Multidisciplinary Applied Research and Innovation Vol. 2 No. 1 (2021) 156-163
© Universiti Tun Hussein Onn Malaysia Publisher's Office

MARI

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/mari
e-ISSN : 2773-4773

Comparison between Causal Model and Time Series Model to Forecast Gold Prices

Norain Ahmad Nordin[1,*], Nuramirah Juma'at[1], Zul Afiq Sazeli[1]

[1]Department of Science and Mathematics, Centre for Diploma Studies, Universiti Tun Hussein Onn Malaysia, Pagoh Higher Education Hub, KM1, Jalan Panchor, Pagoh, 84600 JOHOR

DOI: https://doi.org/10.30880/mari.2020.01.01.016
Received 11 November 2020; Accepted 01 January 2021; Available online 03 February 2021

Abstract: - Gold is the most precious and valuable commodity in the world. It is not only used to make jeweleries but gold also plays an important role in world monetary system. Thus, it is important to predict or to forecast gold prices due to the highly demand of gold in order to plan the right time to buy and invest gold. The current study proposed a univariate model that is Box-Jenkins method and time series regression method as a causal model to forecast gold prices. The data used in this study are daily prices from the 25th of March 2011 to 23th May 2011 and was analyzed using the Statistical Software Minitab Version 15.1.2. A comparison between these two models was made by evaluating the mean absolute deviation (MAD) and mean absolute percentage error (MAPE). The finding suggests that time series regression method produced a smaller value of MAD and MAPE compared to Box-Jenkins ARIMA (0,2,1). This means that in forecasting gold price data, another independent variable that influences them can provide extra information to forecast gold prices more precisely.

Keywords: ARIMA, Box-Jenkins, Gold price, MAD, MAPE, Time series regression

1. Introduction

Gold has a unique properties as a valuable asset. It is highly sought-after precious metal for coinage and jewelries. Furthermore, gold plays important roles in the world monetary system throughout human history and plays a vital role as a store of value especially in times with political and economic uncertainties [1]. Over the last 40 years, there is a significant negative relationship between inflatation and the gold price [2]. Therefore, forecasting or predicting the price of gold has become very significant and important to investors.

According to Sjaastad [3], the price instability in the world gold market was affected by the floating exchange rates among the major currencies. Gold is particularly sensitive to dollar or euro exchange rate [4]. This is due to the import and exports of oil and precious metals come from the United States and the euro land.

Corresponding author: norainn@uthm.edu.my
2021 UTHM Publisher. All rights reserved.

There are many forcast models used in forecasting gold prices. In many research practice, Box-Jenkins modelling is one of the oftenly used univariate time series models because it can be used to analyse almost any set of data [5]. Nambier [6] used ARIMA model in forecasting gold prices. Sharma [7] utilized Box-Jenkins ARIMA to forecast gold price of India and reached a conclusion ARIMA (3,1,3) is the best appropriate model. Artificial Neural Nwtwork (ANN) model and ARIMA model were established by [8] to forecast gold prices and was declared ANN model as efficient model.

Four different methods were used by [9] such as generalized autoregressive conditional heterockedasticity (GARCH) model, Wavelet Neural Network, Trend and Threshold Autoregressive model, and Wavelet Neural Network with Trend and Threshold Autoregressive model for the forecasting of the gold prices. He concluded that Wavelet Neural Network and Wavelet Neural Network with Trend and Threshold Autoregressive model are most efficient. Multiple Liner Regression (MLR) model is proposed by [10] in order to predict the future gold prices.

The prices of gold can be forecasted as a univariate data. However, these prices are also being affected by other independent variable. According to [11], the relationship between gold price, stock returns, exchange rate and oil price were considered. It showed that the exchange rate of United State Dollar directly affects the gold price. The dynamic linkage between the USD and the gold and the crude oil market were examined by [12]. The result revealed positive relationship between gold and oil.

Thus, this study will explore whether causal model or time series model produce more accurate gold prices forecast. Hence, time series regression and Box-Jenkins method will be used in forecasting gold prices versus EUR-USD exchange rate and the performance of the forecasts for both methods will be compared.

2. Forecasting Methods

2.1 Box-Jenkins Method

In the current study, the Box-Jenkins model-building strategy will be utilized. The original Box-Jenkins modeling procedure involves only an iterative three-stages process which are model identification, parameter estimation and diagnostic checking.

Model identification in the Box-Jenkins framework are based on the transformed and differenced data in trying to identify potential ARIMA model that give a good fit to the data. Box Jenkins method was used to represent ARIMA (p, d, q) where d indicates the differentiation levels involved to reach stationary state used in this series. p represents the number of Autoregressive (AR) terms and q represents the number of Moving Average (MA).

The first step is to determine whether the series is stationary or non stationary. If the model is stationary the time series appears to vary about a fixed level by looking at a plot of the series with the sample autocorrelation function. Non stationary time series indicated if the series appears to grow or decline over time and the sample autocorrelations do not trail off to zero rapidly. In order to convert the non stationary series by the data is differenced until the plot series varies about a constant level and sample autocorrelations die out fairly rapidly.

The next step is to estimate the parameters identified in the model. Model fitting consists of finding the best possible estimates for the parameters of the tentatively identified model. There are easier and precise ways to estimate the parameters by using computer packages. Examples of packages that are widely used are Minitab, T-series and R. They provide more accurate values of parameters.

2.2 Time Series Regresssion Method

Observations in different time periods are auto correlated or related in time series regression. This is because the data collected over time tend to have trends, seasonal patterns and so on. This means that, for time series data, the standard regression method cannot be applied because the sample of observations cannot be assumed as independent. Therefore, fitting regression models to the time series data must be handled properly in order to fit the regression models.

The pattern of the series will be first identified in analysing time series regression. In this stage, the error terms, ε_t occurring over time either have positive autocorrelation or negative autocorrelation.

First-order autocorrelation is one type of the positive or negative autocorrelation which is ε_t, the error term in time period *t*, is related to ε_{t-1}, the error term in time period $t - 1$, is given by Eq (1)

$$\varepsilon_t = \rho\varepsilon_{t-1} + a_t, \quad \text{Eq (1)}$$

The correlation coefficient between error terms is denoted as ρ and $a_1, a_2, \cdots$ are values randomly and independently selected from a normal distribution having mean zero and a variance independent of time.

Durbin-Watson (DW) test can be used to test the presence of the first-order autocorrelation**.** DW for positive autocorrelation is commonly used in practice. The DW test for positive autocorrelation is given in Eq (2) as follows:

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}, \quad \text{Eq (2)}$$

where $e_1, e_2, \ldots, e_n$ are time-ordered residuals. The test involves the determination of whether the autocorrelation parameter, ρ is zero. So, the serial correlation can be performed by comparing the calculated value of DW statistic with lower $(d_{L,\alpha})$ and upper $(d_{U,\alpha})$ bounds.

Then autoregressive model is used to eliminate the positive autocorrelation. This can be done in a regression framework, using the dependent variable lagged one or more time periods as the predictor or independent variable. The model Eq (3) for first-order autoregressive process is defined by:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t, \quad \text{Eq (3)}$$

where

$$\varepsilon_t = \rho\varepsilon_{t-1} + a_t.$$

Least squares is used to fit to the data to get the model and hence, the forecasting equation, Eq (4) can be written as

$$\hat{Y}_t = b_0 + b_1 Y_{t-1}. \quad \text{Eq (4)}$$

Then the model is used to forecast the future values.

2.3 Forecastinng Performance Measures

In this study mean absolute percentage error (MAPE) and Mean absolute deviation (MAD) are used to evaluate errors.

Mean absolute percentage error (MAPE) is calculated using the following formula:

$$\text{MAPE} = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right|. \quad \text{Eq (5)}$$

Mean absolute deviation (MAD) is calculated using the following formula:

$$\text{MAD} = \frac{\sum_{t=1}^{n}|e_t|}{n} = \frac{\sum_{t=1}^{n}|Y_t - \hat{Y}_t|}{n}. \quad \text{Eq (6)}$$

3. Results and Discussion

In this research, Box-Jenkins, a univariate model and a causal model called time series regression methods were used to forecast the gold prices. The daily gold prices data was taken from the website www.gold.org while the data of EUR-USD Exchange Rate was taken from www.oanda.com. Data from 25th of March 2011 until 23th of May 2011 were analyzed and the results were discussed. The Statistical Software Minitab Version 15.1.2 and Microsoft Excel are used in analyzing the data.

3.1 Analysis using Box-Jenkins Method and Time Series Regression Method

In this research, the measurements of the daily gold prices data were analysed and written as $y_1, y_2, \cdots, y_{60}$. Since the the gold prices data are volatile, transformation of the data into logarithms form can help to stabilize the variance in a series where the variation changes with the level. The data is plotted using the Minitab software as shown in Figure 1. From the plotting, the presence of upward trend and the series is non stationary as well as exhibited nonseasonal trends can be seen.

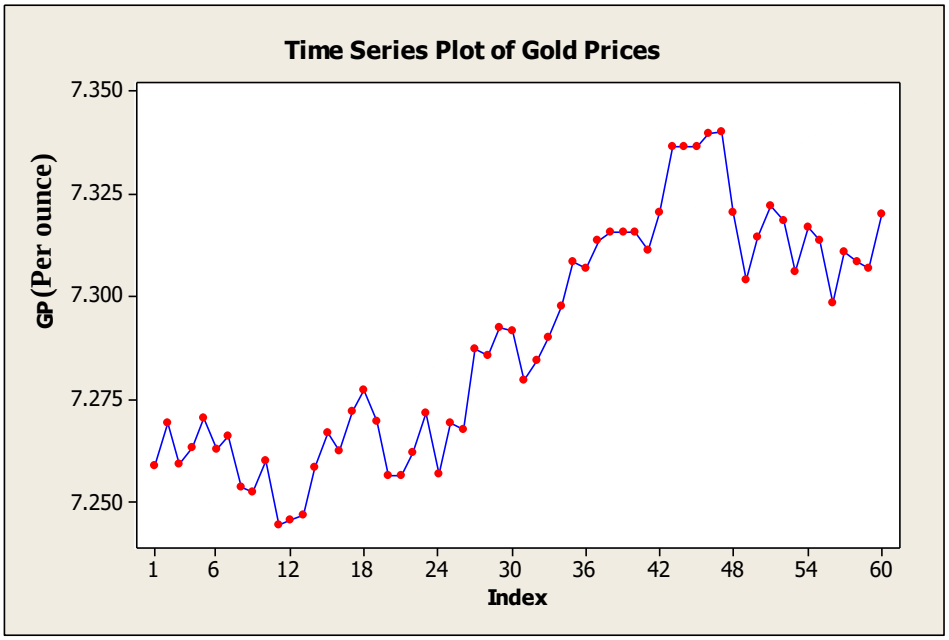

Figure 1: Time series plot of gold prices

At the first stage, the original data were transformed to logarithm form because of the variance of was not constant. The time series data were plotted to determine the stationarity of the data. Then the first difference was made due to the nonstationary data which implied that mean of the data did not fluctuate at constant rate. Autocorrelation function (ACF) and Partial autocorrelation function (PACF) of the data were plotted and being analyzed to determine the p, d, q orders as shown ni Table 1. However, models selected were not significant. Then, the second difference was taken. After the analysis of all the models by looking at the SSE values, ARIMA (0,2,1) was the most suitable model to forecast the gold prices.

Table 1: *t*-ratio value for ARIMA (*p,d,q*)

ARIMA (p,d,q)	*t*-ratio				
	AR I	AR II	AR III	MA I	MA II
(0,2,1)	-	-	-	12.10	-
(1,2,0)	-4.19	-	-	-	-
(1,2,1)	-	-	-	-	-
(2,2,0)	-5.70	-3.59	-	-	-
(2,2,1)	-	-	-	-	
(3,2,0)	-6.23	-4.29	-2.22	-	-
(3,2,1)	-0.93	-1.05	-0.03	11.31	-

In building a time series regression model, two variables is considered that are the gold prices as dependent variable (Y) and the EUR-USD exchange rate as independent variable $(X1)$. The scatter plot of the dependent variable against the independent variable is illustrated in Figure 2. The residual autocorrelation was examined using Durbin-Watson (DW) test and it implied a strong positive relationship between gold prices and EUR-USD exchange rate with the vakue of $r = 0.8136$f. Then, the autocorrelation problem was solved by using autoregressive models. After the autocorrelation has been removed, the model was fitted by using least square method and the model was used to forecast the next ten days values.

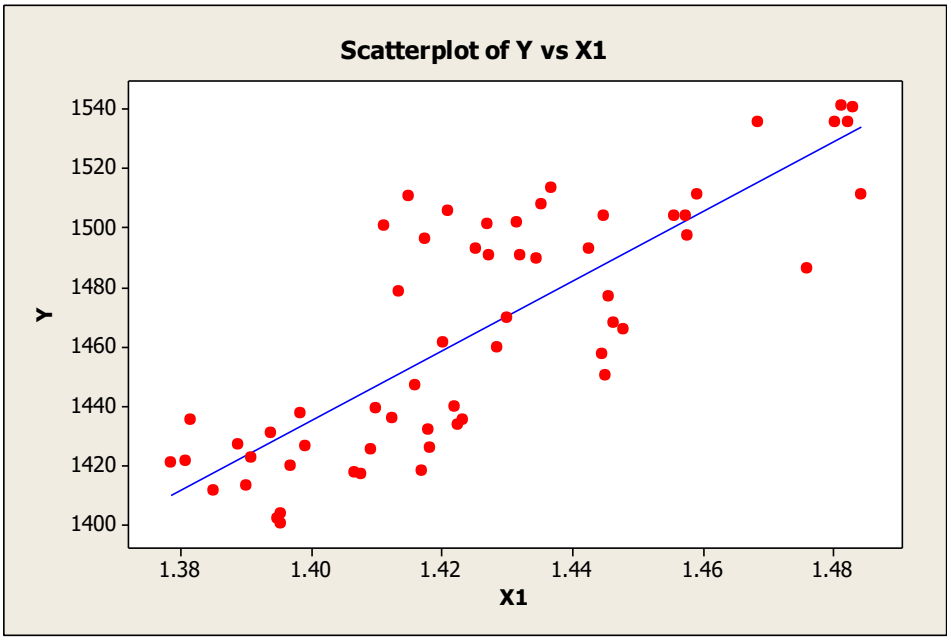

Figure 2: Gold prices versus ERU-USD exchange rate

3.2 Comparison between Box-Jenkins and Time Series Regression

The out-sample gold prices are forecasted by using fiited model obtained from 24th May 2011 to 6th June 2011. In order to check the accuracy of these two models, the actual data and forecasting values of Box-Jenkins model and time series regression are calculated as shown in Table 2 and Table 3 respectively. Two evaluation criteria are then used to compare the two models which are MAPE and

MAD as shown in Table 4. The model that generates the smallest prediction error of MAPE and MAD will be the best forecasting model.

Table 2: The comparison between actual and forecasting value of Box Jenkins

Period (Day)	**Actual Value, Y_t**	**Forecast Value, $\widehat{Y}_t$**	**Absolute Error, $\|e_t\|$**
24/05/2011	1,527.00	1513.00	13.99914
25/05/2011	1,526.30	1516.12	10.12804
26/05/2011	1,518.50	1518.66	0.158471
27/05/2011	1,533.00	1521.77	11.23404
30/05/2011	1,533.00	1524.34	8.663841
31/05/2011	1,536.50	1527.43	9.067984
01/06/2011	1,533.80	1530.03	3.715921
02/06/2011	1,539.50	1533.12	6.379879
03/06/2011	1,540.00	1535.75	4.247616
06/06/2011	1,549.00	1538.83	10.16971

Table 3: The comparison between actual and forecasting value of Time Series Regression

Period (Day)	**Actual Value, Y_t**	**Forecast Value, $\widehat{Y}_t$**	**Absolute Error, $\|e_t\|$**
24/05/2011	1,527.00	1510.16	16.8355
25/05/2011	1,526.30	1525.82	0.4770
26/05/2011	1,518.50	1525.16	6.6587
27/05/2011	1,533.00	1517.76	15.2435
30/05/2011	1,533.00	1531.52	1.4830
31/05/2011	1,536.50	1531.52	4.9830
01/06/2011	1,533.80	1534.84	1.0385
02/06/2011	1,539.50	1532.28	7.2238
03/06/2011	1,540.00	1537.69	2.3145
06/06/2011	1,549.00	1538.16	10.8400

Table 4: MAD and MAPE values of Box-Jenkins and time series regression

Evaluation Criteria	ARIMA (0,2,1)	Time Series Regression
MAD	7.77647	6.70975
MAPE	0.50680	0.437472

In this study, Box-Jenkins ARIMA (0,2,1) produces MAD = 7.77647 and MAPE = 0.50680 with the model $y_t = a_t - 0.9824a_{t-1}$. Meanwhile, time series regression method produces MAD = 6.70975 and MAPE = 0.437472 with the model $\hat{Y}_t = 76.7 + 0.949\ Y_{t-1}$. In conclusion, the time series regression is better method compared to Box-Jenkins ARIMA (0,2,1) in forecasting the gold prices. In time series regression, an independent variable was considered in forecasting the gold prices that might have provided extra information compared to just using a univariate model.

4. Conclusion

In this research, Box-Jenkins, a univariate model and time series regression methods were used to forecast the gold prices versus EUR-USD Exchange Rate. The Statistical Software Minitab Version 15.1.2 and Microsoft Excel were also used in analyzing using time series regression method. Forecast performances were evaluated for both methods by comparing the value of MAD and MAPE. In this research, it can be concluded that time series regression method produced a smaller value of MAD and MAPE compared to Box-Jenkins ARIMA (0,2,1). This means that in forecasting gold price data, another independent variable that influences them can provide extra information to forecast gold prices more precisely. Forecasting literature has suggested that model combination in forecasting can improve forecasts since it captures the unique feature of each model.

Acknowledgement

The authors would also like to thank the Centre for Diploma Studies, Universiti Tun Hussein Onn Malaysia for its support.

References

[1] R. Aggarwal and B. M. Lucey, "Psychological Barriers in Gold Prices", Finan. Econ., vol. 16, pp. 217–230, 2007.

[2] S. Shafiee and E. Topal, "An overview of global gold market and gold price forecasting," Resources Policy, vol. 35, pp. 178-189, 2010.

[3] L. A. Sjaastad, "The price of gold and the exchange rates: once again", Resources Policy, vol 33, pp. 118-124, 2008.

[4] B. M. Lucey and E. Tulley, "The evolving relationship between gold and silver 1978–2002: evidence from a dynamic cointegration analysis: a note", Applied Financial Economics Letters vol 2, pp. 47–53, 2006.

[5] C. Christodoulos, C. Michalakelis and D. Varoutas, "Forecasting with limited data: Combining ARIMA and diffusion models," Technological Forecasting & Social Change, vol 77, no. 4, pp. 558-565, 2010.

[6] G, D.M., G. Nambiar, and R. M, "Forecasting Price And Analysing Factors Influencing The Price Of Gold Using Arima Model And Multiple Regression Analysis," International Journal of Research in Management, Economics and Commerce, vol. 2, no. 11, pp. 548-563, 2012.

[7] R. K. Sharma, "Forecasting Gold price with Box Jenkins Autoregressive Integrated Moving Average Method," Journal of International Economics, vol .7, no. 1, Jan-Jun, pp. 32-61, 2016. The World Gold Counci.

[8] H. Mombeini and A. Yazdani-Chamzini, "Modeling Gold Price via Artificial Neural Network Journal of Economics," Business and Management, vol. 3, no. 7, pp. 699-703, 2015

[9] M. C. Lineesh, K. K. Minu and C. J. John, "Analysis of non stationary nonlinear economic time series of gold price a comprative study", International Mathematical Forum, vol. 5, no. 34, pp. 1673-1683, 2010.

[10] Z. Ismail, A. Yahya, and A. Shabrie, "Forecasting gold prices using multiple linear regression method", American Journal of Applied Sciences, vol. 6, no. 8, pp. 1509-1514, 2009.

[11] K. S. Sujit and B. J. Kumar, "Study on dynamic relationship among gold price, oil price, exchange rate and stock market returns", International Journal of Applied Business and Economic Research, vol. 9, no. 2, pp. 145-165, 2011.

[12] B. Mo, H. Nie and Y. Jiang, "Dynamic linkages among the gold market, US dollar and crude oil market", Phys. A, vol. 491, pp. 984-994, 2018.