Gold Price Prediction Using the Fuzzy Time Series Saxena-Easo Method

Anjas Fernando, Riri Syafitri Lubis, Hendra Cipta

Mathematics Study Program, Faculty of Science and Technology, Universitas Islam Negeri Sumatera Utara, Indonesia

A B S T R A C T

Gold is an investment that has the smallest risk compared to other methods other investments , but every time the price of gold always fluctuates. This fluctuation will make it difficult for investors, the government, or those who need gold price data to see how the prospects for gold investment are going forward. To overcome this, a prediction or forecast is carried out . One of the forecasting methods developed with *fuzzy theory* is the Saxena-Easo *fuzzy time series method* . In this study , this method will be used to predict gold prices in the period 17 August 2021 to 31 December 2021. The purpose of this study was to determine the value of gold prices for the period 17 August 2021 to 31 December 2021 and to determine the accuracy of the Saxena Easo fuzzy time series method. While the parameter used to measure forecasting accuracy is MAPE, if this parameter is in the range of 10% to 20% then the forecasting result is good, and if the parameter is below 10% then the forecast result is very good. The results of forecasting gold prices in the period 17 August – 31 December 2021 have a value that tends to increase and the MAPE obtained is 0.024277% Therefore, it can be said that the prediction results for certain parameters are very good*Keywords:*

Fertility, Contraception, CHAID Method

.

Corresponding Author:

Anjas Fernando,
Mathematics Study Program, Faculty of Science and Technology, State Islamic University of North Sumatra.
Email: anjasfernando24@gmail.com

1. INTRODUCTION

The word investment comes from the English "investment" which means investment. In Arabic, investment is called istitmar which means "to make fruitful", grow and increase in number. Meanwhile, according to capital market terms, investment is the activity of allocating or investing resources (resources) today with the hope that they will be profitable in the future [1] .

. There are many ways that can be done for people who have never made an investment in order to start investing, such as starting a business, stocks, bonds, mutual funds, *real estate* that brings income, saving gold, and others. Among the various ways to make financial investments, saving with gold (precious metal) is easier to do and has the least risk. Gold is an investment that has the least risk of other investment methods such as mutual funds, starting a business, bonds, and others. This is what makes gold the best solution for potential novice investors who want to start but are still unsure where to start to invest.

Seen in a long period of time, the value of gold always rises from time to time, automatically to save gold treasures provide benefits because it is free of inflation. But investing in gold is not without risk. The thing that should be understood from the start is the fact that the price of gold can increase or decrease.

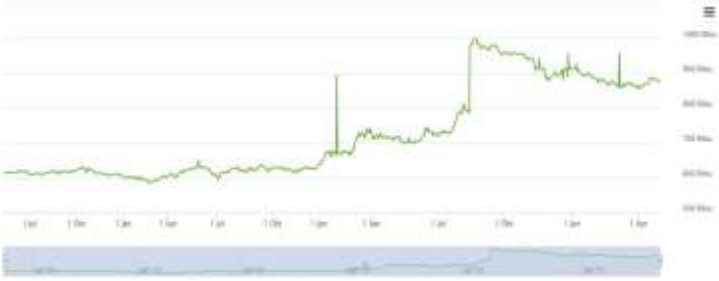

Figure 1 Graph of Gold Price Fluctuations in Rupiah
Source : https://www.pegadaian.co.id/harga

The picture above shows that the price of gold in the rupiah exchange rate has fluctuated over the last 3 years. The fluctuations were very clear at the beginning of 2020 where the price of gold experienced an increase and decrease in a short time, then in the middle of 2020 where the price of gold experienced a very drastic increase.

Uncertain price movements must be communicated to various parties, such as governments, investors, and other entities, so that decisions can be made quickly. This rapid request for information makes forecasting a way to help the parties involved in making better decisions .

The time series data type contains gold price data. Time series data is sequential data over time. For this reason, gold price predictions or forecasts are accurate enough to be profitable according to plan. Various methods for predicting time series have been proposed and used. One of the methods commonly used today is the fuzzy time series model. The fuzzy time series itself was first introduced by Song and Chissom (1993). In this study, the fuzzy time series prediction method is based on the concept of fuzzy logic and is used to predict new student admissions at the University of Alabama.

Research conducted by Meredith Stevenson and John E. Porter (2009) in the development of the fuzzy time series method by modifying the determination of the entire universe converted to a percentage change form, added by Saxena & Easo (2012) who modified the fuzzy time series from Stevenson & Porter (2009) to define fuzzy intervals based on the division of the number of frequencies from each original interval so that it becomes several subintervals. Saxena & Easo (2012) used data from students who attended the University of Alabama from 1971 to 1992. The purpose of this study was to improve the accuracy of the pre-fuzzy time series prediction method, which was done by determining fuzzy intervals based on the frequency division of each initial interval. . This result is lower than other MSE and AFER methods. Therefore, it is concluded that the SaxenaEaso fuzzy time series method has a better level of accuracy [2s] .

Therefore, this study was conducted to calculate the accuracy of the Saxena-Easo *fuzzy time series method* in the movement of gold price data using MAPE . With the hope that this method can provide a small *error value* so that it can be called accurate and feasible in predicting the price of gold. So researchers are interested in conducting a study with the title **" Gold Price Prediction Using the Saxena-Easo *Fuzzy Time Series Method* ".**

2. RESEARCH METHODOLOGY

The type of research used in this research is quantitative research using the Saxena-Easo fuzzy time series method.

a. Types of Research and Data Sources

The type of research used in this study is quantitative research using the Saxena - Easo fuzzy time series method. In this study, the types of data are secondary data, namely data obtained by an organization or company in complete form in the form of publications, and time series, namely data collected from time to time to explain growth or development. [3] . Sources of data obtained through PT. Gallery 24 Pawnshops. Secondary data is in the form of a daily gold price list for the period 2018-2021.

b. Research variable

The variables in this study are gold price data which is converted into rupiah.

c. Research procedure

The procedure for measuring the accuracy of the Saxena-Easo fuzzy time series method is as follows:

1. Collect data,
2. Determine the speaker universe
3. Define the universal set U = [Dmin,Dmax]
4. Fuzzification of data,

According to Wang (1997: 105), fuzzyization is defined as the mapping of a classical set to a *fuzzy set* . The *fuzzy* set that will be used in the previous *defuzzy fication stage is* averaged using the following formula.

$$a_t = \frac{LL_i + UL_i}{2}$$

5. Data defuzzification. This step is the opposite of the *fuzzy fication step* . In this defuzzification step , if the input process is a *fuzzy set* within a certain range, it can be considered as a crisp value . Defuzzification of data was carried out using a modified fuzzy time series forecasting formula from Meredith Stevenson and John E. Porter. Defuzzification of data used the following formula :

$$t_j = \begin{cases} \dfrac{1{,}5}{\dfrac{1}{a_1} + \dfrac{0{,}5}{a_2}}, if\ j = 1 \\ \dfrac{2}{\dfrac{0{,}5}{a_{j-1}} + \dfrac{1}{a_j} + \dfrac{0{,}5}{a_{j+1}}}, if\ 2 \le j \le n-1 \\ \dfrac{1{,}5}{\dfrac{0{,}5}{a_{n-1}} + \dfrac{1}{a_n}}, if\ j = n \end{cases}$$

6. Measure the prediction error using the *mean absolute percentage error* (MAPE) to determine the level of accuracy of the prediction method. The model shows very good performance with MAPE 10% and good performance with MAPE 10% and 20%. The formula for calculating MAPE is:

$$PE_t = \left(\frac{Y_t - \hat{Y}_{(t)}}{Y_t} \right) \times 100\ \%$$

.

$$MAPE = \frac{\sum_{t=1}^{n} |PE_t|}{n}$$

Description :
PE_t= percentage of error period
Y_t = actual data period
$\hat{Y}_{(t)}$= forecast value for the period
 = number of predicted data.

3. RESEARCH RESULT

a. Research result

i. Gold Price Data

Table 4.1 Daily Gold Price Data

Date	Gold price
August 16, 2018	Rp.618.000
August 20, 2018	Rp.618.000
August 21, 2018	IDR 619,000
23 August 2018	IDR 612,000
24 August 2018	IDR 612,000
;	;
August 12, 2021	Rp864,000
August 13, 2021	Rp.869.000
August 14, 2021	IDR 870,000
August 15, 2021	IDR 880,000
August 16, 2021	IDR 880,000

ii. **Calculating *Percentage Change***

Change the actual gold price data that has been obtained in the form of a percentage change in historical data .

$$X_t = \left(\frac{X_t - X_{t-1}}{X_{t-1}} \right) \times 100\%$$

Then the *percentage change* of each data is as follows:

$$X_{20\,Agustus\,2018} = \left(\frac{X_{20\,Agustus\,2018} - X_{16\,Agustus\,2018}}{X_{16\,Agustus\,2018}} \right) \times 100\%$$

$$= \left(\frac{618000 - 618000}{618000} \right) \times 100\%$$
$$= 0$$

$$X_{21\,Agustus\,2018} = \left(\frac{X_{21\,Agustus\,2018} - X_{20\,Agustus\,2018}}{X_{20\,Agustus\,2018}} \right) \times 100\%$$
$$= \left(\frac{619000 - 618000}{618000} \right) \times 100\%$$
$$= 0.161812298$$

;

$$X_{16\,Agustus\,2021} = \left(\frac{X_{16\,Agustus\,2021} - X_{15\,Agustus\,2021}}{X_{15\,Agustus\,2021}} \right) \times 100\%$$
$$= \left(\frac{880000 - 880000}{880000} \right) \times 100\%$$
$$= 0$$

The actual data of the gold price which is transformed in the form of percentage change. The results of the calculation of the smallest percentage change (minimum) of -24,972 and the largest percentage change (maximum) of 34,89426. These limits are used to form a universal set U, namely the lower and upper bounds. The resulting minimum and maximum values are used to define the universe before making predictions. The universal set U can be written as follows:

$$U = [-24.972, 34.89426]$$

iii. Defining the Set of Semes t a

To determine the universal set, first find the number of intervals using the following equation.

$$K = 1 + 3,322 \log n$$
$$K = 1 + 3,322 \log(928)$$
$$K = 1 + 3,322(2,9675479762)$$
$$K = 1 + 9,8581943769$$
$$K = 10,8582 \approx 11$$
$$U = [D_{\min}, D_{\max}]$$
$$U = [-24,972; 34,89426]$$

then the length of the class interval is as follows:

$$R = [D_{\max} - D_{\min}]$$
$$R = [34.89426 - (-24.972)]$$
$$R = 59.86626$$
$$P = \frac{R}{K}$$
$$P = \frac{59.86626}{11} = 5.442388$$

Table 4. 2 Division of the Speaker Universe

interval	Lower limit	Upper limit
a_1	-24,972	-19.529612
a_2	-19.529612	-14.087224
a_3	-14.087224	-8,644836
a_4	-8,644836	-3,202448
a_5	-3,202448	2.23994
a_6	2.23994	7.682328

a_7	7.682328	13.124716
a_8	13.124716	18.567104
a_9	18.567104	24.009492
a_{10}	24.009492	29,45188
a_{11}	29,45188	34.894268

Then according to the distribution of the universe of the speaker above, the division is carried out according to the amount of data that will be divided into several sub intervals, later the sub interval will divide to find the width of the sub interval. The following is the distribution of the universe of speakers which is presented in table 4.3 below:

Table 4. 3 Speaker Universe Repartition

Hose to	interval	Amount of data	Number of Sub-intervals	Sub-interval Width
a_1	[-24,972;-19,529612]	1	1	5.442388
a_2	[-19.529612;-14.087224]	0	0	0
a_3	[-14,087224;-8,644836]	1	1	5,442388
a_4	[-8,644836;-3,202448]	2	2	2,721194
a_5	[-3,202448;2,23994]	913	4	1,360597
a_6	[2,23994;7,682328]	8	3	1,814129333
a_7	[7,682328;13,124716]	1	1	5,442388
a_8	[13,124716;18,567104]	0	0	0
a_9	[18,567104;24,009492]	0	0	0
a_{10}	[24,009492;29,45188]	1	1	5.442388
a_{11}	[29,45188;34,894268]	1	1	5.442388

Based on the table above, it can be shown that there are 4 different frequencies or percentage changes in data, namely 913, 8, 2 and 1. Thus, the interval with the highest first frequency, namely 913 is divided into 4 sub-intervals with the 2nd highest frequency. i.e. 8 will be divided into 3 equal sub-intervals, 3rd highest frequency interval, i.e. 2 will be divided into 2 equal sub-intervals, the 4th highest frequency interval will be divided into equal sub-intervals. Finally, there are 14 subintervals that will be the domain of the fuzzy set formed.

iv. ***Fuzzification***

After taking the interval from the fuzzy set, the next step is to determine the middle value of each interval.

Table 4.4 Results of Repartitioning the Speaker Universe

Fuzzy Set	Lower limit	Upper limit	The midpoint
A1	-24,972	-19,52961	-22.250806
A2	-19.529612	-14.08722	-16.808418
A3	-14.087224	-11,36603	-12.726627
A4	-11,36603	-8,644836	-10.005433
A5	-8,644836	-7.284239	-7.9645375
A6	-7.284239	-5.923642	-6.6039405
A7	-5.923642	-4.563045	-5.2433435

A8	-4.563045	-3,202448	-3.8827465
A9	-3,202448	-1.388319	-2,295383334
A10	- 1,38831866	0.425811	-0.481254
A11	0,425810666	2.23994	1,332875333
A12	2,239939999	7,682328	4.961133999
A13	7.682327999	13,12472	10.403522
A14	13.124716	18.5671	15.84591

The median value is the median value of the interval for each interval. This average will be used in the *defuzzification step* . After knowing the fuzzy set, do *the fuzzyfication* on the percentage change data.

Table 4.5 *Fuzzyfication* results based on *Percentage Change*

Time Data	Gold price	PerChange	Fuzzy
16/08/18	618000	0	-
20/08/18	618000	0	A10
21/08/18	619000	0,161812298	A10
23/08/18	612000	-1,13085622	A10
24/08/18	612000	0	A10
⋮	⋮	⋮	⋮
12/08/21	864000	0	A10
13/08/21	869000	0,578703704	A11
14/08/21	870000	0.115074799	A10
15/08/21	880000	1.149425287	A11
16/08/21	880000	0	A10

v. *Defuzzification*

1. Estimating Percentage Change in Data

Then calculate the average of each class interval, once obtained, then the midpoint is used to predict the percentage change with the triangular membership function using the data in Table 4.4

For example, if you want to predict the percentage change in data in *Fuzzy* $A_j = 1$, it can be calculated as follows:

$$t_j = \frac{1,5}{\frac{1}{a_1} + \frac{0,5}{a_2}}, \text{ for } j = 1$$

$A_1 = a_1$

$$t_1 = \frac{1,5}{\frac{1}{a_1} + \frac{0,5}{a_2}} = \frac{1,5}{\frac{1}{-22,250806} + \frac{0,5}{-16,808418}} = \frac{1,5}{-0,07469}$$

$t_1 = -20,832$

To find the predictive value of the percentage change in data on *fuzzy* $A_j = 2 \le j \le 14-1$ its function is as follows:

$$t_j = \frac{2}{\frac{0,5}{a_{j-1}} + \frac{1}{a_j} + \frac{0,5}{a_{j+1}}}, \text{ for } 2 \le j \le 14-1$$

$A_2 = a_2$

$$t_2 = \frac{2}{\frac{0,5}{a_1}+\frac{1}{a_2}+\frac{0,5}{a_3}} = \frac{2}{\frac{0,5}{a_1}+\frac{1}{a_2}+\frac{0,5}{a_3}}$$

$$t_2 = \frac{2}{\frac{0,5}{-22,250806}+\frac{1}{-16,808418}+\frac{0,5}{-12,726627}} = \frac{2}{-0,12125}$$

$$t_2 = 16,4945$$

As for finding the predictive value of the last data change percentage, namely, *fuzzy* A_j = 14, it is as follows:

$$t_j = \frac{1,5}{\frac{0,5}{a_{n-1}}+\frac{1}{a_n}}, \text{ for } j = 14$$

$$A_{14} = a_{14}$$

$$t_{14} = \frac{1,5}{\frac{0,5}{a_{14-1}}+\frac{1}{a_{14}}} = \frac{1,5}{\frac{0,5}{10,403522}+\frac{1}{15,84591}} = \frac{1,5}{0,111168}$$

$$t_{14} = 13,49304$$

2. Determining the Value of Forecasting Data

After all the data percentage change values are obtained, the next step is to predict the value of the th forecasting data i using the predicted percentage change data values (t_j) using the following equation:

$$F_{(i)} = \left(\frac{t_j}{100}.X_{t-1}\right) + X_{t-1} \quad (2 \le i \le 928 \text{ and } 2 \le t \le 928)$$

3. Predicting Gold Prices based on Forecasting Data Value

The price of gold in the future is predicted with the help of Microsoft Excel using the "TREND" formula based on the forecast data value that has been obtained.

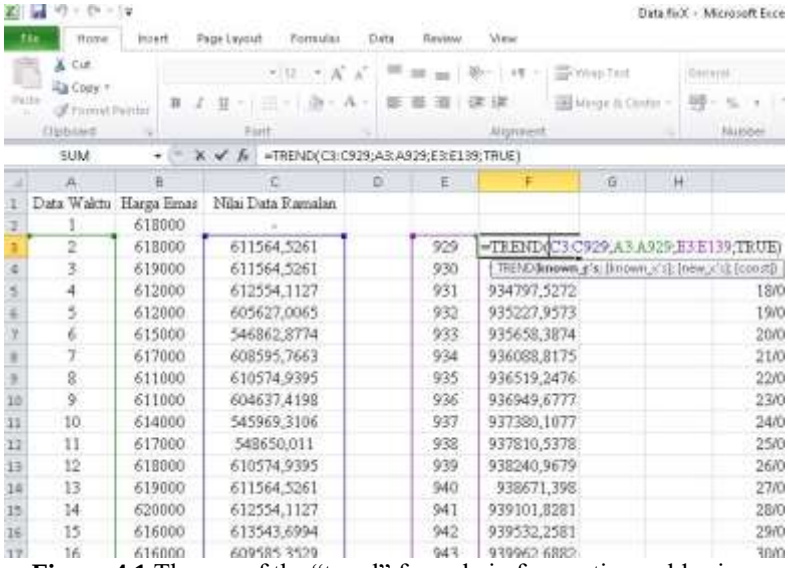

Figure 4.1 The use of the "trend" formula in forecasting gold prices

Based on the picture above, the price of gold after n periods is predicted using the "trend" formula found in Microsoft excel, so the results of forecasting gold prices until the end of 2021 are as follows.

Table 4. 6 Gold Price Forecast Results from August 17, 2021 to the end of 2021

Time Data	Forecasting Price
17/08/2021	933937
18/08/2021	934367
19/08/2021	934768
20/08/2021	935228
21/08/2021	935658
;	;
27/12/2021	990753
28/12/2021	991184
29/12/2021	991614
30/12/2021	992045
31/12/2021	992475

vi. Forecasting Error Measurement

The measurement of forecasting error used to determine how well the forecasting results obtained is *Mean Absolute Percentage Error* (MAPE). *Error* value of forecasting can be calculated using the following equation:

$$MAPE = \frac{\sum_{t=1}^{n} |PE_t|}{n}$$

with , $PE_t = \left(\frac{Y_t - \hat{Y}_{(t)}}{Y_t} \right) \times 100\ \%$

$$PE_2 = \left(\frac{618000 - 611564,5}{618000} \right) \times 100\ \% = 0,010413\%$$

$$PE_3 = \left(\frac{619000 - 611564,5}{619000} \right) \times 100\ \% = 0,012012\%$$

$$PE_4 = \left(\frac{612000 - 612554,1}{612000} \right) \times 100\ \% = -0,00091\%$$

;

$$PE_{928} = \left(\frac{880000 - 870836,2}{880000} \right) \times 100\ \% = 0,010413\%$$

So, $MAPE = \frac{\sum_{t=1}^{n} |PE_t|}{n} = \frac{22,52927}{928}\%$

$$MAPE = 0,024277\%$$

From the results of forecasting, the results obtained from MAPE are 0.024277%

4. CONCLUSION

Forecasting the gold price after the nth period until the end of 2021, the forecast data produced has a value that tends to increase, even almost touching the 1,000,000 figure. This has a great opportunity to happen because based on historical facts from previous years, the trend of gold prices is always rising.

Based on the research results, the Saxena-Easo fuzzy time series method provides excellent forecasting results for gold prices. This can be seen in actual data, with forecast data having prices that are not much different. By using the Saxena-Easo fuzzy time series method, the accuracy of this method also gives a MAPE value of 0.024277%, this means that the Saxena-Easo fuzzy time series method has a very good level of accuracy.

REFERENCES

[1] Sakinah, "Investasi dalam Islam," *Iqtishadia*, vol. 1, no. 2, pp. 248–262, 2014.
[2] M. Stevenson and J. Porter, "Fuzzy time series forecasting using percentage change as the universe of discourse," *Change*, vol. 1971, no. 3.89, pp. 464–467, 1972.
[3] P. Saxena and S. Easo, "Forecasting enrollments based on fuzzy time series with Higher Forecast Accuracy Rate," *Int. J. Comput. Technol. Appl.*, vol. 3, no. 6, pp. 2033–2037, 2012.
[4] L. C. Ramadhani, D. Anggraeni, and A. Kamsyakawuni, "Fuzzy Time Series Saxena-Easo Pada Peramalan Laju Inflasi Indonesia Saxena-Easo," *J. Ilmu Dasar*, vol. 20, no. 1, pp. 53–60, 2019.
[5] Supranto, *Metode Ramalan Kuantitatif untuk Perencanaan Ekonomi dan Bisnis*. Jakarta: PT. Rineka Cipta, 2000.