Fuzzy Time Series and Geometric Brownian Motion in Forecasting Stock Prices in Bursa Malaysia

Nor Hayati Shafii¹, Nur Ezzati Dayana Mohd Ramli², Rohana Alias³, Nur Fatihah Fauzi⁴

¹,²,³,⁴ Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Perlis, 02600 Arau, Perlis, Malaysia

Authors' Email Address: *norhayatishafii@uitm.edu.my, ²dayana.ezzati03@gmail.com, ³annna@uitm.edu.my, ⁴fatihah@uitm.edu.my

Received Date: 30 September 2019
Accepted Date: 6 November 2019

ABSTRACT

Every country has its own stock market exchange, which is a platform to raise capital and is a place where shares of listed company are traded. Bursa Malaysia is a stock exchange of Malaysia and it is previously known as Kuala Lumpur Stock Exchange. All over the world, including Malaysia, it is common for investors or traders to face some loss due to wrong investment decisions. According to the conventional financial theory, there are so many reasons that can lead to bad investment decisions. One of them is confirmation bias where an investor has a preconceived notion about an investment without good information and knowledge. In this paper, we had studied the best way to provide good information for investors in helping them to make the right decisions and not to fall prey to this behavioral miscue.

Two models for forecasting stock prices data are employed, namely, Fuzzy Time Series (FTS) and Geometric Brownian Motion (GBM). This study used a secondary data consisting of AirAsia Berhad daily stock prices for a duration of 20 weeks from January 2015 to May 2015. The 16-week data from January to April 2015 was used to forecast the stock prices for the 4-weeks of May 2015. The results showed that FTS has the lowest values of the Mean Absolute Percentage Error (MAPE) and the Mean Square Error (MSE), which are 1.11% and MYR²0.0011, respectively. For comparison, for GBM, the MAPE is 1.53% and the MSE is MYR²0.0017. The findings implied that the FTS model provides a more accurate forecast of stock prices.

Keywords: Forecasted values, stock market, Fuzzy Time Series, Geometric Brownian Motion

INTRODUCTION

The stock market exchange is one of the most vital components of free-market economy. It allows companies to raise money by offering stock shares and corporate bond. It also acts as a platform to build wealth for investors. Stock price is one of the economic indicators that measures the economic health of a country. Forecasting stock prices is very challenging as its nature is unpredictable and extremely volatile. Its behavior can be influenced by numerous factors such as demand and supply, global economy, politics, market trends and many more. The closing stock price is the price of a stock at the end of a trading day. It is very significant for several reasons; it can determine how well or poorly a stock performs, which is a big deal for not only investors but also financial institutions and other stakeholders. It is also a standard figure watched by an individual or organizations in making decisions about the stock and the company. Wrong decision in selecting the counters may end up in capital loss. Thus, being able to predict the future
of closing stock prices accurately is important to avoid losses and gain more profit. There are many mathematical models introduced by researches in predicting stock prices such as Autoregressive Integrated Moving Average (ARIMA), Artificial Neural Network (ANN), Log Normal Distribution, Geometric Brownian Motion (GBM) and Fuzzy Time Series (FTS). Each of these methods has their own strengths and weaknesses. As a result, it is sometimes hard to pick the best method for predicting future stock prices accurately. There are several researches that have been done on comparison between those listed models to assist an individual and organizations in making decisions but none of them has been done to compare the effectiveness between FTS and GBM. Therefore, this research will provide a basic guideline on the procedure of forecasting future stock prices using GBM and FTS. Their strengths and weaknesses will be listed. In addition, it will be proposed which model is more effective between the two approaches.

DATA AND RESEARCH METHOD

In this study, secondary data is taken from the Bursa Malaysia website. The data consists of daily closing stock prices of AirAsia Berhad for the duration of 20 weeks (five months) from January 2015 to May 2015. A total of 16 weeks data from January to April 2015 is treated as the estimation period which is then be used to forecast the stock prices for the next four weeks of May 2015. The data is chosen just for simulation purposes to demonstrate the accuracy of the methods applied.

Geometric Brownian Motion Approach

Geometric Brownian Motion is a stochastic model of non-negative variation of Brownian Motion. This process only assumes a positive value and is somewhat easy to calculate. This method is one of the mathematical models applied to model stock pricing, natural resources prices and to model the development in demand for services. GBM method is regularly used to model the price movement by estimating its drift and volatility. According to Abidin & Jaffar 2014, this model is very efficient to predict share prices in a short period of investment time and it is suitable for investors who want immediate share prices outlook. Their study shows that GBM is highly accurate model proven by the MAPE value lower than 10% and mentions that GBM can be used to predict the future share prices for the next two weeks duration. This can give enough time for investors to evaluate their decision in making maximum profit.

GBM model deals with randomness, returns, volatility and drift. The proposed GBM in this paper can be summarized as follows:

Step 1: Collection of the historical data of AirAsia Berhad stock prices.

Step 2: Calculation of stock return
\[
R_t = \frac{S_t - S_{t-1}}{S_t}
\]
where:
- \( R_t \): stock return at time \( t \),
- \( S_t \): stock price at time \( t \),
- \( S_{t-1} \): stock price at time \( t - 1 \).

Step 3: Calculation of the drift value (\( \mu \))
\[ \mu = \frac{1}{M \delta t \sum_{t-1}^M R_t} \]  
(1.2)

where:
\( R_t \) : stock return at time \( t \), \( M \) : amount of stock return, \( \mu \) : drift, \( \delta t \) : timesteps.

The drift value is the mean of rate of return at which the price of asset increases as the period of time rises. The \( \delta t \) in the formula denotes the timestep which equals the approximate number of 1/252.

**Step 4:** Calculation of the volatility value (\( \sigma \))

\[ \sigma = \sqrt{\frac{1}{(M-1)\delta t} \sum_{t=1}^M (R_t - \bar{R})^2} \]  
(1.3)

where \( \sigma \) : volatility, \( \bar{R} \) : mean of stock return.

The volatility refers to the movement or fluctuation of the stock prices either increase or decrease.

**Step 5:** Calculation of the stock price forecasting

\[ S_t = S_{t-1} e^{\left( \mu - \frac{1}{2} \sigma^2 \right) \delta t + \sigma \sqrt{\delta t} \epsilon} \]  
(1.4)

where \( S_t \) : stock price at time \( t \), \( S_{t-1} \) : stock price at time \( t-1 \), \( \mu \) : volatility, \( \sigma \) : drift, \( \epsilon \) : any random number from standardized normal distribution.

In this case the period of time \( t \) when the time \( dt = t \) is applied (Abdelmoula, 2006 & Affianti, and Putri, 2018)

**Example calculation of forecasted values for GBM approach**

The value of \( \delta t \) is approximately equals to 1/252. The value of 252 is the approximate number of trading days excluded weekend days and public holidays (Wilmott, 2007). After finding the return, calculate the drift rate, \( \mu \) (growth rate of an asset or expected return) using equation (1.2). Then, the volatility, \( \sigma \) is obtained using (1.3) and equation (1.4) is used to find the forecasted value of stock prices for AirAsia Berhad.
Fuzzy Time Series Forecasting Approach

Fuzzy time series is an idea formulated by Song and Chissom proposed in 1993. It has many applications such as predicting number of students enrollment (Stevenson & Porter, 2009), forecasting wheat production (Stevenson & Porter, 2009), forecasting rainfall distribution of a region (Dani & Sharma, 2013), forecasting short-term electric load proposed by Huang (2015) and etc. According to Hakan et al. (2014), fuzzy time series approach is able to deal with very small data and does not require the linearity assumption. This method makes the process of calculation become straightforward, (Dani & Sharma, 2013) and easy to apply in so many problems.

There are eight steps to be followed. At first, the universal set is defined as \( U = [D_{\text{min}} - D_1, \overline{D}_{\text{max}} + D_2] \) based on the \( D_{\text{max}} \) and \( D_{\text{min}} \). Then, it is converted into some partitions in equal length intervals. In the next step, linguistic values are defined as \( A_i \). After fuzzification, appropriate fuzzy relations which are fuzzy logical relationship (FLRs) and fuzzy logical relationship groups (FLRG) are established between them and in the next step, forecasting are carried out. The proposed FTS in this paper can be summarized as follows:

**Step 1:** Collect the historical data \( D_h \).

**Step 2:** Define the universe of discourse \( U \). Determine the maximum \( D_{\text{max}} \) and the minimum \( D_{\text{min}} \) of the historical stock prices. For easy partitioning of \( U \) positive numbers \( D_1 \) and \( D_2 \) are assigned. The universal set \( U \) is given by \( U = [D_{\text{min}} - D_1, \overline{D}_{\text{max}} + D_2] \).

**Step 3:** Divide the universe of discourse into seven intervals \( U_1 \) until \( U_7 \) with equal length of intervals, \( U_i \). Based on the distribution of actual data, sort \( U_i \) into intervals of distinct length \( v_i \)
Step 4: Establish the fuzzy trapezoidal number. Based on the intervals, \( v_j = [d_1, d_2] \)
\[
\begin{align*}
v_2 &= [d_2, d_3] \\
&\cdots \quad v_m = [d_m, d_{m-1}] 
\end{align*}
\]
that was derived in Step 3, the fuzzy number \( A_1, A_2, A_3, \ldots, A_m \), are defined as follows:
\[
\begin{align*}
A_1 &= [d_0, d_1, d_2, d_3] \\
A_2 &= [d_1, d_2, d_3, d_4] \\
&\cdots \\
A_m &= [d_{m-2}, d_{m-1}, d_m, d_{m+1}].
\end{align*}
\]

Step 5: Identify the fuzzified historical data and assign a corresponding fuzzy set to every value in the data series. For example, if the value of the historical data is allocated in \( v_j \), then it will fit to the fuzzy number \( A_j \). All the data set are arranged based on its fuzzy number.

Step 6: Develop a relationship for all the fuzzified historical data. The relationship is known as FLRs. For example:
\[
A_{y_1} \rightarrow A_{y_2}, \quad A_{y_2} \rightarrow A_{y_3}, \quad \ldots \quad A_{y_y} \rightarrow A_{nx}.
\]

Step 7: Create the FLRG. The FLRs can be arranged into groups based on the same fuzzy number on the left-hand sides of the relationships. Verify the rules for each of the relationship groups shown as follows:
\[
A_k \rightarrow A_j, \quad A_k \rightarrow A_j, \quad A_k \rightarrow A_j.
\]

Step 8: Classify the rule for each fuzzy logical relationship group. There are three different types of rules involved as shown as follows:

**Rule 1:** If the fuzzy logical relationship group of \( A_k \) is empty, \( A_k \rightarrow \phi \) or \( A_k \rightarrow A_k \) then forecasted value \( F_t \) is \( R[NSTFN(A_k)] \).

**Rule 2:** If the fuzzy logical relationship group of \( A_k \) is one to one \( A_k \rightarrow A_j \) then the forecasted value \( F_t \) is \( R[NSTFN(A_j)] \).

**Rule 3:** If the fuzzy logical relationship group of \( A_k \) is one to many \( A_k \rightarrow A_{j_1}, \ A_k \rightarrow A_{j_2}, \ldots, \ A_k \rightarrow A_{j_p} \), then the forecasted value of \( F_t \) is calculated as follows:
\[
F_t = R\left[\frac{NSTFN(A_{j_1}) + NSTFN(A_{j_2}) + \ldots + NSTFN(A_{j_p})}{p}\right].
\]

where \( NSTFN \) is Nearest Symmetric Trapezoidal Fuzzy Number.

Example calculation of forecasted values using Rule 2 and Rule 3 of FTS model

Rule 2: For \( A_{63} \rightarrow A_{69} \)
\[
A_{69} = [2.555, 2.567, 2.579, 2.591]
\]
\[
t_2 = 2.567 \\
t_3 = 2.579 \\
l_1 = (2.567 - 2.555) = 0.012 \\
l_4 = (2.591 - 2.579) = 0.012
\]
\[
\frac{t_4 - t_1}{4} = \left(\frac{0.012 - 0.012}{4}\right) = 0 \quad \frac{t_4 + t_1}{2} = \left(\frac{0.012 + 0.012}{2}\right) = 0.012
\]

\[
\text{NSTFN}(A_{a_0}) = \begin{bmatrix}
\left(t_2 + \frac{t_4 - t_1}{4} - \frac{t_4 + t_1}{2}\right), & \left(t_3 + \frac{t_4 - t_1}{4}\right), & \left(t_3 + \frac{t_4 - t_1}{4}\right), \\
\left(t_3 + \frac{t_4 + t_1}{2} + \frac{t_4 + t_1}{2}\right)
\end{bmatrix}
\]

\[
= \begin{bmatrix}
(2.567 + 0 - 0.012), (2.567 + 0), (2.579 + 0), \\
(2.579 + 0.012 + 0.012)
\end{bmatrix}
\]

\[
= [2.555, 2.567, 2.579, 2.603]
\]

\[
R[\text{NSTFN}(A_{a_0})] = \frac{(2.555 + 2.567 + 2.579 + 2.603)}{4} = 2.576 \approx 2.60
\]

Rule 3: For \(A_{70} \rightarrow A_{70}, A_{73}\)

\(A_{70} = [2.567, 2.579, 2.591, 2.603]\)

\(t_2 = 2.579\) \quad \(t_3 = 2.591\)

\(t_1 = (2.579 - 2.567) = 0.012\) \quad \(t_4 = (2.603 - 2.591) = 0.012\)

\[
\frac{t_4 - t_1}{4} = \left(\frac{0.012 - 0.012}{4}\right) = 0 \quad \frac{t_4 + t_1}{2} = \left(\frac{0.012 + 0.012}{2}\right) = 0.012
\]

\[
\text{NSTFN}(A_{a_0}) = \begin{bmatrix}
\left(t_2 + \frac{t_4 - t_1}{4} - \frac{t_4 + t_1}{2}\right), & \left(t_3 + \frac{t_4 - t_1}{4}\right), & \left(t_3 + \frac{t_4 - t_1}{4}\right), \\
\left(t_3 + \frac{t_4 + t_1}{2} + \frac{t_4 + t_1}{2}\right)
\end{bmatrix}
\]

\[
= \begin{bmatrix}
(2.579 + 0 - 0.012), (2.579 + 0), (2.591 + 0), \\
(2.591 + 0.012 + 0.012)
\end{bmatrix}
\]

\[
= [2.567, 2.579, 2.591, 2.615]
\]

\(A_{73} = [2.603, 2.615, 2.626, 2.638]\)

\(t_2 = 2.615\) \quad \(t_3 = 2.626\)

\(t_1 = (2.615 - 2.603) = 0.012\) \quad \(t_4 = (2.638 - 2.626) = 0.012\)

\[
\frac{t_4 - t_1}{4} = \left(\frac{0.012 - 0.012}{4}\right) = 0 \quad \frac{t_4 + t_1}{2} = \left(\frac{0.012 + 0.012}{2}\right) = 0.012
\]
\[
NSTFN(A_{13}) = \left[ \frac{t_2 + t_4 - t_1}{4} - \frac{t_4 + t_1}{2}, \frac{t_2 + t_4 - t_1}{4} + \frac{t_4 + t_1}{2}, \frac{t_2 + t_4 - t_1}{4} \right] = \left[ (2.615 + 0 - 0.012), (2.615 + 0), (2.626 + 0), (2.626 + 0.012 + 0.012) \right] = \left[ 2.603, 2.615, 2.626, 2.638 \right]
\]

\[
F_i = R\left[ \frac{2.567 + 2.603}{2}, \frac{2.579 + 2.615}{2}, \frac{2.591 + 2.626}{2}, \frac{2.615 + 2.633}{2} \right] = R\left[ 2.585, 2.597, 2.609, 2.633 \right] = \frac{2.585 + 2.597 + 2.609 + 2.633}{4} = 2.606 \approx 2.61
\]

This process is repeated to forecast the stock price values of the remaining fuzzy logical relationship groups.

**Accuracy Test**

In this study, the accuracy of the forecasted value is evaluated using Mean Absolute Percentage Error (MAPE) and Mean Square Error (MSE). The value of MAPE is calculated using equation (1.5) and MSE using equation (1.6).

\[
\varepsilon = \frac{1}{n} \sum_{n=1}^{n} \frac{|X_A - X_P|}{X_A}
\]

\[
MSE = \frac{1}{n} \sum_{n=1}^{n} (X_A - X_P)^2
\]

where

\[ n = \text{number of days}, \quad X_A = \text{actual value of prices}, \quad X_P = \text{predicted value of prices}. \]

The value of MAPE is analyzed based on the scale of judgement as shown in Table 1 while the value of MSE is analyzed based on the lowest value of error.

| Table 1: A scale of judgement of forecast accuracy |
|----------------|---------------------------------|
| MAPE           | Judgement of forecast accuracy  |
| \(\varepsilon < 10\%\) | Highly accurate                  |

246
11% to 20%  Good
21% to 50%  Reasonable
$\varepsilon > 51%$  Inaccurate

These error measurements will distinguish the most accurate model between GBM and FTS.

RESULTS AND DISCUSSION

Table 2 shows the comparison between the actual stock prices and the forecasted stock prices using FTS and GBM models.

<table>
<thead>
<tr>
<th>Date</th>
<th>Actual Price (RM)</th>
<th>Forecasted Price by FTS (RM)</th>
<th>Forecasted Price by GBM (RM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/5/2015</td>
<td>2.24</td>
<td>2.28</td>
<td>2.27</td>
</tr>
<tr>
<td>6/5/2015</td>
<td>2.24</td>
<td>2.24</td>
<td>2.20</td>
</tr>
<tr>
<td>7/5/2015</td>
<td>2.25</td>
<td>2.24</td>
<td>2.26</td>
</tr>
<tr>
<td>8/5/2015</td>
<td>2.23</td>
<td>2.23</td>
<td>2.30</td>
</tr>
<tr>
<td>11/5/2015</td>
<td>2.20</td>
<td>2.25</td>
<td>2.17</td>
</tr>
<tr>
<td>12/5/2015</td>
<td>2.21</td>
<td>2.22</td>
<td>2.25</td>
</tr>
<tr>
<td>13/5/2015</td>
<td>2.24</td>
<td>2.24</td>
<td>2.22</td>
</tr>
<tr>
<td>14/5/2015</td>
<td>2.29</td>
<td>2.24</td>
<td>2.23</td>
</tr>
<tr>
<td>15/5/2015</td>
<td>2.28</td>
<td>2.28</td>
<td>2.29</td>
</tr>
<tr>
<td>18/5/2015</td>
<td>2.28</td>
<td>2.26</td>
<td>2.30</td>
</tr>
<tr>
<td>19/5/2015</td>
<td>2.23</td>
<td>2.26</td>
<td>2.26</td>
</tr>
<tr>
<td>20/5/2015</td>
<td>2.24</td>
<td>2.25</td>
<td>2.25</td>
</tr>
<tr>
<td>21/5/2015</td>
<td>2.16</td>
<td>2.24</td>
<td>2.19</td>
</tr>
<tr>
<td>22/5/2015</td>
<td>2.06</td>
<td>2.07</td>
<td>2.07</td>
</tr>
<tr>
<td>25/5/2015</td>
<td>2.07</td>
<td>2.08</td>
<td>2.05</td>
</tr>
<tr>
<td>26/5/2015</td>
<td>2.06</td>
<td>2.08</td>
<td>2.08</td>
</tr>
<tr>
<td>27/5/2015</td>
<td>2.09</td>
<td>2.08</td>
<td>1.98</td>
</tr>
<tr>
<td>28/5/2015</td>
<td>2.08</td>
<td>2.15</td>
<td>2.04</td>
</tr>
<tr>
<td>29/5/2015</td>
<td>2.19</td>
<td>2.15</td>
<td>2.22</td>
</tr>
</tbody>
</table>

Figure 2 and Figure 3 illustrate the fluctuations of the stock price between the actual and the forecasted values. The horizontal axis of the graph displays the timeframe of day-by-day forecasting period and the vertical axis represents the stock prices over one-month period in May 2015. The movement of the forecasted stock prices from FTS imitates the actual stock prices while the forecasted stock prices from GBM shows higher fluctuations at certain time period. Nevertheless, for overall movement, it shows that most of the forecasted stock prices are closest to the actual stock prices. Hence, to select the best forecasting model, the error values are calculated.
Figure 2: The graph of actual vs forecast stock prices of AirAsia Berhad within four weeks using FTS

Figure 3: The graph of actual vs forecast stock prices of AirAsia Berhad within four weeks using GBM

Table 3 and Table 4 shows the comparison of MAPE and MSE values between GBM and FTS models for four different period of output data which are four weeks, three weeks, two weeks and one week.

**Table 3: Comparison of MAPE values between GBM and FTS models**

<table>
<thead>
<tr>
<th>No. of output data</th>
<th>Geometric Brownian Motion (%)</th>
<th>Fuzzy Time Series (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE 4 weeks of output data</td>
<td>1.53%</td>
<td>1.11%</td>
</tr>
<tr>
<td>MAPE 3 weeks of output data</td>
<td>1.30%</td>
<td>1.00%</td>
</tr>
<tr>
<td>MAPE 2 weeks of output data</td>
<td>1.49%</td>
<td>0.79%</td>
</tr>
<tr>
<td>MAPE 1 weeks of output data</td>
<td>1.68%</td>
<td>0.56%</td>
</tr>
</tbody>
</table>
Table 4: Comparison of MSE values between GBM and FTS model

<table>
<thead>
<tr>
<th>No. of output data</th>
<th>Geometric Brownian Motion (MYR²)</th>
<th>Fuzzy Time Series (MYR²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE 4 weeks of output data</td>
<td>0.0017</td>
<td>0.0011</td>
</tr>
<tr>
<td>MSE 3 weeks of output data</td>
<td>0.0011</td>
<td>0.0010</td>
</tr>
<tr>
<td>MSE 2 weeks of output data</td>
<td>0.0015</td>
<td>0.0008</td>
</tr>
<tr>
<td>MSE 1 weeks of output data</td>
<td>0.0018</td>
<td>0.0004</td>
</tr>
</tbody>
</table>

Three weeks of output data from GBM model has produced the lowest MAPE and MSE values which are 1.3% and MYR² 0.0011, respectively. This indicates that the forecasted values of GBM model are most accurate within 3-week period. Meanwhile for FTS model, the lowest MAPE and MSE are from one week of output data which are 0.56% and MYR² 0.0004 respectively. This indicates that the forecasted values of FTS model are most accurate within one-week period. The biggest values of MAPE and MSE for GBM model are 1.68% and 0.0018 from analysis of output data within one week meanwhile for FTS model are 1.11% and 0.0011 from analysis of output data within four weeks. It is proven that GBM model is not really accurate to be used in forecasting data within one week while FTS is not really accurate to be used in forecasting data within four weeks.

In overall, from the lowest measurement values of MAPE and MSE of both methods, FTS model has smaller values of error measures compared to GBM model. Therefore, from this result, we can conclude that, the FTS model is the more effective way to forecast the closing prices of stock exchange.

CONCLUSION

From the results obtained, several conclusions can be drawn. First and foremost, results showed that FTS is the best model in forecasting a specific stock price in Bursa Malaysia as compared to GBM model since FTS model possesses a smaller MSE and MAPE error measure which are 1.11% and MYR² 0.0011, respectively. Secondly, it is shown that one week of output data of FTS model has the lowest MAPE and MSE values which are 0.56% and MYR² 0.004, respectively. Thus, it can be concluded that the forecasted values are most accurate within one-week period. Other than that, the value of MAPE for FTS model is less than 10% and most of its forecasted stock prices are closest to the actual stock prices. Therefore, this indicates that the accuracy of the forecasted values of output data of this model are highly accurate.

REFERENCES


