The Autoregressive Integrated Moving Average Model in Forecasting Philippine Peso - United States Dollar Exchange Rates

Nhoriel I. Toledo
University of the Philippines, Diliman, Quezon City, Philippines
Centro Escolar University, Mendiola, Manila, Philippines
E-mail: nitoledo@up.edu.ph; nitoledo@ceu.edu.ph

Abstract. It is known that in every quotation of exchange rates, there are two currencies involved. Exchange rate is defined as the price of a unit of foreign currency in terms of a domestic currency. This research aims to provide the best Autoregressive Integrated Moving Average (ARIMA) model that forecasts the Philippine Peso - United States Dollar (PHP-USD) exchange rates using historical average monthly rates from January 2009 to December 2016. Forecasted values from different ARIMA models are compared in terms of errors with actual data values. The best ARIMA model produced is ARIMA(0, 2, 2), which yielded the lowest values for the different performance measures used.

1. Introduction
Exchange rate is significant when comparing local and overseas markets in terms of services, goods, and different financial assets quoted in different currencies. Movements in the exchange rate can affect actual inflation and forecast future price movements. Movements can also affect the external sector since it will give an idea on the impact of foreign trade. Lastly, exchange rates can affect a country’s foreign debt, specifically the cost of servicing such as principal and interest payments. Peso appreciation reduces the amount of peso needed to buy foreign exchange to pay interest and maturing obligations.1

Many studies focus on forecasting exchange rates. Since there are a lot of factors to consider in forecasting these rates, there is still no exact solution or formula on how to forecast it. With the use of different models, the main goal is to find a time-series analysis tool that will give the minimum error on its forecast. A study by Babu and Reddy [1] focuses on forecasting exchange rates using ARIMA, Neural Network and Fuzzy Neuron. They examined the behavior of daily exchange rates from January 2010 to April 2015 of Indian Rupee (INR) against USD, British Pound (GBP), Euro (EUR) and Japanese Yen (JPY). The study resulted to ARIMA as the best model since it yielded the lowest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) scores. Mong Uyen Ngan [2], in his study, showed that ARIMA is suitable for estimating foreign monthly exchange rates in Vietnam from January 2013 to December 2015. Yildiran and Fettahoglu [3] forecasted Turkish Lira - USD exchange rates using ARIMA and concluded that the best model for short-term models, which includes daily exchange rates from August 3, 2016 to March 8, 2017, is ARIMA(2, 1, 0); and for long-term models, which includes daily exchange rates from January 3, 2005 to March 8, 2017, is ARIMA(0, 1, 1).

1 http://www.bsp.gov.ph/downloads/Publications/FAQs/exchange.pdf
A local study in the Philippines by Urrutia, Olfindo and Tampis [4] showed that ARIMA (0, 1, 0) is the best model for forecasting PHP-USD exchange rates using data from the 1st quarter of 2004 to the 4th quarter of 2014. Another study by Martinez and Gaw [5] concluded that ARIMA(1, 1, 2) is the best model for forecasting average monthly exchange rates using data from November 1995 to July 2013.

As studies on forecasting exchange rates in the Philippines is minimal, this research wishes to enrich the ideas and knowledge on how to forecast exchange rates in the Philippines. In particular, this study aims to find the most suited ARIMA implementation in forecasting PHP-USD exchange rates using a more recent monthly average historical data, where seasonality is less observed.

2. ARIMA Model

Forecasting is difficult as time series is stochastic in nature. Hence, we cannot forecast with certainty what will occur in the future. With the assumption of stationarity, the mathematical complexity of the different fitted models is reduced [6].

A stationary time series in general has statistical properties such as the mean, variance, autocorrelation, which are all constant over time. The ARIMA model basically stands for the following:

1. AR or autoregression, which is a model that uses the dependent relationship between an observation and some number of lagged observations.
2. I or integrated, which uses differencing of raw observations in order to stationarize a time series.
3. MA or moving average, which uses the dependency between an observation and a residual error from a moving average model applied to lagged observations.

Each of these components is explicitly specified in the model using a parameter. Note that AR and MA models are two widely used linear models that work on stationary time series, and I is a preprocessing procedure to make the time series stationary, if necessary. Note that both autocorrelation function (ACF) and partial autocorrelation function (PACF) can be used for determining the ARMA model hyperparameters p and q. [7]

| Table 1. Determining p and q of ARMA models |
|---------------------------------------------|
| ACF | AR(p) | MA(q) | ARMA(p,q) |
| PACF | Tails off | Cuts off after lag q | Tails off |
| PACF | Cuts off after lag p | Tails off | Tails off |

Based on Table 1, tails off indicates becoming zero in a gradual process while cuts off indicates becoming zero abruptly. If the PACF of the stationary series displays a sharp cutoff and the lag-1 autocorrelation is positive, that is, if the series appears slightly underdifferenced, then consider adding an AR term to the model. The lag at which the PACF cuts off is the indicated value of p. If the ACF of the stationary series displays a sharp cutoff and the lag-1 autocorrelation is negative, that is, if the series appears slightly overdifferenced, then consider adding an MA term to the model. The lag at which the ACF cuts off is the indicated value of q. Note that the selection for p and q is not unique. Other criterions can be used for choosing p an d q, such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which are both penalized-likelihood criteria. They are used for choosing the best forecast subsets in regression and often used for comparing models, which ordinary statistical tests cannot do. The lower the value of both AIC and BIC, the better the model is.

As we have seen, ARIMA models have numerous parameters and hyperparameters, Box and Jenkins [8] suggested an iterative three-stage approach to estimate an ARIMA model. The following are the procedures:

\[2\text{ https://people.duke.edu/~rmau411arim3.htm}\]
\[3\text{ methodology.psu.edu/resources/AIC-vs-BIC/}\]
1. Model identification. Includes checking of stationarity and seasonality. Differencing is performed, if necessary, to choose the model specification.
2. Parameter estimation. Involves computation of parameters that best fit the selected ARIMA model using MLE or non-linear LS estimation.
3. Model checking. Testing if the obtained model conforms to the specifications of a stationary univariate process, where the residuals should be independent of each other and have constant mean and variance.

The most important thing in model specification is to determine if the time series is stationary or not. Well-established statistical tests like the Kwiatkowski-Phillips-Schmidt-Shin Test can help test stationarity [9].

2.1. Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test
The null hypothesis for the KPSS Test states that the time series is stationary around a level random walk, while its alternative hypothesis states that the time series is non-stationary. Another important requirement for ARIMA models is for the residuals to be independent. With this, the Ljung-Box test will be used.

2.2. Ljung Box Test
The null hypothesis for Ljung Box Test states that the data are independently distributed while its alternative hypothesis states that data are dependent. Aside from independence of the residuals, another requirement for the residuals is to have constant mean and variance following the concept of a normal distribution. The Jarque-Bera Test will be used to test the normality of the residuals.

2.3. Jarque-Bera Test
The null hypothesis for the Jarque-Bera Test states that there is no significant difference between the given distribution and a normal distribution while its alternative hypothesis states that there is a significant difference between the given distribution and a normal distribution.

Once the ARIMA model is specified, the forecasted values will be checked in comparison to the actual values. This research will focus on mean absolute error (MAE), root mean square error (RMSE), mean percentage error (MPE) and mean absolute percentage error (MAPE) as performance measures of our model.

3. Data and Implementation
Data used in this study are average monthly PHP-USD exchange rates from January 2009 to December 2018 obtained from BSP. The division of the training and test sets is adapted from the concept of Pareto distribution which divides the data into 80%-20% ratio, which is applied in different areas of study. Implementation of ARIMA will be done using the R-software.

Step 1: Stationarize the model and determine the order of differencing using the KPSS test.

We test if the given time series data is stationary or not to determine if we need differencing on the data. We will be using the KPSS test for this, with 5% significance level, through the kpss.test() function from the tseries library in R.

| Differentiating | p-value     |
|----------------|------------|
| d = 0          | 0.02447    |
| d = 1          | 0.03216    |
| d = 2          | p > 0.1    |

Table 2. Differencing using the KPSS test

https://betterexplained.com/articles/understanding-the-pareto-principle-the-8020-rule/
Table 2 shows that the p-value is utilized using the KPSS test. Note that for the test, the null hypothesis must not be rejected. That is, the p-value must be greater than $\alpha$, in order to achieve stationarity. Since the p-value is less than $\alpha$, we reject the null hypothesis which shows that the time series is non-stationary. Going with second differencing, with d = 2, the test showed that the p-value is greater than 0.1, which is now greater than 0.05. Hence, we will not reject the null hypothesis. This will now conclude that the data is stationary. In conclusion, for this set of data, we specify d = 2.

Step 2: Determine the possible values of p and q on ARIMA(p,d,q) using ACF, PACF, AIC and BIC.

Figure 1. ACF of Average Monthly PHP-USD Exchange Rates with d = 2

Figure 2. PACF of Average Monthly PHP-USD Exchange Rates with d = 2

After the data is transformed into a stationary series, we will determine the values for p and q using ACF and PACF of the stationarized series. Note that ACF starts with lag-0 and PACF starts with lag-1. Recall that if the ACF of the stationary series displays a sharp cutoff and the lag-1 autocorrelation is negative, then we can consider adding an MA term to the model. Observe in Figure 1 that there is a sharp cutoff at lag-0 and the lag-1 autocorrelation is negative, hence we can add an MA term. Moreover, ACF cuts off at lag-2. So, we have an MA(2) model which indicates q = 2. Recall also that if the PACF of the stationary series displays a sharp cutoff and the lag-1 autocorrelation is positive,
then we can consider adding an AR term to the model. Observe in Figure 2 that the lag-1 autocorrelation is negative. Hence, we cannot consider adding an AR term to the model. With this, we have an AR(0) model which indicates $p = 0$. In conclusion, using the ACF and PACF of the stationarized time series data, the ARIMA model is specified as ARIMA(0, 2, 2). Since the selection of orders $p$ and $q$ is not unique, we can consider other values close to the specified ones. Afterwards, we evaluate AIC and BIC scores of the different models to determine the best among them. We consider the following models as shown in Table 3. Note that these models vary from the specified ARIMA(0, 2, 2) by adding or subtracting 1 from $p$ or $q$. The AIC and BIC values obtained using the arima() function from the forecast library and the BIC() function from the lme4 library in R are given as follows.

**Table 3. AIC and BIC scores of different considered ARIMA models**

| Model           | AIC Score | BIC Score |
|-----------------|-----------|-----------|
| ARIMA (0, 2, 1) | 153.78    | 158.86    |
| ARIMA (0, 2, 2) | 152.57    | 160.20    |
| ARIMA (0, 2, 3) | 154.57    | 164.74    |
| ARIMA (1, 2, 1) | 152.7     | 160.33    |
| ARIMA (1, 2, 2) | 154.57    | 164.74    |
| ARIMA (1, 2, 3) | 156.57    | 169.29    |

Observe in Table 3 that ARIMA(0, 2, 2) and ARIMA(0, 2, 1) have the lowest AIC and BIC scores, respectively. So we consider these two models as our candidates for the best ARIMA model in our study.

Step 3: Determine if the residuals of the candidates for the best ARIMA model forecast are independent and identically normal distribution.

**Table 4. P-value of L-jung Box Test for ARIMA(0, 2, 1) & ARIMA(0, 2, 2) residuals with lags**

| ARIMA Model Residuals with given lag | p-value of L-jung Box Test |
|--------------------------------------|----------------------------|
| ARIMA (0, 2, 1) Residuals with lag = 1 | 0.09315                    |
| ARIMA (0, 2, 1) Residuals with lag = 24 | 0.5828                     |
| ARIMA (0, 2, 2) Residuals with lag = 1 | 0.9379                     |
| ARIMA (0, 2, 2) Residuals with lag = 24 | 0.8571                     |

We check independence of the residuals with the Box.test function from the stats library in R. Based on Table 4, we used the residuals of ARIMA(0, 2, 1) and ARIMA(0, 2, 2) with lag = 1 and lag = 24. We used these lags to check the independence for short-term and long-term lags. Since this research will produce 24 months of forecast, we need to check the independence as well at the same lagged interval. Note that all the p-values are greater than the defined $\alpha = 0.05$. Therefore, the residuals of ARIMA(0, 2, 1) and ARIMA(0, 2, 2) with lag = 1 and lag = 24 are all independently distributed.

**Table 5. P-value of Jarque-Bera Test for ARIMA(0, 2, 1) & ARIMA(0, 2, 2) residuals**

| ARIMA Model Residuals | p-value of Jarque-Bera Test |
|-----------------------|----------------------------|
| ARIMA (0, 2, 1) Residuals | 0.7081                     |
| ARIMA (0, 2, 2) Residuals | 0.6909                     |

To check normality of the residuals, we use the jarque.bera.test function from the tseries library in R. Observe in Table 5 that all p-values are greater than $\alpha = 0.05$, which shows that the residuals of
ARIMA(0, 2, 1) and ARIMA(0, 2, 2) have no significant difference to a normal distribution.

![Figure 3. Forecast plot of ARIMA (0, 2, 1)](image1)

![Figure 4. Forecast plot of ARIMA (0, 2, 2)](image2)

Step 4: Utilize the candidates for the best ARIMA model to forecast, and check their performance measures to identify the best model.

Based on the forecast plots given in Figures 3 and 4, in consistency with the forecasted values, the graph tends to increase. The blue lines on the plots represent the point forecast values while the gray areas represent the 95% confidence level for the range of forecasted values.

After forecasting, we now determine the best ARIMA model by evaluating various performance measures defined.

| ARIMA Model | MAE       | RMSE      | MPE        | MAPE       |
|-------------|-----------|-----------|------------|------------|
| ARIMA (0, 2, 1) | 1.075385126 | 1.231078925 | -2.085059488 | 2.085059488 |
| ARIMA (0, 2, 2) | 0.729677626  | 0.896877497  | -1.412659505  | 1.424301661  |

Observe from Table 6 that ARIMA(0, 2, 2) yields performance measure values that are closer to zero. Thus, we say that the best ARIMA model for forecasting PHP-USD exchange rates, using our data, is ARIMA(0, 2, 2).

The auto.arima function [10] from the forecast library in R, developed by Hyndman and Khandakar, has a similar concept with the one presented in the manual algorithm. The best ARIMA model for our given time series, however, is automatically generated by this function. The following is the algorithm of the auto.arima function.

Note that for the auto.arima function in R, both seasonality and non-seasonality components of the data can be utilized as candidates for best ARIMA models. The model with the lowest AIC out of these variations is the best ARIMA model. Results show that ARIMA(0, 2, 2) is the best model since it has the lowest AIC score of 152.8395. Note that this model is the same ARIMA model obtained in the manual algorithm.

4. Conclusion

This research aimed at finding the best ARIMA implementation on forecasting values of average monthly PHP-USD exchange rates using data from January 2009 to December 2018. USD was utilized since this currency has a great effect on all of the other currencies in the international market. With the concept of ARIMA on forecasting, which focuses on past values including its own lagged forecast errors, we can find the general trend of how a certain currency performs in relation to other currencies. Results show that ARIMA(0, 2, 2) is the best ARIMA model given our data. Forecasted values show that the value of PHP in comparison to USD weakens for the forecast of 2 years time.
One of the limitations of ARIMA as a model for exchange rates is that its only basis are past values. Note that there are lots of fundamental economic variables that determine exchange rates, such as gross national product (GDP), gross domestic product (GDP), consumption, trade balance, inflation rates, interest rates, unemployment, and productivity indexes. Daily exchange rates can also be considered, which may yield different results in comparison to average monthly exchange rates. One can also compare PHP with other currencies. Forecasting models indeed fit the past data well, but it is another issue if these will be replicated in the future. These forecasts might not fit well in the future, but it may help us prepare for future decisions and plans.

5. Acknowledgments
The author would like to acknowledge the guidance of this research adviser, Dr. Jasmin-Mae Santos-Luy from University of the Philippines. This research is also with the guidance of Dr. Avelina Raqueno from Centro Escolar University who helped the author in constructing and finalizing the topic.

References
[1] Babu A and Reddy S 2015 Exchange Rate Forecasting using ARIMA, Neural Network and Fuzzy Neuron Journal of Stock and Forex Trading 4 Issue 3 (Finance and Accounting Institute of Management Technology (IMT), Hyderabad, Telangana, India)
[2] Mong Uyen Ngan, T 2016 Forecasting Foreign Exchange Rate by using ARIMA Model: A Case of VND/USD Exchange Rate Research Journal of Finance and Accounting 7 Issue 12 (School of Economics, Huazhong University of Science and Technology (HUST), Wuhan, P.R. China)
[3] Yildiran, CU and Fettahoglu A 2017 Forecasting USDTRY rate by ARIMA method Cogent Economics and Finance 5 (Accounting and Finance, Kocaeli University, Kocaeli, Turkey)
[4] Urrutia, J, Olfindo ML and Tampis R 2015 Modelling and Forecasting the Exchange Rate of the Philippines: A Time Series Analysis American Research Thoughts 1 Issue 9 (Polytechnic University of the Philippines, Paranaque Campus, Philippines)
[5] Hyndman R and Khandakar Y 2008 Automatic Time Series Forecasting: The forecast Package For R Journal of Statistical Software 27 Issue 3 Retrieved from: http://www.jstatsoft.org/.
[6] Adhikari R 2019 An Introductory Study on Time Series Modeling and Forecasting Retrieved from: https://arxiv.org/ftp/arxiv/papers/1302/1302.6613.pdf.
[7] Zhang, M 2008 Time Series: Autoregressive Models AR, MA, ARMA, ARIMA Lecture Notes University of Pittsburgh Retrieved from: http://people.cs.pitt.edu/milos/courses/cs3750/lectures/class16.pdf
[8] Box G and Jenkins G 1970 Time Series Analysis: Forecasting and Control (Holden-Day, San Francisco, California, United States of America)
[9] Kwiatkowski D, Phillips P, Schimdt P and Shin 1992 Testing the null hypothesis of stationary against the alternative of a unit root Journal of Econometrics 54 159-178 (North-Holland)
[10] Martinez, J and Gaw H 2013 Time-Series Forecasting: Philippine Exchange Rate With Respect to US Dollars De La Salle University (Manila, Philippines)