BAREKENG: Journal of Mathematics and Its Applications
June 2025 Volume 19 Issue 2 Page 1237-1250
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss2pp1237-1250

FORECASTING NICKEL PRICES WITH THE AUTOMATIC CLUSTERING FUZZY TIME SERIES MARKOV APPROACH

M. Al Haris[1*], Wulan Sari[2], Fatkhurokhman Fauzi[3], Muhammad Sam'an[4]

[1,2,3]Department of Statistics, Faculty of Agricultural Science and Technology, Universitas Muhammadiyah Semarang
[4]Department of Informatics, Faculty of Engineering and Computer Science, Universitas Muhammadiyah Semarang
Jln. Kedungmundu No. 18, Semarang, 50273, Indonesia

*Corresponding author's e-mail: * **alharis@unimus.ac.id***

ABSTRACT

Article History:

Received: 8th October 2024
Revised: 3rd February 2025
Accepted: 4th March 2025
Published: 1st April 2025

Keywords:

Automatic Clustering;
Fuzzy Time Series Markov Chain;
Forecasting;
Nickel Price.

Nickel was a critical raw material used in a wide range of industries. The price movement of nickel tends to fluctuate and remain uncertain due to market conditions varying over time. Therefore, forecasting nickel prices was essential to understanding future price movements. In this study, we applied the automatic clustering fuzzy time series Markov chain method. The automatic clustering algorithm generates multiple intervals and fuzzy relations. Subsequently, forecasting was based on these fuzzy relations and a Markov chain transition probability matrix involving three stages to enhance forecast accuracy. We use monthly closing futures nickel price data from January 2009 to May 2024. The accuracy of the forecasting model was measured using the mean absolute percentage error (MAPE). The analysis showed that implementing the automatic clustering fuzzy time series Markov chain method results in excellent forecasting accuracy, with a MAPE value of 1.76% (equivalent to 98.24% accuracy). The predicted nickel price for June 2024 was US$ 19,608.5.

How to cite this article:

M. A. Haris, W. Sar, F. Fauzi and M. Sam'an., "FORECASTING NICKEL PRICES WITH THE AUTOMATIC CLUSTERING FUZZY TIME SERIES MARKOV APPROACH," *BAREKENG: J. Math. & App.*, vol. 19, iss. 2, pp. 1237-1250, 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

Mining is the primary industry that provides essential raw materials for various sectors, including transportation, construction, aerospace, high technology, and energy [1]. One of Indonesia's prominent mining commodities is nickel. Nickel is a reactive element to acids and reacts slowly to air at normal temperature and pressure [2]. Due to these properties, nickel is widely utilized as a raw material in producing batteries, stainless steel, iron-steel alloys, and even military weaponry. Its diverse applications make nickel a crucial resource in industrial activities [3].

As the world's largest producer of nickel, Indonesia can produce 1,600,000 tons of nickel and record export values of US$ 4.13 billion [4]. The high production figures and export values make nickel prices a crucial consideration. World nickel prices are denominated in United States dollars (US) and are constantly fluctuating due to market conditions [5], [6]. These fluctuations create uncertainty in nickel prices, even though relevant parties such as governments, companies, and investors require information about future world nickel prices as a reference for policy-making and decision-making. Therefore, forecasting world nickel prices is essential to understanding nickel price movements in the future [7].

Forecasting is the expected demand prediction for a product within a specific time frame [8], [9], [10]. The fluctuating world nickel prices result in non-stationary data. Therefore, a method that does not rely on stationarity assumptions is needed for forecasting world nickel prices. One such method that avoids stationarity assumptions is fuzzy time series. Fuzzy time series works by retaining historical data and processing it to generate new values in the future without requiring assumptions [8], [11]. This method is also capable of forecasting without prior training. However, its drawback lies in producing only a limited number of fuzzy sets and performing defuzzification in a single step based on the midpoint of intervals without considering state transitions. To address this limitation, a new method called automatic clustering fuzzy time series Markov chain has been developed [12]. The automatic clustering fuzzy time series Markov chain method leverages an automatic clustering algorithm, which creates multiple intervals of varying lengths based on the data used, resulting in numerous fuzzy sets and fuzzy relations [13], [14]. Additionally, the defuzzification process involves three steps based on interval midpoints, fuzzy relations, and the Markov transition probability matrix, considering state transitions to enhance forecasting accuracy [12], [15].

The automatic clustering fuzzy time series Markov chain method has been used to forecast the number of COVID-19 cases in North Sumatra province. This method yielded a mean absolute percentage error (MAPE) of 4.53%, indicating excellent forecasting capability [12]. Similar methods were also employed by [16] and [15]. The results of applying the automatic clustering fuzzy time series Markov chain method in both studies showed good accuracy, with MAPE values remaining below 20%. Additionally, nickel price forecasting was conducted by [5] using the fuzzy Mamdani system method. The analysis revealed that this method performed well in predicting nickel prices, with a prediction error based on MAPE of 7%.

Based on the description above, this research will forecast global nickel prices for the upcoming period in June 2024 using the automatic clustering fuzzy time series Markov chain method. The accuracy of the forecast will be measured using the mean absolute percentage error (MAPE).

2. RESEARCH METHODS

This study utilizes secondary data on monthly global nickel prices (closing prices) from January 2009 to May 2024, comprising 185 observations. The data was obtained from the investing.com website. The variable research was the closing price of global nickel from January 2009 to May 2024 that was indicated in **Table 1**.

Table 1. Monthly Closing Prices of Global Nickel from January 2009 to May 2024

Period	Y_t
January 2009	Y_1
February 2009	Y_2
⋮	⋮
May 2024	Y_{185}

Data source: https://www.investing.com/commodities/nickel?cid=959208

2.1 Fuzzy Time Series

Fuzzy time series is defined as a forecasting method for time series data based on fuzzy principles, where the values of the time series (real numbers) are transformed into fuzzy sets [8], [17], [18]. Each fuzzy set has varying degrees of membership, indicating the extent of truth and error, ranging between 1 (one) and 0 (zero). A fuzzy set A within the universe of discourse U can be expressed as [19]:

$$\boldsymbol{A} = \frac{\mu_A(u_1)}{u_1} + \frac{\mu_A(u_2)}{u_2} + \cdots + \frac{\mu_A(u_n)}{u_n} \tag{1}$$

The expression $\mu_{\boldsymbol{A}}(\boldsymbol{u_i})$ represents the membership degree of the element $\mu_{\boldsymbol{i}}$ in the fuzzy set $\boldsymbol{A_i}$, where $\mu_{\boldsymbol{A}}(\boldsymbol{u_i}) \in [\boldsymbol{0}, \boldsymbol{1}]$ for $\boldsymbol{1} \leq \boldsymbol{i} \leq \boldsymbol{n}$. The value of $\mu_{\boldsymbol{A}}(\boldsymbol{u_i})$ can be interpreted as follows:

$$\mu_A(u_i) = \begin{cases} 1 & i = j \\ 0{,}5 & , if\ i = j - 1\ or\ j + 1 \\ 0 & the\ others \end{cases} \tag{2}$$

Fuzzy sets are represented as the fundamental concept in fuzzy time series based on the following definition [14]:

Definition 1

Assume $Y(t)$ (where $t = 1,2, \ldots$) represents a subset of real numbers. It is defined as the universe of discourse containing fuzzy sets $f_i(t)$ (where $t = 1,2, \ldots$). If $F(t)$ is the union of $f_1(t), f_2(t), \ldots$ then $F(t)$ is expressed as a fuzzy time series concerning $Y(t)$ (where $t = 1,2, \ldots$).

Definition 2

Assume that $F(t)$ occurs due to $F(t-1)$, i.e., $F(t-1) \rightarrow F(t)$. Consequently, there exists a fuzzy relationship between $F(t)$ and $F(t-1)$, represented by the following equation:

$$F(t) = F(t-1)^{\circ} R(t, t-1) \tag{3}$$

Given that R represents the combination of fuzzy relations and is known as the first-order model of $F(t)$, where "∘" denotes the max-min composition operator and $R(t, t-1)$ is the relation matrix that describes the fuzzy relationship between $F(t)$ and $F(t-1)$.

Definition 3

Assume $F(t)$ as a fuzzy time series, and $R(t, t-1)$ represents the first-order relationship of $F(t)$. If $R(t, t-1) = R(t-1, t-2)$ is independent of time t (i.e., it does not depend on the difference between time t and $t-1$), then $F(t)$ is called a time-invariant fuzzy time series.

Definition 4

When $F(t) = A_j$ and $F(t-1) = A_i$, the relationship between $F(t)$ and $F(t-1)$ is termed a fuzzy logical relationship (FLR), denoted as $A_i \rightarrow A_j$. Here, A_i represents the left-hand side (LHS), and A_j represents the right-hand side (RHS) of the FLR.

Definition 5

Assume there exist FLR $A_i \rightarrow A_{j1}$, $A_i \rightarrow A_{j2}$, ..., $A_i \rightarrow A_{jn}$ that share the same fuzzy set at state A_i. In this case, a fuzzy logic relationship group (FLRG) is formed, denoted as $A_i \rightarrow A_{j1}, A_{j2}, \ldots, A_{jn}$.

2.2 Automatic Clustering

Automatic clustering classifies numerical data based on proximity, where smaller distances indicate higher similarity between data points [13]. The automatic clustering algorithm is computed through the following stages:

a. Create a numerical data sequence from the smallest to the largest values ($\boldsymbol{d_1}, \boldsymbol{d_2}, \dots, \boldsymbol{d_n}$), assuming that there are no repeated values in the data. Referring to this sequence, calculate the average difference in the formation of intervals for time series data (average_diff) based on the following equation:

$$average\_diff = \frac{\sum_{i=0}^{n-1}(d_{i+1} - d_i)}{n-1} \tag{4}$$

The average_diff represents the mean distance between numerical data points after sorting.

b. Set the first numerical data point (the smallest data value) as the current cluster. Based on the value of average_diff, determine whether the next numerical data point should be added to the current cluster or form a different cluster, following the principles outlined below:

Principle 1: Suppose the current cluster represents the first cluster, containing only one numerical data point ($\boldsymbol{d_1}$), and assume that $\boldsymbol{d_2}$ is the next numerical data point. This can be displayed as $\{\boldsymbol{d_1}\}, \boldsymbol{d_2},, \boldsymbol{d_3}, \dots,, \boldsymbol{d_i},, \boldsymbol{d_n}$.

If the value of $\boldsymbol{d_2} - \boldsymbol{d_1} \leq \boldsymbol{average\_diff}$, then $\boldsymbol{d_2}$ is included as a member of the current cluster along with $\boldsymbol{d_1}$. If not, create a new cluster where $\boldsymbol{d_2}$ becomes a member.

Principle 2: Suppose the current cluster is not the first cluster and the current cluster contains only numerical data, denoted as $\boldsymbol{d_j}$. Assume that $\boldsymbol{d_k}$ represents the nearest numerical data to $\boldsymbol{d_j}$ and $\boldsymbol{d_i}$ represents the numerical data with the highest value in the previous cluster. This can be displayed as $\{\boldsymbol{d_1}\}, \dots, \{\dots, \boldsymbol{d_i}\}, \{\boldsymbol{d_j}\}, \boldsymbol{d_k}, \dots, \boldsymbol{d_n}$.

If the value of $\boldsymbol{d_k} - \boldsymbol{d_j} \leq \boldsymbol{average\_diff}$ and the value of $\boldsymbol{d_k} - \boldsymbol{d_j} \leq \boldsymbol{d_j} - \boldsymbol{d_i}$, then $\boldsymbol{d_k}$ is placed in the current cluster that contains $\boldsymbol{d_j}$. Otherwise, form a new cluster where $\boldsymbol{d_k}$ becomes a member.

Principle 3: Suppose the first cluster is not the current cluster, and there exists numerical data greater than 1 in the current cluster. Assume that the largest numerical data in the current cluster is $\boldsymbol{d_i}$ and $\boldsymbol{d_j}$ represents the numerical data closest to $\boldsymbol{d_i}$. This can be displayed as $\{\boldsymbol{d_1}\}, \dots, \{\dots, \boldsymbol{d_i}\}, \boldsymbol{d_j}, \dots, \boldsymbol{d_n}$.

If the value of $\boldsymbol{d_j} - \boldsymbol{d_i}$ is less than or equal to average_diff and also less than or equal to cluster_diff, then $\boldsymbol{d_j}$ is placed as a member of the current cluster along with $\boldsymbol{d_i}$. Otherwise, form a new cluster that includes $\boldsymbol{d_j}$. The value of cluster_diff can be calculated based on the following equation:

$$cluster\_diff = \frac{\sum_{i=1}^{n-1}(c_{i+1} - c_i)}{n-1} \tag{5}$$

The cluster_diff represents the mean of the current cluster, where $\boldsymbol{c_1}, \boldsymbol{c_2}, \dots, \boldsymbol{c_n}$ represent the data within the current cluster.

c. Referring to the clustering results from the previous step, adjust the content of each cluster by considering the following principles:

Principle 1: If a cluster consists of more than 2 data points, retain the data with the smallest and largest values while discarding the other data.

Principle 2: If a cluster consists of 2 data points, retain all the data.

Principle 3: If a cluster contains only one numeric data point, denoted as $\boldsymbol{d_q}$, then remove $\boldsymbol{d_q}$ and replace it with two values $\boldsymbol{d_q - average\_diff}$ representing the smallest value and $\boldsymbol{d_q + average\_diff}$ representing the largest value within the cluster. However, the cluster needs further adjustment in the following situations:

Situation 1: If the situation occurs in the first cluster, remove the value $\boldsymbol{d_q - average\_diff}$ and retain $\boldsymbol{d_q}$.

Situation 2: If the situation occurs in the last cluster, remove the value $\boldsymbol{d_q + average\_diff}$ and retain $\boldsymbol{d_q}$.

Situation 3: If the value of $\boldsymbol{d_q - average\_diff}$ is less than the smallest value in the previous cluster, then disregard Principle 3.

d. Assume that step 3 yields the following clustering results:

$$\{\boldsymbol{d_1}, \boldsymbol{d_2}\}, \{\boldsymbol{d_3}, \boldsymbol{d_4}\}, \{\boldsymbol{d_5}, \boldsymbol{d_6}\} \ \dots, \{\boldsymbol{d_r}\}, \{\boldsymbol{d_s}, \boldsymbol{d_t}\}, \dots, \{\boldsymbol{d_{n-1}}, \boldsymbol{d_n}\}$$

i. The first cluster $\{\boldsymbol{d_1}, \boldsymbol{d_2}\}$ is formed as the interval $[\boldsymbol{d_1}, \boldsymbol{d_2})$.

ii. If the current cluster consists of $\{\boldsymbol{d_k}, \boldsymbol{d_1}\}$ and the current interval is $[\boldsymbol{d_1}, \boldsymbol{d_2})$, then:

1. If the value of $\boldsymbol{d_j} \geq \boldsymbol{d_k}$, form the interval $[\boldsymbol{d_j}, \boldsymbol{d_1})$ and designate it as the current interval. Additionally, assign the next cluster $\{\boldsymbol{d_m}, \boldsymbol{d_n}\}$ as the current cluster.
2. If the value of $\boldsymbol{d_j} < \boldsymbol{d_k}$, replace the current cluster $\{\boldsymbol{d_k}, \boldsymbol{d_1}\}$ with the interval $[\boldsymbol{d_i}, \boldsymbol{d_j})$. Then, create a new interval $[\boldsymbol{d_j}, \boldsymbol{d_k})$ based on the intervals $[\boldsymbol{d_i}, \boldsymbol{d_j})$ and $[\boldsymbol{d_k}, \boldsymbol{d_1})$. Set $[\boldsymbol{d_k}, \boldsymbol{d_1})$ as the current interval, and designate the next cluster $\{\boldsymbol{d_m}, \boldsymbol{d_n}\}$ as the current cluster. If$[\boldsymbol{d_i}, \boldsymbol{d_j})$ is the current interval and $\{\boldsymbol{d_k}\}$ is the current cluster, modify the current interval from $[\boldsymbol{d_i}, \boldsymbol{d_j})$ to $[\boldsymbol{d_i}, \boldsymbol{d_k})$. Then, assign the following cluster as the current cluster.

e. Based on the intervals formed from the previous step, create new intervals where each interval is divided into $\boldsymbol{p}$ sub-intervals, with $\boldsymbol{p \geq 1}$. Then, calculate the midpoint and length of each interval.

2.3 Markov Chain

Markov chain analysis is a technique that studies the current characteristics of a variable based on its past behaviour to predict its future behaviour. Markov chains yield probabilistic model information (stochastic processes), which is used to support decision-making. Therefore, this analysis is descriptive in nature and not an optimization technique [14], [20], [21]. If at the time $\boldsymbol{n}$, the stochastic process $\{\boldsymbol{Y_n}, \boldsymbol{n = 1, 2, \dots}\}$ is in state $\boldsymbol{i}$, then it can be written as $\boldsymbol{Y_n = i}$. The subsequent stochastic process is in state $\boldsymbol{j}$ with a probability $\boldsymbol{P_{ij}}$ can be expressed as:

$$P\{Y_{n+1} = j \mid Y_{n-1} = i_{n-1}, \dots, Y_1 = i_1, Y_0 = i_0\} = \boldsymbol{P_{ij}} \tag{6}$$

For every state $\boldsymbol{i_0}, \boldsymbol{i_1}, \dots, \boldsymbol{i_{n-1}}, \boldsymbol{i_n}, \boldsymbol{j}$ and $\boldsymbol{n \geq 0}$. This process is called a Markov chain, where the current event $\boldsymbol{Y_n}$ determines the probability of the next event $\boldsymbol{Y_{n+1}}$. The value $\boldsymbol{P_{ij}}$ represents the probability of transitioning from state $\boldsymbol{i}$ to state $\boldsymbol{j}$. The probability is positive, and the transition process moves to a specific state, resulting in the following equation:

$$\boldsymbol{P_{ij}} \geq 0, i, j \geq 0; \sum_{j=0}^{\infty} \boldsymbol{P_{ij}} = 1, i = 1,2, \dots \tag{7}$$

Suppose $\boldsymbol{P}$ is the matrix of one-step transition probabilities, denoted as $\boldsymbol{P_{ij}}$, then:

$$\boldsymbol{P} = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix}. \tag{8}$$

2.4 Automatic Clustering Fuzzy Time Series Markov Chain

Automatic clustering fuzzy time series Markov chain is an extension of the fuzzy time series method that determines intervals using automatic clustering algorithms and incorporates Markov chains during the defuzzification process. The addition of Markov chains aims to generate the highest probabilities using transition probability matrices, thereby enhancing the accuracy of forecasting results. The steps for applying the automatic clustering fuzzy time series Markov chain method are as follows [12], [15]:

a. Apply the automatic clustering algorithm to form intervals.

b. Arrange the data sequence starting from the smallest value to the largest value, and defining the universe of discourse $\boldsymbol{U} = [\boldsymbol{Y_{min}}, \boldsymbol{Y_{max}}]$ based on the available time series data.

c. Defining fuzzy sets to obtain membership values for each fuzzy set from time series data using **Equation (1)**.

d. Fuzzifying the existing time series data into fuzzy sets (linguistic variables). If a time series data point falls within the interval $\boldsymbol{u_i}$, it will be fuzzified into the fuzzy set $\boldsymbol{A_i}$.

e. Forming fuzzy logical relationships (FLR) and fuzzy logical relationship groups (FLRG) based on the results of fuzzification. Assuming that the fuzzified data at a time $\boldsymbol{t}$ and $\boldsymbol{t+1}$ are denoted as $\boldsymbol{A_i}$ and $\boldsymbol{A_j}$, respectively, the FLR '$\boldsymbol{A_i} \to \boldsymbol{A_j}$' is formed, where $\boldsymbol{A_i}$ represents the current state time series data and $\boldsymbol{A_j}$ represents the subsequent state time series data. The resulting FLR are then used to create FLRG by including FLR that share the same current state (left-hand side) within the same FLRG.

f. Constructing the Markov transition probability matrix based on FLRG. Assume that the Markov probability matrix is of dimension $\boldsymbol{p} \times \boldsymbol{p}$, where $\boldsymbol{p}$ represents the number of fuzzy sets. The transition probability values are calculated using the following formula:

$$\boldsymbol{P_{ij}} = \frac{r_{ij}}{r_i} \tag{9}$$

with $\boldsymbol{P_{ij}}$ denoted transition probability from state $\boldsymbol{A_i}$ to $\boldsymbol{A_j}$, $\boldsymbol{r_{ij}}$ denoted the number of transitions from state $\boldsymbol{A_i}$ to $\boldsymbol{A_j}$, $\boldsymbol{r_i}$ denoted the number of data points in $\boldsymbol{A_i}$.

g. Defuzzification of data based on FLR, FLRG, and the Markov transition probability matrix results in an initial forecast value. The initial forecast value is determined according to the following rules:

Rule 1: If FLRG $\boldsymbol{A_i}$ is an empty set ($\boldsymbol{A_i} \to \emptyset$) for the data at time t, then the forecasted value $\boldsymbol{\hat{Y}_t}$ is taken from $\boldsymbol{m_{i(t)}}$ which represents the midpoint value of interval $\boldsymbol{u_i}$.

Rule 2: If FLRG $\boldsymbol{A_i}$ represents a one-to-one mapping ($\boldsymbol{A_i} \to \boldsymbol{A_j}$, where $\boldsymbol{P_{ik} = 0}$ and $\boldsymbol{P_{ik} = 1}$, $\boldsymbol{k \neq j}$)), then the forecasted value $\boldsymbol{\hat{Y}_t}$ is taken from $\boldsymbol{m_j}$, which corresponds to the median of interval $\boldsymbol{u_i}$, using the following equation:

$$\hat{Y}_t = m_j P_{ij} = m_j \tag{10}$$

Rule 3: If FLRG $\boldsymbol{A_i}$ represents a one-to-many mapping ($\boldsymbol{A_k} \to \boldsymbol{A_1}, \boldsymbol{A_2}, \dots, \boldsymbol{A_n}, \boldsymbol{k = 1, 2, , \dots, n}$), assuming that the data set ($\boldsymbol{Y_{t-1}}$) at time ($\boldsymbol{t-1}$) belongs to state $\boldsymbol{A_k}$, the forecasted value $\boldsymbol{\hat{Y}_t}$ is formulated as follows:

$$\hat{Y}_t = m_1 P_{k1} + m_2 P_{k2} + \cdots + m_{k-1} P_{k(k-1)} + Y_{(t-1)} P_k + m_{k+1} P_{k(k+1)} + \cdots + m_n P_n \tag{11}$$

Given that $\boldsymbol{m_1}, \boldsymbol{m_2}, \dots, \boldsymbol{m_{k-1}}, \boldsymbol{m_{k+1}}, \dots, \boldsymbol{m_n}$ are the medians of $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_{k-1}}, \boldsymbol{u_{k+1}}, \dots, \boldsymbol{u_n}$, we substitute $\boldsymbol{m_k}$ into $\boldsymbol{Y_{t-1}}$ to obtain information from state $\boldsymbol{A_k}$ at time $(\boldsymbol{t-1})$.

h. Calculating adjustment values is done to minimize forecast errors. The calculation is based on the following rules:

Rule 1: Let's assume that state $\boldsymbol{A_i}$ is related to state $\boldsymbol{A_j}$. Starting from state $\boldsymbol{A_i}$ at time $\boldsymbol{t-1}$, where $\hat{\boldsymbol{Y}}_{(\boldsymbol{t-1})} = \boldsymbol{A_i}$, a transition occurs to state $\boldsymbol{A_j}$ at time $\boldsymbol{t}$. As a result, the adjustment value $\boldsymbol{D_t}$ is calculated using the following equation:

$$D_{t1} = \pm \frac{1}{2} \tag{12}$$

with $\boldsymbol{l}$ represents the length of the interval.

Rule 2: Let's assume that state $\boldsymbol{A_i}$ occurs at time $\boldsymbol{t-1}$, where $\hat{\boldsymbol{Y}}_{(\boldsymbol{t-1})} = \boldsymbol{A_i}$. Then, a forward transition to state $\boldsymbol{A_{i+s}}$ occurs at time $\boldsymbol{t}$, resulting in the calculation of the adjustment value $\boldsymbol{D_t}$ as following equation:

$$D_{t1} = \pm \left(\frac{l}{2}\right) s \tag{13}$$

where *s* represents the number of transition movements.

i. Calculating the final forecast value, we use the following equation:

$$\hat{Y}'_t = \hat{Y}_t + D_{t1} + D_{t2} = \hat{Y}_t \pm \left(\frac{l}{2}\right) \pm \left(\frac{l}{2}\right) v \tag{14}$$

2.5 Evaluation of Forecast Accuracy

Evaluation of forecast accuracy uses the mean absolute percentage error (MAPE) value. The MAPE value is calculated based on the following equation [19], [22]:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{(Y_t - \hat{Y}'_t)}{(Y_t)} \right| \times 100\% \tag{15}$$

where n represents the number of periods, Y_t represents the original data for period t, and $\hat{Y}'_t$ represents the final forecasted value. The MAPE values are categorized into four criteria, as shown in Table 2.

Table 2. The Criteria for the MAPE Value

MAPE Values	Criteria
<10%	Very Good (Highly Accurate)
>10% - 20%	Good (Accurate)
>20% - 50%	Fairly Good (Moderately Accurate)
>50%	Not Good (Inaccurate)

3. RESULTS AND DISCUSSION

The historical data for global nickel futures closing prices within the time range from January 2009 to May 2024 reveals that the highest recorded price occurred in March 2022 at US$ 32,107, while the lowest price was observed in May 2016 at US$ 8,435. The average price range for global nickel stands at US$ 16,577.14. The visualization of the global nickel price plot from January 2009 to May 2024 is shown in **Figure 1**.

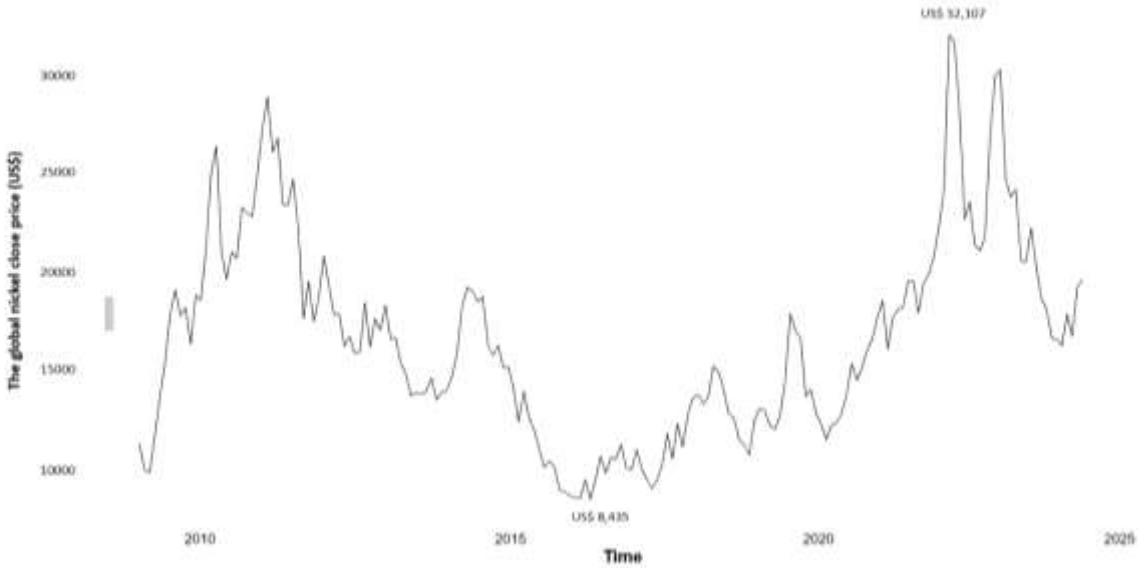

Figure 1. The Time Series Plot of the Global Nickel Price from January 2009 to May 2024

The time series plot of global nickel prices in **Figure 1** reveals a pattern that changes over time. These fluctuations are influenced by several factors, such as market demand and supply, commodity availability, exchange rate variations, and the global economic and political situation **[23]**. These changes indicate significant price movements characterized by fluctuations. Notably, the highest spike occurred from February to March 2022, with an increase of US$ 7,825, while the steepest decline happened from May to June 2022, resulting in a decrease of US$ 5,694. Based on the results of the analysis, the automatic clustering fuzzy time series Markov chain method is suitable for forecasting global nickel prices.

3.1 Automatic Clustering

The application of automatic clustering begins by sorting numerical data from the smallest value to the largest. **Table 3** shows the result of sorting the monthly closing prices of global nickel from January 2009 to May 2024 in ascending order.

Table 3. The Data is Sorted from Smallest to Largest

No	Data	No	Data	No	Data
1	8,435	4	8,620	7	8,970
2	8,490	5	8,820	⋮	⋮
3	8,520	6	8,900	180	32,107

Based on the sorted data, the next step is to find the value of average_diff using **Equation (4)**. According to **Equation (4)**, the calculated average_diff is 132.25. The sorted data along with this average_diff value is then used to form clusters.

The formation of clusters begins by designating the first data point (the smallest) as the current cluster. Referring to the value of average_diff, determine whether the next data point will be included in the current cluster or if a new cluster should be formed, following the principles outlined in steps 1, 2, and 3. In principle 3, additional criteria are based on the cluster_diff value, which is calculated using **Equation (5)** and varies for each cluster. The results of cluster formation are presented in **Table 4**.

Table 4. Clusters Result and Their Constituent Members

Cluster	Members	Cluster	Members	Cluster	Members
1	8,435	2	8,820		9,445
	8,490		8,900		9,450
	8,520		8,970	⋮	⋮
	8,620	3	9,390	98	32,107

Based on the obtained clusters in Table 4, the next step involves adjusting the cluster members according to the principles outlined in steps 1, 2, and 3. The adjusted cluster assignments are shown in Table 5.

Table 5. Results of Cluster Member Adjustments

Cluster	Members	Cluster	Members	Cluster	Members
1	8,435 8,620	3	9,390 9,450	⋮	⋮
2	8,820 8,970	4	9,765 9,800	98	31,974.75 32,107

The next step is to form intervals based on adjustment clusters. The intervals that are formed are shown in Table 6.

Table 6. Interval

Interval
u_1 = [8,435 ; 8,620)
u_2 = [8,620 ; 8,820)
u_3 = [8,820 ; 8,970)
⋮
u_{155} = [31,974.75 ; 32,107]

The formed intervals are then divided into p sub-intervals. In this study, the value of p used is 1, so the value of the sub-interval is the same as the interval. These intervals are then calculated for their length and midpoints. The results of the length and midpoint calculations of the intervals are shown in Table 7.

Table 7. Intervals, Interval Lengths, and Interval Midpoints

Intervals	Interval lengths	Interval midpoints
u_1 = [8,435 ; 8,620)	185	8,527.5
u_2 = [8,620 ; 8,820)	200	8,720
u_3 = [8,820 ; 8,970)	150	8,895
⋮	⋮	⋮
u_{155} = [31,974.75 ; 32,107]	132.25	32,040.87

3.2 Fuzzy Time Series Markov Chain

The initial step in applying the fuzzy time series Markov chain to forecasting global nickel futures is to form the universal set $\boldsymbol{U} = [Y_{min}, Y_{max}]$. Based on the available data, the universal set $\boldsymbol{U}$ = [8,435 ; 32,107] is obtained.

The next step is to define the fuzzy sets $\boldsymbol{A_i}$. The intervals obtained in the previous step are used to form fuzzy sets (linguistic variables), with 155 intervals, resulting in 155 fuzzy sets. According to Equation (1), the membership values of the fuzzy sets range from 0, 0.5, and 1. Below are the equations of the formed fuzzy sets.

$$A_1 = \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \cdots + \frac{0}{u_{154}} + \frac{0}{u_{155}}$$

$$A_2 = \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \cdots + \frac{0}{u_{154}} + \frac{0}{u_{155}}$$

$$\vdots$$

$$A_{155} = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \cdots + \frac{0.5}{u_{154}} + \frac{1}{u_{155}}$$

The following process is to perform fuzzification of time series data. This step aims to convert the time series data of global nickel prices (numeric) into fuzzy variables (linguistic) in the form of intervals. The results of the fuzzification of global nickel price data are shown in **Table 8**.

Table 8. Fuzzification of Time Series Data

No	Period	World Nickel Prices (US)	Fuzzification
1	January 2009	11,300	A_{21}
2	February 2009	9,925	A_8
3	March 2009	9,800	A_7
⋮	⋮	⋮	⋮
185	May 2024	19,642	A_{101}

The fuzzification results are then used to form FLR, which are fuzzy sets representing the relationship between each data point $\boldsymbol{t}$ and $\boldsymbol{t+1}$ (time $\boldsymbol{t}$ ⟶ time $\boldsymbol{t+1}$). The FLR obtained from the global nickel price data from January 2009 to May 2024 are shown in **Table 9**.

Table 9. Fuzzy Logic Relationship (FLR)

No	Period	FLR
1	January 2009 ⟶ February 2009	$A_{21} \longrightarrow A_8$
2	February 2009 ⟶ March 2009	$A_8 \longrightarrow A_7$
3	March 2009 ⟶ April 2009	$A_7 \longrightarrow A_{24}$
⋮	⋮	⋮
184	April 2024 ⟶ May 2024	$A_{97} \longrightarrow A_{101}$

The formed FLR are then grouped into FLRG, where each FLR with a similar current state (left-hand side) is combined into the same FLRG. The results of the FLRG grouping are shown in **Table 10**.

Table 10. Fuzzy Logic Relationship Group (FLRG)

No	FLR	FLRG
1	A_1	A_1, A_1, A_5, A_5
2	A_2	A_1
3	A_3	A_2, A_4
⋮	⋮	⋮
155	A_{155}	A_{153}

The next step is to form the Markov transition probability matrix. The Markov transition shows the interrelation between states (fuzzy sets). The values of the Markov transition probability matrix depend on the FLRG obtained from the previous step. The resulting matrix has an order of $\mathbf{155 \times 155}$ according to the number of fuzzy sets and is calculated using **Equation (9)**. Below are the results of the Markov transition probability matrix.

$$P = \begin{bmatrix} 1/2 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Calculating the initial forecast values is based on FLR, FLRG, and the Markov transition probability matrix, which are subsequently calculated using **Equation (10)** and **Equation (11)**. Forecasting data at time *t* is based on data at time $\boldsymbol{t-1}$, so January 2009 data (*t*) cannot be forecasted as there is no *t-1* data. For example, the February 2009 data have FLR $\boldsymbol{A_{21}} \longrightarrow \boldsymbol{A_8}$ and FLRG $\boldsymbol{A_{21}} \longrightarrow \boldsymbol{A_8}, \boldsymbol{A_{10}}$. The initial forecast value is calculated as:

$$\boldsymbol{\hat{Y}_2 = m_8 P_{21/8} + m_{10} P_{21/10} = (9,862.5)\left(\frac{1}{2}\right) + 9,987.5\left(\frac{1}{2}\right) = 9,925}$$

The obtained initial forecast values are shown in **Table 11**.

Table 11. Initial Forecast Values

No.	Period	Actual Data	Initial forecast values
1	January 2009	11,300	*
2	February 2009	9,925	9,925
3	March 2009	9,800	9,782.5
⋮	⋮	⋮	⋮
184	April 2024	19,642	19,608.5

The automatic clustering fuzzy time series Markov chain method has steps for forecast adjustment to minimize errors from the initial forecast values. The calculations are done using **Equation (12)** and **Equation (13)**, resulting in the outcomes shown in **Table 12**.

Table 12. Adjusted Forecast Values

No.	Current State	Next State	Adjusted forecast values
1	*	21	*
2	21	8	-697.4286
3	8	7	0
⋮	⋮	⋮	⋮
185	97	101	0

The next step is to calculate the final forecast results based on the initial forecast values and the adjusted forecast values. The calculation is based on **Equation (14)** and yields the final forecast results, as shown in **Table 13**.

Table 13. Final Forecast Values

No	Period	Actual Data	Initial forecast values	Adjusted forecast values	Final forecast values
1	January 2009	11,300	*	*	*
2	February 2009	9,925	9,925	-697.4286	9,228.57
3	March 2009	9,800	9,782.5	0	9,782.5
⋮	⋮	⋮	⋮	⋮	⋮
184	April 2024	19,642	19,608.5	0	19,608.5

3.3 Evaluation of Forecast Accuracy

The forecast results for global nickel prices using the automatic clustering fuzzy time series Markov chain method are then visualized and compared with the actual data. The plot comparing actual data, initial forecasts, and final forecasts is shown in **Figure 2**.

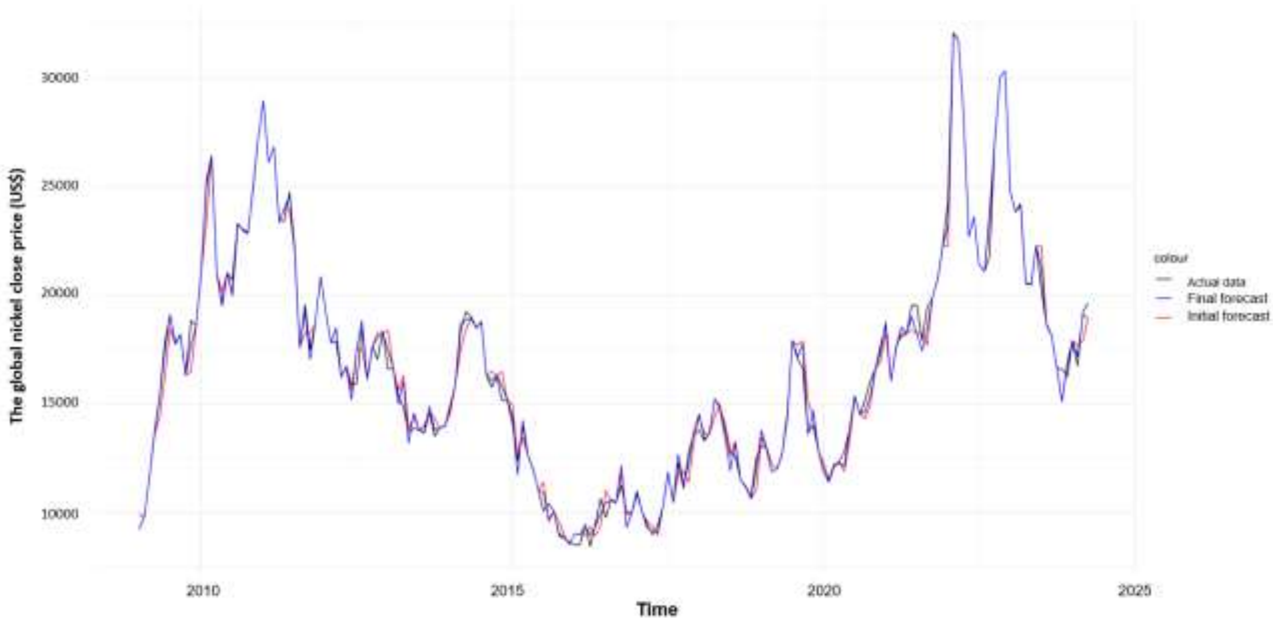

Figure 2. Plot of Actual Data, Initial Forecasts, and Final Forecasts

Figure 2 shows that the forecast results for global nickel prices using the automatic clustering fuzzy time series Markov chain method, both initial and final forecasts, have values close to and patterns that follow the actual data. The accuracy evaluation of global nickel price forecasts using the automatic clustering fuzzy time series Markov chain method is calculated based on the MAPE value using **Equation (15)**. The obtained MAPE values are shown in **Table 14**.

Table 14. Evaluation of Forecast Accuracy

forecast values	MAPE
Initial forecast values	2.55%
Final forecast values	1.76%

The analysis results show that all obtained MAPE values are less than 10%, indicating that the automatic clustering fuzzy time series Markov chain method accurately forecasts global nickel prices. The initial forecast was calculated using FLR, FLRG, and the Markov transition probability matrix, resulting in a MAPE of 2.55% (accuracy 97.45%). The final forecast was obtained from a combination of initial and adjustment values (number of state transitions) to achieve more accurate predictions. The final estimates show the smallest MAPE value, 1.76% (accuracy 98.24%), indicating that three-stage defuzzification with adjustment values can enhance forecasting accuracy. Forecast the global nickel price for the next period, i.e., June 2024 (t), where the fuzzification in May 2024 (t-1) is at A_{101} which has no transition ($A_{101} \rightarrow \emptyset$). Thus, the calculation of the global nickel price forecast for June 2024 (t) is $\hat{Y}'(186) = m_{101}$ = 19,608.5.

4. CONCLUSIONS

The global nickel futures market from January 2009 to May 2024 shows dynamic patterns influenced by market forces, commodity availability, exchange rates, and global economic and political conditions. Implementing the automatic clustering fuzzy time series Markov chain method forecasts global nickel prices effectively. Automatic clustering enhances the model's ability to detect patterns within the data, resulting in more precise and accurate forecasts. At the same time, the Markov chain effectively handles the uncertainty and probabilistic nature of time series data. This effectiveness is demonstrated by a prediction error based on a MAPE value of 1.76%, indicating a prediction accuracy of 98.24%. Therefore, this method is highly suitable for forecasting global nickel prices. The forecasted global nickel price for June 2024 using the automatic clustering fuzzy time series Markov chain method is US$ 19,608.5.

REFERENCES

[1] M. Pouresmaieli, M. Ataei, A. Nouri Qarahasanlou, and A. Barabadi, "INTEGRATION OF RENEWABLE ENERGY AND SUSTAINABLE DEVELOPMENT WITH STRATEGIC PLANNING IN THE MINING INDUSTRY," *Results Eng.*, vol. 20, no. August, p. 101412, 2023, doi: 10.1016/j.rineng.2023.101412.

[2] S. Salinita and A. Nugroho, "LATERITIC NICKEL ORE MODELLING FOR RESERVES ESTIMATION AT PT. ANUGERAH TOMPIRA NIKEL IN THE MASAMA AREA, BANGGAI REGENCY," *J. Teknol. Miner. dan Batubara*, vol. 10, no. 2, pp. 54–68, 2014, doi: https://doi.org/10.30556/jtmb.Vol10.No2.2014.737.

[3] D. D. Radhica and R. A. A. Wibisana, "PROTEKSIONISME NIKEL INDONESIA DALAM PERDAGANGAN DUNIA," *Cendekia Niaga*, vol. 7, no. 1, pp. 74–84, 2023, doi: 10.52391/jcn.v7i1.821.

[4] Kementerian Perdagangan Republik Indonesia, "TRADE POST: TRANSISI ENERGI DAN PENINGKATAN PERMINTAAN KOMODITAS LOGAM SERTA POSISI INDONESIA SEBAGAI PRODUSEN NIKEL DUNIA," *Badan Kebijakan Perdagangan*, vol. II, pp. 34–39, 2023. [Online]. Available: https://bkperdag.kemendag.go.id/media_content/2023/06/tradepost__20230627112551tradeposte-magazineedisijuni2023.pdf

[5] Aprisal and A. M. Abadi, "FUZZY SISTEM: ESTIMASI HARGA NIKEL DUNIA," *JUMLAHKU J. Mat. Ilm. STKIP Muhammadiyah Kuningan*, vol. 5, no. 1, pp. 48–58, 2019, doi: https://doi.org/10.33222/jumlahku.v5i1.583.

[6] Nurbaiti, M. S. Boedoyo, and P. Yusgiantoro, "PENGELOLAAN NIKEL INDONESIA TERHADAP PERTAHANAN NEGARA DAN KETAHANAN ENERGI," *J. Ketahanan Energi*, vol. 8, no. 2, pp. 33–52, 2022, [Online]. Available: https://jurnalprodi.idu.ac.id/index.php/KE/article/view/4465

[7] S. Swarup and G. S. Kushwaha, "NICKEL AND COBALT PRICE VOLATILITY FORECASTING USING A SELF-ATTENTION-BASED TRANSFORMER MODEL," *Appl. Sci.*, vol. 13, no. 8, 2023, doi: 10.3390/app13085072.

[8] E. N. Sofiyanti, S. Ulinuha, R. Okiyanto, M. Al Haris, and R. Wasono, "PERAMALAN HARGA EMAS MENGGUNAKAN METODE FUZZY TIME SERIES CHEN DALAM INVESTASI UNTUK MEMINIMALISIR RISKO," *J. Math. Comput. Stat.*, vol. 7, no. 1, pp. 55–66, 2024, doi: https://doi.org/10.35580/jmathcos.v7i1.1955.

[9] R. Wasono, Y. Fitri, and M. Al Haris, "FORECASTING THE NUMBER OF AIRPLANE PASSENGERS USING HOLT

WINTER'S EXPONENTIAL SMOOTHING METHOD AND EXTREME LEARNING MACHINE METHOD," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 18, no. 1, pp. 0427–0436, 2024, doi: 10.30598/barekengvol18iss1pp0427-0436.

[10] I. F. Amri, W. Sari, V. A. Widyasari, N. Nurohmah, and M. Al Haris, "THE ARIMA-GARCH METHOD IN CASE STUDY FORECASTING THE DAILY STOCK PRICE INDEX OF PT. JASA MARGA (PERSERO)," *Eig. Math. J.*, vol. 7, no. 1, pp. 25–33, 2024, doi: 10.29303/emj.v7i1.174.

[11] Y. Ekananta, L. Muflikhah, and C. Dewi, "PENERAPAN METODE AVERAGE-BASED FUZZY TIME SERIES UNTUK PREDIKSI KONSUMSI ENERGI LISTRIK INDONESIA," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 3, pp. 1283–1289, 2018, [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/1126

[12] A. R. Siregar, R. F. Sari, and R. Widyasari, "FORECASTING THE NUMBER OF COVID-19 SUFFERERS IN NORTH SUMATRA USING THE AUTOMATIC CLUSTERING FUZZY TIME SERIES MARKOV CHAIN METHOD," *J. Math. Sci. Comput. with Appl.*, vol. 2, no. 1, pp. 46–54, 2021, [Online]. Available: https://pcijournal.org/index.php/jmscowa/article/view/48

[13] Y. P. Anggodo and W. F. Mahmudy, "PERAMALAN BUTUHAN HIDUP MINIMUM MENGGUNAKAN AUTOMATIC CLUSTERING DAN FUZZY LOGICAL RELATIONSHIP," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 2, p. 94, 2016, doi: 10.25126/jtiik.201632202.

[14] Y. Alyousifi, M. Othman, A. Husin, and U. Rathnayake, "A NEW HYBRID FUZZY TIME SERIES MODEL WITH AN APPLICATION TO PREDICT PM10 CONCENTRATION," *Ecotoxicol. Environ. Saf.*, vol. 227, p. 112875, 2021, doi: 10.1016/j.ecoenv.2021.112875.

[15] E. Haryono, A. Widodo, and S. Abusini, "KAJIAN MODEL AUTOMATIC CLUSTERING-FUZZY TIME SERIES-MARKOV CHAIN DALAM MEMPREDIKSI DATA HISTORIS JUMLAH KECELAKAAN LALU LINTAS DI KOTA MALANG," *J. Sains Dasar*, vol. 2, no. 1, pp. 63–71, 2013, doi: 10.21831/jsd.v2i1.3365.

[16] I. N. Hidayati, M. Al Haris, and T. W. Utami, "METODE AVERAGE BASED FUZZY TIME SERIES MARKOV CHAIN PADA DATA LAJU INFLASI DI INDONESIA AVERAGE," in *Prosiding Seminar Nasional UNIMUS*, 2022, pp. 580–597. [Online]. Available: https://prosiding.unimus.ac.id/index.php/semnas/article/view/1211/1208

[17] S. Damayanti, S. Yosmar, and N. Afandi, "IMPLEMENTATION OF FUZZY TIME SERIES CHEN FOR FORECASTING INDONESIAN OIL AND GAS IMPORTS VALUE," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 17, no. 2, pp. 0685–0694, 2023, doi: 10.30598/barekengvol17iss2pp0685-0694.

[18] S. Damayanti, J. Rizal, S. Yosmar, N. Afandi, and V. Acnesya, "EARTHQUAKE FREQUENCY DATA MODELING IN MENTAWAI USING FUZZY TIME SERIES LEE AND FUZZY TIME SERIES TSAUR," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 18, no. 1, pp. 0281–0294, 2024, doi: 10.30598/barekengvol18iss1pp0281-0294.

[19] D. Devianto, K. Ramadani, Maiyastri, Y. Asdi, and M. Yollanda, "THE HYBRID MODEL OF AUTOREGRESSIVE INTEGRATED MOVING AVERAGE AND FUZZY TIME SERIES MARKOV CHAIN ON LONG-MEMORY DATA," *Front. Appl. Math. Stat.*, vol. 8, 2022, doi: 10.3389/fams.2022.1045241.

[20] X. Zeng, L. Shu, and J. Jiang, "FUZZY TIME SERIES FORECASTING BASED ON GREY MODEL AND MARKOV CHAIN," *IAENG Int. J. Appl. Math.*, vol. 46, no. 4, pp. 464–472, 2016, [Online]. Available: https://www.iaeng.org/IJAM/issues_v46/issue_4/IJAM_46_4_08.pdf

[21] R. A. Kafi, Y. R. Safitri, Y. Widyaningsih, and B. D. Handari, "COMPARISON OF WEIGHTED MARKOV CHAIN AND FUZZY TIME SERIES MARKOV CHAIN IN FORECASTING STOCK CLOSING PRICE OF COMPANY X," in *AIP Conference Proceedings*, 2019. doi: 10.1063/1.5132460.

[22] A. B. Salsabila, F. Firdaniza, B. N. Ruchjana, and A. S. Abdullah, "PYTHON SCRIPT FUZZY TIME SERIES MARKOV CHAIN MODEL FOR FORECASTING THE NUMBER OF DISEASES COCOA PLANT IN BENDUNGAN DISTRICT," *Int. J. Data Netw. Sci.*, vol. 7, no. 2, pp. 627–636, 2023, doi: 10.5267/j.ijdns.2023.3.009.

[23] N. Didenko, "MODELING THE GLOBAL NICKEL MARKET WITH A TRIANGULAR SIMULTANEOUS EQUATIONS MODEL," *Int. J. Syst. Assur. Eng. Manag.*, vol. 11, no. S1, pp. 119–129, May 2020, doi: 10.1007/s13198-019-00936-0.