E-ISSN 2621–086X

Indonesian Journal of Applied Statistics

Vol. 7, No. 1, pp. 41-49, May 2024
Journal homepage: https://jurnal.uns.ac.id/ijas/index
doi: 10.13057/ijas.v7i1.54309

Forecasting on Closing Stock Price Data Using Fuzzy Time Series

Sri Subanti*, Asti Rahmaningrum
Study Program of Statistics, Universitas Sebelas Maret, Surakarta, Indonesia

***Corresponding author:** sri_subanti@yahoo.co.id

Submitted: 11-Aug-2021 **Revised:** 7-Feb-2024 **Accepted:** 2-Sep-2024

Abstract. *The stock prices move up and down during trading time which is obtained from time series data. Investors need to estimate the fluctuation of stock prices in the future day to make the best investment decision. Fuzzy time series can be used as an alternative by investors in making stock price predictions. The advantage of this forecasting method compared to others is that it can formulate a problem based on expert knowledge or empirical data. This research aims to apply fuzzy time series in estimating the future value of closing stock price on the LQ45 Index. Three different methods will be applied to the data which are Chen, Lee, and Cheng. The data of the LQ45 Index will be obtained during the period of January, 4th until April 30th, 2021. The LQ45 index is chosen by many investors because it has high returns. All three model were applied and has a different rule in the calculation stage. The results show that all three models give different forecasting values and different performance of accuracy. The Lee method has the lowest values of accuracy, meanwhile the Cheng method has the highest value of accuracy. It can be concluded that Lee method is the best model indicated by the lowest value of RMSE, MAD, and MAPE for estimating the closing stock price of the LQ45 index.*

Keywords: *Stock Price, Fuzzy Time Series, LQ45 Index*

1. INTRODUCTION

Stock in the form of paper explains the ownership of securities which is traded in the stock market [1]. The interaction of the seller and the buyer causes stock price movement, moreover, the closing stock price refers to the last price during trading day. The stock price movement has to be considered in order to minimize risk. Investors can estimate the value of future prices using forecasting analysis. Forecasting is a technique method to estimate value in the future by analyzing the pattern of historical data [2]. Forecasting is very useful in determining investment decision so that investors can decide whether to sell or buy stock on the next trading.

One of the forecasting types is fuzzy time series which is applied to time series data. The difference among other forecasting methods is that the fuzzy time series can be applied to linguistic values. Fuzzy time series was introduced by Song and Chissom in 1993 and its method has developed a lot. The riyal-rupiah exchange rate has been forecast using fuzzy time series method [3], moreover, fuzzy time series have been used in many areas such as meteorology [4] and agriculture [5]. In investment areas, fuzzy time series can be used to estimate the future value of stock prices. Investment decision-making utilized time series data to estimate their business plan because the pattern of changes in historical data is believed to be repeated in the present.

Furthermore, forecasting analysis can be applied to time series data or is known as time series analysis. Time series analysis aims to identify the components factors that can affect the value in a number series [6]. It is done by breaking down the component of the historical data and projecting them for future values or forecasting. The stock price movement during trading obtained daily time series data of closing stock price which is recorded after trade closed. The performance of stock price in the market will be measured with stock price index. One of the stock indexes is the LQ45 which has stocks with high returns, therefore many investors choose. LQ45 index consists of 45 stocks with high liquidity and large market capitalization [7].

The research aim is to forecast closing stock price data of the LQ45 Index using fuzzy time series method. The advantage of fuzzy time series is that it does not require any assumption in the calculation process. Moreover, it has various models where the difference is on the defuzification process. This research will apply three different models that are Chen, Lee, and Cheng as development models of fuzzy time series. Previously, this research will determine the same interval length for the three models using a frequency distribution table. This research also determines the best model by calculating the model accuracy on each result. The results of the research are expected to be used as consideration in making investment decision by investors, both short-term and long-term.

2. THEORETICAL BASIC

2.1 Time Series

A time series is a set of data that is arranged in order of time, for an example in the form of daily, weekly, monthly, and yearly. Time series are used as a reference in predicting what changes will occur in the next period based on patterns that occur in the present and the past. Therefore, time series is widely used in forecasting analysis such as stock prices, money supply, exchange rate, consumer price index, population growth index, and so on.

2.2 Fuzzy Time Series

Forecasting based on fuzzy time series is a forecasting method to capture the patterns of historical data as a basis for projecting future data [8]. In general, the step of fuzzy time series analysis according to Song and Chissom [9] are:

Step 1. Define the Universe of Discourse $U = Dmin - Z_1,\ Dmax + Z_2$ where $Dmin$ is the smallest value of the data, $Dmax$ is the largest value of the data and Z is any positive value.

Step 2. Forming the interval of U with the same length. Determining the length interval of class can be done by frequency distribution table as follows Equation (1).

$$The\ number\ of\ class = 1 + 3.322 \log (n) \tag{1}$$

Furthermore the length intervals can be obtained by:

$$interval\ length = \frac{the\ number\ of\ class}{Dmin - Dmax}$$

Step 3. Define the degree of fuzzy set membership as written below Equation (2).

$$\begin{aligned} A_1 &= \frac{a_{11}}{u_1} + \frac{a_{12}}{u_2} + \frac{a_{13}}{u_3} + \cdots + \frac{a_{1m}}{u_m} \\ A_2 &= \frac{a_{21}}{u_1} + \frac{a_{22}}{u_2} + \frac{a_{23}}{u_3} + \cdots + \frac{a_{2m}}{u_m} \\ &\vdots \end{aligned} \tag{2}$$

$$A_k = \frac{a_{k1}}{u_1} + \frac{a_{k2}}{u_2} + \frac{a_{k3}}{u_3} + \cdots + \frac{a_{km}}{u_m}$$

Step 4. Fuzzify the historical data and transform into fuzzy set.
Step 5. Forming the fuzzy logical relationship (FLR) and fuzzy logical relationship group (FLRG) denoted by $A_j \rightarrow A_k$, where A_j is the current state and A_k is the next state. The notation $A_j \rightarrow A_k$ means that if the fuzzified of class i is A_j then the fuzzified of class $i+1$ is A_k.
Step 6. Defuzzyfication, where the quantifiable result in the form of a fuzzy set is produced.

Fuzzy Time Series Chen
The formation of FLRG in fuzzy time series Chen does not require repetition. For example if $A_j \rightarrow A_{k1}, A_{k2}, \ldots, A_{kn}$ then the FLRG only counted once or $A_j \rightarrow A_k$. The defuzzification of chen is defined in Equation (3).

$$\hat{Y}_t = \frac{m_i + \ldots + m_n}{n} \tag{3}$$

where $\hat{Y}_t$ is the forecast value at period *t* and m_i is the midpoint of U at class i.

Fuzzy Time Series Lee
The formation of FLRG in Lee does require repetition so that if $A_j \rightarrow A_{k1}, A_{k2}, \ldots, A_{kn}$ the FLRG is $A_j \rightarrow nA_k$. The defuzification of lee is defined in Equation (4).

$$\hat{Y}_t = \frac{p(i)m_i}{\Sigma p} + \cdots + \frac{p(n)m_n}{\Sigma p} \tag{4}$$

where $\hat{Y}_t$ is the forecast value at period t , $p(i)$ is number of repetition at class i, m_i is the mid point of U at class i.

Fuzzy Time Series Cheng
The formation of FLRG in Cheng does require repetition so that if $A_j \rightarrow A_{k1}, A_{k2}, \ldots, A_{kn}$ the FLRG is $A_j \rightarrow nA_k$. The defuzzification of chen stated that the forecast value at period t is the forecast value at $(t-1)$ The defuzification of cheng is defined in Equation (5).

$$\hat{Y}_t = \frac{p(i)m_i}{\Sigma p} + \cdots + \frac{p(n)m_n}{\Sigma p} \tag{5}$$

where $\hat{Y}_t$ is the forecast value at period t, p is number of repetition, $m_{i(t-1)}$ is the mid point of U at class i in period $(t-1)$.

2.3 Forecasting Model Accuracy

The measurement of the forecasting accuracy is calculated by comparing the actual value with the forecast value. Forecasting results can be said to be good if it were close to the actual values. The forecasting model accuracy can be calculated using [10].

Mean Absolute Percentage Error (MAPE)
Determination of MAPE as follows Equation (6):

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \frac{|F_t - X_t|}{X_t} \tag{6}$$

Where $F_t - X_t = e_t$ is the error value at period t, F_t is the forecast value at period t, X_t is the actual value at period t, and n is the number of observations.

Mean Absolute Deviation (MAD)
Determination of MAD as follows Equation (7):

$$MAD = \frac{1}{n}\sum_{t=1}^{n} |F_t - X_t| \tag{7}$$

where $F_t - X_t = e_t$ is the error value at period t, F_t is the forecast value at period t, X_t is the actual value at period t, and n is the number of observations.

Root of Mean Squared Error (RMSE)
Determination of RMSE as follows Equation (8):

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(F_t - X_t)^2}{n}} \tag{8}$$

where $F_t - X_t = e_t$ is the error value at period t, F_t is the forecast value at period t, X_t is the actual value at period t, and n is the number of observations.

3. METHODS

The secondary data derived from yahoo.finance.com used in this research that is daily closing stock price data of the LQ45 index for the period of January 4th until April 30th 2021, therefore 82 data. The steps of analysis are as follows:

i. Construct the descriptive statistics from the data obtained
ii. Forming the universe of discourse or U with the same length using the frequency distribution table.
iii. Determine the fuzzy set on U
iv. Fuzzifying the daily closing stock price of the LQ45 index
v. Determine the fuzzy logical relationship (FLR) and fuzzy logical relationship group (FLRG)
vi. Defuzzifying using Chen, Lee, and Cheng.
vii. Calculate the forecasting model accuracy
viii. Conclusions.

4. RESULTS AND DISCUSSION

The descriptive statistics of the data has the lowest price on April, 13th otherwise the highest price is on January, 20th 2021. The mean of the data on period observed is 941.659 and has the standard deviation at 33.364 as shown in Table 1.

Table 1. The descriptive statistics of the data

Data	Min	Max	Mean	Std. Deviation
The daily closing stock price of LQ45 Index on January, 4th -April, 30th 2021	882	1015	941.659	33.364

Figure 1 show the time series plot of the LQ45's daily closing stock price for the period of January 4th until April 30th 2021. According to Figure 1 it can be stated that the closing stock price of the LQ45 during the observed period tends to trend down from January to May 2021.

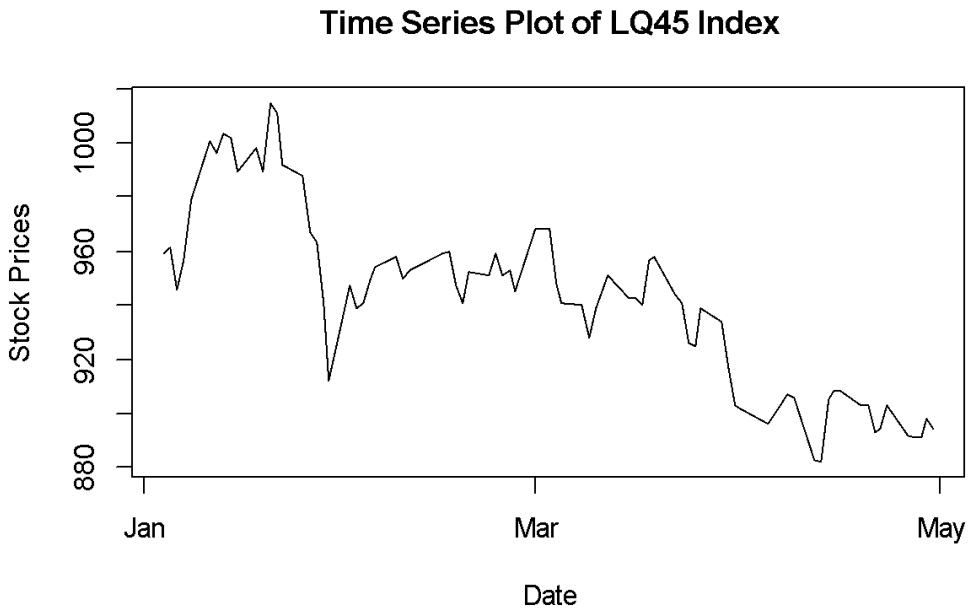

Figure 1. Time series plot of the LQ45 index

Using equation (1) the data was formed into 8 classes with each class length is 18, then the Universe of Discourse can be obtained below:

$$U_1 = [882, 899], U_2 = [900, 917], U_3 = [918, 935], U_4 = [936, 953],$$
$$U_5 = [954, 971], U_6 = [972, 989], U_7 = [990, 101], U_8 = [1008, 103].$$

Fuzzy set defined based on the equation (2) as follows:

$$A_1 = \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8},$$
$$A_2 = \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8},$$
$$A_3 = \frac{0}{u_1} + \frac{0.5}{u_2} + \frac{1}{u_3} + \frac{0.5}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8},$$
$$A_4 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0.5}{u_3} + \frac{1}{u_4} + \frac{0.5}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8},$$
$$A_5 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0.5}{u_4} + \frac{1}{u_5} + \frac{0.5}{u_6} + \frac{0}{u_7} + \frac{0}{u_8},$$
$$A_6 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0.5}{u_5} + \frac{1}{u_6} + \frac{0.5}{u_7} + \frac{0}{u_8},$$
$$A_7 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0.5}{u_6} + \frac{1}{u_7} + \frac{0.5}{u_8},$$
$$A_8 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0.5}{u_7} + \frac{1}{u_8},$$

Next is fuzzyification step where each observation is classified into classes based on fuzzy set. Then forming fuzzy logical relationship (FLR) that can be obtained by connecting A_i at period t with A_i at period $(t+1)$. The result of fuzzification and the FLR shown in Table 2.

Table 2. The result of fuzzification and the FLR

Date	Actual Enrollment	Fuzzified Enrollment	Fuzzy Logical Relationship (FLR)
Jan 4, 2021	959	A_5	$A_5 \rightarrow A_5$
Jan 5, 2021	961	A_5	$A_5 \rightarrow A_4$
Jan 6, 2021	946	A_4	$A_4 \rightarrow A_5$
Jan 7, 2021	957	A_5	$A_5 \rightarrow A_6$
Jan 8, 2021	979	A_6	$A_6 \rightarrow A_7$
Jan 11, 2021	1001	A_7	$A_7 \rightarrow A_7$
Jan 12, 2021	996	A_7	$A_7 \rightarrow A_7$
Jan 13, 2021	1003	A_7	$A_7 \rightarrow A_7$
Jan 14, 2021	1002	A_7	$A_7 \rightarrow A_6$
Jan 15, 2021	989	A_6	$A_6 \rightarrow A_7$
…	…	…	…
April 26, 2021	892	A_1	$A_1 \rightarrow A_1$
April 27, 2021	891	A_1	$A_1 \rightarrow A_1$
April 28, 2021	891	A_1	$A_1 \rightarrow A_1$
April 29, 2021	898	A_1	$A_1 \rightarrow A_1$
April 30, 2021	894	A_1	$A_1 \rightarrow A_1$

The fuzzy relationship group (FLRG) formed by grouping fuzzy sets which have the same current state then grouped into one group at the next state. The FLRG formed by two types that are the FLRG without repetition and The FLRG with repetition as shown in Table 3.

Table 3. The result of FLRG and FLRG with repetition

Group	Fuzzy Logical Relationship Group (FLRG)	Fuzzy Logical Relationship Group (FLRG) With Repetition
Group 1	$A_1 \rightarrow A_1, A_2$	$A_1 \rightarrow A_2, A_1, A_2, A_1, A_2, A_1, A_1, A_1$ $, A_1, A_1$
Group 2	$A_2 \rightarrow A_1, A_2, A_4$	$A_2 \rightarrow A_4, A_2, A_2, A_1, A_2, A_2, A_2, A_1$ $, A_2, A_2, A_2, A_1, A_1, A_2$
Group 3	$A_3 \rightarrow A_2, A_3, A_4$	$A_3 \rightarrow A_4, A_3, A_4, A_2$
Group 4	$A_4 \rightarrow A_2, A_3, A_4, A_5$	$A_4 \rightarrow A_5, A_2, A_4, A_4, A_4, A_5, A_4, A_5$ $, A_4, A_4, A_4, A_4, A_4, A_5, A_4, A_4, A_3, A_4$ $, A_4, A_4, A_4, A_5, A_4, A_3, A_3$
Group 5	$A_5 \rightarrow A_4, A_5, A_6$	$A_5 \rightarrow A_5, A_4, A_6, A_5, A_4, A_5, A_4, A_5, A_5$ $, A_4, A_4, A_4, A_5, A_5, A_4, A_5, A_4,$

Group 6	$A_6 \rightarrow A_5, A_7, A_8$	$A_6 \rightarrow A_7, A_7, A_8, A_5$
Group 7	$A_7 \rightarrow A_6, A_7$	$A_7 \rightarrow A_7, A_7, A_7, A_6, A_6, A_6$
Group 8	$A_8 \rightarrow A_7, A_8$	$A_8 \rightarrow A_8, A_7$

Then, the FLRG with repetition can be written in the form of matrix as shown in Table 4.

Table 4. The matrix of FLRG with repetition

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
A_1	7	3	-	-	-	-	-	-
A_2	4	9	-	1	-	-	-	-
A_3	-	1	1	2	-	-	-	-
A_4	-	1	3	16	5	-	-	-
A_5	-	-	-	8	8	1	-	-
A_6	-	-	-	-	1	-	2	1
A_7	-	-	-	-	-	3	3	
A_8	-	-	-	-	-	-	1	1

The next step is the defuzzification of forecasting values using Chen, Lee, and Cheng method. Based on equations (3), (4), and (5) the defuzzification of Chen, Lee, and Cheng is calculated in Table 5.

Table 5. The defuzzification of forecasting values using Chen, Lee, and Cheng method

Chen		
Fuzzification	FLRG	Defuzzification
A_4	$A_4 \rightarrow A_2, A_3, A_4, A_5$	$A_4 = \frac{908.5 + 926.5 + 944.5 + 962.5}{4} = 935.5$
Lee		
Fuzzification	FLRG	Defuzzification
A_4	$A_4 \rightarrow A_5, A_2, A_4, A_4, A_4, A_5, A_4, A_5, A_4, A_4, A_4, A_4, A_4, A_5, A_4$ $A_4, A_3, A_4, A_4, A_4, A_4, A_5, A_4, A_3, A_3$	$A_4 = \frac{5}{25}(962.5) + \frac{1}{25}(908.5) + \frac{16}{25}(944.5) + \frac{3}{25}(926.5) = 944.5$
Cheng		
Fuzzification	FLRG	Defuzzification
A_4	$A_5 \rightarrow A_5, A_4, A_6, A_5, A_4, A_5, A_4, A_5, A_5, A_4, A_4, A_4, A_5, A_5, A_4$ A_5, A_4	$A_5 = \frac{8}{17}(962.5) + \frac{8}{17}(944.5) + \frac{1}{17}(980.5) = 955.088$

Table 6 show the defuzzification result using Chen, Lee, and Cheng of the data obtained.

Table 6. The result of defuzzification

Date	LQ45	Fuzzification	Chen	Lee	Cheng
04-Jan-21	959	A_5	962.5	955.088	
05-Jan-21	961	A_5	962.5	955.088	955.088
06-Jan-21	946	A_4	935.5	944.5	955.088
07-Jan-21	957	A_5	962.5	955.088	944.5
08-Jan-21	979	A_6	992.5	994	955.088
11-Jan-21	1001	A_7	989.5	989.5	994
12-Jan-21	996	A_7	989.5	989.5	989.5
13-Jan-21	1003	A_7	989.5	989.5	989.5
14-Jan-21	1002	A_7	989.5	989.5	989.5
15-Jan-21	989	A_6	992.5	994	989.5
…	…	…	…	…	…
26-Apr-21	892	A_1	899.5	895.9	905.929
27-Apr-21	891	A_1	899.5	895.9	895.9
28-Apr-21	891	A_1	899.5	895.9	895.9
29-Apr-21	898	A_1	899.5	895.9	895.9
30-Apr-21	894	A_1	899.5	895.9	895.9

The model accuracy test can be applied to the forecast value in determining what method gives best result between Chen, Lee, and Cheng. Using equations (6), (7), and (8) the results of model accuracy are shown in the Table 7.

Table 7. The result of model accuracy

Method	$RMSE$	MAD	$MAPE$
Chen	8.972	7.744	0.825
Lee	6.237	5.051	0.533
Cheng	11.358	8.524	0.903

5. CONCLUSION

In this research the three different methods of Fuzzy Time Series which are Chen, Lee, and Cheng applied to daily closing stock price data of the LQ45 Index. The results are all three methods give different forecasting values. Based on the calculation of accuracy, Lee method has the lowest values of $RMSE$, MAD, and $MAPE$ which are 6.237, 5.051, 0.533% respectively. Meanwhile, the Cheng method showed the highest values of accuracy which are 11.358, 8.524, 0.903% respectively. It can be concluded that Lee method gives the best result for forecasting of the LQ45 stock price index.

REFERENCES

[1] T. Darmadji and H. M. Fakhruddin, *Pasar modal di Indonesia, edisi ketiga*. Jakarta, IDN: Salemba Empat, 2012.

[2] A. Aswi and S. Sukarna, *Analisis deret waktu aplikasi dan teori*. Makassar, IDN: Andhira Publisher, 2006.

[3] D. Permana and I. A. Fitri, "Application of fuzzy time series-markov chain method in forecasting data of exchange rate Riyal-Rupiah," in *Journal of Physics: Conf. Series* **1554** 0120055, 2020, pp. 1-7, doi:10.1088/1742-6596/1554/1/012005

[4] A. Arnita, N. Afnisah, and F. Marpaung, "A comparison of the fuzzy time series methods of Chen, Cheng and markov chain in predicting rainfall in Medan," in *Journal of Physics: Conf. Series* **1462** 012044, 2020, pp. 1-11, doi:10.1088/1742-6596/1462/1/012044.

[5] I. P. Rifki, D. Yudhiantoro, and E. Wahyurini, "Fuzzy time series model Cheng untuk meramalkan volume hasil panen pada Tanaman Garut," *Telematika*, vol. 4, no. 1, pp. 11-17, 2020.

[6] B. Render, R. M. Stair Jr., and M. E. Hanna, *Quantitative analysis for management*, Ed 8th, Inc. New Jersey, USA: Pearson Education, 2003.

[7] Bursa Efek Indonesia, www.idx.co.id accessed on June, 2021.

[8] W. D. Sugianto, A. M. Soleh, and F. M. Afendi, "Forecasting simulation with ARIMA and combination of Stevenson Porter Cheng fuzzy time series," *International Journal of Computer Science and Network*, vol. 6, no. 6, pp. 806-811, 2017.

[9] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series part I," *Fuzzy Sets and Systems*, vol. 54, no. 1, pp. 1-9, 1993.

[10] P. C. Chang, Y. W. Wang, and C. H. Liu, "The development of a weighted evolving fuzzy neural network for PCB sales forecasting," *Expert Systems with Applications*, vol. 32, no. 1, pp. 86-96, 2007.