

Research Article

Average-Based Fuzzy Time Series Markov Chain Based on Frequency Density Partitioning

Susilo Hariyanto,¹ Y. D. Sumanto,¹ Siti Khabibah,¹ and Zaenurrohman²

¹Department of Mathematics, Diponegoro University, Semarang 50275, Indonesia ²Magister Program of Mathemathics, Diponegoro University, Semarang 50275, Indonesia

Correspondence should be addressed to Susilo Hariyanto; susilohariyanto@lecturer.undip.ac.id

Received 26 November 2022; Revised 4 January 2023; Accepted 12 May 2023; Published 29 May 2023

Academic Editor: Wei-Chiang Hong

Copyright © 2023 Susilo Hariyanto et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fuzzy time series (FTS) is one of the forecasting methods that has been developed until now. The fuzzy time series is a forecasting method that uses the concept of fuzzy logic, which Song and Chissom first introduced. The fuzzy time series (FTS) Markov chain uses the Markov chain in defuzzification. The determination of the length of the interval in the fuzzy time series plays an important role in forming a fuzzy logic relationship (FLR), and this FLR will be used to determine the forecasting value. One method that can be used to determine the interval length is average-based. However, several studies use partitioning based on frequency density to obtain the optimal interval length to get better forecasting accuracy. This study combines the fuzzy time series Markov chain, Average-based fuzzy time series, and Fuzzy time series based on frequency density partitioning to become average-based fuzzy time series Markov chain based fuzzy time series fuzzy time series Markov chain method. This method is implemented in forecasting the Indonesian Islamic stock index (ISSI) for the selected period. The calculation of the accuracy level using the mean square error (MSE) and the mean average percentage error (MAPE) shows that the fuzzy Markov chain-based fuzzy time series based on the frequency density partition has a high level of accuracy in forecasting.

1. Introduction

A time series is defined as a collection of observations or observations made sequentially over time. Usually, observations in time series are not independent or can be said to be correlated. Thus, the order of the observations becomes important. This results in statistical procedures and techniques based on independent assumptions being no longer valid; thus, different methods and approaches are needed. Time series analysis aims for depiction, exposure, prediction, and monitoring [1].

Forecasting predicts a variable's values based on the known values of that variable or related variables [2]. The rationale for the time series is that current observations depend on previous observations. Therefore, many types of forecasting use time series data, including the fuzzy time series method, smoothing, average, and moving average, and others. The forecasting method using the fuzzy logic concept, hereinafter known as fuzzy time series, was first proposed by Song and Chissom in 1993. Song and Chissom used time-invariant and time-variant methods in forecasting. As a result, several fuzzy time series (FTS) methods have been developed, including Chen [3, 4], Chen and Hsu [5], weighted [6], backpropagation [7], multiple-attribute [8], percentage change [9], and the Markov chain [10].

The Markov chain is a stochastic process where future events only depend on today's events and do not depend on past conditions. The Markov chain is defined by a transition opportunity matrix that contains information regulating the system's movement from one state to another [10]. In fuzzy time series, the Markov chain is used in the defuzzification stage [10]. Defuzzification is the calculation step of fuzzy time series forecasting based on a fuzzy logical relationship group (FLRG). In FLRG, there is a relationship between the current and next states. The current state is the value that will be calculated as the predicted value, and the next state is the data used to get the value in the current state. Therefore, the relationship between the current state and the next state in the FLRG is considered a conditional process in line with the basic principles of the Markov chain method [10]. The Markov chain is used in several fields, one of which is the research of Indriyani and Pratiwi, et al. [11], which uses the Markov chain in the measles spread pattern. The Markov chain method is also used by Prasetya and Ferdian [12] in scheduling Oerlikon machine maintenance to optimize maintenance costs and time. The results show that calculations with the Markov chains produce more optimal time and costs.

Ruey-Chyn Tsaur conducted a study by combining the fuzzy time series method with the Markov chain concept to predict the Taiwan currency exchange rate against the dollar. The method used is known as the fuzzy time series Markov chain. The results obtained are the fuzzy time series Markov chain that has a better level of accuracy than the fuzzy time series [10]. Many studies have used the FTS-Markov chain including Dinatha et al. [13] who used the FTS-Markov chain to predict export profits. In addition, Hidayah and Sugiman [14] and Mangkunegara and Yerizon [15] both use the FTS-Markov chain to forecast exchange rates. Their research [13–15] produced a small error value, but in determining the FLR, there are still shortcomings and will produce different forecasting results because the length of the interval is determined according to the perception of each researcher.

In the fuzzy time series method, the determination of the interval length does not have a definite formula in its calculation; the interval is formed depending on the researcher [16], even though the determination of the length of the interval is very influential on the formation of a fuzzy logical relationship (FLR) which will result in differences in the results of forecasting calculations [17].

One method that can be used to determine the length of the interval is the average-based model which was introduced by Xihao and Yimin [17]. This average-based fuzzy time series uses an average-based method in determining the length of the interval. Research by Xihao and Yimin [17] also shows that the use of the average-based fuzzy time series method produces better forecasts than Chen's fuzzy time series method. Wuryanto and Puspita [18] and Kumar N. and Kumar H. [19] also used the average-based FTS model to predict the development of confirmed cases of COVID-19. Research [18, 19] has a small error value, but the interval length is still less than optimal because each class interval has the same interval length regardless of the frequency in each class.

Furthermore, Chen and Hsu [5] developed Chen's fuzzy time series, by repartitioning based on frequency density to predict the number of applicants at the University of Alabama. Chen and Hsu redivided intervals before the fuzzification process. Chen and Hsu's research shows that after the redivided interval, the fuzzy time series gets a better accuracy value than other existing fuzzy time series. Irawanto et al. [20] used frequency density-based partitioning for stock index forecasting. Wulandari et al. [21] used frequency density partitioning for forecasting the production of petroleum which resulted in a small error value.

Based on the description above, the researcher is interested in examining the average-based fuzzy time series Markov chain based on frequency density partitioning (FDP). Because based on the introduction above that the researchers have read, there has been no research that has conducted redivided intervals based on frequency density in the average-based fuzzy time series Markov chain method, and because frequency density partitioning produces subintervals, which is based on empirical analysis, these subintervals cause the fuzzy numbers to get closer to the crisp value. This idea is explained in the following chart in Figure 1.

This method is applied to forecast the Indonesian Islamic stock index (ISSI). Furthermore, to see the level of accuracy of the method, the mean square error (MSE) and the mean average percentage error (MAPE) are used. Then, for comparison, the researcher used Chen's fuzzy time series.

2. Basic Knowledge

2.1. Fuzzy Time Series. The definition of fuzzy time series was first introduced by Song and Chissom [16]. Let U be the universe of discourse, with $U = \{u_1, u_2, \dots, u_n\}$ on a fuzzy set A_i defined as

$$A_{i} = \frac{f_{A}(u_{1})}{u_{1}} + \frac{f_{A}(u_{2})}{u_{2}} + \dots + \frac{f_{A}(u_{n})}{u_{n}},$$
(1)

where f_A is the membership of the fuzzy set A_i , u_k is an element of the fuzzy set A_i , and $f_A(u_k)$ shows the degree of membership of u_k in A_i , where $k = 1, 2, 3, \dots, n$.

Definition 1. We assume $Y(t)(t = \dots, 0, 1, 2, \dots)$ subset of U, which becomes the universe discourse, where the fuzzy set A_i has been defined previously. Then, if F(t) is a collection of A_i $(i = 1, 2, \dots, n)$, then F(t) is called fuzzy time series on $Y(t)(t = \dots, 0, 1, 2, \dots)$ [5].

Definition 2. If F(t) is caused by F(t-1), then the relation in the first order rmodel F(t) can be stated as follows [3]:

$$F(t) = F(t-1)OR(t, t-1),$$
(2)

where "O" is the Max–Min composition operator and R(t, t-1) is a relation matrix to describe the fuzzy relationship between F(t-1) and F(t).

2.2. Average-Based Algorithm. Average-based algorithm is an algorithm that can be used to set the interval length that is determined at the initial stage of forecasting when using fuzzy time series. The steps of the average-based algorithm are as follows [17]:

(a) Determine the absolute difference (lag) between data *n* + 1 and data *n* with the formula:



FIGURE 1: Flow chart idea of average-based fuzzy time series Markov chain based on modified frequency density partitioning.

$$\log = |\text{Data } n + 1 - \text{Data } n|. \tag{3}$$

(b) Determine the length of the interval

Length of interval =
$$\left(\frac{\text{total lag}}{\text{numbers of data}}\right)$$
: 2. (4)

- (c) Based on the interval length obtained from step (b), determine the basis value of the interval length according to Table 1
- (d) The length of the interval is then rounded up according to the interval basis table

2.3. Frequency Density Partition. Chen and Hsu [5] developed a fuzzy time series by redivided intervals based on frequency density. In his research, after partitioning the universe discourse into *n* intervals of equal length, subpartition the intervals of the same length based on frequency density. By rule,

- (a) the interval with the first densest frequency is divided into 4 subintervals
- (b) the interval with the second densest frequency is divided into 3 subintervals
- (c) the interval with the third densest frequency is divided into 2 subintervals

2.4. Markov Chain. Markov first introduced the Markov chain in 1906. The Markov chain analysis is a method that studies the properties of the past to estimate the properties

of these variables in the future. Conceptually, the Markov chain can be described by assuming $\{X_n, n = 0, 1, 2, \dots\}$ as a finite stochastic process or the probability value can be calculated. The set of probability values of the stochastic process is denoted by the set of positive integers $\{0, 1, 2, \dots\}$.

If $X_n = i$, then this process occurs in *i* when *n*, assuming that whenever this process occurs in state *i*, on a point of probability P_{ij} who will move to the state *j*. Thus, it can be written as follows:

$$P\{X_{n+1} = j | X_{n-1} = i_{n-1}, \cdots, X_1 = I_1, X_0 = i_0\} = P_{ij}.$$
 (5)

For all states $i_0, i_1, \dots, i_{n-1}, i, j, n \ge 0$. This process is called the Markov chain.

The above equation is interpreted in the Markov chain as a conditional distribution of the future state X_{n+1} obtained from the previous state X_0, X_1, \dots, X_{n-1} and the current state X_n and does not depend on the previous state but depends on current state.

The value of P_{ij} represents the probability of the transition process from *i* to *j*. Because the probability value is always positive and the transition process moves, then P_{ij} ≥ 0 and $i, j \geq 0$, sum $P_{ij} = 1, j = 1, \dots, \infty$, and $i = 0, 1, \dots$ Let *P* be the transition probability matrix P_{ij} ; then, it can be denoted in the following equation [10]:

$$P = \begin{bmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ P_{20} & P_{21} & P_{22} & \cdots \\ \vdots & & \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}.$$
 (6)

TABLE 1: Basis mapping table.

Range	Basis
0,1 - 1,0	0,1
1,1 – 10	1
11 - 100	10
101 – 1000	100

The Markov chain process in the fuzzy time series used is a transition probability matrix. The transition probability matrix is used as the basis for forecasting calculations. The probability from the current state to the next state is obtained from the FLRG. State transition probabilities are written as follows [10]:

$$P_{ij} = \frac{M_{ij}}{M_i}, \, i, j = 1, 2, \cdots, n,$$
(7)

where P_{ij} is the transition probability of state A_i to state A_j one step, M_{ij} is the number of transitions from state A_i to state A_j one step, and M_i is the amount of data included in the state A_i .

The probability matrix *P* of all states is dimension $n \times n$, with *n* being the number of fuzzy sets, and can be written as follows [10]:

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix}.$$
 (8)

2.5. Average-Based Fuzzy Time Series Markov Chain Based on Frequency Density Partitioning. Average-based fuzzy time series Markov chain based on frequency density partition is a forecasting method using fuzzy time series combined with the Markov chain and using average-based as interval determination which is then redivided interval based on frequency density.

2.6. Forecasting Error Measurement. The reliability of a forecast can be determined by looking at the mean square error (MSE) and mean average percentage error (MAPE); these are the MSE and MAPE formulas [1]:

MSE =
$$\frac{\sum_{t=1}^{n} (Y_t - F'_t)^2}{n}$$
, (9)

MAPE =
$$\frac{1}{n} \sum_{t=1}^{n} \frac{|Y(t) - F'(t)|}{Y(t)} \times 100\%$$
, (10)

where Y_t is the actual data period t, F'_t is the t period forecasting value, and n is the predictable amount of data.

3. Materials and Methods

3.1. Data Collection. The data used in the study is the weekly data on the Indonesian Sharia Stock Price Index (ISSI) for the period June 2019–May 2021, which was obtained from the http://yahoo.finance.com site. The results of the forecast-ing test are then validated using the MSE and MAPE values. Furthermore, it is compared with Chen's fuzzy time series.

3.2. Forecasting Method. This study uses an average-based fuzzy time series Markov chain based on frequency density partitioning. The difference between this study and the previous research lies in the use of density partitioning frequencies in the average fuzzy-based Markov time series chain which will make the fuzzy values closer to the crisp values, so that the forecasting values have a good accuracy value. The steps are as follows [5, 10, 17, 22]:

- (a) Define the universe of discourse
- (b) Divide the universe of discourse into intervals using the average-based method
- (c) Distribute all research data into intervals
- (d) Determine the frequency density
- (e) Perform redivided intervals based on frequency density
- (f) Define a fuzzy set. Let A_1, A_2, \dots, A_n be a fuzzy set that has a linguistic value from a linguistic variable; the definition of fuzzy set A_1, A_2, \dots, A_n in the universe of discourse U is as follows:

$$A_{1} = \frac{1}{u_{1}} + \frac{0.5}{u_{2}} + \frac{0}{u_{3}} + \frac{0}{u_{4}} + \dots + \frac{0}{u_{n}},$$

$$A_{2} = \frac{0.5}{u_{1}} + \frac{1}{u_{2}} + \frac{0}{u_{3}} + \frac{0}{u_{4}} + \dots + \frac{0}{u_{n}},$$

$$\vdots$$

$$A_{n} = \frac{0}{u_{1}} + \frac{0}{u_{2}} + \frac{0}{u_{3}} + \dots + \frac{0.5}{n-1} + \frac{1}{u_{n}},$$
(11)

where $u_i(i = 1, 2, \dots, n)$ is an element of the universe of discourse (U)

- (g) Fuzzification of historical data
- (h) Define FLR
- (i) Define FLRG
- (j) Determine the transition probability with the formula:

$$P_{ij} = \frac{M_{ij}}{M_i}, i, j = 1, 2, \cdots, n,$$
 (12)

and determine the probability matrix of all states of the FLRG with dimensions of $n \times n$, where n is the number of fuzzy sets that can be written as follows:

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix}.$$
 (13)

(k) Calculate the initial forecast value using a probability matrix with the following rules:

Rule 1: if the FLRG A_i is one to one (e.g., $A_i \longrightarrow A_k$, where $P_{ik} = 1$ and $P_{ij} = 0, j \neq k$), the forecast F(t) is m_k which is the middle value of u_k with the following equation:

$$F(t) = m_k P_{ik} = m_k. \tag{14}$$

Rule 2: if the FLRG A_i is one too many (e.g., $A_j \rightarrow A_1$, A_2, \dots, A_n . $j = 1, 2, \dots, n$), when Y(t-1) at time (t-1) is included in state A_j , then the forecasting F(t) is

$$\begin{split} F(t) &= m_1 P_{j1} + m_2 P_{j2} + \dots + m_{j-1} P_{j(j-1)} \\ &+ Y(t-1) P_{jj} + m_{j+1} P_{j(j+1)} + \dots + m_n P_{jn}, \end{split} \tag{15}$$

where $m_1, m_2, \dots, m_{j-1}, m_{j+1}, \dots, m_n$ is the middle value $u_1, u_2, \dots, u_{j-1}, u_{j+1}, \dots, u_n$ and Y(t-1) are state values A_j at time t-1

Rule 3: if the FLRG A_i is empty $(A_i \rightarrow \emptyset)$, forecast value F(t) is m_i which is the middle value of u_i with the following equation:

$$F(t) = m_i. \tag{16}$$

- (l) Adjusting the trend of forecasting values with the following rules:
- (i) If state A_i communicates with A_i, starting from state A_i at time t − 1 expressed as F(t − 1) = A_i and undergoing an increasing transition to state A_j at the time t where (i < j), then the adjustment value is

$$D_{t1} = \left(\frac{l}{2}\right),\tag{17}$$

where *l* is the basis interval

(ii) If state A_i communicates with A_i , starting from state A_i at the time t - 1 expressed as $F(t - 1) = A_i$ and experiencing a decreasing transition to state A_j at the time *t*, where (i > j), the adjustment value is

$$D_{t1} = -\left(\frac{l}{2}\right). \tag{18}$$

(iii) If state A_i at the time t - 1 is expressed as $F(t - 1) = A_i$ and undergoes a jump forward transition to state A_{i+s} at the time t, where $(1 \le s \le n - i)$, then the adjustment value is

$$D_{t2} = \left(\frac{l}{2}\right)s,\tag{19}$$

where *s* is the number of forwarding jumps

(iv) If state A_i at the time t - 1 is as $F(t - 1) = A_i$ and undergoes a jump-backward transition to state $A_{i-\nu}$ at the time t, where $(1 \le \nu \le i)$, then the adjustment value is

$$D_{t2} = -\left(\frac{l}{2}\right)v,\tag{20}$$

where v is the number of jumps backward

(m) Determine the final forecast value based on the adjustment of the trend of the forecasting value.

If FLRG A_i is one to many and state A_{i+1} can be accessed from state A_i , where state A_i is related to A_i , then the forecasting result becomes $\dot{}(t) = F(t) + D_{t1} + D_{t2} = F(t) + (l/2) + (l/2)$. If FLRG A_i is one to many and state A_{i+1} can be accessed from A_i , where state A_i is not related to A_i , then the forecasting values become $F'(t) = F(t) + D_{t2} = F(t) + (l/2)$. If FLRG A_i is one to many and state A_{i-2} can be accessed from state A_i , where A_i is not related to A_i , then the forecasting result is $F'(t) = F(t) - D_{t2} = F(t) - (l/2) \times 2 = F(t) - l$. If vis a jump step, the general form of the forecast is

$$F'(t) = F(t) \pm D_{t1} \pm D_{t2} = F(t) \pm \left(\frac{l}{2}\right) \pm \left(\frac{l}{2}\right) \nu.$$
 (21)

4. Results and Discussion

Average-based fuzzy time series Markov chain based on frequency density partitioning was tested on the Indonesian Sharia Stock Price Index (ISSI) forecasting to see whether this method could optimize the interval on the FLR to produce more accurate forecasts.

In forecasting using an average-based fuzzy time series Markov chain based on frequency density partitioning, the first step is to collect ISSI historical data obtained as many as 104 data. The data is used to determine the universe of discourse. Furthermore, dividing into several subsets using an average based is as follows:

- (a) Determine the smallest values (D_{\min}) and greatest value (D_{\max}) obtained $D_{\min} = 123.78$ and $D_{\max} = 192.86$ and the value of $D_1 = 0.78$ and $D_2 = 0.14$ so that it can be defined U = [123.78 0.78, 192.86 + 0.14] = [123, 193]
- (b) Calculate the absolute difference between the data D_n and D_{n+1} using equations $\log D_n = |D_{n+1} D_n|$; as an example, data to 1 is as follows:

$$\begin{split} & \log D_1 = |D_{n+1} - D_n|, \\ & \log D_1 = |183,8 - 182,76|, \\ & \log D_1 = |1,04|, \\ & \log D_1 = 1,04. \end{split}$$

The calculation above is also used for the second data and so on; the total difference from the data is 304.62. Furthermore, the difference in the data is calculated on average using the equation $\sum |D_{n+1} - D_n|/\text{lots of data} = 304.62/103$, so the obtained average absolute difference is 2.96.

- (c) The result is then divided by 2 to get 1.48
- (d) The value of 1.48 is then determined using Table 1, and the basis for the length of the interval is 1
- (e) U can be a partition into the same interval length, namely, u₁, u₂, u₃, u₄, u₅, ..., u₆₆, u₆₇, u₆₈, u₆₉, u₇₀; successively, the value for each interval is

$$u_{1} = [123, 124] \qquad \vdots \\ u_{2} = [124, 125] \qquad u_{66} = [188, 189] \\ u_{3} = [125, 126] \qquad u_{67} = [189, 190] \\ u_{4} = [126, 127] \qquad u_{68} = [190, 191] \\ u_{5} = [127, 128] \qquad u_{69} = [191, 192] \\ \vdots \qquad u_{70} = [192, 193] \end{cases}$$
(23)

The next step is to distribute the data to each interval and determine the frequency density; we get u_{64} as the first densest frequency with a frequency of 7, which is further divided into 4 subintervals ($u_{64,1} = [186, 186.25]$, $u_{64,2} = [$ 186.25, 186.5], $u_{64,3} = [186.5, 186.75]$, and $u_{64,4} = [186.75,$ 187]), the second most populous is u_{65} divided into 3 subintervals ($u_{65,1} = [187, 187.33]$, $u_{65,2} = [187.33, 187.67]$, and $u_{65,3} = [187.67, 188]$), the third most populous is u_{26} and u_{60} divided into 2 subintervals ($u_{26,1} = [148, 148.5]$ and $u_{26,2} = [148.5, 149]$ and $u_{60,1} = [182, 182.5]$ and $u_{60,2} = [182.5,$ 183], respectively) and eliminates intervals that have no frequency, i.e., u_2 , u_3 , u_4 , u_5 , u_6 , u_7 , u_8 , u_9 , u_{11} , u_{12} , u_{13} , u_{17} , u_{19} , u_{25} , u_{32} , u_{33} , u_{35} , u_{39} , u_{41} , u_{42} , u_{43} , u_{44} , u_{45} , and u_{48} . Next, look for the middle value (m_i) of each interval; we get table middle value at Table 2. Table 2 shows the mean value of each interval that has been repartitioned based on frequency density. This middle value is used to calculate the initial forecast value on the fuzzy time series. The following is an example of calculating the mean value $u_1(m_1)$:

$$m_{1} = \frac{(\text{lower limit } u_{1} + \text{upper limit } u_{1})}{2},$$

$$m_{1} = \frac{(123 + 124)}{2},$$

$$m_{1} = \frac{247}{2},$$

$$m_{1} = 123.5.$$
(24)

Furthermore, defining fuzzy sets, fuzzy sets that can be formed from the universe of conversation are 53 fuzzy sets. Based on equation (8), the fuzzy set formed is as follows:

$$\begin{split} A_{1} &= \left\{ \frac{1}{u_{1}} + \frac{0.5}{u_{10}} + \frac{0}{u_{14}} + \frac{0}{u_{15}} + \dots + \frac{0}{u_{68}} + \frac{0}{u_{69}} + \frac{0}{u_{70}} \right\}, \\ A_{2} &= \left\{ \frac{0.5}{u_{1}} + \frac{1}{u_{10}} + \frac{0.5}{u_{14}} + \frac{0}{u_{15}} + \dots + \frac{0}{u_{68}} + \frac{0}{u_{69}} + \frac{0}{u_{70}} \right\}, \\ A_{3} &= \left\{ \frac{0}{u_{1}} + \frac{0.5}{u_{10}} + \frac{1}{u_{14}} + \frac{0.5}{u_{15}} + \dots + \frac{0}{u_{68}} + \frac{0}{u_{69}} + \frac{0}{u_{70}} \right\}, \\ &\vdots \\ A_{56} &= \left\{ \frac{0}{u_{1}} + \frac{0}{u_{10}} + \frac{0}{u_{14}} + \frac{0}{u_{15}} + \dots + \frac{1}{u_{68}} + \frac{0.5}{u_{69}} + \frac{0}{u_{70}} \right\}, \\ A_{57} &= \left\{ \frac{0}{u_{1}} + \frac{0}{u_{10}} + \frac{0}{u_{14}} + \frac{0}{u_{15}} + \dots + \frac{0.5}{u_{68}} + \frac{1}{u_{69}} + \frac{0.5}{u_{70}} \right\}, \\ A_{53} &= \left\{ \frac{0}{u_{1}} + \frac{0}{u_{10}} + \frac{0}{u_{14}} + \frac{0}{u_{15}} + \dots + \frac{0}{u_{68}} + \frac{0.5}{u_{69}} + \frac{1}{u_{70}} \right\}. \end{split}$$

$$(25)$$

The next step is to perform fuzzification; the data from the fuzzification results are presented in Table 3.

Table 3 shows the results of ISSI weekly data fuzzification; fuzzification is performed to convert firm values into fuzzy values. An example of the fuzzification process for data on June 9, 2019 (t = 1), is 182.76 entered in the interval $u_{60,2} = [182.5, 183]$. Next, the formed fuzzy set $u_{60,2}$ has a membership degree of 1 when it is in the fuzzy set A_{38} , so that for the 9 June 2019 data, the fuzzified data obtained is A_{38} .

After fuzzification is obtained, the next step is to determine FLR and FLRG, presented in Tables 4 and 5.

Table 4 shows that data 1 is fuzzified at A_{38} and the second data is fuzzified at A_{39} so that the FLR is $A_{38} \longrightarrow A_{39}$. FLR plays an important role because it is FLR that is used to determine forecasting values.

TABLE 2: Middle value.

<i>u</i> _i	m_i	u_i	m _i
<i>u</i> ₁	123,5	÷	÷
u_{10}	132,5	u_{67}	189,5
u_{14}	136,5	u_{68}	190,5
<i>u</i> ¹⁵	137,5	u_{69}	191,5
<i>u</i> ₁₆	138,5	u_{70}	192,5

TABLE 3: Fuzzification results.

t	Actual data	Fuzzy data	Т	Actual data	Fuzzy data
1	182,76	A ₃₈	:	:	
2	183,8	A_{39}	101	175,12	A_{30}
3	186,01	A_{42}	102	170,19	A_{25}
4	186,66	A_{44}	103	171,94	A_{26}
5	186,18	A_{42}	104	176,52	A_{31}

TABLE 4: FLR.

Data order	FLR	Data order	FLR
1-2	$A_{38} \longrightarrow A_{39}$	÷	:
2-3	$A_{39} \longrightarrow A_{42}$	101-102	$A_{30} \longrightarrow A_{25}$
3-4	$A_{42} \longrightarrow A_{44}$	102-103	$A_{25} \longrightarrow A_{26}$
4-5	$A_{44} \longrightarrow A_{42}$	103-104	$A_{26} \longrightarrow A_{31}$

TABLE 5: FLRG.

Current state	Next state	Current state	Next state
A_1	$(1)A_2$:	÷
A_2	$(1)A_{3}$	A_{51}	$(1)A_{49},(1)A_{50}$
A_3	$(1)A_4,(1)A_9$	A_{52}	$(1)A_{51}$
A_4	$(1)A_{5}$	A_{53}	$(1)A_{52},(1)A_{53}$

Based on Table 5, all FLR formed in Table 4 are grouped into interconnected FLRG, for example, FLRG on A_3 , where A_3 is the current state and has a relationship to $A_3 \longrightarrow A_4$ and $A_3 \longrightarrow A_9$. These 2 FLR are grouped into 1 FLRG, namely, $A_3 \longrightarrow A_4$, A_9 .

The next step is to calculate the initial forecast; FLRG is used to form a Markov chain and transition probability matrix. In this study, a transition probability matrix of the order 53 × 53 was formed, where each element is a probability value obtained from equation (7), for example, the calculation of the elements of the probability matrix for $A_i \rightarrow A_j$, with i = 11 and $j = 1, 2, 3, \dots, n$, because state A_i transitions to another state 3 times, namely, 2 times to state A_{11} and 1 time to state A_{13} ; then, $P_{11,11} = 2/3 = 0.667$ and $P_{11,13} = 1/3$ = 0.333. The same method is used to determine each ele-

TABLE 6: Initial forecasting results (F(t)).

Period	Actual data	F(t)	Period	Actual data	F(t)
09/6/19	182,76	Na	÷	÷	:
16/6/19	183,8	180,17	23/5/21	171,94	171,5
23/6/19	186,01	183,22	30/5/21	176,52	176,5

TABLE 7: Final forecast value.

Periode	Y(t)	F'_t	Periode	Y(t)	F'_t
09/6/19	182,76	NA	:	:	:
16/6/19	183,8	180,42	02/5/21	174,82	178,89
23/6/19	186,01	184,72	09/5/21	175,12	179
30/6/19	186,66	188,31	16/5/21	170,19	168
07/7/19	186,18	186,53	23/5/21	171,94	172
14/7/19	189,01	189,06	30/5/21	176,52	179

ment on the probability matrix so that the matrix can be as follows:

	0	1	0	0		0	0	0	0]	
	0	0	1	0		0	0	0	0	
	0	0	0	0.5		0	0	0	0	
	0	0	0	0		0	0	0	0	
<i>P</i> =	:	÷	:	÷	÷	÷	÷	÷	:	
	0	0	0	0		0	0.25	0	0	
	0	0	0	0		0.5	0	0	0	
	0	0	0	0		0	1	0	0	
	0	0	0	0		0	0	0.5	0.5	
										(26)

After the probability matrix element values are obtained, the next step is to calculate the initial starting value using formulas (14), (15), and (16). For example, for June 16, 2019 (t = 2), the data seen is the previous week's data, namely, June 9, 2019 (t = 1), where the state transitions from $A_{38} \rightarrow A_{39}$; then, the forecast calculation $F(t) = m_{28}P_{38,28} + Y_1P_{38,38} + m_{39}P_{38,39} = 173.5(0.33) + 182.76(0) + 183.5(0.67) = 180.17$; the same steps are also used for $t = 3, 4, 5, \cdots 104$. The summary of the initial forecasting results is shown in Table 6.

Table 6 shows the initial forecasting values for the period 16 June 2019 to 30 May 2021. These initial forecasting values were obtained from the defuzzification results of the FLRG group.

The next step is to adjust the forecasting trend. For example, the adjustment value for June 16, 2019, the next state is A_{39} , and the current state is A_{38} ; then, the adjustment calculation uses the forecast adjustment rule point c with equation (19) $D_{t2} = (l/2)s = (0.5/2)1 = 0.25$. For the



FIGURE 2: Graph of comparison of actual data with Chen and average-based fuzzy time series Markov chain based on FDP.

calculation of other forecasting value adjustments, use equations (17), (18), (19), and (20).

After the adjustment value is obtained, then calculate the final forecast value. For the calculation of the adjusted forecast value, follow the existing rules in equation (21). For example, calculations for adjusted forecast values $F'_2 = F_2 \pm D_{t2} = 180.17 + 0.25 = 180.42$; by doing the same way, the summary of the final forecasting results is shown in Table 7.

Table 7 shows the results of the final forecast that has made some adjustments. This final forecasting value is obtained from the sum of the initial forecasting value with the adjustment value. The initial forecasting value that has made several adjustments produces a final forecasting value that is closer to the actual data. The comparison of the final forecasting value with actual data can be seen in Figure 2.

Figure 2 shows a comparison of forecasting results using an average-based Markov Chain based on frequency density partitioning. The graph in blue shows the actual data, and the graph in orange shows the results of forecasting using an average-based fuzzy time series Markov chain based on frequency density partitions. Average-based fuzzy time series Markov chain based on frequency density partitioning shows a pattern that is almost the same as the actual data, although the resulting forecasting value is not the same as the actual data; the pattern of forecasting values uses the average-based fuzzy time series Markov chain based on density partitioning the frequency that follows the pattern of the actual data.

The last step is to calculate the forecast accuracy value using MSE and MAPE. For the MSE and MAPE values, this calculation uses formulas (9) and (10), respectively; the average-based fuzzy time series method based on FDP is 5.76 and 1.04%. This shows good forecasting performance on the average-based fuzzy time series Markov chain based on frequency density partitioning as shown in Figure 2, where the forecasting value is closer to the actual value. The good accuracy value on the fuzzy time series Markov chain is obtained due to the use of frequency density partitions which produce subintervals so that the fuzzy values can be close to the crisp values. This is also supported by the use of average-based method to determine the optimal interval length.

5. Conclusions

Forecasting using the average-based fuzzy time series Markov chain based on the frequency density partition (FDP) has a good accuracy value; this can be seen from the MSE and MAPE values of 5.76 and 1.04%, respectively. This is a good accuracy value because the Markov chain fuzzy time series uses an average-based method to determine the length of the interval so that the length of the interval used is not just the perception of the researcher. The length of this interval is then partitioned based on frequency density to obtain a more optimal interval length. Determination of the length of the interval on the FTS-Markov chain plays an important role in forming a fuzzy logical relationship (FLR), and this FLR is used to determine the forecast value.

Data Availability

The table data used to support the findings of this study are included within the supplementary information file and also within the article and are also available online at https://g.co/finance/ISSI:IDX?window=5Y.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This research was funded in addition to the state budget (Selain APBN) of Diponegoro University in 2022 (grant number 215/UN7.A/HK/VII/2022 and SPK Number 569-113/UN7.D2/PP/VII/2022). We would like to thank the chancellor of Diponegoro University and all parties involved in this paper.

References

- [1] G. Kitagawa, *Introduction to Time Series Modeling*, Chapman & Hall/CRC, Boca Raton, 2010.
- [2] S. Makridakis and S. C. Wheelwright, "Metode dan Aplikasi Peramalan," in *Terjemahan Hari Suminto*, Binarupa Aksara, 2nd edition, 1999.
- [3] S. Chen, "Forecasting enrollments based on fuzzy time series," *Fuzzy Sets and Systems*, vol. 81, no. 3, pp. 311–319, 1996.
- [4] S. Chen, "Temperature prediction using fuzzy time series," IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics), vol. 30, no. 2, pp. 263–275, 2000.
- [5] S. Chen and C. Hsu, "A new method to forecasting enrollments using fuzzy time series," *International Journal of Applied Science and Engineering*, vol. 2, no. 3, pp. 234–244, 2004.
- [6] H.-K. Yu, "Weighted fuzzy time series models for Taiex forecasting," *Physica A: Statistical Mechanics and its Applications*, vol. 349, no. 3-4, pp. 609–624, 2005.
- [7] K. Huanrg, "The application of neural networks to forecast fuzzy time series," *Physica A: Statistical Mechanics and its Applications*, vol. 363, no. 2, pp. 481–491, 2006.
- [8] C. H. Cheng, G. W. Cheng, and J. W. Wang, "Multi-attribute fuzzy time series method based on fuzzy clustering," *Physica A: Statistical Mechanics and its Applications*, vol. 34, no. 2, pp. 1235–1242, 2008.
- [9] M. Stevenson and J. E. Porter, "Fuzzy time series forecasting using percentage change as the universe of discourse," *World Academy of Science, Engineering and Technology*, vol. 55, pp. 154–157, 2009.
- [10] R. C. Tsaur, "A fuzzy time series- Markov chain model with an application to forecast the exchange rate between the Taiwan and US dollar," *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 7B, pp. 4931–4942, 2012.
- [11] R. Indriyani and R. H. Pratiwi, "Penerapan Model epidemi discrate Time Markov chian Susceptible Vaccinated Infected Recovered (studi kasus pada polapenyebaran penyakit campak)," *ProSANDIKA*, vol. 3, no. 2, 2022.
- [12] A. W. Prasetya and R. Ferdian, "Penerapan Metode Markov Chain dalam Penjadwalan Perawatan Mesin Oerlicon untuk mengoptimalkan Biaya dan Waktu Perawatan di PT Dirgantara Indoneisia," *Sytax Literate; Jurnal Ilmiah Indonesia*, vol. 7, no. 2, 2022.
- [13] R. Dinatha, A. Uperiati, and D. A. Purnamasari, "Prediksi Keuntungan Ekspor dengan Metode Fuzzy Time Series Model Markov Chain," *Student Online Journal (SOJ) UMRAH-Teknik*, vol. 2, no. 2, pp. 551–564, 2022.
- [14] D. A. Hidayah and S. Sugiman, "Peramalan Nilai Tukar Rupiah terhadap Dollar Amerika dengan Metode Fuzzy Time Series Markov Chain," *Unnes Journal of Mathematics*, vol. 10, no. 2, pp. 85–96, 2021.

- [15] P. Mangkunegara and Y. Yerizon, "Fuzzy Time Series Markov Chain dalam Meramalkan Nilai Tukar Mata Uang Antara Ringgit Malaysia dengan Rupiah," *Journal of Mathematics UNP*, vol. 5, no. 2, 2020.
- [16] 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.
- [17] S. Xihao and L. Yimin, "Average-based fuzzy time series models for forecasting Shanghai compound index," *World Journal of Modelling and Simulation*, vol. 4, no. 2, pp. 104– 111, 2008.
- [18] E. D. Wuryanto and N. V. I. Puspita, "Model Average Based-Fuzzy Time Series Markov Chain untuk Prediksi Perkembangan Terkonfirmasi Positif Covid-19," *Jurnal Informatika Upgris*, vol. 7, no. 2, 2021.
- [19] N. Kumar and H. Kumar, "A novel hybrid fuzzy time series model for prediction of Covid-19 infected cases and deaths in India," *ISA Transactions*, vol. 124, pp. 69–81, 2022.
- [20] B. Irawanto, R. W. Ningrum, R. Wulandari, B. Surarso, and Farikhin, "Frequency density-based partitioning (FDP) for forecasting IHSG," *Journal of Physics: Conference Series*, vol. 1321, no. 2, article 22081, 2019.
- [21] R. Wulandari, B. Surarso, and B. Irawanto, "First-order fuzzy time series based on frequency density partitioning for forecasting production of petroleum," *IOP Conference Series: Materials Science and Engineering*, vol. 846, no. 1, article 12063, 2020.
- [22] S. Hariyanto, Z. Zaenurrohman, and T. Udjiani, "Average based-FTS Markov chain based on a modified frequency density partitioning to predict Covid-19 in central Java," *Cauchy: Jurnal Matematika Murni dan Aplikasi*, vol. 7, no. 2, pp. 231–239, 2022.