data points that have log return *x* further from value zero tend to fall above the curve.

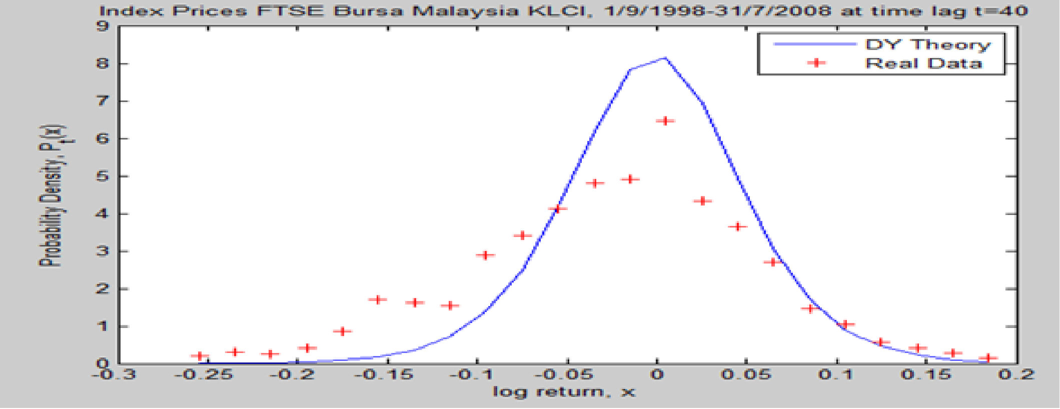

Figure 4: Theoretical curve and the probability distribution of log returns for historical prices of FTSE Bursa Malaysia KLCI at a delay of time, $t = 40$

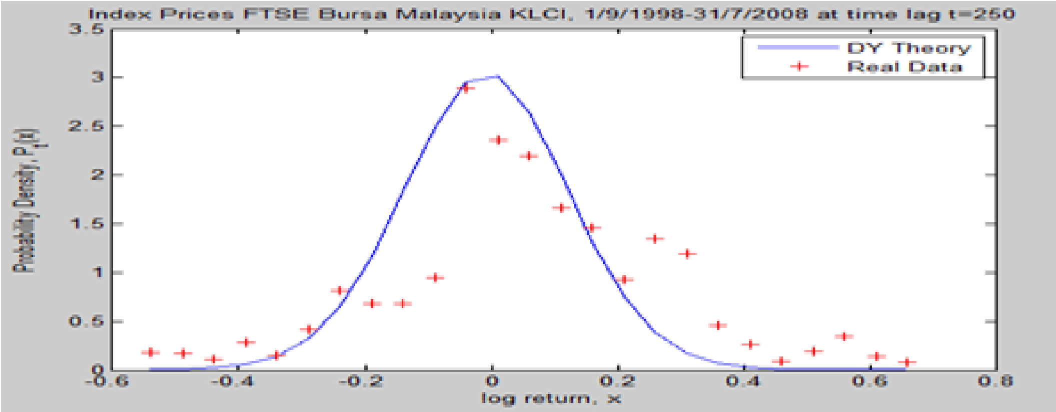

Figure 5: Theoretical curve and the probability distribution of log returns for historical prices of FTSE Bursa Malaysia KLCI at a delay of time, t = 250

From Figure 4 and 5, there are more real data points which deviate from the curve. The real data points have flatter shape than curve of the theoretical curve. The data points that have log return x nearer to value zero tend to fall below the curve. Whereas, the data points that have log return x further from value zero tend to fall above the curve. However, the shape of real data points tends to shift to the left.

In summary, only Figure 2 show quite close match between the theoretical curve and real data.

3.4 Hurst Exponent Estimation in FTSE Bursa Malaysia KLCI

Figure 6 showed the pox plot of the Hurst Exponent. The slope in the graph represents the Hurst exponent *H*.

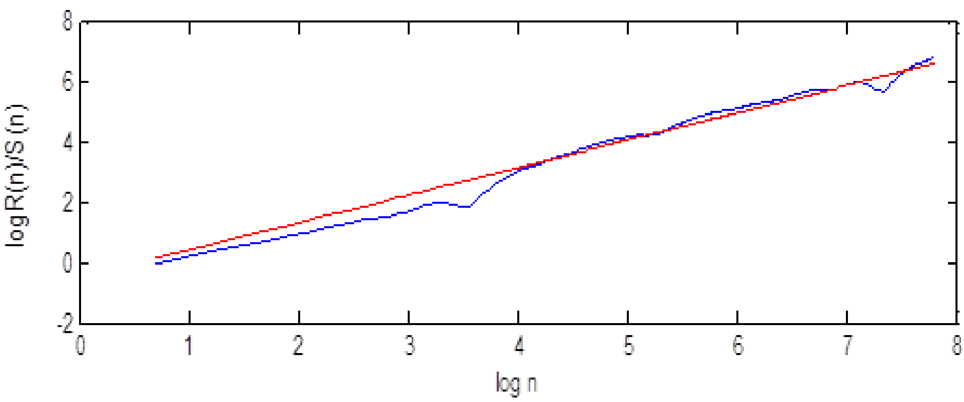

Figure 6: Pox plot of Hurst Exponent

From figure 6, the Hurst Exponent is estimated to be 0.9038. This indicates that the data set had long dependence range toward the time series. It can be categorised in case (i) which is $0.5 < H \leq 1$. It shows that the time series have long-term positive autocorrelation. This means that a high value in the series will probably be followed by another high value and that the values along time into the future will also tend to be high. This series is considered to be persistent. The strength of this persistent trend depends on how close H is to 1.

4 Discussion and Conclusion

For the comparison between the analytic formula that derived by Dragulescu and Yakovenko [2] with the probability distribution of log returns for historical prices of FTSE Bursa Malaysia KLCI, the curves given by the analytic formula are much more flatter than the shape constituted by the real data points at a delay of time, t = 40 and 250. Furthermore, most of the real data points also do not lie nicely on the curves, the greater the time lag, the greater the deviation noticed. However, the probability distribution of log returns for prices of FTSE Bursa Malaysia KLCI fitted the theoretical curve better at a delay of time, t= 1 and 20. As t increases, the greater the deviation noticed.

The Hurst exponent estimation showed that the index prices of FTSE Bursa Malaysia KLCI have long range dependence time series. Thus, it is reasonable that the probability distribution of log returns for historical prices of FTSE Bursa Malaysia KLCI does not lie nicely for the greater time lag.

The probability distribution of the index prices of FTSE Bursa Malaysia KLCI will help investors to have an idea on the probability of getting certain returns after a certain period. Since FTSE Bursa Malaysia KLCI comprises the largest 30 companies listed on the Main Board, it will more aptly define market activities in Malaysian stock market. The result would provide a guideline to the investors to plan a good strategy for their investment based on the result of this study.

We conclude that, the analytical formula that was derived by Dragulescu and Yakovenko [2] does not agree very well with the probability distribution for log returns of index prices FTSE Bursa Malaysia KLCI which is from 1st September 1998 until 31st July 2008. By looking at Figure 1, there is a strong downward trend in 2000 but still we are trying to fit the data using a constant positive growth rate μ. However, the constant positive growth rate μ is not suitable to describe the growth rate of the data. As mentioned in the paper of Dragulescu and Yakovenko [2], the growth rates in 1998, 2000 and 2008 were rather negative values than positive values. The Hurst exponent showed that the data set has long term memory which is persistent. Thus, it is reasonable that the analytical formula does not agree very well with the probability distribution for log returns of index prices FTSE Bursa Malaysia KLCI during the period. This is due to the exclusion of long memory in the model. It can be seen for Heston model in equation (2) and (3) which averaged the long time variance, θ. For future research, this study will consider long memory in Heston model which implies a new technique in parameters estimation need to be developed.

Acknowledgements

The authors are thankful to the Ministry of Education (Higher Education) and Universiti Teknologi Malaysia for the financial support under vote 07J62 in carrying out this research.

References

[1] Heston, S. L. A closed-form solution for options with stochastic volatility with applications to bond and currency options. The Review of Financial Studies. 1993. Vol 6. Issue 2. 327-343.

[2] Dragulescu, A. A. and Yakovenko, V. M. Probability distribution of returns in the heston model with stochastic volatility. Quantitative Finance. Institute of Physics Publishing. 2002. Vol 2. 443-453.

[3] Silva, A. C. and Yakovenko, V. M. Comparison between the probability distribution of returns in the Heston model and empirical data for stock indexes. Physica A. 2003. Vol 324. 303 - 310. Elsevier.

[4] Hughes, H. Crony capitalism and the East Asian currency and financial ‘crises’. 1999 Policy. 3-5. Spring.

[5] Goh, S. K. and Michael Lim, M. H. The impact of the global financial crisis : the case of Malaysia. TWN Global Economy Series. Third World Network. Jutaprint. 2010.

[6] The World Bank. Malaysia Economic Monitor. 2001. 61483.

[7] Hurst, H. Long term storage capacity of reservoirs. Transactions of The America Society of Civil Engineering. 1951. 116:770-799.

[8] Mandelbrot, B.B. Limit theorems on the self-normalised range for weakly and strongly dependent processes. Z. Wahrscheinlichkeitstheorie verw. Gebiete. 1975. 31:271-285.

[9] Gardiner, C. W. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. Berlin: Springer. 1993.

[10] Picchini, U. SDE Toolbox : Simulation and estimation of stochastic differential equations with MATLAB. 2007. http://sdetoolbox.sourceforge.net

[11] Maruyama, G. Continuous Markov processes and stochastic equations. Rendiconti del Circolo Matematico di Palermo. 1955. Vol 4. Issue 1. 48 - 90. Springer-Verlag.

[12] Mandelbrot, B. B. and Taqqu, M. S. Robust R/S analysis of long run serial correlation. Proceedings of the 42nd session of the International Statistical Institute, Manila. 1979. Bulletin of the I.S.I. 48 (Book 2) 69-104.

[13] Beran, J. *Statistics for Long-Memory Processes*. New York: Monographs on Statistics and Applied Probability 61, Chapman and Hall. 1994.

MATEMATIKA, 2014, Volume 30, Number 1a, 44-58

MATEMATIKA, 2014, Volume 30, Number 1a, 44-58