UNIVERSITI TEKNOLOGI MARA
Cawangan Kedah
Kampus Sungai Petani

2021 iCMS
INTERNATIONAL CONFERENCE ON COMPUTING, MATHEMATICS AND STATISTICS

e-PROCEEDINGS

of The 5th International Conference on Computing, Mathematics and Statistics (iCMS2021)

4-5 August 2021

Driving Research Towards Excellence

e-Proceedings of the 5th International Conference on Computing, Mathematics and Statistics (iCMS 2021)

Driving Research Towards Excellence

Editor-in-Chief: Norin Rahayu Shamsuddin

Editorial team:
Dr. Afida Ahamad
Dr. Norliana Mohd Najib
Dr. Nor Athirah Mohd Zin
Dr. Siti Nur Alwani Salleh
Kartini Kasim
Dr. Ida Normaya Mohd Nasir
Kamarul Ariffin Mansor

e-ISBN: 978-967-2948-12-4
DOI

Library of Congress Control Number:

Publication by
Department of Mathematical Sciences
Faculty of Computer & Mathematical Sciences
UiTM Kedah

TABLE OF CONTENT

PART 1: MATHEMATICS

PART 2: STATISTICS

PART 3: COMPUTER SCIENCE & INFORMATION TECHNOLOGY

PART 4: OTHERS

MODELLING MALAYSIAN GOLD PRICES USING BOX-JENKINS APPROACH

Isnewati Ab Malek[1], Dewi Nur Farhani Radin Nor Azam[2], Dinie Syazwani Badrul Aidi[3] and Nur Syafiqah Sharim[4]
[1, 2, 3, 4] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Negeri Sembilan, Kampus Seremban, 70300 Seremban, Negeri Sembilan, Malaysia
([1] isnewati@uitm.edu.my, [2] j.dewifarhani@gmail.com, [3] diniesyazwanie256@gmail.com, [4] syaaaafiqahs@gmail.com)

Nowadays, gold is an excellent choice of investment for many reasons. It can be used as a hedge against inflation, the function of money and it will always be valuable because of rarity. The Malaysian Kijang Emas is Malaysia's official gold bullion coin and is minted by Malaysia's Royal Mint. This study aims to describe the trend of Kijang Emas and to find the best-fitted model of the ARIMA model in modelling volatile data. The general finding of this study is that the Kijang Emas prices indicate the presence of an upward trend, and no seasonality component exists in the data series. In estimating the parameters for the Box-Jenkins ARIMA model, Maximum Likelihood Estimation (MLE) is used. The modelling performance of ARIMA is evaluated by using the value of Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). In terms of forecasting performance, ARIMA (2,1,1) is the more appropriate model for forecasting the future Kijang Emas prices because it has the smallest value of RMSE and MAE.
***Keywords*:** Box-Jenkins, ARIMA, Gold, Kijang Emas, Stationary

1. Introduction

In 2018, with the gold prices volatile market conditions, investor turn their cash investments with gold by either purchasing jewellery or gold bullion coin. This is to ensure that the investor does not lose their purchasing power in the forthcoming days. The increasing demand for gold initiates the price of gold to rise. Kijang Emas Gold Bullion Coins is an alternative form of investment. The Kijang Emas is minted by Malaysia's Royal Mint and is sold by Malayan Banking Berhad. Tun Dr Mahathir Mohamad, the former Prime Minister of Malaysia, released the gold bullion on 17 July 2001. The gold bullion coins come in three sizes; 1 oz, ½ oz and ¼ oz. Price movement for gold is determined by the international gold market.

The gold prices are known for their volatility, which is a condition where the conditional variance changes between extremely high and low values (Miswan et al., 2013). According to Choong et al. (2012) gold is one of the best ways to save for the future and prepare for the worst. Gold prices can be very high at some point and can be very low. Some people are making money by buying gold at a low price and selling it later at a higher price. However, they cannot predict accurately the right time to buy and sell the gold due to inconsistent gold prices.

Based on the previous study, Ho et al. (2017) said that the ARIMA model is the best forecasting technique when involving time-series data. In addition, the study was done by Tripathy and Naliniprava (2017) mentioned that one of the widely used models for predicting the gold price nowadays is the ARIMA model by assuming the future values of time series have a functional relationship with current and past values. Moreover, in the Ali et al. (2016) study, the Box-Jenkins methodology has been applied for forecasting the daily gold price from the USA GOLD website. At the first difference, it is known that the data is stationary by using the Line Diagram, Correlogram and ADF Test. Both models for ARIMA (0,1,1) and (1,1,0) have very close values of AIC and BIC to each other after model estimation. As a result, it indicates that ARIMA (0,1,1) is more appropriate model than ARIMA (1,1,0) by comparing the values of MAE, MAPE and RMSE.

A study was done by Guha and Bandyopadhyay (2016) also use the Box-Jenkins method to predict the future values of gold prices. The estimated ARIMA models, which is ARIMA (1,1,1) and ARIMA (0,1,1) are selected. The result showed that ARIMA (1,1,1) was selected as the best model because it has the lowest value of AIC and BIC.

In 2020, the world faced a shocking situation caused by a coronavirus (Covid-19) that affected the whole world which caused deflection. This causes many consequences that affect people, economies, the environment, and daily life. Yousef and Shehadeh (2020), indicate that the number of Covid-19 global cases has positively impacted gold price. In addition, they said that the global recession likely to be caused by the Covid-19 pandemic may mean investors will continue to seek refuge in gold for some time to come. As the world economy fell, people worried that coronavirus might worsen in the future, so they overcame it by investing in gold for their safe assets. Consequently, the demand for gold can continue to rise, at the same time pushing its price upwards until a vaccine appears to stabilize the world economy (Grima et al., 2020).

Due to the volatility of gold prices, it is a good choice to predict gold prices as it can help those people who are planning to invest in gold in the future. Therefore, this study is important to describe Kijang Emas's trend and find the best-fitted model to forecast the prices.

2. Methodology

2.1 Data Description

This data is available on the Central Bank of Malaysia (BNM) website. This study focuses on the buying price of bullion coin Kijang Emas for 1 troy ounce collected in daily terms starting from 1st September 2016 until 30th September 2020.

2.2 Box-Jenkins Methodology

Box-Jenkins's method is a widely used process to find out the best model for time series data. According to Lazim (2013), the Box-Jenkins approach is synonymous with the general ARIMA modelling. This method is used to fulfil the second objective which is to determine the best fitted ARIMA model of Kijang Emas price. The term ARIMA is in the short, stands for the combination that comprises the Autoregressive Integrated Moving Average, model. The model, thus, obtained is represented in a general term as ARIMA *(p, d, q)* where the symbol 'd' denotes the number of times the variable buying prices need to be differenced to achieve stationary. A simple model case ARIMA (1,1,1) can be written as,

$$w_t = \mu + \emptyset_1 w_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t \quad (1)$$

where $w_t = y_t - y_{t-1}$ represents the first difference of the buying price series and is assumed stationary. In (1), the values of p = 1, d = 1 and q = 1. The values of p and q were the number of significant spikes in the Partial Correlation Function (PACF) and Autocorrelation Function (ACF), respectively.

To choose the best ARIMA model, some statistical measures were applied. Some of the common statistical measures used to validate the best ARIMA models are the Akaike's Information Criteria (AIC) and the Bayesian Information Criteria (BIC). The AIC was implemented to compare distinct possible models and discover which one is the best fit for the data (Bevans, 2020). Meanwhile, the BIC aimed to choose a model that achieves the most accurate out-of-sample forecast by stabilizing between the models' complexity and goodness of fit (Hyndman, 2018). The lower the value of AIC and BIC, the model is said to be the best ARIMA model.

Finally, the best model is selected based on the results of comparing their respective measures in which the model that produced the smallest value of Root Mean Square Error (RMSE). The model is ready to be used for forecasting when all criteria are fulfilled, and the model is significant.

3. Result and Discussions

The results and analysis of the research will be discussed in this section. The process of selecting the best parameter is done by using EViews software.

3.1 The Overall Trend of Malaysian Gold Prices

At the initial stage, a simple data investigation was conducted to understand the basic pattern of the series and hence to identify any characteristic existing. Figure 1 shows the historical plot of the buying price for 1 troy ounce of Kijang Emas Gold Bullion Coins. The graph indicates an upward trend over the four years which began a steady climb in 2019 to 2020 for Kijang Emas. The average 1 troy ounce gold price is RM 5404 in 2016 and jumped to RM 7492 in 2020, which about 38.6% of increment.

Figure 1: Historical Plot of Buying Price of Kijang Emas Bullion Gold Coins

3.2 Analysis of ARIMA (*p, d, q*) Model

The application of the Box-Jenkins lies in the assumption that the data series is stationary. A series is said to be stationary if it does not show an upward or downward trend over time. If the assumption is not met, then the necessary procedures are performed to achieve stationary in the series. A series can be made stationary by taking the differencing on the data set. Differencing is the process of removing the trend pattern from the actual data. Data of gold price buying price is plotted in Figure 1, and the plot is clearly indicating that it is not stationary because there is an upward trend in the series. Besides, the stationary of a time series model also can be determined by statistical testing. The appropriate statistical test for stationary testing is the unit root test.

Table 1: Unit Root Test of Actual Data Series

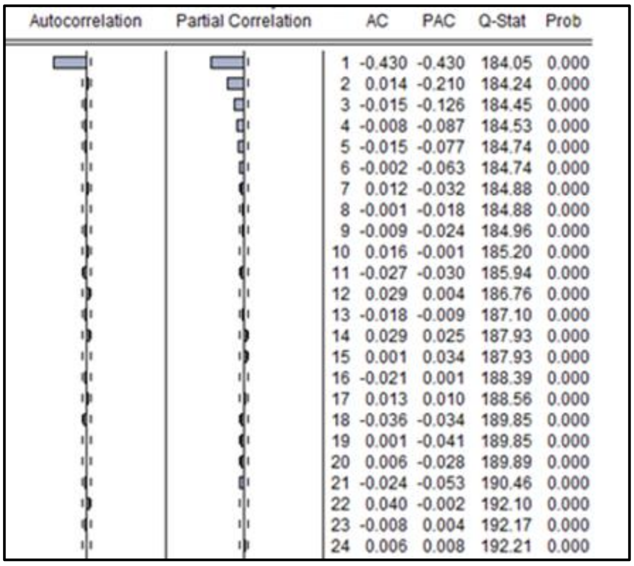
	t-Statistic	**Probability value**
Augmented Dickey-Fuller test statistic	-1.435817	0.8501

The H_0 (null hypothesis) stated that the data is not stationary. The decision rule involve is rejected H_0 if the probability value less than $\alpha = 0.05$. The p-value $= 0.8501$ which is greater than α, the null hypothesis has failed to reject. Hence the data is not stationary. Since the data is not stationary, the difference between the current value of Kijang Emas price, y_t, and the preceding value of Kijang Emas price, y_{t-1} is taken to fulfil the main assumption of the Box-Jenkins methodology.

Table 2: The Unit Root Test After First Order Differencing

	t-Statistic	**Probability value**
Augmented Dickey-Fuller test statistic	-21.79584	0.0000

Table 2 shows the output of the Unit Root Test after the first order differencing. From the table, it shows that the probability value is 0.0000 which is less than $\alpha = 0.05$. Back to the hypothesis statement, p-value less than α leads to rejecting the null hypothesis and it results that the data does not have a unit root and is stationary. The order of difference is defined as the number of times the series needs to be differenced to achieve stationary. After the data is stationary, the order of *d = 1* for the ARIMA *(p, 1, q)* model. The order of *p* and *q* are chosen based on observing the lags of PACF and ACF, respectively. Figure 2 shows the correlogram after performing the first-order differencing.



Lag	AC	PAC	Q-Stat	Prob
1	-0.430	-0.430	184.05	0.000
2	0.014	-0.210	184.24	0.000
3	-0.015	-0.126	184.45	0.000
4	-0.008	-0.087	184.53	0.000
5	-0.015	-0.077	184.74	0.000
6	-0.002	-0.063	184.74	0.000
7	0.012	-0.032	184.88	0.000
8	-0.001	-0.018	184.88	0.000
9	-0.009	-0.024	184.96	0.000
10	0.016	-0.001	185.20	0.000
11	-0.027	-0.030	185.94	0.000
12	0.029	0.004	186.76	0.000
13	-0.018	-0.009	187.10	0.000
14	0.029	0.025	187.93	0.000
15	0.001	0.034	187.93	0.000
16	-0.021	0.001	188.39	0.000
17	0.013	0.010	188.56	0.000
18	-0.036	-0.034	189.85	0.000
19	0.001	-0.041	189.85	0.000
20	0.006	-0.028	189.89	0.000
21	-0.024	-0.053	190.46	0.000
22	0.040	-0.002	192.10	0.000
23	-0.008	0.004	192.17	0.000
24	0.006	0.008	192.21	0.000

Figure 2: Correlogram After First Order Differencing

After a stationary condition has been achieved, the next stage is to perform model identification. The process of identifying the suitable model for the data series involved the analysis of the ACF and PACF as shown in Figure 2. Based on the ACF diagram, there is one significant spike at lag 1 that determine the order for MA, *q=1*. On the other hand, the PACF shows several spikes, the most significant at lag 1 (exceeding the standard error line), and the spike at lag 2, 3, 4 and 5 that suggest the order for AR, *p=5*.

In addition, confidence limits also can be used to determine the parameters for ARIMA (*p, d, q*). By using the confidence limit formula which is $= \pm \frac{2}{\sqrt{n}}$, the confidence limit for this study is $(-0.064, +0.064)$. The lags that are outside of the confidence limit are considered significant.

Table 3: The ACF and PACF of ARIMA

Lags	Autocorrelation Function (ACF)	Partial Autocorrelation Function (PACF)
1	-0.430	-0.430
2	0.014	-0.210
3	-0.015	-0.126
4	-0.008	-0.087
5	-0.015	-0.077
6	-0.002	-0.063
7	0.012	-0.032
8	-0.001	-0.018
9	-0.001	-0.024
10	-0.016	-0.001

Based on the ACF values in Table 3, there is only one significant spike at lag 1. While in PACF, there are five significant spikes at lag 1 until 5 since the values are outside the value of -0.0064. By referring to Figure 2 and Table 3, the following five models have been identified and estimated using EViews software. The models are ARIMA (1,1,1), ARIMA (2,1,1), ARIMA (3,1,1), ARIMA (4,1,1) and ARIMA (5,1,1). To determine which of the models fits the best, two criteria will be used, that is AIC and BIC, and the results are summarized in Table 4.

Table 4: The AIC and BIC of the ARIMA Model

Model	Akaike Info Criterion (AIC)	Bayesian Information Criteria (BIC)
ARIMA (1,1,1)	12452.44	12467.15
ARIMA (2,1,1)	12452.42	12472.03
ARIMA (3,1,1)	12472.03	12478.78
ARIMA (4,1,1)	12456.22	12485.63
ARIMA (5,1,1)	12457.86	12492.16

According to AIC, ARIMA (2,1,1) has the lowest value. Whereas, based on the BIC, ARIMA (1,1,1) is considered to be the best model since it has the lowest value. As mentioned earlier, the lower the value of AIC and BIC, the model is said to be the best ARIMA model. To validate the above result is correct, further analysis using the error measures as a comparison was performed and the result is shown in Table 5.

Table 5: The RMSE and MAE of the ARIMA Model

ARIMA Model	Root Mean Square Error (RMSE)	Mean Absolute Error (MAE)
ARIMA (1,1,1)	128.2214	50.7927
ARIMA (2,1,1)	128.0901	50.6864

Based on the smallest value of RMSE and MAE, the resulting point towards ARIMA (2,1,1). On these results, ARIMA (2,1,1), is therefore the winner and be proposed as the best forecasting model.

4. Conclusion

The Kijang Emas price data examined in this study can be characterized with the ARIMA (2,1,1) model. ARIMA (1, 1,1) and ARIMA (2, 1, 1) were selected based on five different model parameters, as it provides the best model that meets all the criteria of the fit statistics. However, the lower value of RMSE and MAE for ARIMA (2,1,1), when compared to that of ARIMA (1,1,1), showed that ARIMA (2,1,1) is the more appropriate model in predicting the future values of Kijang Emas price.

References

Ali, Asad, Muhammad Iqbal Ch, Sadia Qamar, Noureen Akhtar, Tahir Mahmood, Mehvish Hyder, and Muhammad Tariq. (2016). Forecasting of daily gold price by using box-jenkins methodology. *International Journal of Asian Social Science*, 6(11):614-24.

Bank Negara Malaysia. (n.d.). Retrieved from https://www.bnm.gov.my/kijang-emas-prices.

Choong, P. S., Kwoo, P. Y., Piong, C. K., & Wong, W. X. (2012). Determinants of gold price: using simple and multiple linear regression. [Unpublished master's thesis]. Universiti Tunku Abdul Rahman.

Grima, S., Dalli Gonzi, R., & Thalassinos, E. (2020). The impact of COVID-19 on Malta and it's economy and sustainable strategies. *Journal of Corporate Governance, Insurance, and Risk Management*, 7(1):53-73.

Guha, B., & Bandyopadhyay, G. (2016). Gold price forecasting using ARIMA Model. *Journal of Advanced Management Science*, 4(2).

Ho, Thanh Tri, Dao Phan, Van Ninh Nguyen, and Juraj Sipko. (2017). Application of ARIMA model to forecast gold price in Vietnam. *Proceedings The 11th International Days of Statistics and Economics*, 469-77.

Khan, M. M. A. (2013). Forecasting of gold prices (Box-Jenkins Approach). *International Journal of Emerging Technology and Advanced Engineering*, *3*(3):662–670.

Lazim, M. (2013). *Introductory business forecasting. A practical approach 3rd edition.* Penerbit Press, Universiti Teknologi Mara.

Miswan, Nor Hamizah, Pung Yean Ping, and Maizah Hura Ahmad. (2013). On parameter estimation for Malaysian gold prices modelling and forecasting. *International Journal of Mathematical Analysis*, 7(21–24):1059–68.

Tripathy, Naliniprava. 2017. Forecasting gold price with auto regression integrated moving average model. *International Journal of Economics and Financial*, 7(4):324-29.

Yousef, I., Shehadeh, E., et al. (2020). The impact of the covid-19 on gold price volatility. *International Journal of Economics & Business Administration (IJEBA)*, *8*(4):353– 364.

2021 ICMS

INTERNATIONAL CONFERENCE ON COMPUTING, MATHEMATICS AND STATISTICS

e ISBN 978-967-2948-12-4

9 7 8 9 6 7 2 9 4 8 1 2 4