GOLD PRICE PREDICTION IN INDONESIA BASED ON INTEREST RATE USING DISTRIBUTED LAG ALMON TRANSFORMATION

Nanda Yumna Aqilah¹*, Sekti Kartika Dini ²

¹,² Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Islam Indonesia
Jl. Kaliurang Km 14.5, Sleman, Yogyakarta, 55584, Indonesia

Corresponding author’s e-mail: *20611001@students.ui.ac.id

ABSTRACT

Gold is valued for its safety and profitability, driven by steady price changes and influenced by interest rates. Accurately predicting gold prices is very important to make the right investment decisions. This study aims to build a gold price model in Indonesia using the Almon transformation lag distribution and see gold price predictions based on the model that has been built. We used the data on gold prices and interest rates from January 2016 to December 2023. Based on the results of the analysis, the best Almon transformation model used in this study is the Almon model with a maximum lag length of 16 and the second polynomial degree. The prediction results have a MAPE of 16.49%, which shows that the Almon model can predict gold prices well for one year. This study contributes to the understanding of gold price dynamics amid economic variations. However, limitations in the model assumptions should be considered.

Received: 27th February 2024
Revised: 15th April 2024
Accepted: 10th July 2024
Published: 1st September 2024

Keywords:
Distributed-lag Almon Transformation; Gold Price Prediction; Interest Rate.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License.

How to cite this article:

Copyright © 2024 Author(s)
Journal homepage: https://oj3.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

Gold or precious metals are often considered safer than other investment instruments [1]. The public has started recognizing the benefits of gold investing in Indonesia since it is an easy investment for people at all social levels. Gold or precious metals can not only be owned by someone with a significant income. However, they can also be owned by anyone from various walks of life. Gold investment tends to be safe and profitable because the change in gold prices is stable and even increases yearly [2]. Moreover, gold's ability to maintain value over time, even during economic uncertainties, reinforces its status as a reliable investment choice for many Indonesians.

One long-term investment option is gold since its price will rise steadily. However, it often fluctuates within a short period. Several factors affect changes in gold prices, including changes in interest rates. Interest rates are one of the factors that have a significant influence on changes in gold prices. According to research by Wicaksono [3], interest rates significantly negatively affect gold prices by 21.9%. It shows that if there is an increase in interest rates in Indonesia, the price of gold will fall. Gold investors will profit less if they sell their gold when interest rates rise. Therefore, investors need to know the prediction of gold prices in the future by paying attention to interest rates, which are one of the factors that influence it.

Prediction is a method used to estimate a value in the future using historical data [4]. Time series analysis can be used to determine or forecast future conditions based on a set of data over a certain period in the past. This is based on the fact that the situation or the past greatly influences human behavior, so the time factor plays an important role [5]. This methodological approach captures the influence of past events on current and future behaviors. It enables analysts to forecast potential scenarios with greater accuracy. Analysts can refine their forecasts by incorporating variables such as economic indicators and social factors, providing valuable insights into future conditions.

Time series analysis can be done by formulating the time series into a classical regression model to analyze the behavior of time series data. Regression models involving current independent variables and dependent variables influenced by previous time-independent variables are distributed lag models. This model uses the most recent period and the impact of changes in the independent variable on the dependent variable within the same observation period when working with time series data [6].

The distributed-lag model is widely used in econometric analysis and investment. In investing, it is necessary to pay attention to the economic situation in the past and the factors that affect the investment to obtain future profits. Generally, the lag-distribution model is used to analyze long-term investment [7]. The lag distribution model can be determined using the Almon approach. Shirley Almon assumes that the coefficients $\beta$ will increase and decrease following the shape of a polynomial of the degree corresponding to $I$, the lag length. It is in line with the price of gold, which fluctuates daily. This method is preferred over others because it effectively captures the changing patterns of coefficients $\beta$ according to a polynomial shape determined by the lag length ($I$). Statistically, the Almon approach integrates selected variables to allow for a more precise estimation of the impact of lagged variables on the dependent variable. This approach can mitigate potential biases, enabling a more profound analysis of cause-and-effect relationships in time series data, such as daily fluctuations in gold prices.

Many researchers have carried out research using distributed lag Almon transformation in various problem areas, such as in agriculture conducted by Pratami, Sudarno, & Ispriyanti [8] on rice production forecasting in Central Java; the study used the distributed lag method of Koyck and Almon transformation, Almon transformation being the best method compared to Koyck transformation with a MAPE value of 1.79%. Furthermore, research related to the money supply and rupiah exchange rate using the distributed lag method of Koyck and Almon transformation conducted by Lihawa, Resmawan, Isa, & Nashar [9], the study showed that the Koyck transformation is better than the Almon transformation shown by the SIC value with an SIC value of 17.1150. Almon's distributed lag transformation method was applied in research conducted by Virgantari & Rahayu [10] to examine the effect of the Operating Cost and Operating Income (BOPO) ratio on the Ratio on Asset (ROA) of Bank Rakyat Indonesia, the results of this study obtained an Almon transformation model that has a coefficient of determination ($R^2$) of 0.75 and there is no autocorrelation. Research using Autoregressive Distributed Lag (ARDL) with the Koyck and Almon method approach was also conducted by Paramitha [11] to predict the profitability of Bank Rakyat Indonesia; the research obtained the best Almon model with a lag length of 4 and second-degree polynomial. In these studies, researchers have yet to find research that uses the gold price variable for the Almon Method. Therefore, this research aims to predict the price of gold using the Almon Method with the independent variable used, namely interest rates.
Based on this description, researchers have not found research using Almon’s distributed lag transformation method with the gold price variable. Therefore, in this study, research will be conducted to predict gold prices using the distributed lag method of Almon transformation with the independent variable, namely interest rates.

2. RESEARCH METHODS

This study uses secondary data on gold prices and interest rates in Indonesia from January 2016 to December 2023, sourced from the websites of the World Gold Council [12], Badan Pusat Statistik [13], and Bank Indonesia [13]. The method used is the Almon transformation distributed lag model.

2.1 Distributed-Lag Model

The distributed-lag model is a regression model that uses time series data. This regression model includes present and past values as independent variables (X). Lag is a term that describes the time it takes for variable X to affect the dependent variable (Y) [14]. Thus, the general model of lag-distribution is as follows.

\[ Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_k X_{t-k} + u_t \]  \hspace{1cm} (1)

\[ Y_t \] is impacted by the coefficient value \( \alpha \), the value of X at time \( t \) \( (X_t) \), the value of X at one time before \( (X_{t-1}) \), the value of X at two times before \( (X_{t-2}) \) to the value of X at k times before \( (X_{t-k}) \), and the value of the bias \( (u_t) \).

There are two types of lag-distribution models: infinite and finite.

1. Infinite lag-distribution model
   If the lag length is not defined, it is called the infinite lag-distribution model [5]. The infinite-lag distribution model is written in Equation (2).

\[ Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + u_t \]  \hspace{1cm} (2)

2. Finite lag-distribution model
   The finite lag-distribution model is when the lag length has been set for \( k \) periods [5]. The finite-lag distribution model is written in Equation (3).

\[ Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_k X_{t-k} + u_t \]  \hspace{1cm} (3)

2.2 Almon’s Approach to the Lag Distribution Model

The finite lag model approximates the parameter estimate of the lag-distributed model, i.e.

\[ Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_k X_{t-k} + u_t \]  \hspace{1cm} (4)

Which can be abbreviated as follows:

\[ Y_t = \alpha + \sum_{i=0}^{k} \beta_i X_{t-i} + u_t \]  \hspace{1cm} (5)

Based on the Weierstrass theorem, Almon assumes that \( \beta_i \) can be estimated by a polynomial of degree corresponding to \( i \), i.e., the lag length. When \( \beta \) follows a second-degree polynomial, the model can be written as follows:

\[ \beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 \]  \hspace{1cm} (6)

Assuming that \( \beta \) follows the second degree, then by substituting Equation (6) into Equation (5), Equation (7) is obtained.

\[ Y_t = \alpha + \sum_{i=0}^{k} (\alpha_0 + \alpha_1 i + \alpha_2 i^2) X_{t-i} + u_t \]
\[ Y_t = \alpha + \alpha_0 \sum_{i=0}^k X_{t-i} + \alpha_1 \sum_{i=0}^k iX_{t-i} + \alpha_2 \sum_{i=0}^k i^2 X_{t-i} + u_t \]  
\text{(7)}

If defined as follows:
\[ Z_{0t} = \sum_{i=0}^k X_{t-i} \]
\[ Z_{1t} = \sum_{i=0}^k iX_{t-i} \]
\[ Z_{2t} = \sum_{i=0}^k i^2 X_{t-i} \]  
\text{(8)}

Then Equation (7) can be written as follows:
\[ Y_t = \alpha + \alpha_0 Z_{0t} + \alpha_1 Z_{1t} + \alpha_2 Z_{2t} + u_t \]  
\text{(9)}

The coefficients of Equation (7) can be estimated using the least squares method. If the errors meet the assumptions of the classical linear model, estimation \( \hat{\alpha} \) and \( \hat{\alpha}_i \) will have desirable traits. After estimation \( \hat{\alpha} \) obtained from Equation (9), the coefficient \( \hat{\beta}_i \) can be estimated using Equation (6) as follows:
\[ \hat{\beta}_0 = \hat{\alpha}_0 \]
\[ \hat{\beta}_1 = \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 \]
\[ \hat{\beta}_2 = \hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2 \]
\[ \hat{\beta}_3 = \hat{\alpha}_0 + 3\hat{\alpha}_1 + 9\hat{\alpha}_2 \]
\[ \ldots \]
\[ \hat{\beta}_k = \hat{\alpha}_0 + k\hat{\alpha}_1 + k^2\hat{\alpha}_2 \]  
\text{(10)}

So, the lag distributed estimation model of the Almon Model is as follows:
\[ \hat{Y}_t = \hat{\alpha} + \hat{\beta}_0 X_t + \hat{\beta}_1 X_{t-1} + \hat{\beta}_2 X_{t-2} + \ldots + \hat{\beta}_k X_{t-k} \]  
\text{(11)}

Three steps must be taken to estimate the lag distribution using the Almon approach [5].

1. Determine the maximum length of the lag.
2. Once the lag length (k) is determined, the polynomial degree (m) must be determined. The degree should generally be smaller than the lag length (k).
3. The variable Z can be formed Once m and k have been determined. Z is a linear combination of the X variables, so multicollinearity among the Z variables is possible.

2.3 Best Model Selection

One of the best models from several models that have been tested is selected using the Akaike Information Criterion (AIC) value [15]. The model with the lowest AIC value among all the models that have been created is the best. The AIC equation is as follows [16].
\[ k \ln AIC = \frac{2k}{n} + \ln \left( \frac{\sum_{i=1}^n \hat{u}_i}{n} \right) \]  
\text{(12)}

With \( k \) is the number of parameters estimated in the regression model, \( n \) is the number of observations and \( \hat{u}_i \) is the residuals in the regression model.

2.4 Evaluation of Prediction Results

The most commonly used measure to evaluate forecast results is Mean Absolute Percentage Error (MAPE). The calculation that shows the absolute error value of the prediction data compared to the actual
data is called MAPE [17]. MAPE calculation can be done by first calculating the error of each prediction result. Percentage Error (PE) is the percentage error of a prediction computed using the following equation.

\[
PE_i = \frac{y_i - \hat{y}_i}{y_i} \times 100\%
\]  

(13)

With \( \hat{y}_i \) is the prediction result in period i, and \( y_i \) is the actual data in period i.

After the PE is known from all prediction results, the MAPE calculation uses Equation (14).

\[
MAPE = \frac{1}{n} \sum |PE_i|
\]  

(14)

With \( n \) is the number of observations.

Forecasting results are said to be good if the MAPE value obtained is getting smaller. The criteria for the MAPE can be seen in Table 1[18].

<table>
<thead>
<tr>
<th>MAPE</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;10%</td>
<td>Excellent forecasting capability</td>
</tr>
<tr>
<td>10% - 20%</td>
<td>Good forecasting capability</td>
</tr>
<tr>
<td>20% - 50%</td>
<td>Reasonable forecasting capability</td>
</tr>
<tr>
<td>&gt; 50%</td>
<td>Bad forecasting capability</td>
</tr>
</tbody>
</table>

3. RESULTS AND DISCUSSION

3.1 Data Description

A descriptive analysis is used to see the development of gold prices and interest rates in Indonesia from January 2016 to December 2023, as presented in Figure 1.

**Figure 1** (a) Gold price movement chart; (b) BI rate movement chart

*Figure 1* (a) shows that the price of gold from January 2016 to December 2023 fluctuates. The cost of gold fluctuates in a short period, but in the long term, the price of gold has increased. Gold prices increased slowly in 2020, where the average increase in 2020 was 1.8%. Gold prices increased due to the COVID-19 outbreak and the US-China trade conflict [19]. *Figure 1* (b) shows that the BI interest rate fluctuated from January 2016 to December 2023. In January 2016, the BI interest rate had the highest value from January 2016 to December 2023, with a value of 7.25%, and the lowest BI interest rate occurred from February 2021 to July 2022, with a value of 3.5%. In July 2018, the BI interest rate increased slowly until July 2019 and then decreased in August 2019. Bank Indonesia increased the BI rate to strengthen the attractiveness of domestic financial assets [20].
3.2 Almon Transformation Distributed Lag Model Estimation

Equation (1) is used in modeling the Almon transformation. The first step is to determine the maximum length of lag (k) and polynomial degree (m) that will be used to determine the value of \( Z_{mt} \). This study will use the maximum lag length of k = 14 to k = 16 with the second polynomial degree. Then, three models will be formed. The AIC value between the Almon models with a maximum lag length of k = 14, k = 15, and k = 16 with the second polynomial degree is utilized to identify the Almon model that best fits this investigation. The parameter estimation results of all models can be seen in Table 1.

Table 2. Almon model parameter estimation

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameters</th>
<th>Coefficient</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>1268980</td>
<td>2172.37</td>
<td></td>
</tr>
<tr>
<td>( Z_{0t} )</td>
<td>-38597</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Z_{1t} )</td>
<td>16101</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Z_{2t} )</td>
<td>-1204</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>1306897.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 shows that the AIC value of the Almon model with a maximum lag length of k = 16 with a second polynomial degree is smaller than the maximum lag length of k = 14 and k = 15 with a second polynomial degree. Therefore, the best Almon model for this study is the Almon model with a maximum lag length of k = 16 second polynomial degree. The maximum lag length k = 16 means that the length of time required by variable X to affect the dependent variable (Y) is 16 periods. Based on the parameter estimation results listed in Table 1, the Almon transformation model formed is as follows.

\[
\hat{Y}_t = 1341619.2 + 34703.9 Z_{0t} + 12108.6 Z_{1t} + 790.2 Z_{2t}, \tag{15}
\]

There are two parameter significance tests: the F and the t-test. The F test is conducted to see the effect of all independent variables on the dependent variable. The results of the F test with a significance level of five percent show a p-value of 1.747 x 10^{-05}; which is less than the significance level. To determine whether the model is workable or whether at least one independent variable significantly affects interest rates. The t-test is conducted to see the impact of each independent variable on the dependent variable. The t-test with a significance level of five percent shows that all variables \( Z_{0t}, Z_{1t}, Z_{2t} \) have a p-value that is less than the significance level. To conclude that each independent variable significantly affects the dependent variable.

There are four classical assumptions, namely the assumption of normally distributed residuals, homoscedasticity, non-autocorrelation, and non-multicollinearity. The results of the classical assumption test are presented in Table 3.

Table 3. Classical Assumption Test Results

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Test Method</th>
<th>p-value</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assumption of normally distributed residuals</td>
<td>Jarque-Bera</td>
<td>0.140</td>
<td>residuals are normally distributed</td>
</tr>
<tr>
<td>Assumption of homoscedasticity</td>
<td>Breusch Pagan</td>
<td>0.226</td>
<td>Homoscedasticity occurs</td>
</tr>
</tbody>
</table>
Table 3, using a significance level of 5%, shows that the assumptions of normally distributed residuals and homoscedasticity assumptions are met. In contrast, the assumptions of non-autocorrelation and non-multicollinearity assumptions are not met. In the assumption of non-multicollinearity, the VIF value on all variables \((Z_{y1}, Z_{y2}, Z_{y3})\) is more than 10, which indicates a multicollinearity problem in all variables \(Z_{mt}\). As said by Gujarati [5], the variable \(Z_{mt}\) is very likely to experience multicollinearity because \(Z_{mt}\) is a linear combination of the variable \(X_t\). According to Gujarati & Porter (In Lihawa, Resmawan, Isa, & Nashar [9]), in the Almon transformation, \(Y_t\) is regressed with the constructed \(Z\) variable so that it is not the original \(X\) variable. Therefore, the classical assumptions have been met in this modeling.

Based on the estimation results of the Almon transformation model in Equation (15), the distributed lag model estimate obtained is as follows.

\[
\hat{Y}_t = 1341619.2 - 34700X_t - 23400X_{t-1} - 13600X_{t-2} - 5490X_{t-3} + 1090X_{t-4} + 6090X_{t-5} + 9500X_{t-6} + 11300X_{t-7} + 11600X_{t-8} + 10300X_{t-9} + 7360X_{t-10} + 2880X_{t-11} - 3190X_{t-12} + 10800X_{t-13} + 20100X_{t-14} + 30900X_{t-15} + 43300X_{t-16} 
\]

The distributed lag model shows that the current gold price \((Y_t)\) is influenced by the interest \((X_t)\) of the current period and the previous periods up to 16 periods. In the current period \((X_t)\), the interest rate has a negative coefficient of -34700, that a one-unit increase in the current interest rate will cause a decrease in the price of gold in the current period by 34700. The coefficients for interest rates from previous periods negatively influence the current gold price. It implies that if interest rates rise, the price of gold tends to fall.

### 3.3 Gold Price Prediction

Based on the distributed lag model estimation results in Equation (16), gold price predictions are made for one year ahead, namely January 2024 to December 2024. The following are the results of gold price predictions for one year.

<table>
<thead>
<tr>
<th>Month</th>
<th>Gold Price Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>749624.5</td>
</tr>
<tr>
<td>February</td>
<td>743581.8</td>
</tr>
<tr>
<td>March</td>
<td>749967.1</td>
</tr>
<tr>
<td>April</td>
<td>765224.3</td>
</tr>
<tr>
<td>May</td>
<td>785797.6</td>
</tr>
<tr>
<td>June</td>
<td>808130.9</td>
</tr>
<tr>
<td>July</td>
<td>828668.3</td>
</tr>
<tr>
<td>August</td>
<td>843853.8</td>
</tr>
<tr>
<td>September</td>
<td>850131.5</td>
</tr>
<tr>
<td>October</td>
<td>835269.5</td>
</tr>
<tr>
<td>November</td>
<td>807217.3</td>
</tr>
<tr>
<td>December</td>
<td>762023.9</td>
</tr>
</tbody>
</table>

The prediction results in Table 4 show a pattern of gold price trends characterized by a sustained rise over nine months, followed by a decline from October to December. This pattern reflects seasonal variations and market dynamics that affect gold prices, especially in response to economic factors such as interest rates. The Mean Absolute Percentage Error (MAPE) calculation yields a value of 16.49%. This value can be
categorized in the 'good' category, as listed in Table 1, indicating that the distributed lag model with Almon transformation can be used to predict gold prices well.

The decline in gold prices at the end of the year reflects the impact of higher interest rates, making gold less attractive as a hedge. As explained by the lag distribution model that has been obtained, gold prices tend to fall when interest rates rise. This finding is consistent with research conducted by Wicaksono [3], who found that interest rates significantly negatively affect gold prices. It is by economic theory which suggests that economic factors, including interest rates, play an important role in determining gold prices. Analysis of the MAPE value of 16.49% shows that the prediction results provide a relatively accurate estimate of gold price movements. Although there is still an error in the prediction, this MAPE value is within the acceptable range for gold price analysis and forecasting in a dynamic market context.

4. CONCLUSIONS

They are based on the final distributed lag Almon transformation model with the best model selection using AIC. The model with the lowest AIC value is the Almon transformation distributed lag model with a maximum lag length of 16 and the second polynomial degree used to predict gold prices for one year. The accuracy of the prediction results of gold prices during the period January 2024 to December 2024 using the distributed lag model with Almon transformation is categorized as good, with a MAPE value of 16.49%. Suggestions for further research can use other methods and involve other factors such as inflation, dollar exchange rates, and so on to predict gold prices. They can also pay attention to geopolitical conditions and changes in market behavior when researching gold prices.

REFERENCES


GOLD PRICE PREDICTION IN INDONESIA BASED ON INTEREST RATE USING ...