BAREKENG: Journal of Mathematics and Its Applications
June 2025 Volume 19 Issue 2 Page 1365-1378
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss2pp1365-1378

STOCK PRICE FORECASTING USING FUZZY C-MEANS AND TYPE-2 FUZZY TIME SERIES

Rineka Brylian Akbar Satriani[1], Farikhin[2*], Bayu Surarso[3]

[1,2,3]*Mathematics Department, Faculty of Science and Mathematics, Universitas Diponegoro*
Jln. Prof. Soedarto, SH, Tembalang, Semarang, Jawa Tengah, 50725, Indonesia

*Corresponding author's e-mail: *farikhin@live.undip.ac.id*

Article History:

Received: 17th November 2024
Revised: 2nd February 2025
Accepted: 5th March 2025
Published: 1st April 2025

Keywords:

Forecast;
Fuzzy C-Means;
Type-2 Fuzzy Time Series.
.

ABSTRACT

Stock prices have unstable movements, so forecasting is needed to decide to invest appropriately according to the strategy. Fuzzy Time Series (FTS) uses fuzzy sets to forecast future time series values using historical data. However, interval partitioning in FTS needs to be considered as it can affect the forecasting results. FCM is applied to solve the problem of interval assignment in the universe of discourse. It allows the evaluation of the distribution of historical data and forming intervals of different sizes. Type 2 Fuzzy Time Series (T2FTS) is an extension of FTS to improve forecasting performance and refine fuzzy relationships. This research aims to improve forecasting accuracy using the Fuzzy C-Means (FCM)-T2FTS combination. This research uses daily data on BBRI stock prices from January 2023 to May 2024, with the variables used being close, high, and low prices. The results showed that determining the interval length using unequal length is more efficient than fixed interval length and can improve model performance, demonstrated from the MAPE values of T2FTS and FCM-T2FTS, which are 2.09% and 1.97%, respectively, the difference between the two MAPEs, is 0.12%. Hence, FCM-T2FTS is 12% more efficient than T2FTS. Therefore, FCM-T2FTS can improve forecasting accuracy.

How to cite this article:

R.B.A. Satriani, Farikhin and B. Surarso., "STOCK PRICE FORECASTING USING FUZZY C-MEANS AND TYPE-2 FUZZY TIME SERIES," *BAREKENG: J. Math. & App.*, vol. 19, iss. 2, pp. 1365-1378, June, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Stocks are company assets that have a price [1]. Stocks have uncertain movements, so one way to minimize the risk of investing in stocks is to look at time series data patterns of stock price fluctuations. Securities indicate that an individual or a legal entity owns the issuing company. Stocks are open for everyone to invest in the capital market. Stock prices include close, low, and high prices. Data on stock prices can be in the form of daily, monthly, and annual data. The platform used to accumulate money is the stock market. In the stock market, companies that issue shares earn profits that are used for business and shareholders benefit from the purchase of shares through the value of shares and dividends obtained from the company [2]. Stock prices have unstable movements and have an impact on the economy. Bank Rakyat Indonesia (BRI) is one of the major government banks in Indonesia. BRI stock-coded BBRI is included in stocks in great demand by investors because the stocks are included in the LQ45 index, which has a good reputation.

Fuzzy Time Series (FTS) uses fuzzy sets to forecast future time series values using their historical values. FTS is growing rapidly as a result of current scientific advances. The development can be felt in several fields, such as finance, science, and society. Research [3] and [4] explain the time series forecasting model on student enrollment. In the financial sector, the model can be used to forecast the price of a stock because the stock pattern is not fixed, so it is necessary to do forecasting to determine the right strategy for the future. Research conducted by [5] on the Taiwan Futures Exchange (TAIFEX) using a combined FTS provides good model results.

FTS has several models, such as the Chen, Cheng, and Markov Chain Models. The Chen model has the best results among other models in rainfall forecasting [6]. Chen's model is more efficient than Song and Chissom, using simpler arithmetic operations than max-min composition [7]. However, the Chen model still has shortcomings that need to be improved, such as the lack of partition and interval partition. The interval partition has an essential influence on forecasting because it can affect the accuracy of a forecast [8]. Therefore, FCM can be applied to improve forecasting performance, which helps overcome the interval length, and research shows that FCM provides good accuracy results [8].

FCM is a clustering technique that divides data into groups to assist in identifying trends and patterns in the data. It is also applied to solve the interval division problem in the universe of discourse. It allows for considering the distribution of historical data and the formation of different size intervals [8]. Real-time series are usually randomly distributed, which encourages researchers to find methods that can divide the universe of discourse into intervals that have unequal lengths. Determining the length of intervals with different sizes is used so that all data does not fall into the same interval, causing the concept of fuzzification to be reduced. Therefore, determining the number of intervals can help overcome long-term and short-term distributions in FCM [9].

Over time, FTS experienced many developments by combining it with other methods to improve forecasting accuracy, such as FCM [8], [10], [11], FCM and ANN [12], FCM and hesitant fuzzy set [13], firefly clustering [14], heuristic model [15], and others. In addition to integrating with other methods, FTS experienced developments such as Type 2 Fuzzy Time Series (T2FTS) [16]. Huarng is a development of the Chen model FTS to improve forecasting performance and refine the fuzzy relationship of FTS by forecasting the Taiwan Stock Index (TAIEX) and applying T2FTS; it is found that forecasting using T2FTS has better results than FTS [16]. T2FTS gives good results when compared to Type-1 FTS [17].

Based on this explanation, the novelty of this research is the combination of FCM in T2FTS. This research aims to improve BBRI stock price forecasting by modifying T2FTS and FCM, where FCM is used to strengthen the weakness of Chen's model, which is then applied to T2FTS to determine the interval. The FCM-T2FTS combination is compared with T2FTS to observe the changes due to modifications in determining the interval. Accuracy calculations are performed using MAPE (Mean Absolute Percentage Error) as a tool in financial analysis to choose the best model.

2. RESEARCH METHODS

2.1 Type-2 Fuzzy Time Series (T2FTS)

Based on [16], T2FTS is the extension of the FTS Chen model or can be referred to as Type 1 to improve forecasting performance and refine fuzzy relations. Generally, FTS only uses one variable in forecasting. Because using one variable means using only a small portion of the existing data, it misses important information that can help in improving accuracy. T2FTS adds more observational data to produce richer fuzzy relations and better accuracy than FTS [18]. T2FTS uses union and intersection operators in forecasting.

Definition 1. [16] *Operators defined to calculate the relationship between two FLRGs*

$$\vee(LHS_d, LHS_e) = RHS_d \cup RHS_e \quad (1)$$

$$\wedge(LHS_d, LHS_e) = RHS_d \cap RHS_e \quad (2)$$

Where operator $\cup$ is union dan $\cap$ is intersection. LHS_d and RHS_d are LHS and RHS from each FLRG.

Definition 2. [16] Union and intersection operators for $\vee_m$ dan $\wedge_m$ are defined by **Equation (3)** and **Equation (4)**:

$$\vee_m(LHS_c, LHS_d, LHS_e, \dots) = \vee \dots (\vee(\vee(LHS_c, LHS_d), LHS_e), \dots), \quad (3)$$

$$\wedge_m(LHS_c, LHS_d, LHS_e, \dots) = \wedge \dots (\wedge(\wedge(LHS_c, LHS_d), LHS_e), \dots) \quad (4)$$

where $LHS_c, LHS_d, LHS_e, \dots$ and $RHS_c, RHS_d, RHS_e, \dots$ represents LHS and RHS of FLRG $c, d, e, \dots$.

Definition 3. [16]

a. *If $\vee_m(LHS_c, LHS_d, LHS_e, \dots) = \emptyset$, then let $\vee_m(LHS_c, LHS_d, LHS_e, \dots) = LHS_x$.*
   where LHS_x is obtained from FLRG constructed from Type 1.
b. *If $\wedge_m(LHS_c, LHS_d, LHS_e, \dots) = \emptyset$, then let $\wedge_m(LHS_c, LHS_d, LHS_e, \dots) = LHS_x$.*
   where LHS_x is obtained from FLRG constructed from Type 1.

T2FTS algorithm based on [16], [19] can be written as follows:

a. Determine the universe of discourse.
b. Determine the interval.
c. Define the fuzzy set.
d. Fuzzification.
e. Determine FLR and FLRG.
f. Determine Type 2 observation.
g. Categorize the out-of-sample observations in the FLRG and obtain a forecast.
h. Apply the operators on FLRG to all observations.
i. Defuzzification.
j. Calculate the forecasts on T2FTS.

2.2 Fuzzy C-Means (FCM)

FCM is a clustering method that has a basic idea of the initial data in the form of input data $\boldsymbol{X} = \{\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, \dots, \boldsymbol{x_n}\}$. FCM divides the data into fuzzy sets by minimizing the sum of squared errors of each group presented in **Equation (5)** as follows:

$$J_m(X, V, U\_FCM) = \sum_{i=1}^{c} \sum_{k=1}^{n} (u_{ik})^m d^2(x_k, v_i) \quad (5)$$

where,

c = number of clusters in X; $2 \leq c \leq n$,

m = fuzziness parameter; $1 \leq m \leq \infty$,

U_FCM = fuzzy c-partition of X,

$d(x_k, v_i)$ = distance between object(x_k) and cluster center(v_i),

u_{ik} = membership value of the kth data to cluster-i.

Algorithm of FCM as shown in [20]:

a. Determine the number of clusters (c) with c $(2 \leq c \leq n)$, fuzziness parameter (m), maximum iteration, error (ϵ), initials of objective function$(P_0) = 0$, and iteration $(t = 1)$.

b. Initialize the membership matrix U_FCM by using a random value.

c. Determine the cluster center

$$v_i = \frac{\sum_{k=1}^{n} (u_{ik})^m x_k}{\sum_{k=1}^{n} (u_{ik})^m} \tag{6}$$

d. Calculate the objective function at the t-th iteration using Equation (5).

e. Determine the new membership matrix

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m-1}}} \tag{7}$$

f. If $\|u_{ik}(t+1) - u_{ik}(t)\| < \varepsilon$, then done, but if not then back to step 3.

b. ε as a boundary determinant, where if the value of $\|u_{ik}(t+1) - u_{ik}(t)\| < \varepsilon$ then the iteration can be stopped because it has reached convergence.

2.3 Algorithms

According to the description above, the calculation uses the T2FTS and FCM-T2FTS combination. Algorithms can be obtained, shown in the following Figure 1.

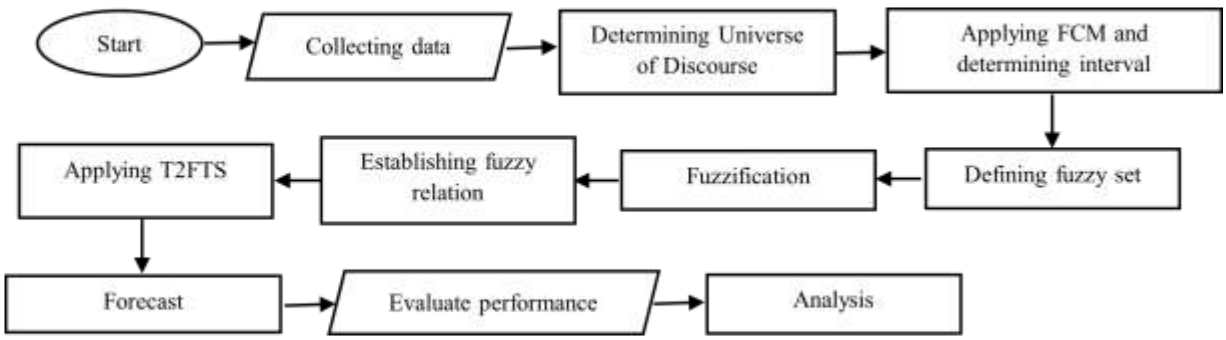

Figure 1. FCM-T2FTS Algorithms

Based on Figure 1, the first step that needs to be done is to collect BBRI stock price data through the Yahoo finance website. The stock price includes close, high, and low prices. Next, divide the data into training data and testing data. Determine the interval using the FCM and Chen models, then define the fuzzy sets and perform fuzzification. In the established fuzzy relation stage, FLR and FLRG are generated. Applying T2FTS, determining Type 2 observation data, namely testing data. Categorize observations into FLRG using the previously obtained testing data and apply FLRG operators, namely union and intersection operators. Perform defuzzification and get the forecast. T2FTS and the FCM-T2FTS combination have differences only in the formation of intervals.

Next, performance is evaluated using MAPE (Mean Absolute Percentage Error), which measures how accurate a forecasting model is by calculating the average percentage error between the actual value and the predicted value. MAPE is calculated using **Equation (8)**.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{X_t - \hat{X}_t}{X_t}\right| \times 100\% \tag{8}$$

MAPE has criteria and can be applied to evaluate the accuracy of a forecasting model **[21]**, criteria are shown in **Table 1**.

Table 1. MAPE Accuracy Value

MAPE	Criteria
<10%	Very Good
10% - 20%	Good
20% - 50%	Reasonable
>50%	Poor

2.4 Data Source

The data for this research are BBRI daily stock close, low, and high prices from 02 January 2023 until 31 May 2024 with a total of 330 data. Close is the final stock price, low is the lowest, and high is the highest on that day. Data is taken from the Yahoo finance website. In the process, the data is categorized into training and testing data. The training data used is close data from 02 January until 29 December 2023 with 239 data. The testing data used is close prices from 01 January until 31 May 2024 with a total of 91 data, high from 01 January until 31 May 2024 with 91 data, and low prices from 01 January until 31 May 2024 with 91 data. Data processed using Python programming language with Google Collaboratory.

3. RESULTS AND DISCUSSION

3.1 Collecting Data

The research begins with data collection, the basis for subsequent analysis. Yahoo! Finance, a website that provides comprehensive stock market information, is a reliable data source. The next stage categorizes the data into training data (Type 1 observation) and testing data (Type 2 observation). **Figure 2** shows the distribution of these two data sets.

Figure 2. BBRI Stock Price Data

Figure 2 shows the division and distribution of the two data sets, training data and testing data. The training data used on Type 1 includes daily close stock price from 02 January until 29 December 2023 with 239 data. The testing used for Type 2 observation and the data includes daily close stock price, high price, and low price from 02 January until 31 May 2024 with a total of 91 data. The blue line is the close price training data. The close price testing data is shown with the green line, and the red and yellow dotted lines show the high and low price testing data.

3.2 Determining Universe of Discourse (U)

The determination of the universe of discourse and intervals uses all close price data from 01 January 2023 to 31 May 2024. Obtained $D_{min} = 4380$, $D_{max} = 6400$, $D_1 = 84$ and $D_2 = 67$. Using **Equation (9).**

$$U = [D_{min} - D_1, D_{max} + D_2] \quad (9)$$

$$U = [4296\,, 6467]$$

Therefore, the universe of discourse is $[4296\,, 6467]$.

3.3 Applying Fuzzy C-Means and Determining Interval

Determine cluster number $(c) = 9$, $m = 2$, maximum iteration $= 1000$, error $(\epsilon) = 0.0001$, and initials of the objective function$(P_0) = 0$. Next, assuming that the total of all data is equal to 1, each membership value is initially set using a random number distributed between zero and one. The initialization results in iteration 1 are presented in **Table 2**.

Table 2. Membership Matrix Initialization Iteration 1

Cluster	**Close Price**				
	4870	**4850**	⋯	**4410**	**4380**
C1	0.1105	0.1106	⋯	0.1111	0.1111
C2	0.1235	0.1230	⋯	0.1722	0.1170
⋯	⋯	⋯	⋯	⋯	⋯
C8	0.1217	0.1212	⋯	0.1164	0.1162
C9	0.0968	0.0974	⋯	0.1040	0.1042
Total	**1**	**1**	⋯	**1**	**1**

The cluster center can be found using **Equation (10)** as follows,

$$v_i = \frac{\sum_{k=1}^{n} (u_{ik})^m x_k}{\sum_{k=1}^{n} (u_{ik})^m} \quad (10)$$

With v_i as the i-th cluster center, u_{ik} as membership value of the kth data to the i-th cluster, x_k which represents the kth data, n which is the total amount of data close price from 02 January 2023 until 31 May 2024, namely $n = 330$, and m is a fuzziness parameter that has a value of 2. Determining the cluster center using FCM and assisted with Python so that the cluster center in iteration 1 is presented in **Table 3.**

Table 3. Center Cluster Iteration 1

Cluster	**Cluster Center**
C_1	5284.201
C_2	5285.937
C_3	5289.142
⋯	⋯
C_8	5339.880
C_9	5370.292

Next, determine the new membership and count until the value is obtained $\|u_{ik}(t+1) - u_{ik}(t)\| < \varepsilon$, where $\|u_{ik}(t+1) - u_{ik}(t)\|$ convergence criteria to stop iteration in FCM and ε is the error threshold with $\varepsilon = 0.0001$. The calculation stops at the iteration 111 because the value has converged with $0.000089 < \varepsilon$. Thus, the cluster center obtained at the iteration 111 is presented in **Table 4**.

Table 4. Cluster Center Iteration 111

Cluster	Cluster Center
C_1	4545.958
C_2	4740.232
C_3	4863.945
…	…
C_8	5743.021
C_9	6166.130

Next, determine the interval after getting the cluster center, and the iteration stops at iteration 111. This step uses the midpoint of the cluster centers as the interval boundary to determine the interval [10]. For example, if cluster 1 is 4545.958 and cluster 2 is 4740.232, then the midpoint of the centers is 4659.573. Thus, the first interval is (4296, 4643.095), where 4296 is the lower bound of the universe of discourse obtained previously. This method helps to find the correct number of intervals according to the range and variation of the time series. By fixing the interval size using cluster centres, the data distribution in each interval becomes evener. The intervals obtained are shown in **Table 5**.

Table 5. Interval and Midpoint

u_i	Interval	Midpoint
u_1	(4296, 4643.095)	4469.5475
u_2	(4643.095, 4802.0885)	4722.5918
u_3	(4802.0885, 4991.5245)	4896.8065
...	…	…
u_8	(5663.8435, 5954.5755)	5809.2095
u_9	(5954.5755, 6467)	6210.7878

3.4 Defining Fuzzy Set

In this step, each A_i is defined over the interval $u_1, u_2, u_3, \dots, u_9$. So that the degree of membership can be obtained as follows:

$$A_1 = {}^{1}/_{u_1} + {}^{0,5}/_{u_2} + {}^{0}/_{u_3} + \cdots + {}^{0}/_{u_9}$$
$$A_2 = {}^{0,5}/_{u_1} + {}^{1}/_{u_2} + {}^{0,5}/_{u_3} + \cdots + {}^{0}/_{u_9}$$
$$A_3 = {}^{0}/_{u_1} + {}^{0,5}/_{u_2} + {}^{1}/_{u_3} + \cdots + {}^{0}/_{u_9}$$
$$\dots$$
$$A_9 = {}^{0}/_{u_1} + {}^{0}/_{u_2} + {}^{0}/_{u_3} + \cdots + {}^{1}/_{u_9}$$

3.5 Fuzzification

The next step is to fuzzify the data. Fuzzification is converting data into a fuzzy set based on the results of the previous steps. For example, using data training close prices taken, 02 January 2023 donated $x_1 = 4870$ in $u_3 = [4744.888, 4969.333]$, which means that the value of x_1 can be converted into a liguistic value, namely A_3 because the degree of membership in u_3 is 1. Applies to all data at this fuzzification stage. The fuzzification training data result is shown in **Table 6**, with t denoting the data sequence at time t.

Table 6. Fuzzification Data

Date	t	Close Price	Fuzzification
02/01/2023	1	4870	A_3
03/01/2023	2	4850	A_3
04/01/2023	3	4770	A_2
…	..	…	…
28/12/2023	238	5725	A_7
29/12/2023	239	5725	A_7

3.6 Establishing Fuzzy Relationships

In this step, the Fuzzy Logical Relationship (FLR) and Fuzzy Logical Relationship Group (FLRG) are formed. Based on the fuzzification result, FLR indicates a relationship between two consecutive fuzzy sets. In stock price analysis, FLR shows how the stock price in the current state affects the stock price in the next state. For example, the current state is 03 January 2023 with a price 4850 (A_3) and the data of the following state is 04 January 2023 with a price 4770 (A_2), so the FLR obtained is $A_3 \rightarrow A_2$. Apply to all the data, and the FLR results are shown in **Table 7.**

Table 7. Fuzzy Logical Relationship (FLR)

Date	t	Close Price	Fuzzification	FLR
02/01/2023	1	4870	A_3	-
03/01/2023	2	4850	A_3	$A_3 \rightarrow A_3$
04/01/2023	3	4770	A_2	$A_3 \rightarrow A_2$
…	..	…	…	
28/12/2023	238	5725	A_7	$A_6 \rightarrow A_7$
29/12/2023	239	5725	A_7	$A_7 \rightarrow A_7$

Based on the FLR, then formed into FLRG. FLRG is a grouping of FLR. For example, several data show the same form of change, so the data are grouped. There is data that each has an FLR $A_3 \rightarrow A_3$, $A_3 \rightarrow A_2$, and $A_3 \rightarrow A_4$ because they have the same relationship, so they are grouped together and the FLRG is $A_3 \rightarrow A_2, A_3, A_4$. The result of FLRG is shown in **Table 8**.

Table 8. Fuzzy Logical Relationship Group (FLRG)

FLRG	
A_1	$\rightarrow A_1, A_2$
A_2	$\rightarrow A_1, A_2, A_3$
A_3	$\rightarrow A_2, A_3, A_4$
…	…
A_7	$\rightarrow A_6, A_7, A_8$
A_8	$\rightarrow A_7, A_8$

For A_9, there is no FLRG because the daily close stock prices in the training data are not in the interval 9, so the daily stock prices in the testing data that are in the interval 9 are empty sets. Furthermore, it uses testing data as an additional variable that will be used in the T2FTS process. The testing data used close, high, and low daily prices from 02 January until 31 May 2024.

3.7 Applying T2FTS

3.7.1 Determining Type 2 Observation

At this stage, fuzzification is carried out for BBRI stock price testing data. The type 2 observations are close price (X), high price (Y), and low price (Z). Type 2 observations are shown in **Table 9**.

Table 9. Type 2 Observation Data Testing

Date	t	X_t	Y_t	Z_t
02/01/2024	240	5675	5675	5625
03/01/2024	241	5600	5650	5600
…	…	…	…	…
29/05/2024	329	4410	4500	4400
30/05/2024	330	4380	4480	4310

3.7.2 Categorize the Out-of-Sample Observations in the FLRG

Furthermore, fuzzification is carried out on each testing data, including close (X), high (Y), and low (Z) prices. After fuzzification, each data is categorized into the corresponding FLRG in **Table 8**. This categorization defines the pattern of relationships between data, a reliable method that assists in forecasting based on the previously formed relationship. The fuzzification results are shown in **Table 10**.

Table 10. Fuzzification T2FTS

Date	t	X_t	Fuzzification	Y_t	Fuzzification	Z_t	Fuzzification
02/01/2024	240	5675	A_8	5675	A_8	5625	A_7
03/01/2024	241	5600	A_7	5650	A_7	5600	A_7
04/01/2024	242	5700	A_8	5700	A_8	5575	A_8
…	…	…	…	…	…	…	…
29/05/2024	329	4410	A_1	4500	A_1	4400	A_1
30/05/2024	330	4380	A_1	4480	A_1	4310	A_1

3.7.3 Apply the operators on FLRG to all observations

After obtaining the results of fuzzification of each variable, forecasting is carried out by applying the union $\vee_m$ and intersection $\wedge_m$ operators. These operators are used to calculate forecasting using the fuzzification results of the previous step. So, for example, forecasting on 04 January 2024 (t = 242), then in calculating it uses the fuzzification result when 03 January 2024 (t = 241).

$X_{241}: A_7 \to A_6, A_7, A_8$; $Y_{241}: A_7 \to A_6, A_7, A_8$; $Z_{241}: A_7 \to A_6, A_7, A_8$

Operator $\vee_m$ (union multiple)

$$\vee_m(LHS_c, LHS_d, LHS_e, \dots) = (RHS_c \cup RHS_d \cup RHS_e \cup \dots)$$

Then for $t = 242$

$$\vee_m(A_7, A_7, A_7) = \{A_6, A_7, A_8\} \cup \{A_6, A_7, A_8\} \cup \{A_6, A_7, A_8\} = \{A_6, A_7, A_8\}$$

So, the forecasting value is in one of the two fuzzy sets, A_7, A_8 and A_9.

Operator $\wedge_m$ (intersection multiple)

$$\wedge_m(LHS_c, LHS_d, LHS_e, \dots) = (RHS_c \cap RHS_d \cap RHS_e \cap \dots)$$

Then for $t = 242$

$$\wedge_m(A_7, A_7, A_7) = \{A_6, A_7, A_8\} \cap \{A_6, A_7, A_8\} \cap \{A_6, A_7, A_8\} = \{A_6, A_7, A_8\}$$

So, the forecasting value is in one of the two fuzzy sets, A_6, A_7, and A_8

Then calculated for all data testing. Obtained the results of applying the operator as shown in **Table 9.**

3.7.4 Defuzzification

After getting the results in the calculation of the two operators, then defuzzification is carried out using **Equation (11)**.

$$Defuzzification_t = \frac{\sum_{z=1}^{j} m_{qz}}{j} \tag{11}$$

$defuzzification_k\ (t)$ is the defuzzification forecasting of Type 2 observations, dan m_{qz} is the midpoint and there are total observations is j. Suppose at $t = 242$, then

a. Operator union $\vee$

$$Defuzzification_{t(union)} \quad \frac{\sum_{z=1}^{j} m_{qz}}{j}$$

$$Defuzzification_{242(union)} = \frac{m_6 + m_7 + m_8}{3}$$

$$= \frac{5425.37925 + 5584.5415 + 6210.78775}{3} = \frac{16819.13203}{3}$$

$$= 5606.37675$$

b. Operator intersection $\wedge$

$$Defuzzification_{t(intersection)} = \frac{\sum_{z=1}^{j} m_{qz}}{j}$$

$$Defuzzification_{242(intersection)} = \frac{m_6 + m_7 + m_8}{3}$$
$$= \frac{5425.37925 + 5584.5415 + 6210.78775}{3}$$
$$= \frac{16819.13203}{3} = 5606.37675$$

Calculations were carried out on all data and the results of these calculations are shown in **Table 9**.

3.8 Type 2 Forecasting

Then, forecasting is carried out using the defuzzification operator union and intersection results. Type 2 Forecasting step using **Equation (12)**.

$$\hat{X}_t = \frac{Defuzzification_{t(union)} + Defuzzification_{t(intersection)}}{2} \tag{12}$$

where $\hat{X}_t$ is type 2 forecasting in period t, $Defuzzification_{t(union)}$ is the defuzzification result using the union operator in period t, dan $Defuzzification_{t(intersection)}$ is the defuzzification result using the intersection operator in period t.

After obtained $Defuzzification_{242(union)} = 5606.37675$ and $Defuzzification_{242(intersection)} = 5606.37675$, then $\hat{X}_{242}$

$$\hat{X}_t = \frac{Defuzzification_{t(union)} + Defuzzification_{t(intersection)}}{2}$$

$$\hat{X}_{242} = \frac{Defuzzification_{242(union)} + Defuzzification_{242(intersection)}}{2}$$
$$= \frac{5606{,}37675 + 5606{,}37675}{2} = \frac{11212.7535}{2} = 5606.37675$$

So, the forecasting on 04 January 2024 is 5606.37675.

Result of $Defuzzification_{t(union)}$, $Defuzzification_{t(intersection)}$, and $\hat{X}_t$ shown in **Table 11.**

Table 11. Forecast Result

t (Date)	Data	FLRG		$\vee_m$	$\wedge_m$	Defuzzification		$\hat{X}_t$
						Union	Intersection	
240 (02/01/ 2024)	X	A_8	$\to A_7, A_8$	$\{A_7, A_8\}$	$\{A_7, A_8\}$	5695.8755	5695.8755	5695.8755
	Y	A_8	$\to A_7, A_8$					
	Z	A_8	$\to A_7, A_8$					
241 (03/01/ 2024)	X	A_8	$\to A_7, A_8$	$\{A_6, A_7, A_8\}$	$\{A_7, A_8\}$	5606.37675	5695.8755	5651.6261
	Y	A_8	$\to A_7, A_8$					
	Z	A_7	$\to A_6, A_7, A_8$					
242 (04/01/ 2024)	X	A_7	$\to A_6, A_7, A_8$	$\{A_6, A_7, A_8\}$	$\{A_6, A_7, A_8\}$	5606.37675	5606.37675	5606.3768
	Y	A_7	$\to A_6, A_7, A_8$					
	Z	A_7	$\to A_6, A_7, A_8$					
…	…	…	…	…	…	…	..	…

t (Date)	Data	FLRG		$\vee_m$	$\wedge_m$	Defuzzification		$\hat{X}_t$
						Union	Intersection	
330 (30/05/ 2024)	X	A_1	$\rightarrow A_1, A_2$	$\{A_1, A_2\}$	$\{A_1, A_2\}$	4596.0696	4596.0696	4596.0696
	Y	A_1	$\rightarrow A_1, A_2$					
	Z	A_1	$\rightarrow A_1, A_2$					

Then, the forecasting results obtained in the next period on 3 June 2024 amounted to 4596.0696.

3.9 T2FTS

As a comparison, calculations are made on T2FTS, as follows:

The universe of discourse is divided into intervals of equal length. Many intervals can be obtained using the Sturges formula as follows [6].

$k = 1 + 3.322 \log(n)$, with k being the number of intervals and n being the number of data, then

$k = 1 + 3.322 \log(n) = 1 + 3.322 \log(330) = 1 + 8.367 = 9.367 \approx 9$. Hence, the number of intervals obtained is 9.

Next, determine the length of the interval (p).

$p = \frac{D_{max} - D_{min}}{k} = \frac{6400-4380}{9} = 224.444$

Then the intervals formed are shown in Table 12.

Table 12. Interval using Chen Model

u_i	Interval	Midpoint
u_1	(4296, 4520.444)	4408.222
u_2	(4520.444, 4744.888)	4632.666
u_3	(4744.888, 4969.333)	4857.111
…	…	…
u_8	(5867.111, 6091.555)	5979.333
u_9	(6091.555, 6467)	6279.278

The next step is the same calculation as FCM-T2FTS. The results of forecasting using T2FTS are shown in Table 13.

Table 13. Forecast Result T2FTS

Date	t	Forecast
02/01/2024	240	5642.667
03/01/2024	241	5586.556
04/01/2024	242	5586.556
…	…	…
29/05/2024	329	5576.556
30/05/2024	330	4520.444

Then, the forecasting results obtained in the next period on 3 June 2024 amounted to 4520.444.

3.10 Evaluate Performance

Equation (8) uses the Mean Absolute Percentage Error (MAPE) to determine the accuracy of the prediction. By comparing the predicted and actual values, MAPE calculates the percentage error; a smaller value indicates that the model is more precise. Based on the forecasting results using T2FTS and FCM-T2FTS, the MAPE values are 2.09% and 1.97%, respectively. The accuracy test results have a minimal difference. By the MAPE criteria in Table 1, the MAPE value obtained by both is less than 10%, indicating that it has outstanding accuracy and is included in the excellent category.

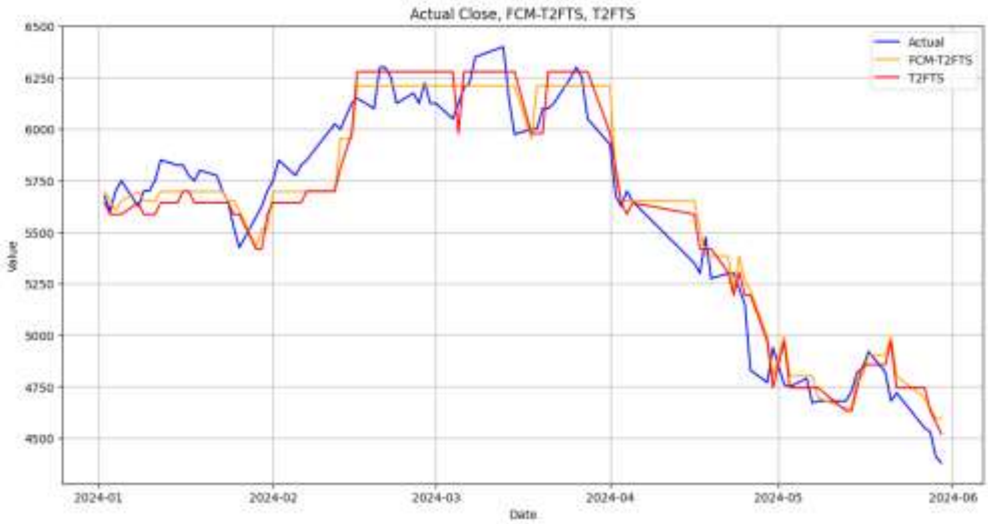

Figure 3. Comparison T2FTS with Combination FCM-T2FTS

Based on **Figure 3**, it can be seen that the forecasting results of the two models for the period January to May 2024 have a slight difference. The blue graph shows the actual data, the orange graph shows the FCM-T2FTS forecasting results, and the red graph shows the T2FTS forecasting results. Considering that both models provide excellent values, it is necessary to calculate the efficiency to determine the best model. The difference between the two models is 0.12%, and by using T2FTS as a reference, the efficiency is obtained by 12%. The model shows that the FCM-T2FTS combination model is more efficient by 12% compared to T2FTS. Thus, the FCM-T2FTS combination provides better results than T2FTS. As in **[9]**, which also provides results that the combination of FCM and FTS provides good results because FCM is a clustering method that is useful for overcoming interval partitions, and in **[14]**, T2FTS provides good results also by refining fuzzy relations. So, the combination of FCM and T2FTS also can improve accuracy results for stock price forecasting.

4. CONCLUSIONS

This research develops T2FTS combined with FCM, where FCM determines the interval. The data used is BRI daily stock prices from January 2023 to May 2024 obtained from the Yahoo Finance site, which is accessible to the public. This proposed model aims to improve the accuracy of results in stock price forecasting. Based on the results obtained, the MAPE value of T2FTS and the FCM-T2FTS combination provides excellent accuracy results, namely 2.09% and 1.97%, respectively. The difference between the two models is 0.12%, so FCM-T2FTS is 12% more efficient than T2FTS for forecasting BRI stock prices, and the FCM-T2FTS combination can improve forecasting accuracy results. FCM can improve forecasting, especially in interval partitioning, because the results show that determining interval length using unequal length is more efficient than fixed interval length and can enhance the model's performance.

ACKNOWLEDGMENT

This research was partially supported by a grant from the Faculty of Science and Mathematics Diponegoro University 2024. Number Contract: 27.III.G/UN7.F8/PP/II/2024.

REFERENCES

[1] S. D. Laskarjati and I. S. Ahmad, "PERBANDINGAN PERAMALAN HARGA SAHAM MENGGUNAKAN AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) DAN FUZZY TIME SERIES MARKOV CHAIN (STUDI KASUS: SAHAM PT INDOFOOD CBP SUKSES MAKMUR TBK)," *Jurnal Sains dan Seni ITS*, vol. 11, no. 6, pp. 397–404, 2022, doi: 10.12962/j23373520.v11i6.91417.

[2] T. Julian, T. Devrison, V. Anora, and K. M. Suryaningrum, "STOCK PRICE PREDICTION MODEL USING DEEP LEARNING OPTIMIZATION BASED ON TECHNICAL ANALYSIS INDICATORS," *Procedia Comput Sci*, vol. 227, pp. 939–947, 2023, doi: 10.1016/j.procs.2023.10.601.
[3] Q. Song and B. S. Chissom, "FORECASTING ENROLLMENTS WITH FUZZY TIME SERIES - PART I," *Fuzzy Sets Syst*, vol. 54, no. 1, pp. 1–9, 1993, doi: https://doi.org/10.1016/0165-0114(93)90355-L.
[4] Q. Song and B. S. Chissom, "FORECASTING ENROLLMENTS WITH FUZZY TIME SERIES - Part II," *Fuzzy Sets Syst*, vol. 62, no. 1, pp. 1–8, 1994.
[5] N. Van Tinh and N. C. Dieu, "A NEW HYBRID FUZZY TIME SERIES FORECASTING MODEL BASED ON COMBINING FUZZY C-MEANS CLUSTERING AND PARTICLE SWAM OPTIMIZATION," *Journal of Computer Science and Cybernetics*, vol. 35, no. 3, pp. 267–292, Aug. 2019, doi: 10.15625/1813-9663/35/3/13496.
[6] 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: Conference Series*, Institute of Physics Publishing, Mar. 2020, pp. 1–11. doi: 10.1088/1742-6596/1462/1/012044.
[7] S.-M. Chen, "FORECASTING ENROLLMENTS BASED ON FUZZY TIME SERIES," *Fuzzy Sets Syst*, vol. 81, no. 3, pp. 311–319, 1996.
[8] S. T. Li, Y. C. Cheng, and S. Y. Lin, "A FCM-Based Deterministic Forecasting Model For Fuzzy Time Series," *Computers and Mathematics with Applications*, vol. 56, no. 12, pp. 3052–3063, Dec. 2008, doi: 10.1016/j.camwa.2008.07.033.
[9] R. M. Pattanayak, S. Panigrahi, and H. S. Behera, "HIGH-ORDER FUZZY TIME SERIES FORECASTING BY USING MEMBERSHIP VALUES ALONG WITH DATA AND SUPPORT VECTOR MACHINE," *Arab J Sci Eng*, vol. 45, no. 12, pp. 10311–10325, Dec. 2020, doi: 10.1007/s13369-020-04721-1.
[10] K. Bisht and A. Kumar, "A METHOD FOR FUZZY TIME SERIES FORECASTING BASED ON INTERVAL INDEX NUMBER AND MEMBERSHIP VALUE USING FUZZY C-MEANS CLUSTERING," *Evol Intell*, vol. 16, no. 1, pp. 285–297, Feb. 2021, doi: 10.1007/s12065-021-00656-0.
[11] S. Iqbal, C. Zhang, M. Arif, M. Hassan, and S. Ahmad, "A NEW FUZZY TIME SERIES FORECASTING METHOD BASED ON CLUSTERING AND WEIGHTED AVERAGE APPROACH," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 5, pp. 6089–6098, 2020, doi: 10.3233/JIFS-179693.
[12] E. Egrioglu, C. H. Aladag, and U. Yolcu, "FUZZY TIME SERIES FORECASTING WITH A NOVEL HYBRID APPROACH COMBINING FUZZY C-MEANS AND NEURAL NETWORKS," *Expert Syst Appl*, vol. 40, no. 3, pp. 854–857, Feb. 2013, doi: 10.1016/j.eswa.2012.05.040.
[13] H. Qu, M. Sun, J. Zhao, and Y. Zhang, "FUZZY TIME SERIES FORECASTING BASED ON HESITANT FUZZY SETS AND FUZZY C-MEANS CLUSTERING," in *International Conference on Communication Technology Proceedings, ICCT*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020, pp. 1514–1518. doi: 10.1109/ICCT50939.2020.9295742.
[14] R. W. Ningrum, B. Surarso, Farikhin, and Y. M. Safarudin, "FORECASTING JAKARTA COMPOSITE INDEX (IHSG) BASED ON CHEN FUZZY TIME SERIES AND FIREFLY CLUSTERING ALGORITHM," *J Phys Conf Ser*, vol. 983, no. 1, pp. 1–6, Apr. 2018, doi: 10.1088/1742-6596/983/1/012055.
[15] K. Huarng, "HEURISTIC MODELS OF FUZZY TIME SERIES FOR FORECASTING," *Fuzzy Sets Syst*, vol. 123, pp. 369–386, 2001, doi: https://doi.org/10.1016/S0165-0114(00)00093-2.
[16] K. Huarng and H. K. Yu, "A TYPE 2 FUZZY TIME SERIES MODEL FOR STOCK INDEX FORECASTING," *Physica A: Statistical Mechanics and its Applications*, vol. 353, no. 1–4, pp. 445–462, Aug. 2005, doi: 10.1016/j.physa.2004.11.070.
[17] A. Nur Hanifah, W. Sulandari, H. Pratiwi, P. Studi Statistika, and U. Sebelas Maret, "PERBANDINGAN PROYEKSI HARGA SAHAM MENGGUNAKAN FUZZY TIME SERIES TIPE 1 DAN 2," *Matematika dan Statistika*, vol. 8, 2023.
[18] Abhishekh, S. S. Gautam, and S. R. Singh, "A NEW TYPE 2 FUZZY TIME SERIES FORECASTING MODEL BASED ON THREE-FACTORS FUZZY LOGICAL RELATIONSHIPS," *International Journal of Uncertainty, Fuzziness and Knowldege-Based Systems*, vol. 27, no. 2, pp. 251–276, Apr. 2019, doi: 10.1142/S0218488519500120.
[19] A. A. Pramesti, W. Sulandari, S. Subanti, and Y. Yudhanto, "FORECASTING THE COMPOSITE STOCK PRICE INDEX USING FUZZY TIME SERIES TYPE 2," *Radiant*, vol. 4, no. 2, pp. 118–133, Aug. 2023, doi: 10.52187/rdt.v4i2.167.
[20] A. Sofro, R. A. Riani, K. N. Khikmah, R. W. Romadhonia, and D. Ariyanto, "ANALYSIS OF RAINFALL IN INDONESIA USING A TIME SERIES-BASED CLUSTERING APPROACH," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 18, no. 2, pp. 0837–0848, May 2024, doi: 10.30598/barekengvol18iss2pp0837-0848.
[21] P.-C. Chang, Y.-W. Wang, and C.-H. Liu, "THE DEVELOPMENT OF A WEIGHTED EVOLVING FUZZY NEURAL NETWORK FOR PCB SALES FORECASTING," *Expert Syst Appl*, vol. 32, no. 1, pp. 86–96, Jan. 2007, doi: 10.1016/j.eswa.2005.11.021.