Prediction of Rupiah Against US Dollar by Using ARIMA

Adiba Qonita, Annas Gading Pertiwi, Triyanna Widiyaningtyas
Electrical Engineering Department
Universitas Negeri Malang
Malang, Indonesia
qonitadiba@gmail.com, anheeys@gmail.com, triyannaw.ft@um.ac.id

***Abstract*—The currency exchanges rate is one of the most important things in the economy. The currency exchange rate is needed in the business word for example, investment and profit assessment. Prediction of rupiah rate is done to get the price of the rupiah against US dollar in the future to be used as consideration in decision-making, thereby reducing the risk of loss. Therefore, we need a method that can help in making business decisions about when to make the right trades with a high degree of accuracy. This study aims to predict the value of rupiah against US dollar by using ARIMA (Autoregressive Integrated Moving Average). This study uses four stages, including (1) the preparation of the dataset, (2) preprocessing of data, (3) the use of ARIMA models, (4) test accuracy. The data used for the test is the data rate from January 4th 2010 until June 24th 2016. The result showed that ARIMA method has an accuracy rate of 98.74%. Based on the result, it can be concluded that the development of the predictive value of the rupiah against the US dollar using ARIMA method was accurate to use.**

Keywords*—*exchange rate; dataset; ARIMA; USD

I. Introduction

The economic condition in Indonesia is still less stable, causing the exchange rate of US dollar against rupiah soared. This is due largely to international transactions in Indonesia country uses US dollars. Unstable economic conditions may reduce the amount of capital investment by foreign investor's retrogressive development in Indonesia because foreign investors are very large role in the economic growth rate in Indonesia [1].

In terms of buying and selling, prediction is chosen as the one of the important thing in selecting market to invest [2]. Every prediction that came out very calculated accuracy. The fact that the number of manual prediction errors is influenced by many factors, one of them is caused by human calculation error. This of course would be fatal to investors. In addition, the manual prediction will wasting time, thought and effort when the data that's used are large. Therefore, we need a method that can overcome the problems that cannot be solved manually.

Exchange rate is the price of a domestic currency against foreign currencies. Foreign exchange rate system will depend on the nature of the market. In a free market, the exchange rate will change according to changes in demand and supply [3]. Foreign exchange rates can be classified into selling rate, buying rate, and the middle rate. Buying rate is the exchange rate used when money dealers buy foreign currency. Selling rate is rate at which a bank selling the foreign currency. Then, the middle exchange rate is the value of the sum of the buying and selling rates were halved [4].

Periodic data (time series) is a set of observational data obtained overtime. In general, data collecting is done in a certain period. Periodic data can be used as a basis for decision-making as well as forecasting [5].

Autoregressive Integrated Moving Average (ARIMA) model is a model that is fully ignoring independent variable in forecasting [6]. The ARIMA model is one of forecasting technique of time series which is only based on the observed behavior of variable data. ARIMA models completely ignore the independent variable for this model using the present value and past values of the dependent variables to produce accurate short-term forecasting. ARIMA model is a combination of the Autoregressive (AR) model and Moving Average (MA) models.

Some relevant studies have been developed, such as research Arantika Cecilya et all on 2009 has examined "The US Dollar Exchange Rate Predictions Against Rupiah Using Genetic Algorithm and Elman Recurrent Neural Network". The application of Genetic Algorithm and Elman Recurrent Neural Network requires a lot of use of a combination of genetic parameters for testing the system to get good results [7].

Research that been done by Ahmad Amiruddin Anwary on 2011 has examined "Prediction Rupiah against the US dollar, using the method of Fuzzy Time Series" with a accuracy level of the predicted results measured by the Average Forecasting Error Rate (AFER) value. The study used data with a small amount, so that the prediction become less than optimal [8].

The research by Yulina Mahena et al (2015) has developed "The World Gold Price Predictions Of Gold Stock Investment Decision Support Using Data Mining Techniques". The study is quite accurate, but it would be better if you use an additional attribute such a request (demand), inventories (supply), and the condition of the world economy that can support more accurate results [9].

Alexander Setiawan et al have developed "Application Cosmetics Sales Forecasting using ARIMA Method" on 2015.

978-1-5386-0549-3/17/$31.00 ©2017 IEEE

This research resulted in the value of Mean Square Error (MSE) is small, but the results of forecasting are less accurate compared with data of the company [10].

Febi Satya P. has developed "The use of ARIMA Method to forecast Short Term Load Electricity Consumption" on 2015. The study resulted in the value of Mean Absolute Percentage Error (MAPE) is quite large, that is 6.03% [11].

There are a lot of algorithms that can be used to predict time series data other than ARIMA, such as Fuzzy Time Series. Fuzzy Time Series can predict time series data that has either short or long range and able to do calculation with string variables, but it's hard to find the parameter [8]. Other than that, Neural Network algorithms able to learn data's patterns and giving the best solutions from data that hasn't known yet, but it's hard to determine the amount of layers and neurons and the time used to give the parameter's configuration is relative long. Lots of training data are needed for good outcomes, while the more training data, the longer training processes took [12]. ANFIS Algorithms is also good at predicting time series data and big scale calculations, but have high complexity on computation and isn't really good for long range prediction [13].

Based on the background of the problem, it will be developed "Prediction of Rupiah against US Dollar by Using ARIMA". The purpose of this study is to predict the exchange rate of Rupiah against US dollar by using ARIMA method to determine the major currency exchange rate US dollars in the future as a material of consideration in making decision on transaction of business. Prediction of Rupiah against US dollar using ARIMA method is expected to help the parties concerned to predict the market quickly, easily and accurately with the predicted result that have high accuracy values.

The researchers used ARIMA method that produces an accurate short-term forecasting. In this study, researchers will only predict Rupiah against the US dollar because the dollar is relatively stable in the economy and more frequently used for international transactions in Indonesia.

II. Method

This study uses four stages, including (1) preparation of the dataset, (2) preprocessing of data, (3) the use of ARIMA method, and (4) test accuracy.

A. Preparation of The Dataset

The data used is the data of Rupiah against US dollar which is available on the official website of Bank Indonesia http://www.bi.go.id. The dataset consist of four attributes, which consists of the value, selling rate, buying rate and date. Data in the form of time series data are taken from the date of January 4, 2010 until the date of June 24, 2016 with a number of records as many as 1595 records. Figure 1 below is a dataset to see the exchange rate rupiah against the dollar.

EXCHANGE RATES ON TRANSACTION
CURRENCIES USD
Grafik Time Series

Sell	Buy	Date
13,168.00	13,036.00	22 Jul 2016
13,188.00	13,056.00	21 Jul 2016
13,166.00	13,034.00	20 Jul 2016
13,151.00	13,021.00	19 Jul 2016
13,178.00	13,046.00	18 Jul 2016
13,151.00	13,021.00	15 Jul 2016
13,153.00	13,023.00	14 Jul 2016
13,160.00	13,030.00	13 Jul 2016

Fig. 1. Dataset (source: Bank Indonesia)

Dataset's graphic display of Rupiah against the US Dollar are shown in Figure 2. Variable used in this research is the buying rate of the Rupiah against the US dollar.

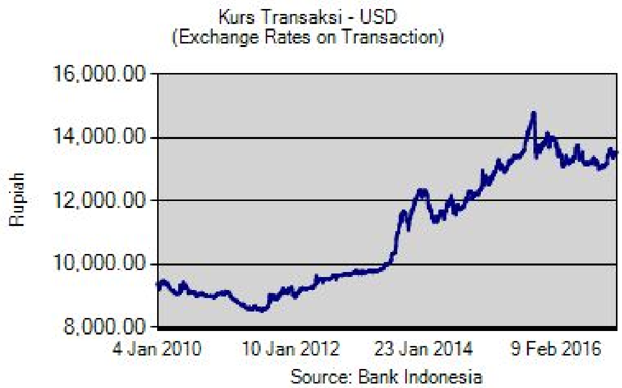

Fig. 2. Dataset Graphic

B. Preprocessing Data

Before the classification process carried out, it is necessary to do the data preprocessing stage first. This preprocessing stage includes (a) reduction of attributes, (b) data cleaning, (c) load data, (d) factors to numeric), (e) factors to date, and (f) stationarity testing.

Attribute reduction is the process of removing non-essential attribute of the data. In this case, the researchers decided to reduce the attribute value and the selling rate. So, the experiment is using just two remaining attributes, that is buying rate and date attribute. Buying rate attribute chosen because buying rate become a model for the bank or money changer to determine the selling rate. While the attributes of the date chosen for the time series data. The dataset's attributes after reduction is shown in Figure 3.

Buy	Date
13.411,00	2017-01-04
13.418,00	2017-01-03
13.369,00	2016-12-31
13.369,00	2016-12-30
13.406,00	2016-12-29
13.380,00	2016-12-28
13.369,00	2016-12-27
13.403,00	2016-12-23
13.368,00	2016-12-22
13.406,00	2016-12-21

Fig. 3. The dataset's attribute after reduction

Data cleaning is the process of changing the data format into a .csv format. Initial dataset used is still an .xlsx file, so as to be processed needs to be converted into the .csv format.

Load data is the process of loading the .csv format data. The following results displayed after the data load process successfully performed.

```
> data <- read.csv("E:/usd.csv")
> summary(data)
      Buy                  Date     
 Min.   : 8418   2010-01-04:   1  
 1st Qu.: 9069   2010-01-05:   1  
 Median : 9890   2010-01-06:   1  
 Mean   :10843   2010-01-07:   1  
 3rd Qu.:12914   2010-01-08:   1  
 Max.   :14654   2010-01-11:   1  
                 (Other)   :1719  
```

Fig. 4. The result of load data

Factor to Numeric is the process of data conversion from factor type into numeric type of data to be processed.

Attributes buying rate in the dataset is in the form factor of data, so it's need to be converted into numeric data types so that the data can be processed. Buying rate of data is changed so that no point marks that divide the started amount worth thousands. While the comma separator decimals is converted into a dot. The results of the dataset after the change factor to the numeric data types is shown in Figure 5.

	Buy	Date
1	13411	2017-01-04
2	13418	2017-01-03
3	13369	2016-12-31
4	13369	2016-12-30
5	13406	2016-12-29
6	13380	2016-12-28
7	13369	2016-12-27
8	13403	2016-12-23
9	13368	2016-12-22
10	13406	2016-12-21

Fig. 5. The result of dataset after change factor to numeric

Factor to Date is the process of converting the date data into factor. Attributes that changed in this process is the attribute date.

Dataset attribute date form is in factor type, so it is necessary to change the data type into data to be used as a reference time series changes. The date format used is yyyy-mm-dd. The following result shows the dataset after the changes in the data type factor to date.

```
Data
data          1725 obs. of 2 varial
   Buy : num 13411 13418 13369 1336
   Date: Date, format: "2017-01-04"
```

Fig. 6. The result of dataset after change factor to numeric

Stationarity test is a test to determine whether there isn't an increase or decrease in the data. Data transformation into stationer data by doing differencing, which calculates the change or difference in value of observation.

In the dataset's graph, there's increase and decrease in the data, so that data needs to be changed by differencing so the data became stationary. One of the ways is with differencing stochastic process that is subtracting time series dataset with its unit's squared root. For example, a time series data have squared root equation as follows $Y_t = y_{t-1} + \mu_t$

So, the stochastic difference is:

$$\Delta Y_t = y_t - y_{t-1} = \mu_t \quad (1)$$

The amount of differencing process that has been done will determine ordo of coefficient d on ARIMA. If the data been through differencing process *d* times to be stationer, so that data is called non stationer homogeny level *d* ARIMA(0,d,0). On current prediction, the dataset is stationer that can be seen on the following graph.

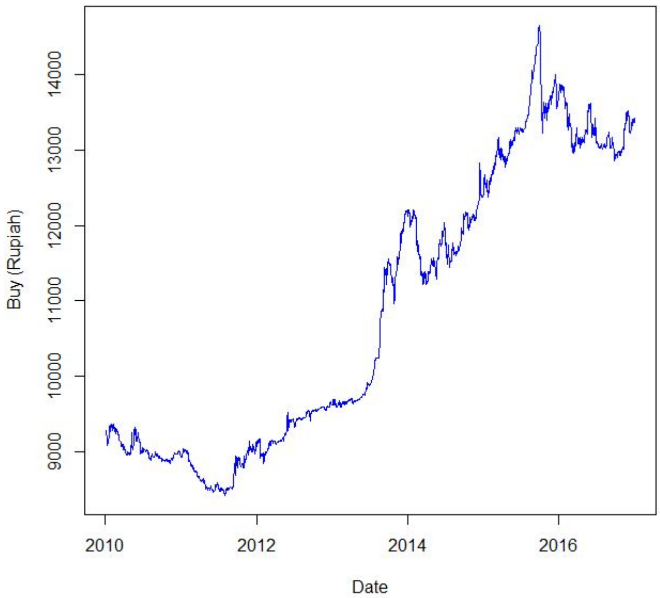

Fig. 7. Stasioner dataset

The following shows the result of differencing, which express the data stationary.

```
> adf.test(data$Buy)

        Augmented Dickey-Fuller Test

data:  data$Buy
Dickey-Fuller = -1.1488, Lag order = 11, p-value =
0.9147
alternative hypothesis: stationary
```

Fig. 8. Differencing Process

C. ARIMA (Autoregresif Integrated Moving Average) Model

Autoregressive Integrated Moving Average (ARIMA) model is a model that is fully ignoring independent variable in forecasting [6]. ARIMA suits observation of time series that are statistically correlated with each other or dependent [14].

ARIMA methods were divided into three groups, namely: autoregressive (AR) method, moving average (MA) method, and the mixed model autoregressive moving average (ARMA), which has the characteristics of the first two models.

Autoregressive calculation can be performed to determine the appropriate model with time series, determine the value of the order of p (determine the length of the equation that is formed), and can estimate the value of autoregressive coefficient ($Ø_1$, $Ø_2$, ... , $Ø_p$). Autoregressive mode with order AR (p) or ARIMA (p, 0, 0) is expressed as follows:

$$X_t = \mu + Ø_1 X_{t-1} + Ø_2 X_{t-2} + ... + Ø_p X_{t-p} + e_t \qquad (2)$$

where:

$Ø_p$ = Autoregressive parameters to *p*

e_t = *White Noise,* error value at time *t*

μ = *independent variable*

The independent variable is a row of similar value of the variable in period t last. While e_t is an error that describes the random noise that can not be explained by the model [15].

Another model of ARIMA model is a moving average (MA) denoted in the MA (q) or ARIMA (0, 0, q). MA (q) is a model that sees the movement of variable by the past's residual, which is written in the following equation.

$$X_t = \mu + e_t - Ø_1 e_{t-1} - Ø_2 e_{t-2} - ... - Ø_q e_{t-q} \qquad (3)$$

where:

$Ø_q$ = Parameter of Moving Average

e_{t-k} = White noise / error at time *t-k*

μ = constant

The equation above shows that the values of *Xt* depending on the value of the previous error of the value of the variable itself. Approach between the autoregressive and moving average are required autocorrelation measurement between successive values of *Xt*. The model of the moving average measure of autocorrelation between error value [15].

The combination of AR and MA model will form a new model, namely ARMA (autoregressive moving average) with order ARMA (*p,q*) The general form of the ARMA equation is a combination of the AR and MA equation are denoted as follows.

$$X_t = \mu + Ø_1 X_{t-1} + Ø_2 X_{t-2} + ... + Ø_p X_{t-p} + e_t - Ø_1 e_{t-1} - Ø_2 e_{t-2} - ... - Ø_q e_{t-q} \qquad (4)$$

D. Test Accuracy

Test accuracy can indicate closeness results of measurement to the true value. A measurement system can be accurate and precise, or accurate but not precise, or precise but innacurate or imprecise and inaccurate [16].

The test accuracy using Mean Absolute Percent Error (MAPE), which is an alternative method to evaluate the error based on the percentage of the absolute errors. MAPE counts the deviation between actual data with forecasting value and then it counts the average of percentages. MAPE equation can be written as follows.

$$MAPE = \frac{\Sigma(\text{The sum of error})}{\text{The total of data}} 100\% \qquad (5)$$

III. Result and Discussion

A. The test prediction with ARIMA

The test predictions done by defining the number of days that would be predicted. In this test, researchers predicted buying rate for the next 30 days. The result of the prediction is shown in Figure 9.

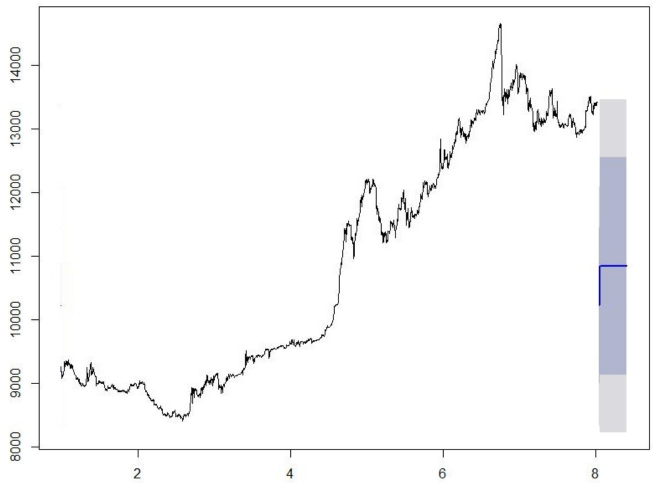

Fig. 9. The result of 30 day chart predictions

The result of the predictive value of Rupiah against US dollar by using ARIMA can be seen in Figure 10.

Fig. 10. The prediction result of Rupiah against US Dollar in 2016

where,
Actual = The actual exchange rate
Rate = The tendency of the exchange rate
Forecast = The result of prediction
Lower bound = Minimum limit
Upper Bound = Maximum limit

On the data on June 24, 2016 above, the value of predicted value, the lower and upper is not known because prediction calculation begin on June 25, 2016. While on June 25, 2016 below, the actual value and the trend is unknown because there has been no actual data to the exchange rate on June 24, 2016. The data on June 25, 2016 to 29 days ahead is the result of the prediction. Table 1 below is a detail of the graph the predicted outcome of Rupiah against US dollar using ARIMA method.

TABLE I. THE PREDICTION RESULT USING ARIMA

	actual	trend	pred	lower	upper	date
1724	9261	10205.937	NA	NA	NA	2010-01-05
1725	9283	9917.934	NA	NA	NA	2010-01-04
1726	NA	NA	10221.49	12089.66	8353.321	2017-01-05
1727	NA	NA	10843.21	13457.87	8228.562	2017-01-06
1728	NA	NA	10843.21	13457.87	8228.562	2017-01-07
1729	NA	NA	10843.21	13457.87	8228.562	2017-01-08
1730	NA	NA	10843.21	13457.87	8228.562	2017-01-09
1731	NA	NA	10843.21	13457.87	8228.562	2017-01-10

B. Accuracy Test

Accuracy test using MAPE method produces MAPE value of 1.259442. So that the accuracy of test results in this study has an accuracy of 98.74%. The test result of MAPE is shown as follows.

```
> accuracy(arma.preds, myTs_test)
               ME     RMSE      MAE      MPE     MAPE
Test set 106.5404 154.4969 115.8949 1.154695 1.259442
```

Fig. 11. Accuracy test

IV. CONCLUSSION

Based on the results we concluded that the best model of ARIMA method to predict the value of the rupiah against the US dollar is the ARMA model (2,2).

ARMA model (2,2) has a MAPE test results for 1.259442 with an average accuracy of 98,74%. Thus, we can conclude that ARIMA is a feasible method to predict the value of rupiah against the US dollar, where the results of study stated that the rupiah exchange rate against the dollar for 30 days from June 25, 2016 decreased slightly.

References

[1] D. Ernita, S. Amar, and E. Syofyan, “Analisis Pertumbuhan Ekonomi, Investasi, dan Konsumsi Di Indonesia,” J. Kaji. Ekon., vol. I, no. 02, pp. 176–193, 2013.

[2] R. H. Kusumodestoni, “Prediksi Forex Menggunakan Model Neural Network,” vol. 6, no. 2, pp. 205–210, 2015I.S. Jacobs and C.P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.

[3] N. Muhammad, “Metode Penelitian,” in Metode Penelitian, Jakarta, Indonesia, 1988, ch. 1, pp. 5.

[4] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[5] J. Hanke and D. Wichern, Business forecasting, 1st ed. [Chennai]: Pearson, 2015.

[6] Hendranata, Anton. 2003. ARIMA (Autoregressive Moving Average). Manajemen Keuangan Sektor Publik FEUI.

[7] A. Cecilya, R. N. Dayawati, and A. T. Wibowo, “Prediksi Nilai Tukar Us Dollar Terhadap Rupiah Menggunakan Algoritma Genetika Dan Elman Recurrent Neural Network.,” 2009.

[8] A. A. Anwary, “Prediksi Kurs Rupiah Terhadap Dollar Amerika Menggunakan Metode Fuzzy Time Series,” Universitas Diponegoro, 2011.

[9] Y. Mahena, M. Rusli, and E. Winarso, “Prediksi Harga Emas Dunia Sebagai Pendukung Keputusan Investasi Saham Emas Menggunakan Teknik Data Mining,” vol. 2, no. 1, pp. 36–51, 2015.

[10] A. Setiawan, A. Wibowo, and S. WIjaya, “Aplikasi Peramalan Penjualan Kosmetik dengan Metode ARIMA,” Konf. Nas. Sist. Inf., pp. 1–10, 2013.

[11] F. Satya, “Penggunaan Metode ARIMA Untuk Prakiraan Beban Konsumsi Listrik Jangka Pendek,” 2015.

[12] Setiawan, W. (2008). Tiruan Multilayer Feedforward Network 108–113.

[13] Fitriah, A., & Abadi, A. M. (2011). T – 2 Aplikasi Model Neuro Fuzzy Untuk Prediksi Tingkat Inflasi Di Indonesia, (1997).

[14] Makridakis, Spyro. , Steven C. Wheelwright, dan Victor E. McGee. Metode dan Aplikasi Peramalan, Jakarta: Erlangga, 1999.

[15] G. J. D. a. R. G. Box, Time Series Analysis Forecasting and Control, New Jersey: John Wiley & Sons Inc Publication, 2008. J. R. Taylor, Introduction to Error Analysis : The Study of Uncertainties in Physical Measurements, University Science Books, 1999.

[16] Khadafy, Al Riza & Wahono, Romi S. (2015). Penerapan CART untuk Mengurangi Data Noise pada Klasifikasi Multi Kelas dengan Decision Tree. Journal of Intelligent Systems, Vol. 1, No. 2