



Comparison Of Jenkins Box Method And Multiple Linier Regression In Predicting The Noble Metal Price

Bayu Gunara¹, Yennimar²

^{1,2}Teknik Informatika,

^{1,2}Universitas Prima Indonesia, Jl. Sekip Sei Kambing Medan 20111, Indonesia

Email : bgunara1@gmail.com

ARTICLE INFO

Article history:

Received: 25/01/2020

Revised: 07/02/2020

Accepted: 18/02/2020

Keywords:

gold investment, forecasting, Multiple Linear Regression, Box-Jenkins, ARIMA, hybrid methods.

ABSTRACT

Nowadays, gold investment practitioners generally use instinct and guess in investing in gold. This is certainly a problem because it has a high error margin. To solve these problems, the forecasting process can be carried out. To be able to forecast gold prices with low error rates, various studies have been conducted. The Box-Jenkins method performs better than other methods in predicting the price of gold, because the Box-Jenkins method applies forecasting by relying on the historical statistics of gold prices beforehand. The Box-Jenkins method is an iterative of choosing the best model for the stationary series of a group of linear time series models called the ARIMA (Autoregressive Integrated Moving Average) model. However, the ARIMA method is a complex method and is not easy to use and requires a long execution time to obtain forecasting results with a high degree of accuracy. To improve the accuracy of prediction results from ARIMA, the ARIMA method can be combined with the multiple regression method into a hybrid method. The Multiple Linear Regression Method is a mathematical technique that minimizes the difference between the actual value and the predicted value. The results of this study are an application of forecasting the price of gold using the ARIMA method and Multiple Linear Regression. The application also provides a facility to test the results of the methods used. Based on the results of testing the accuracy of the prediction results from the hybrid method with 30 data = 48%, 60 data = 40%, and 118 data = 40.81%.

Copyright © 2020 Jurnal Mantik.
All rights reserved.

1. Introduction

Gold is a precious metal that is often used as a medium of exchange in trade and as a financial standard in various countries. To get optimal profits, people who want to invest in gold certainly expect to get a low price at the time of purchase and an expensive price at the time of sale. Between 1975 and 1979, for example, gold price fluctuations ranged from \$ 121.00 - \$ 236.10. In 1982 to 2005, the price range of gold was between \$ 200 - \$ 400. In 2016, the highest selling price of gold reached \$ 1365 [3]. In the current conditions, gold investment practitioners generally use instinct and guess in making gold investments [4]. Investors in gold usually buy gold when the price of gold is considered low enough in the hope of reselling it when the price of gold is high. This is certainly a problem because this method has a high margin of error, where the price of gold that is falling can continue to fall to a much lower point, this can actually cause large losses to gold investors if they have to sell the gold when the price of gold is much lower than the price when it was bought [4]. The economic factors that affect the price of gold, include. To solve these problems, the forecasting process can be done (forecasting). Forecasting is a function of management to support the decision making process. The forecasting process is also described as an estimation process for a situation that is not yet known [2]. To be able to forecast gold prices with low error rates, various studies have been conducted, Yuan (2012) uses an algorithm Projection Pursuit to find non-linear variables and construct a Backpropagation (BP) Neural Network to predict the price of gold [2]. Other research on correlation analysis and concludes that international business, politics, market





conditions, especially commodity markets, consumer spending trends, and inflation are factors that have a major influence on the movement of gold prices. Meanwhile, the Box-Jenkins method has better performance than other methods of forecasting prices gold, This is because the Box-Jenkins method applies forecasting by relying on historical gold price statistics, the more historical gold price data available, the forecast results will also be better and have a lower error rate. The Box-Jenkins method is an iterative of choosing the best model for series the stationary of a group of models time series linear called the ARIMA model (Autoregressive Integrated Moving Average). This method assumes that the value series produced by the process stochastic (random) with a form that can be explained [5]. According to Madhur Srivastava, et. al (2010), the ARIMA method is a complex method and is not easy to use and requires a long execution time to obtain forecasting results with a high degree of accuracy.

According to [6], to improve the accuracy of forecasting results from ARIMA, the ARIMA method can be combined with the method multiple regression. Madhur Srivastava, et. al. (2010) put forward a hybrid method between methods Regression and ARIMA in predicting the number of errors (bug) which will appear in the making software. Multiple Linear Regression can be used for interactions between variables dependent with several variables independent. Metode Multiple Linear Regression is a mathematical technique that minimizes the difference between the actual value and the estimated value [1]. Based on testing conducted by PT Madhur Srivastava, et. al. (2010), obtained information that the intermediate hybrid method ARIMA dan regression has a higher level of accuracy when compared to the ARIMA method only.

Based on the description above, this research will make an application of forecasting the price of gold with a hybrid method that combines the methods Box Jenkins Approach and Multiple Linear Regression to help people who want to invest in gold know the movement of gold prices in the future. Gold investment will be better if the investment knows the right time to buy gold when prices are low and sell gold when prices are high so that it can minimize losses and optimize the benefits of gold investment. With the application of this deepening is expected to overcome the problems that exist in gold investment.

2. Method

The system development method used is SDLC (System Development Life Cycle) or often referred to as the waterfall method.

The steps in this methodology are:

a) Needs Analysis

All needs will be defined individually. The analysis is done by asking the user what is needed in the development of the system, especially historical data on gold prices. Analysis of functional needs of the system created using the help of use case diagrams and analysis of non-functional requirements of the system using the framework of Performance, Information, Economics, Control, Efficiency, Service (PIECES)

b) Process Analysis

At this stage, the work process of the Hybrid method will be described using a flowchart. After that, the calculation process of the Hybrid method will be elaborated in the process of forecasting the price of gold using a simple example.

c) System design

At this stage, a comprehensive design of the system will be built. This design includes system modeling using use case diagrams, database design which includes table design and inter-table relations using Microsoft SQL Server 2014 applications, user interface design that includes output design and input design. The application used to design is Balsamiq Mockup.

d) Writing programs or coding (coding)

After all needs are defined, the next step is to translate the results of the design process into a form of computer program that is understood by computer machines. The algorithm used to predict is the algorithm Box Jenkins Approach and Multiple Linear Regression. Making coding of this forecasting application will use the Microsoft Visual C # 2015 programming language.

e) Testing the program (testing)

At this stage a trial is carried out on the program to look for abnormalities (bugs) that can still occur in the program. The testing process will use a data set of gold prices (closing price per day in the afternoon, taken from the website price-emas.org) starting in 2017 hingga 2018 dan will be





compared to the forecast results obtained with the actual price of gold in a certain period. From the results of this test, it will be known the accuracy of the results of forecasting the price of gold from the hybrid model between the Box Jenkins Approach method and Multiple Linear Regression. The testing process of the accuracy of the hybrid model between the Box Jenkins Approach algorithm and the Multiple Linear Regression will use the confusion matrix method for the values of p , q and $d = 1$. This testing process will use real gold price data as a comparison with the forecast price.

To improve the accuracy of the prediction results from the Box Jenkins (ARIMA) method, the ARIMA method will be combined with the Regression method into a hybrid model. The work process of this Hybrid method can be described in the form of a flowchart as shown in the following figure:

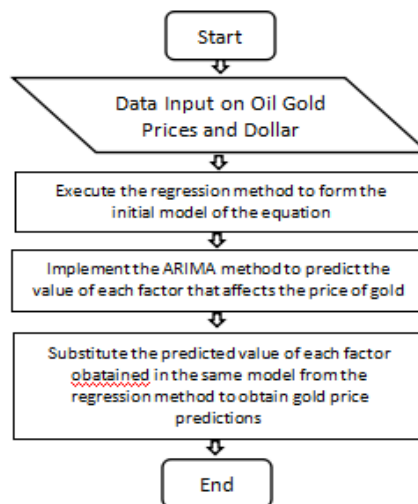


Fig 1. Flowchart of the Hybrid Method Process.

The work process of the linear regression method can be described in the form of a flowchart as shown in the following figure:

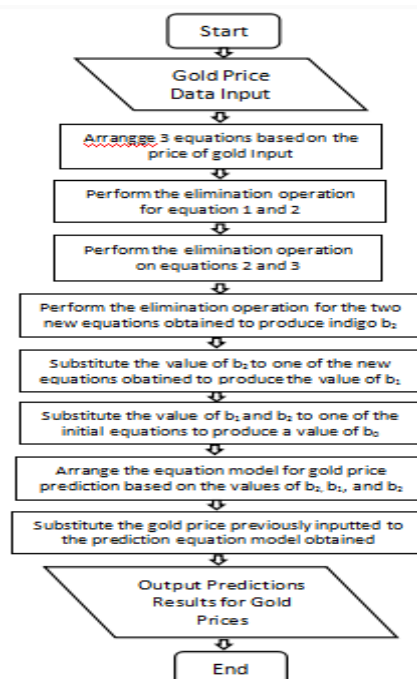


Fig 2. Flowchart of the Multiple Linear Regression Process Process



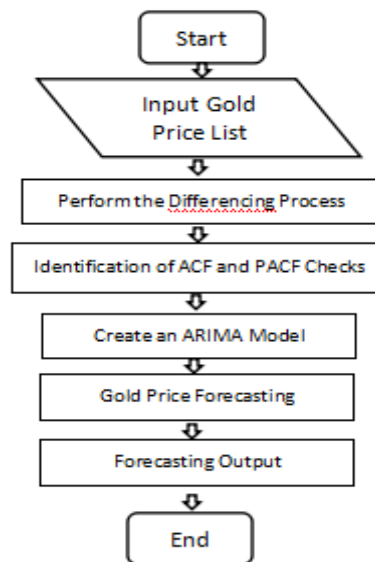


Fig 3. Flowchart of the ARIMA Method.

The working procedure of the ARIMA method can be described in the form of a flowchart as shown in the following figure :

3. Result And Discussion

3.1. Result

When you first run the gold price forecasting application with a hybrid model between the Box Jenkins Approach method and Multiple Linear Regression, the form that will appear first is the 'Login' form, which can be seen in Figure 4 below:

User ID	<input type="text" value="admin"/>
Password	<input type="password" value="*****"/>
<input type="button" value="Login"/> <input type="button" value="Batal"/>	
<input type="button" value="Login as Guest"/>	

Fig 4. Login Form.

If the user wants to fill in historical data on gold prices, oil prices and dollar exchange rates, the user must log in as an admin user. The user must enter a valid user ID and password, in order to use the admin section of the software.

If the login process is successful, then the system will display the Main form as shown in the following image:





Fig 5. Display Main Form.

In this Main form, there are several menus that can be accessed, such as:

- a) Main Menu, which serves to provide an interface to fill in the main data required by the system. This main menu has several sub menus, namely:
 - a. Gold Price Input sub menu, which functions to display the Gold Price Input form
 - b. Dollar Exchange Input sub menu, which functions to display the Dollar Exchange Price Input form.
 - c. Oil Price Input sub menu, which functions to display the Oil Price Input form..
 - d. Gold Price Forecast menu, which functions to forecast gold prices by displaying the Gold Price Input form.
- b) Report Menu, which functions to display the forecasting information using the Hybrid method.
- c) Testing Menu, which functions to carry out the testing process by displaying the Test Results form.
- d) Add New User Menu, which functions to display the Add User form.

In the 'Main Menu', there are several sub menus that can be accessed, such as:

- a) Gold Price Input Sub Menu, which functions to display the Gold Price Input form as shown in the following image:

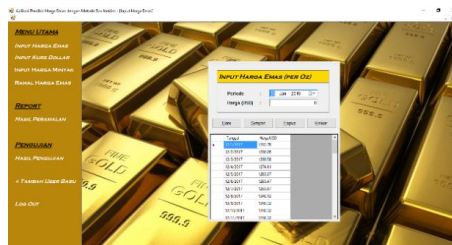


Fig 6. Form Input Gold Prices.

In this Gold Price Input form, users can enter historical details of the gold price, such as period and price (USD). After completing the historical details of the gold price, the user can click the Save button to save the data to the Kurs_Emas table in the database. Data stored in that database will be displayed in the Gold Price Data table.

- b) Dollar Exchange Input Sub Menu, which is performed to display the Dollar Exchange Data Input form as shown in the following figure:



Fig 7. Input Dollar Exchange Forms.

In this Dollar Exchange Input form, users can enter historical details of the dollar exchange rate, such as the date and price of the exchange rate conversion. After completing the details of the dollar exchange rate data, the user can click the Save button to save the data to the Kurs_Dollar table in the database. Data stored in that database will be displayed in the Dollar Exchange Data table.

- c) Oil Price Input Data Sub Menu, which functions to display the Oil Price Input Data form as shown in the following image:





Fig 8. Oil Price Input Forms.

On this Oil Price Input Data form, users can enter data on the historical details of oil prices, such as the date and price of oil. After completing the detailed oil price data, the user can click the Save button to save the data to the Oil table in the database. Data stored in that database will be displayed in the Oil Data table.

- d) Gold Price Prediction Sub Menu, which functions to display the Gold Price Forecasting form as shown in the following image:

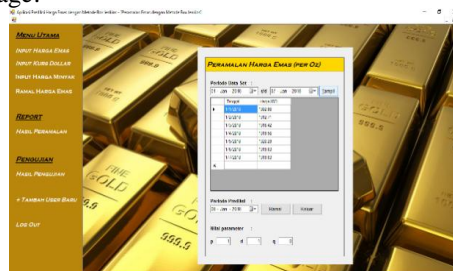


Fig 9. Form Forecasting Gold Prices.

In this Gold Price Forecasting form, users can enter the initial data needed in the forecasting process, such as the data set period that will be used in the forecasting process and the period that will be predicted. In addition, the user can also determine the parameter values p , d and q manually. These three parameter values are used in working procedures of the ARIMA method. After the user enters all the initial data needed, the user can click the Prediction button to start the forecasting process.

In this Forecasting Form, there are two main buttons that can be accessed namely:

- Forecasting Results button used to display all the forecasting results steps.
- The Previous Step button is used to display the previous calculation step.
- The Next Step button is used to display the next calculation step.

Add User Sub Menu, which functions to display the Add User form as shown in the following image:

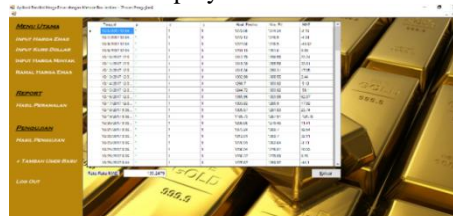


Fig 11. Form Add User

User data entered will be used when you want to log into the system. The details of user data entered include data user names, passwords.

The testing process can be done by clicking on the Test Results menu, so the system will display the Test Results form as shown in the following image:





Fig 12. Form Testing Results

In this form the average error will be calculated from the results of tests conducted by the user. The testing process will use the mean absolute error (MAE) method.

As shown in the picture above, based on testing using the MAE method performed on 30 data, 60 data and 118 data, obtained information that the average value of the error from the MAE method is 135.2479.

3.2. Discussion

Meanwhile, to test the accuracy of the predicted results of gold prices rising or falling, the Confusion Matrix method will be used, along with the details:

Based on the results of tests conducted on the software, the following information can be obtained:

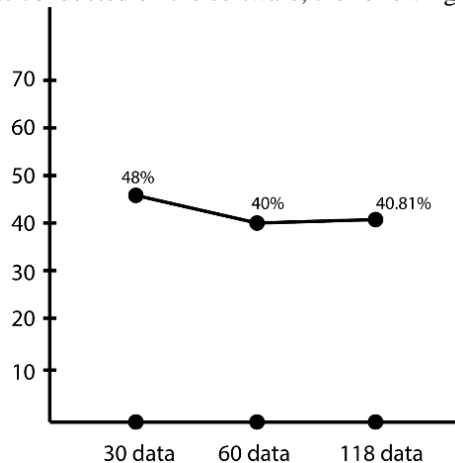


Fig 14. Outline of Test

- By using the MAE method to calculate the difference in the price of gold between the real value and the predicted value, information is obtained that the average error for 118 data obtained is equal to 135.2479
- Dengan By using the confusion matrix method to calculate the accuracy of the predicted results up or down, obtained information that the accuracy of the prediction results is equal to 40.81% for the dataset totaling 118 pieces of data.
- Hybrid method can be used to predict gold prices increase or decrease, but the level of accuracy is less satisfactory. To improve the accuracy of the predicted results, other supporting factors must be included that affect the price of gold.
- Hybrid method has a higher level of accuracy in predicting gold prices in the short term, but the level of accuracy is less good in predicting gold prices in the long run.

4. Conclusion

From the discussion in the previous chapters, finally the research in this final project can be drawn several conclusions, among others.

- Based on the results of tests conducted, information is obtained that the hybrid method can be used to forecast gold prices with accurate gold price forecasting which reaches 40.81% from 118 test data. This means that the gold price forecasting results from the Hybrid method can be used as a reference only when investing in gold.
- he hybrid method is suitable for use in forecasting the price of gold for the short term with an accuracy of 48% from 30 test data.





5. Reference

- [1] Bodie Z, Kane A, Marcus A. J. (2009). **Investments: eight edition**. McGraw Hill. USA.
- [2] Ismail, Z., A. Yahya dan A. Shabri. (2009). **Forecasting Gold Prices Using Multiple Linear Regression Method**, American Journal of Applied Sciences 6 (8): 1509-1514, ISSN 1546-9239, Science Publications.
- [3] Kitco. (2017). Kitco.com, **Gold Prices History**.
- [4] Mohamad Heykal, Erlin. (2011). **Analisis Kinerja Perdagangan Kontrak Berjangka Logam Mulia pada Periode Oktober 2009 - Desember 2009**, BINUS BUSINESS REVIEW Vol. 2 No. 1 Mei 2011: 181-191.
- [5] Puji Catur Siswipraptini, Sri Rahayu. (2016). **Aplikasi Simulasi dan Pemodelan Menggunakan Metode Linier Trend pada CV. Bina Multi Barokah**, P-ISSN 2085-4315 / E-ISSN 2502-8332.
- [6] Srivastava, M., D. Badal, R. K. Jain. (2010). **Regression and ARIMA Hybrid Model for New Bug Prediction**, Madhur Srivastava et al. / (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 08, 2010, 2622-2628.
- [7] Yuan, J., F. Z. Guang, Q. M . Fang. (2012). **Coal Calorific Value Prediction Based on Projection Pursuit Principle**, TELKOMNIKA Indonesian Journal of Electrical Engineering Vol.10, No.6, October 2012, pp. 1287~1292.
- [8] Anita. (2015). **Analisis Komparasi Investasi Logam Mulia Emas Dengan Saham Perusahaan Pertambangan Di Bursa Efek Indonesia 2010-2014**. Jurnal Bisnis dan Manajemen. Volume 5 No.2.
- [9] Tanti Octavia, Yulia, Lydia, **Aplikasi Simulasi dan Pemodelan Menggunakan Metode Linier Trend pada CV. Bina Multi Barokah**, Seminar Nasional Informatika 2013 (semnasIF 2013), UPN "Veteran" Yogyakarta 2013.
- [10] Chandra, A. (2017). **Estimasi Trip Frequency dengan Menggunakan Model Multiple Linear Regression**, Jurnal Metris 18 (2017) 25–28,

