

A Comparative Study Between Holt's Double Exponential Smoothing and Fuzzy Time Series Markov Chain in Gold Price Forecasting

Jasmani Bidin¹, Sharifah Fhahriyah Syed Abas^{2*}, Noorzila Sharif³,
Che Afif Azhan Che Muhammad Fahimi⁴, Ku Azlina Ku Akil⁵

^{1,2,3,4,5} Faculty of Computer & Mathematical Sciences,
Universiti Teknologi MARA Perlis Branch, Arau Campus, 02600 Arau, Perlis, Malaysia

Corresponding author: * sfhahriyah@uitm.edu.my

Received Date: 10 August 2022

Accepted Date: 1 September 2022

Published Date: 30 September 2022

HIGHLIGHTS

- Fuzzy Time Series Markov Chain and Holt's Double Exponential Smoothing are only two of the forecasting models applied and compared in this case study.
- Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) are used to compare the performance of the two models in predicting gold price.
- Fuzzy Time Series Markov Chain is found to be a better method in predicting gold price as it produces smaller values of RMSE and MAPE.
- It is advisable for future researchers to expand the study by using other forecasting techniques in order to find the best model in predicting gold price and it may assist the investors for better planning.

ABSTRACT

Gold price is important to a country's economy as it can be used as a hedge against inflation especially during financial turmoil. Besides, the gold price also has an impact on the stock market price. As an investor, to make a good investment plan, information regarding the fluctuation price of gold is necessary to minimize the risk. Therefore, this study proposes to compare two of the forecasting models, namely Holt's Double Exponential Smoothing and Fuzzy Time Series Markov Chain to forecast the price of gold. Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) are used to determine a better forecasting model with smaller error. Initially, the data price of gold is analysed by using Durbin Watson Test to check the suitability of the data for time series analysis. The finding of this study shows that Fuzzy Time Series Markov Chain is more accurate in predicting gold price as compared to Holt's Double Exponential Smoothing because it produces smaller values of RMSE and MAPE.

Keywords: Gold Price, Holt's Double Exponential Smoothing, Fuzzy Time Series Markov Chain, RMSE, MAPE

INTRODUCTION

Gold has been used as money for exchange purposes. Therefore, each piece of gold has its value in the monetary system. However, the value of gold itself has dominant power over the economy as a tool in the monetary system. In the nineteenth and twentieth centuries, gold acted as a cash guarantee for issue for



banknotes (Uzun & Kirai, 2017). Later, in Malaysia, many people buy gold either for savings or long-term investment instruments because the price of gold itself will either increase or decrease over time. Thus, buying gold is not just for accessory but also as a medium to gain some income. Most people who know how good gold can be in terms of investment will buy a gold bar because the purity level of the gold can affect the selling value of the gold. The purity of the gold is measured by karat, which is pure gold with 24 karats.

Two main factors contribute to determine the price of gold: which are demand and supply. These factors control the fluctuation of gold price. Other factors that influence the change in the price of gold are inflation rates and currency variation. The fear and uncertainty in the global economy will affect the swing of gold price, turning it to be the most attractive asset for all investors. In other words, the price of gold is the mirror of the world economic situation (Ghalayini & Farhat, 2020).

A good forecasting model will produce better forecasting output with minimal error. It may assist the investors to make better decision making. For example, the forecasting of return on gold price for a shorter period of time can help provide valuable information to an investor about the movement of gold price for short and long-term buying and selling strategies. The volatility of gold prices can be predicted more precisely, and it is beneficial for commodity markets and the global economy (Uzun & Kiral, 2017).

Gold price plays an important role in country's economies as it is used as hedging tool against inflation. It is a type of asset that is negatively correlated with another asset or portfolio. Adding a certain percentage of gold in investment portfolio may assist in decreasing the level of risk during financial turmoil since it is not affected by Consumer Price Index (Shakil et al., 2018). In a study of S&P500 stock market index (GSPC), the gold price has the highest impact on the stock market price in long-run and short-run, compared to other variables such as oil price. As an implication, investors should react against changes in the gold price (Gokmenoglu & Fazlollahi, 2015). Due to that reason, the prediction of gold price accurately is important for investors, portfolio managers and policy makers. The prospective investors should consider gold in their portfolios as a store of value and a diversification tool and cautious of the price fluctuation (Chaku et al., 2022).

Many studies have been conducted to forecast gold price and various forecasting models have been used to find the best model. In a study done by Taufik (2020), Fuzzy Time Series Markov Chain is the best forecasting model as compared to Brown's Double Exponential Smoothing and Holt's Double Exponential. It is due to smaller values of the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE). Besides, Abdelkader (2020) compared two forecasting methods namely Fuzzy Auto-Regressive Integrated Moving Average (FARIMA) with Auto-Regressive Integrated Moving Average (ARIMA) as a control group to predict the price of gold for the year 2019. Based on Mean Square Error (MSE) and RMSE, the study showed that FARIMA is a better method. Another comparative study done by Suranart (2014) to compare three forecasting methods which are Neural Network (NN), Radial Basis Function Network (RBF) and Support Vector Regression (SVR). The result based on Mean Absolut Deviation (MAD), MSE, Mean Forecast Error (MFE) and MAPE and the result shows that SVR becomes the most reliable forecasting model followed by RBF and NN.

Besides those three comparative studies, there are many other studies related to individual forecasting techniques for gold price. Recently, Chaku et al. (2022) used Seasonal Autoregressive Integrates Moving Average (SARIMA) model to forecast the gold price based on the observation of monthly gold price data from January 2015 until December 2020. This study used the data to fit the seasonal model. The forecast average monthly gold prices are generated from January 2021 to December 2025. The study found that there is no significant difference between the actual gold prices against predicted values. Guha (2016)



applied ARIMA to forecast gold price based on monthly data from November 2003 to January 2014. In this study, six different parameters were chosen to be tested and ARIMA model (1, 1, 1) provided the best model which satisfies all five criteria of fit statistics; RMSE, MAPE, the Mean Absolute Error (MAE), Bayesian information criterion (BIC) and Lungs Box Q statistics. Artificial Neural Network (ANN) was used to anticipate gold prices (Chukwudike et al., 2020). This method has been successfully fitted to the data series chosen. Among those seventeen ANN structures that have been suggested, ANN (2-6-1) was the best structure since it has the least error in MSE and MAE.

This study is only designed to compare between Holt's Double Exponential Smoothing model and Fuzzy Time Series Markov Chain Model in predicting price of gold. To select a better forecast model with least error, MAPE and RMSE are used to analyse the forecast outputs. Initially, the historical data of gold price is analysed using Durbin Watson Test in order to check the suitability of the data for time series analysis. Although this study uses the same method as done by Taufik et al. (2020) but the data used are different in terms of time setting and location. They use daily price of gold in Indonesia from May to July 2020 while this study used monthly price of gold in Malaysia from January 2016 to December 2020. The result cannot be assumed the same. Therefore, this study aims to determine suitable method for gold price prediction in Malaysia.

METHODOLOGY

The models used in this study are Fuzzy Time Series Markov Chain and Holt's Double Exponential Smoothing. The model formulation and calculation for both models' solutions will be discussed throughout this study. Both results will be evaluated using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) to determine a better model. In this study, monthly data of gold prices from January 2016 until December 2020 are collected from Index Mundi website page.

Method of Data Analysis

Before doing further analysis, the collected data must be checked whether there is any missing data. After that, the data should be tested to see the compatibility of the time series analysis applied to the price of gold monthly data. For that reason, the Durbin-Watson Test must be applied using the formula below:

$$DW = 2(1 - p(1)) \quad (1)$$

where $p(1)$ = first-order autocorrelation of the data

If the value of the Durbin Watson is between 0 and 1.5, it shows that the data is considered dependent on time and suitable for time series analysis. The data will then be used to create a model using the Fuzzy Time Series Markov Chain and Holt's Double Exponential Smoothing.

Fuzzy Time Series Markov Chain

Markov Chain method is different from the classic Fuzzy Time Series, and this method has a transition metrics statistic concept in the forecasting calculation. The formula used in this method is shown in equation (2) below:

$$U = [D_{\min} - D_1, D_{\max} + D_2] \quad (2)$$



The first step is to define the universe of discourse U by determining the minimum value (D_{\min}) and maximum value (D_{\max}) for the time series data, x_t . Meanwhile, D_1 and D_2 are any two positive integers to enable the intervals of data are divided appropriately and evenly.

Then, the universe of discourse U will be divided into equal intervals and specify its length as illustrate in equation (3) below.

$$l = \frac{(D_{\max} + D_2) - (D_{\min} - D_1)}{n} \quad (3)$$

where l = the length of an interval

n = the number of intervals

Next step, for each interval, the midpoint will be calculated and labelled as u_1, u_2, u_3 until u_n .

The fuzzy sets A_1 until A_n are defined as represented in equation (4):

$$\begin{aligned} A_1 &= \left[\frac{1}{u_1}, \frac{0.5}{u_2}, \frac{0}{u_3}, \dots, \frac{0}{u_n} \right], \\ A_2 &= \left[\frac{0.5}{u_1}, \frac{1}{u_2}, \frac{0.5}{u_3}, \dots, \frac{0}{u_n} \right] \\ &\vdots \\ A_n &= \left[\frac{0}{u_1}, \frac{0}{u_2}, \dots, \frac{0.5}{u_{n-1}}, \frac{1}{u_n} \right] \end{aligned} \quad (4)$$

Then, each data of gold price, x_t will be fuzzified into a related fuzzy set.

Next, Fuzzy logic relationship (FLR) is defined as $A_i \rightarrow A_j$ if $F(t-1) = A_i$ and $F(t) = A_j$ are the relationship between two consecutive observations. From FLR, Fuzzy Logical Relationship Group (FLRG) is formed. Then, Markov Transition Probability Matrix with the size of $n \times n$ is obtained by using equation (5).

$$P_{ij} = \frac{m_{ij}}{m_i} \quad (5)$$



If $P_{ij} > 0$, the A_j is accessible from state A_i . Hence, we will defuzzify the forecast result from the probability matrix. Finally, the tendency of the forecasting result needs to be adjusted and the calculation of the final output can be performed.

Holt's Double Exponential Smoothing

The equations for Holt's Double Exponential Smoothing are:

$$S_t = \alpha X_t + (1 - \alpha)(S_{t-1} + b_{t-1}) \quad (6)$$

$$b_t = \beta(S_t - S_{t-1}) + (1 - \beta)b_{t-1} \quad (7)$$

$$F_{t+m} = S_t + mb_t \quad (8)$$

Equation (6) is the exponentially smoothed series S_t in which α and β are the parameters to be determined with values from 0 to 1. It is used to calculate the successive exponential smoothed value used in the trend estimate equation. Next, equation (7) is the trend estimate equation which will be calculated by taking the difference between two successive exponential smoothed value, $S_t - S_{t-1}$. The calculation result produces an estimate for a trend in the data.

Furthermore, both equation (6) and (7) will be used to forecast using equation (8). The trend estimate will be multiplied by the number of periods to be forecast, m and the results will then be added to the exponentially smoothed series, S_t to eliminate the randomness from the data. Finally, the forecasting results are generated for the model evaluation to compare both methods.

Model Evaluation Methods

The evaluation of the result from each method is done using RMSE and MAPE, given by equation (9) and (10) below:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y(t) - F(t))^2} \quad (9)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y(t) - F(t)|}{Y(t)} \times 100 \quad (10)$$

The smaller value of RMSE indicates that the model is better. As for MAPE, if the result is less than 10 percent, it is of excellent significance level. If the result is 10 percent until 20 percent, it is of good significance level. Next, if the result is 20 percent until 50 percent, it is moderate significance level. While for low significance level, the result is greater than 50 percent.

FINDINGS AND DISCUSSIONS



The data are analysed using Holt's Double Exponential Smoothing and Fuzzy Time Series Markov Chain. Those two models were tested using Durbin Watson Test in equation (1) and the value of first-order autocorrelation, p is 0.982299492 and simulated using Microsoft Excel. The calculation of result obtained is as follows:

$$\begin{aligned} DW &= 2(1 - p(1)) \\ &= 2(1 - 0.982299492) \\ &= 0.035401 \end{aligned}$$

The result above shows that the data is suitable for time series analysis because the value of Durbin Watson's, $DW = 0.035401$ lies between 0 and 1.5. Thus, it indicates that the data is dependent on time.

For a start in applying the Fuzzy Time Series Markov Chain, the calculation to find the universe of discourse, U is as follows:

$$\begin{aligned} U &= [D_{\min} - D_1, D_{\max} + D_2] \\ &= [4770.08 - 270, 8249.85 + 251] \\ &= [4500.08, 8500.85] \end{aligned}$$

So, the intervals for the Fuzzy sets will start at 4500.08 and end at 8500.85. In order to determine the number of intervals, the Sturges formula is being applied as follows:

$$\begin{aligned} k &= 1 + 3.322 \log n \\ &= 1 + 3.322 \log 60 \approx 7 \end{aligned}$$

where k = number of classes and n = total of classes data

Then, the fuzzy set of seven intervals are generated as shown in Table 1.

Table 1: Fuzzy Sets Intervals

Fuzzy Set	Interval		
A_1	4500.08	-	5,071.62
A_2	5071.62	-	5,643.16
A_3	5643.16	-	6,214.70
A_4	6214.70	-	6,786.23
A_5	6786.23	-	7,357.77
A_6	7357.77	-	7,929.31
A_7	7929.31	-	8,500.85

Table 2 shows that each historical data of gold price is fuzzified and fuzzy logical relation.

Table 2: Fuzzification and fuzzy logical relation (FLR)



Date	Price per Troy ounce (RM)	Fuzzification	FLR		
Jan-16	4770.08	A1	N/A	-	A1
Feb-16	5023.91	A1	A1	-	A1
Mar-16	5084.95	A2	A1	-	A2
Apr-16	4850.36	A1	A2	-	A1
May-16	5100.76	A2	A1	-	A2
Jun-16	5220.86	A2	A2	-	A2
Jul-16	5376.05	A2	A2	-	A2
⋮	⋮	⋮	⋮	⋮	⋮

Fuzzy logical relation group (FLRG) is created and the result is shown in Table 3.

Table 3: Table for Fuzzy Logical Relationship Group (FLRG)

FLRG	
A1	→ A1, A2
A2	→ A1, A2, A3
A3	→ A3, A4
A4	→ A3, A4, A5
A5	→ A5, A6
A6	→ A6, A7
A7	→ A6, A7

Using Equation (5), the Markov Transition Probability Matrix with the size of 7×7 is obtained, shown in Table 5.

Table 5: The Markov Transition Matrix

	A1	A2	A3	A4	A5	A6	A7
A1	4/7	3/7	0	0	0	0	0
A2	1/17	31/34	1/34	0	0	0	0
A3	0	0	1/2	1/2	0	0	0
A4	0	0	1/5	3/5	1/5	0	0
A5	0	0	0	0	1/2	1/2	0
A6	0	0	0	0	0	4/5	1/5
A7	0	0	0	0	0	1/2	1/2

Finally, the forecast value is calculated and the result is represented in Table 6.



Table 6: The result for Fuzzy Logical Relationship Group value and forecasting value

Date	Value of FLRG	Forecast
Jan-16	0	0
Feb-16	5021.783	5021.783
Mar-16	5166.829	5452.598
Apr-16	5092.179	4806.409
May-16	5067.658	5353.427
Jun-16	5106.594	5106.594
...
Jul-20	7567.832	7567.832
Aug-20	7941.448	8227.217
Sep-20	7946.696	7946.696
Oct-20	7810.691	7524.922
Nov-20	7957.12	7957.12
Dec-20	7790.792	7790.792

The first step in calculating Holt's Double Exponential Smoothing is by using the exponentially smoothed series to calculate the successive exponential smoothed value. The equation being used can be referred in equation (6). Then, trend estimate equation is used by taking the difference between two successive exponential smoothed value and the result produces an estimate for a trend in the data. We used Microsoft Excel with Solver for this method to find the alpha and beta values with the minimum RMSE and MAPE.

As a result, the value of α is 1, and the value of β is 0.05235. The forecasting values in Table 7 below shows the result based on those α and β .

Table 7: Table of Forecast value for price of Gold by using Holt's Double Exponential Smoothing

Date	Price per Troy (RM)	Forecast
Jan-16	4,770.08	4770.08
Feb-16	5,023.91	4770.08
Mar-16	5,084.95	5037.20
Apr-16	4,850.36	5100.74
May-16	5,100.76	4853.04
Jun-16	5,220.86	5116.41
...



Jul-20	7,873.04	7488.82
Aug-20	8,249.85	7975.95
Sep-20	7,977.84	8367.10
Oct-20	7,892.63	8074.71
Nov-20	7,684.72	7979.97
Dec-20	7,550.69	7756.61

Both set of forecasting values are compared by using RMSE and MAPE to identify which one of them will have a smaller value.

Table 8: RMSE and MAPE for both models

	Root Mean Square Error	Mean Absolute Percentage Error
Holt's Double Exponential Smoothing	173.26	2.26%
Fuzzy Time Series Markov Chain	161.20	2.18%

Based on Table 8 above, the result shows that Fuzzy Time Series Markov Chain have a smaller value for both RMSE and MAPE. Thus, Fuzzy Time Series Markov Chain is considered as a better method to forecast the price of gold compared to Holt's Double Exponential Smoothing.

CONCLUSION AND RECOMMENDATIONS



This study analyses two methods of forecasting, which are Holt's Double Exponential Smoothing and Fuzzy Time Series Markov Chain, by comparing their performances regarding the gold price. The Durbin Watson test shows that the data are dependent on time, which means the data are suitable for time series forecasting analysis.

The two forecasting results are tested by using RMSE and MAPE to show a better model for forecasting the price of gold. The results of both statistic measurements must have the lower value, and for Mean Absolute Percentage Error, the value must be below 10 percent. Based on both criteria, the model with a better result is the Fuzzy Time Series Markov Chain with $RMSE = 161.20$ and $MAPE = 2.18\%$. This result is quite similar with the result obtained from Taufik (2020) even though both studies are using difference data duration. The finding shows that Fuzzy Time Series Markov Chain is a better model in predicting price of gold in Malaysia as compared to Holt's Double Exponential Smoothing

This study applies only two methods; Holt's Double Exponential Smoothing and Fuzzy Time Series Markov Chain to predict the price of gold. It is advisable for future researchers to expand the study by using other forecasting techniques such as ARIMA, Multiple Linear Regression and Brown's Double Exponential Smoothing as well as other non-statistical model such as ANN. In order to choose the most accurate method, many accuracy tests can be applied other than RMSE and MAPE, such as MAE and BIC. Besides, this study could be extended for broader topic such as analysis of the factors that affect the price of gold so that different perspectives and ideas can be discussed through those topics.

CONFLICT OF INTEREST DISCLOSURE

The authors have no conflicts of interest to declare.

REFERENCES

- Abdelkader, S., Mohammed, M., & Hacen, K. (2020). Fuzzy Auto-Regressive Integrated Moving Average (FARIMA) Model for Forecasting the Gold Prices. *Journal of Smart Economic Growth*, 5(1), 1-5.
- Chaku, S. E., Gabriel, F. G., Abdulrazaq A. A., Aliyu, S. A., Adehi M. U. & Timnan, B. N. (2022). Time Series Modeling and Forecasting of Gold Prices on International Financial Markets. *American Journal of Sciences and Engineering Research*, 5(2), 76-86.
- Chukwudike, N., Ugoala, C. B., Maxwell, O., Okezie, U.-I., Bright, O., & Henry, U. (2020). Forecasting Monthly Prices of Gold Using Artificial Neural Network. *Journal of Statistical and Econometric Methods*, 9(3), 2241–0376. http://www.scienpress.com/Upload/JSEM/Vol9_3_2.pdf
- Gokmenoglu, K. K., & Fazlollahi, N. (2015). The Interactions among Gold, Oil, and Stock Market: Evidence from S&P500. *Procedia Economics and Finance*, 25, 478-488.
- Guha, B., & Bandyopadhyay, G. (2016). Gold price forecasting using ARIMA model. *Journal of Advanced Management Science*, 4(2), 117-121.
- Ghalayini, L., & Farhat, S. (2020), Modeling and Forecasting Gold Prices. *Journal of Economics and Business*, 3(4), 1708-1729.



Shakil, M. H, Mustapha I. M., Tasnia M., Saiti B, (2018). Is gold a hedge or safe haven? An application of ARLD Approach. *Journal of Economics, Finance and Administrative Science*, 23(44), 60-76.

Suranart, K., Kiattisin, S. & Leelasantitham, A. (2014). Analysis of comparisons for Forecasting Gold Price using Neural Network, Radial Basis Function Network and Support Vector Regression. *International Conference on Information and Communication Technology, Electronic and Electrical Engineering Proceeding Series*, 1-5. <https://doi.org/10.1109/JICTEE.2014.6804078>

Taufik, M., Afrah, A. S., Sintiya, E. S., & Hariyanto, D. (2020). A comparative study of time-series models for forecasting the Indonesian gold price. *ACM International Conference Proceeding Series*, 79–83. <https://doi.org/10.1145/3427423.3427438>

Uzun, B., & Kiral, E.(2017). Application of Markov Chains-Fuzzy states to gold price. *Procedia Computer Science*, 120, 365–371. <https://doi.org/10.1016/j.procs.2017.11.251>

