Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 127(3) (2019) 225-247 EISSN 2392-2192

Application out-of-sample forecasting in model selection on Nigeria exchange rate

Akpensuen Shiaondo Henry*, **K. E. Lasisi**, **E. A. Akpan** and **A. A. Gwani**

Department of Mathematical Sciences, Abubakar Tafawa Balewa University, Bauchi, Nigeria

*E-mail address: henroo85@rocketmail.com

ABSTRACT

In time series, several competing models may adequately fit a given set of data. At times choosing the best model may be easy or difficult. However, there are two major model selection criteria; it could be either in-sample or out-of-sample forecasts. This study was necessitated because Empirical evidence based on out-of-sample model forecast performance is generally considered more trustworthy than evidence based on in-sample model performance which can be more sensitive to outliers and data mining. And also the fact that Out-of-sample forecasts also better reflect the information available to the forecaster in real time was also an added motivation. Hence this study considered data from Nigeria exchange rate (Naira to US Dollar) from January 2002 to December 2018 comprising 204 observations. The first 192 observations were used for model identification and estimation while the remaining 12 observations were holdout for forecast validation. Three ARIMA models; ARIMA (0, 1, 1), ARIMA (1, 1, 2) and ARIMA (2, 1, 0) were fitted tentatively. Base on in-sample information criteria ARIMA (0, 1, 1) was the best model with minimum AIC, SIC and HQ information criteria. However, on the basics of out-of-sample forecast evaluation using RMSE, MSE, MAE, and MAPE, ARIMA (2, 1, 0) perform better than ARIMA (0, 1, 1). The implication of this study is that, a model that is best in the in-sample fitting may not necessary give a genuine forecasts since it is the same data that is used in model identification and estimation that is also use in forecast evaluation.

***Keyword*:** ARMA model, In-sample forecasting, Model selection and evaluation, Out-sample forecasting, Exchange Rate

(Received 04 April 2019; Accepted 22 April 2019; Date of Publication 23 April 2019)

1. INTRODUCTION

Time series modeling is a dynamic research area which has attracted attentions of research Community over last few decades. The main aim of time series modeling is to carefully collect and rigorously study the past observations of a time series to develop an appropriate model which describes the inherent structure of the series. This model is then used to generate future values for the series, i.e. to make forecasts. Time series forecasting thus can be termed as the act of predicting the future by understanding the past [15]. Due to the indispensable importance of time series forecasting in numerous practical fields such as business, economics, finance, science and engineering, and many more [18], proper care should be taken to fit an adequate model to the underlying time series. It is obvious that a successful time series forecasting depends on an appropriate model fitting. A lot of efforts have been done by researchers over the years for the development of efficient models to improve the forecasting accuracy. As a result, various important time series forecasting models have evolved over the years.

One of the most popular and frequently used stochastic time series models is the Autoregressive Integrated Moving Average (ARIMA), [2] model. The basic assumption made to implement this model is that the considered time series is linear and follows a particular known statistical distribution, such as the normal distribution. ARIMA model has subclasses of other models, such as the Autoregressive (AR) [6], Moving Average (MA) [2] and Autoregressive Moving Average *(*ARMA*)* [2] models.

Although the essence of modeling in time series is to use the model to forecast future, forecasting is also useful in model selection and evaluation when there are several competing models. These competing models are compared in terms of goodness of fit (in sample fit) and forecasting power (out of sample forecast) [9]. The in sample fit criteria for selection of model are Alkaike's information criteria (AIC), Schwartz's information criteria, and Hannan and Quinn information criteria [17].

On the other hand, [4] provides an excellent opportunity to look at what is called out of sample behaviour of time series data. That is, a time series will provide forecast of new future observation which can be check against what is actually observed. Meanwhile the out of sample forecast is accomplished when the data used for constructing the model are different from that used in forecasting evaluation. That is, the data is divided into two portions. The first portion is for model construction and the second portion is used for evaluating model performance with possibility of forecasting new observations which can be checked against what is observed [4 & 16]. Hence in this study we dwell on model section based on the ability to of the model to forecast future values using the data on Naira to dollar exchange rate.

In Nigeria several authors have applied in- sample information criteria for model selection on Naira to US Dollar exchange rate. Given the advantages of out-of-sample model selection over in-sample model selection this work will be an improvement on studies of [1, 5, 8, 11, 12, 14] and several others.

2. MATERIALS AND METHODS

2. 1. An Autoregressive Process AR (p)

Autoregressive models are based on the idea that the current value of the series $\mathrm{E_t}$ can be explained as a function of p past values, E_{t-1}, E_{t-2},. . . E_{t-p}, where p determines the number of steps into the past needed to forecast the current value. An autoregressive model of order p, abbreviated AR (p), can be written as:

$$E_t = \boldsymbol{\varphi_1} E_{t-1} + \varphi_2 E_{t-2} + \cdots + \varphi_p E_{t-p} + \varepsilon_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots[1]$$

where E_t is a stationary series, φ_1, φ_2---φ_p are parameters of AR. Unless otherwise stated, we assume that ε_t is a Gaussian white noise series with mean zero and variance σ_ε^2. The highest order p is refer to as the order of the model.

The model in lag operator is specify as:

$$(1 - \varphi_1 B + \varphi_2 B^2 - \cdots - \varphi_p B^2) E_t = \varepsilon_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots[2]$$

where the lag backshift operator B is defined as $B^2 E_t = E_{t-p}$, p = 0, 1, 2, …..
More precisely we express the model as: $\varphi(B) E_t = \varepsilon_t$

The autoregressive operator $\varphi(B)$ is define as $\varphi(B) = 1 - \varphi_1 B + \varphi_2 B^2 - \cdots - \varphi_p B^p$

2. 2. The Moving Average Process

The moving average process of order q or MA (q) is define as

$$E_t = \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [3]$$

where θ_1, θ_2,…θ_p are parameters of MA, ε_t is white noise (error term)

2. 3. Autoregressive Moving Average (ARMA) process

A natural extension of pure autoregressive and pure moving average processes is the mixed autoregressive moving average (ARMA) processes, which includes the autoregressive and moving average as special cases (Wei 2006). A process $\{FDI_t\}$ is an ARMA (p,q) process if $\{FDI_t\}$is stationary and if for every t,

$$\varphi(B) E_t = \theta(B) \varepsilon_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots[4]$$

$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 \ldots - \ldots \varphi_p B^p$ is the autoregressive coefficient polynomial
$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 \ldots - \ldots \theta_p B^p$ is the moving average coefficient polynomial

2. 4. Autoregressive Integrated Moving Average (ARIMA) Model

Box, Jenkins and Reinsel (2008) considered the extension of ARMA model to deal with homogeneous non-stationary time series in which E_t, itself is non-stationary but its d^{th} difference is a stationary ARMA model. Denoting the d^{th} difference of FDI_t by

$$\varphi(B) = \theta(B) \nabla^d E_t = \theta(B) \varepsilon_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots[5]$$

where $\varphi(B)$ is the non-stationary autoregressive operator such that the roots of $\varphi(B)$ = 0 are and the remainder lie outside the unit circle. $\theta(B)$ is stationary moving average operator.

2. 5. Box-Jenkins Methodology

After describing various time series models, the next issue to our concern is how to select an appropriate model that can produce accurate forecast based on a description of historical pattern in the data and how to determine the optimal model orders. Box and Jenkins (1973) developed a practical approach to build ARIMA model, which best fit to a given time series and also satisfy the parsimony principle. Their concept has fundamental importance on the area of time series analysis and forecasting.

The BoxJenkins methodology does not assume any particular pattern in the historical data of the series to be forecasted. Rather, it uses a three step iterative approach of ***model identification, parameter estimation,*** and ***diagnostic checking*** to determine the best parsimonious model from a general class of ARIMA models. This three step process is repeated several times until a satisfactory model is finally selected. Then this model can be used for forecasting future values of the time series.

2. 6. Model Identification

Model identification involves examining the given data by various methods, to determine the values of p, q and d. The values are determined by using autocorrelation function (ACF) and partial autocorrelation function (PACF). This can be done by observing the graph of the data or autocorrelation functions [7]. For any ARIMA process, the theoretical (PACF) has non-zero partial autocorrelation at 1,2…p lags and partial autocorrelation at all lags. While the theoretical (ACF) has non-zero autocorrelation at all lags. The non-zero lags of sample PACF and ACF are tentatively accepted as the p and q parameters [14]. For a non- stationary series the data is differenced to make stationary. The number of times the series is differenced determines the order of d. Therefore for a stationary data d = 0 and ARIMA (p, d, q) can be written as ARIMA (p, q).

2. 7. Model Estimation

After an optimal model has been identified, the model estimation methods make it possible to estimate simultaneously all the parameters of the process, the order of integration coefficient and parameters of an ARMA structure. There many methods of estimating parameters of linear time series models but for the purpose of this study we shall consider the maximum likelihood method. Only two methods of maximum likelihood estimation considered in time series analysis are discussed.

Given: $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_p)$, $\mu = E(X_t)$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_q)$ and $\sigma_e^2 = E(e_t^2)$ from observations of the causal $ARMA(p,q)$ process defined by

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + a_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \qquad [6]$$

(i) Conditional Maximum Likelihood Estimation

For an $ARMA(p,q)$ model (3.36), the joint probability density function of $e = \{e_1, e_2, \dots, e_n\}'$ is given by

$$P(\boldsymbol{a}|\boldsymbol{\varphi}, \mu, \theta, \sigma_e^2) = (2\pi\sigma_e^2)^{\frac{2}{n}} exp\left\{-\frac{1}{2\sigma_e^2}\sum_{i=1}^{n} e_t^2\right\} \qquad [7]$$

rewriting [6] as

$$e_t = \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + X_t - \varphi_1 X_{t-1} - \varphi_2 X_{t-2} - \cdots - \varphi_p X_{t-p} \qquad [8]$$

Let $\boldsymbol{X} = (X_1, X_2, \dots, X_n)'$ and assume initial conditions $\boldsymbol{X}_* = (X_{1-p}, \dots, X_{-2}, X_{-1}, X_0)'$ and $\boldsymbol{e}_* = (e_{1-p}, \dots, e_{-2}, e_{-1}, e_0)'$ are known. The conditional log-likelihood function is

$$lnL_*(\boldsymbol{\varphi}, \mu, \theta, \sigma_e^2) = -\frac{2}{n}(2\pi\sigma_e^2) - \frac{S_*(\boldsymbol{\varphi},\mu,\theta)}{2\sigma_e^2} \qquad [9]$$

where $S_*(\boldsymbol{\varphi}, \mu, \theta) = \sum_{i=1}^{n} e_t^2\,(\boldsymbol{\varphi}, \mu, \theta | \boldsymbol{X}_*, \boldsymbol{a}_*, \boldsymbol{X})$ [10]
is the conditional sum of squares function. The quantity of $\widehat{\boldsymbol{\varphi}}, \hat{\mu}$, and $\hat{\theta}$ which maximize (10) are called the conditional estimators. Because $L_*(\boldsymbol{\varphi}, \mu, \theta, \sigma_e^2)$ involves the data only through $S_*(\boldsymbol{\varphi}, \mu, \theta)$, these estimators are the same as the conditional least squares estimators obtained from minimizing the conditional sum of squares function $S_*(\boldsymbol{\varphi}, \mu, \theta)$. The estimator $\hat{\sigma}_e^2$ of σ_e^2 is obtained from

$$\hat{\sigma}_e^2 = \frac{S_*(\boldsymbol{\varphi},\mu,\theta)}{n-p-q-1} \qquad [11]$$

where $n-p-q-1$ (degrees of freedom) equals the number of terms used in the sum of $S_*(\boldsymbol{\varphi}, \mu, \theta)$ minus the number of parameters' estimator.

(ii) Unconditional Maximum Likelihood Estimation

Box, Jenkins and Reinsel (2008) suggest the following unconditional log-likelihood function;

$$lnL_*(\boldsymbol{\varphi}, \mu, \theta, \sigma_e^2) = -\frac{2}{n}(2\pi\sigma_e^2) - \frac{S_*(\boldsymbol{\varphi},\mu,\theta)}{2\sigma_e^2} \qquad [12]$$

where

$$S(\boldsymbol{\varphi}, \mu, \theta) = [E(e_t | \boldsymbol{\varphi}, \mu, \theta, \boldsymbol{X})]^2 \qquad [13]$$

is the conditional sum of squares function.
where $E(e_t | \boldsymbol{\varphi}, \mu, \theta, \boldsymbol{X})$ is the conditional expectation of e_t given $\boldsymbol{\varphi}, \mu, \theta$, and $\boldsymbol{X}$.

The quantities, $\widehat{\boldsymbol{\varphi}}, \hat{\mu}$, and $\hat{\theta}$ that minimize function (3.26) are called unconditional maximum likelihood estimators and are equivalent to the unconditional least squares estimators obtained by minimizing (3.27). In practice, the summation in (3.27) is approximated by a finite form

$$S(\boldsymbol{\varphi}, \mu, \theta) = \sum_{i=-M}^{n} [E(e_t | \boldsymbol{\varphi}, \mu, \theta, \boldsymbol{X})]^2 \qquad [14]$$

where M is sufficiently large integer.

The estimator $\hat{\sigma}_e^2$ of σ_e^2 is obtained from

$$\hat{\sigma}_e^2 = \frac{S_*(\boldsymbol{\varphi}, \mu, \theta)}{n}$$

2. 8. Model Verification

The last step in Box-Jenkins methodology is model verification or model diagnosis. The conformity of white noise residual of the model fit will be judged by plotting the ACF and PACF of the residual to see whether it does not have any pattern or we perform Ljung-Box Test on the residual.

The test hypothesis:

Ho: There is no serial correlation

H1: There is serial correlation

The test statistics of the Ljung-Box

$$LB = n(n+2)\sum_k^m \frac{\rho_k^2}{n-k} \qquad \text{Distributed to } x_{m-e}^2 \qquad \ldots\ldots\ldots[15]$$

where n is the sample size, $m = lag\ length$ and p is the sample autocorrelation coefficient.

The decision: if LB is less than critical value of x^2, then we do not reject the null hypothesis. This means that a small value of Ljung-Box statistic will be in support of no serial correlation or i.e the errors are normally distributed. This is concern about model accuracy.

2. 9. Information Criteria

In- sample Information Criteria

Given multiple competing models, we decide upon a final model which is one popular method to use a model selection criteria Alkaike's Information criteria (AIC), Schwartz Information criteria and the Hannan Quinn criteria (HQC) which attempts to choose a model that adequately describes the data but in the most parsimonious way as possible or minimizing the number of parameters, for example AR (3) model doesn't outperform AR (2) model by a certain predefined quantity or criteria, than AR (2), the most parsimonious model is chosen. In general, the model chosen is the one that minimizes the respectively criteria scores

2. 9. 1. Akaike's Information Criterion

Akaike's information criterion (AIC) originally proposed by Akaike, attempts to select a good approximating model for inference based on principle of parsimony. AIC proposes the use of relative entropy or Kull black-Libeler (K-L) information as fundamental basis for model selection. A suitable estimator of the relative K-L information is used and involves two terms.

The first term is a measure of lack of fit while the second is a penalty for increasing size of the model, assuming parsimony in the number of parameters. The AIC criterion to minimized is

$$AIC(n) = \log(\sigma^2) + \frac{2n}{T} \qquad \ldots[16]$$

where n is the dimensionality of the model σ^2 is the maximum likelihood estimate of the white noise variance, and T is the sample size

2. 9. 2. Schwartz Bayesian Information Criterion

The Bayesian information criteria (BIC) originally proposed by Swartz was derived in a Bayesian context and is dimensional consistent in that it accept to consistently estimate the dimension of the true model. It assumes a true model exist in the set of candidate models, therefore, requires a large sample size to be effective.

The BIC criteria to be minimized is

$$BIC(n) = \log(\sigma^2) + \frac{nlog(T)}{T} \quad \ldots [17]$$

where n is the dimensionality of the model, σ^2 is the maximum likelihood estimate of the white noise variance and T is the sample size.

2. 9. 3. Hannan-Quinn Criterion

The Hanna – Quinn criteria originally proposed by Hanna and Quinn was derived from the law of iteration logarithm, it is another dimension consistent model and only differs from AIC and BIC with respect to the penalty term.

The HQ criteria to be minimized is

$$HQ(n) = \log(\sigma^2) + \frac{2nlog(T)}{T} \quad \ldots [18]$$

where n is the dimensionality of the model, σ^2 is the maximum likelihood estimate of the white noise variance and T is the sample

2. 10. Augmented Dickey Fuller test

The test was introduced by Dickey and Fuller (1979) to test for the presence of unit root(s). the regression model for the test is given as

$$\Delta E_t = \rho E_{t-1} + \beta X_{t-1} + \delta_1 \Delta E_{t-1} + \delta_2 E_{t-2} + \cdots + \delta_p E_{t-p} + e_t \quad \ldots[19]$$

The hypothesis testing

$H_{0:} = \rho = 0$(the series contain unit root)

$H_1: \rho < 0$ (the series is stationary)

where

ΔE_t is the difference series
E_{t-1} is the immediate previous observation
$\delta_1, \ldots, \delta_p$ is the coefficient of the lagged differenced term up to p
X_{t-1} Is the optimal exogenous regresses which may be constant or constant trend
ρ And β parameters to be estimated

2. 11. Measure of Predictive Accuracy

The common measures of predictive accuracy that have been used to evaluate the forecast performance of a single model in statistics are the root mean square error (MSE), mean square error (MSE) mean absolute error (MAE) and mean percentage absolute error (MPAE). The measures are computed as follow

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} e_i^2 \quad \text{...[21]}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \quad \text{... [22]}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |e_i| \quad \text{... [23]}$$

$$\text{MAPE} = \left(\frac{1}{n}\sum_{i=1}^{n} |e_i|\right) \times 100\% \quad \text{... [24]}$$

where e_i is the forecast error and n is the number of forecast error

3. RESULTS AND DISCUSSION.

In this section we shall use the monthly official exchange rate in Nigeria to identify and estimate ARIMA model that adequately represents the series and use some diagnostic tests to evaluate the model. The data set is from Nigeria official exchange rate for the Naira to US Dollar from January 2002 to December 2018. Gretl and E-views are statistical software's used for data analysis.

3. 1. Graphical Representation of the exchange rata data (E_t).

The upward and downward movements of the exchange rate series in Figures 1 & 2 indicate that the series are not stationary.

3. 2. Unit root test

The Augmented Dickey Fuller (ADF) test was used to check the stationary of the exchange rate series. The result of the ADF test is shown in the table below.

Table 1. Levels: constant and linear trend

Variable	ADF	Critical Values		Conclusion
E_t	-1.612344	-4.007084	@1%	Not Stationary
		-3.433651	@5%	Not Stationary
		-3.140697	@10%	Not Stationary

Table 2. Levels: No constant and linear trend

Variable	ADF	Critical Values		Conclusion
E_t	1.160226	-2.577190	@1%	Not Stationary
		-1.942508	@5%	Not Stationary
		-1.615589	@10%	Not Stationary

We observe from Figures 1 and 2 and Tables 1 and 2 that the series is not stationary at levels. However, stationarity was achieve after the first difference as shown in Figure 3 and Tables 3 and 4.

Table 3. ADF test After First Difference: with constant and Linear Trend

Variable	ADF	Critical Values		Conclusion
E_t	-20.49616	-4.007084	@1%	Stationary
		-3.433651	@5%	Stationary
		-3.140697	@10%	Stationary

Table 4. ADF test After First Difference: with constant and Linear Trend

Variable	ADF	Critical Values		Conclusion
E_t	-20.36980	-2.577190	@1%	Stationary
		-1.942508	@5%	Stationary
		-1.615589	@10%	Stationary

3. 2. ARIMA modeling of Naira to Dollar exchange rate (E_t)

It should be noted that, even if an ARIMA model has been correctly identified and give good result, this does not mean that it is the only model that can be considered, various model should be identified and tested. With regard to Naira to Dollar exchange rate **(E_t)**, three ARIMA models, ARIMA (0, 1, 1), ARIMA (2, 1, 0), and ARIMA (1, 1, 2) were identified and fitted tentatively as shown in Table 5

Table 5. ARIMA Models for Naira to Dollar Exchange rate.

Model	Parameter	Estimate	s.e	Z-ratio	p-value	AIC	BIC	HNQ
ARIMA (0,1,1)	$\boldsymbol{\theta_1}$	**0.620838**	**0.0530209**	**11.71**	**<0.0001**	**1157.559**	**1164.063**	**1160.193**
ARIMA (2,1,0)	φ_1	0.620684	0.0697471	8.899	<0.0001	1162.856	1173.612	1166.807
	φ_1	−0.254996	0.0695209	−3.668	0.0002			
ARIMA (1,1,2)	φ_1	0.0840018	0.784031	0.1071	<0.00001	1160.533	1173.542	1165.803
	θ_1	0.581727	0.783838	0.7422	<0.0001			
	θ_2	0.0210098	0.491990	0.04270	<0.00001			

3. 3. Model Diagnostic Checking or Evaluation

We use diagnostic test of the model residuals to check if the model has adequately fitted the series. First we plot the ACF and the PACF of the standardized residuals to visually see if there exists serial correlation. Next we performed the Ljung- Box test for the three competing models to check if there exists serial correlation in the residual. The ACF and PACF plots of the residuals from ARIMA (0, 1, 1), ARIMA (2, 1, 0) and ARIMA (1, 1, 2) shows that all correlations are within the threshold limits indicating that the residuals are white noise. This can be seen in Figures 5, 6 and 7. A Ljung-Box test for the three competing model returns p-values greater than the critical value at 5%, this also suggests that the residuals are white noise and that the models are adequate. This can be seen in Table 7

Table 6. Ljung – Box Residual autocorrelation test result for the three competing models

Model	**Test Statistic**	**P-value**
ARIMA (0,1,1)	3.12449	0.9890
ARIMA (2,1,0)	5.63517	0.8449
ARIMA (1,1,2)	2.6449	0.9767

3. 5. Model forecast Evaluation and selection.

So far we have three competing models that are adequate and our interest is towards selecting a model that best represents the series (E_t) in terms of forecasting new future observation (out- of – sample forecast). Given the series E_t which consist of 204 observations, we split the series into two portions. The first 192 observations are used for estimating the three models and the remaining 12 observations are used for the out – of – sample forecast. The graph of forecasts from ARIMA (0,1,1), ARIMA (2, 1,0) and ARIMA (1,1,2) models as shown in Figures 8, 9 and 10 respectively, indicate that the forecasts are close to the actual values that are all within 95% interval forecasts.

Base on in-sample information criteria ARIMA (0, 1, 1) is the best model. However, comparing the out-of-sample forecasts of the competing models using Root Mean Square Error (RMSE), Mean Square Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) as the measures of accuracy, ARIMA (2, 1, 0) appears to have minimum RMSE, MSE, MAE, MAPE therefore, gives the best representation of the out-of-sample forecasts of E_t. (Refer to Table 8)

Hence our estimated ARIMA (2, 1, 0) is specify as follows

$$E_t = 0.620684E_{t-1} - 0.254996E_{t-2}$$

s.e (0.0697471) (0.0695209)

Z-ratio (8.899) (−3.668)

P-value (<0.0001) (0.0002)

[Excerpts from Table 6]

Table 7. Measures of Predictive Accuracy.

Model	RSME	MSE	MAE	MAPE
ARIMA(0, 1, 1)	0.5767	0.3325	0.5344	0.1747
ARIMA(2,1,0)	0.4851	0.2353	0.4995	0.1498
ARIMA(1, 1,2)	0.5345	0.2858	0.4995	0.1633

4. CONCLUSIONS

It is always argued that a model that is best in the in-sample fitting may not necessarily give more genuine forecasts since it is the same set of data that used for model estimation is also used for forecast evaluation. By implication, model selection based on in-sample criteria such as Akaike information criteria, Schwarz information criteria and Hannan Quinn information criteria may not provide more genuine forecasts. Also, a model selected on the basis of in-sample criteria does not give information about the future observations. Hence, in this study, we dwell completely on the performance of the out- sample forecast to enhance the selection of a model. The out-sample forecasting is advantageous over the in-sample forecasting in that, the model selection is based on how best the forecasts perform and able to provide information about future observations. Furthermore, to measure the accuracy of the out-sample forecasts of the competing models; we apply the RMSE, MSE, MAE and MAPE. The smaller the errors, the better the forecasting power of that model.

In relation to this work we considered data on official Nigeria exchange rate of Naira to US Dollar rate from January 2002 to December 2018. Three competing models were identified for the exchange rates series. Diagnostic checking revealed that all the competing models adequately represent the exchange rates series. However on the basis of out-of-sample model selection and evaluation; ARIMA (2, 1, 0) appeared to be the best performing out-of-sample forecasting model with minimum RMSE, MSE, MAE and MAPE for the series. The implication of this study is that, time series forecasting is not only meant for predicting future observations but also in aiding model selection and evaluation on Nigeria exchange rate.

References

[1] Bala, D.A. and Asemota, J.O. Exchange Rate Volatility in Nigeria: application of GARCH model with exogenous break. *CBN journal of applied statistics* 4(1) (2013) 89-116

[2] Box G.E.P and Jenkins G. Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco (1970).

[3] Brockwell P.J and Davis R.A, Introduction Time Series and Forecasting, 2nd Edition, Springer Publication, (2003).

[4] Chatfield, C. Time Series Forecasting. 5th Ed. New York: Chapman and Hall CRC

[5] Ette Harrison Etuk, Forecasting Nigerian Naira-US Dollar Exchange Rates by a Seasonal Arima Model. *American Journal of Scientific Research* Issue 59 (2012), pp. 71-78

[6] Hipel K.W and McLeod A.I. Time Series Modelling of Water Resources and Environmental Systems. Amsterdam, Elsevier I (1994)

[7] Makrisdakis, S & S.C Wheelwright[ed] (1984). Forecasting Methods for Management. John Wiley and San Chichester) 250-290

[8] Muhammed M.B and Abdulmuahymin S.A. Modelling the Exchange Ability of Nigerian currency (Naira) with respect to US Dollar. *International Journal of Scientific and Engineering Research* 7(7) (2016) 86-104

[9] Moffat, I. U. & Akpan, E. A. (2014). Time Series Forecasting: A Tool for Out- Sample Model and evaluation. *American Journal of Scientific and Industrial Research* 5(6), 185-194

[10] Moosa, E. A. (2008). Forecasting the Chinese Yuan – US Dollar exchange rate under the New Chinese Exchange Rate Regime. *International Journal of Business and Economics* 7 (1): 23-35

[11] Nwankwo S.C. (2014). Autoregressive Integrated Moving Average (ARIMA) Model for Exchange Rate (Naira to dollar). *Academic Journal of Interdisciplinary Studies*, 3(4) (2014) 182-195

[12] Newaz M.K. Comparing Performance of Time Series Model for Forecasting Exchange Rate. *BRAC University Journal* 2 (2008) 55-65

[13] Tindaon, S (2015). Forecasting the NTD/USD exchange rate using Autoregressive Model, Department of International Business, CYCU, Taiwan

[14] Oladejo, M. O. & Abdullahi, B. (2015). A Suitable Model for the Forecast of Exchange Rate in Nigeria (Naira versus US Dollar). *International Journal of Science and Research* 4(5), 2669-2676

[15] Osbuohien I. O and Edokpa I.W. Forecasting Exchange Rate Between the Nigeria Naira and the US Dollar using ARIMA model. *International Journal of Engineering Science Invention*, 2(4) (2013) 16-22

[16] Raicharoen T, Lursinsap C and Sanguanbhoki, P. Application of Critical Support Vector Machine to Time Series Prediction", Circuits and Systems, 2003. ISCAS '03. Proceedings of the 2003 International Symposium on Volume 5, 25-28 May, 2003, pages: V-741-V-744.

[17] Wei, W. W. S. Time Series Analysis Univariate and Multivariate Methods. (2nd Ed.). New York: Adison Wesley, (2006) 33-59.

[18] Zhang G.P. A Neural Network Ensemble Method with Jittered Training Data for Time Series Forecasting. *Information Sciences* 177 (2007) 5329-5346

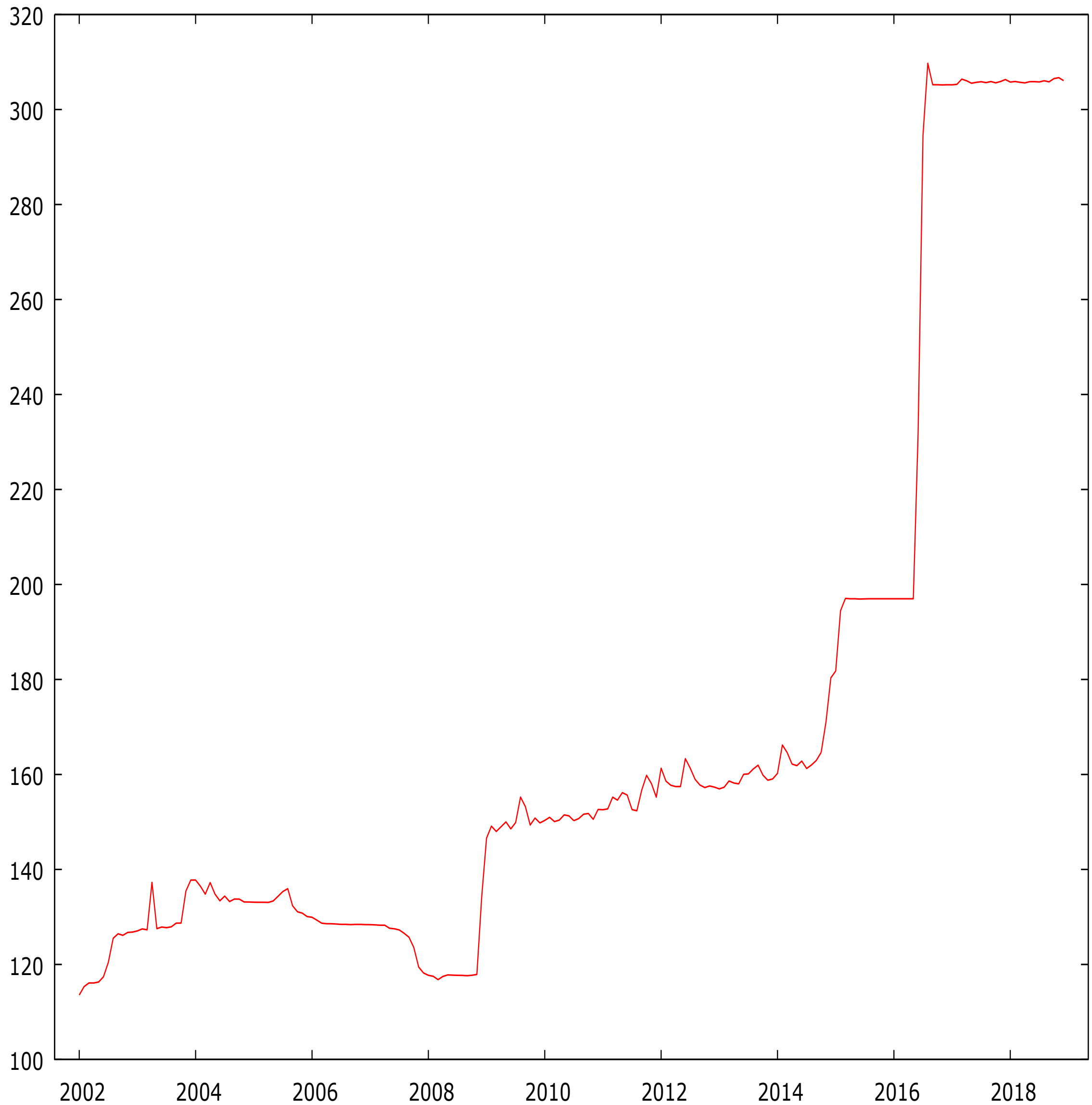

Figure 1. Time plot of the monthly exchange rate of Naira to Dollar from January 2002 to December 2018

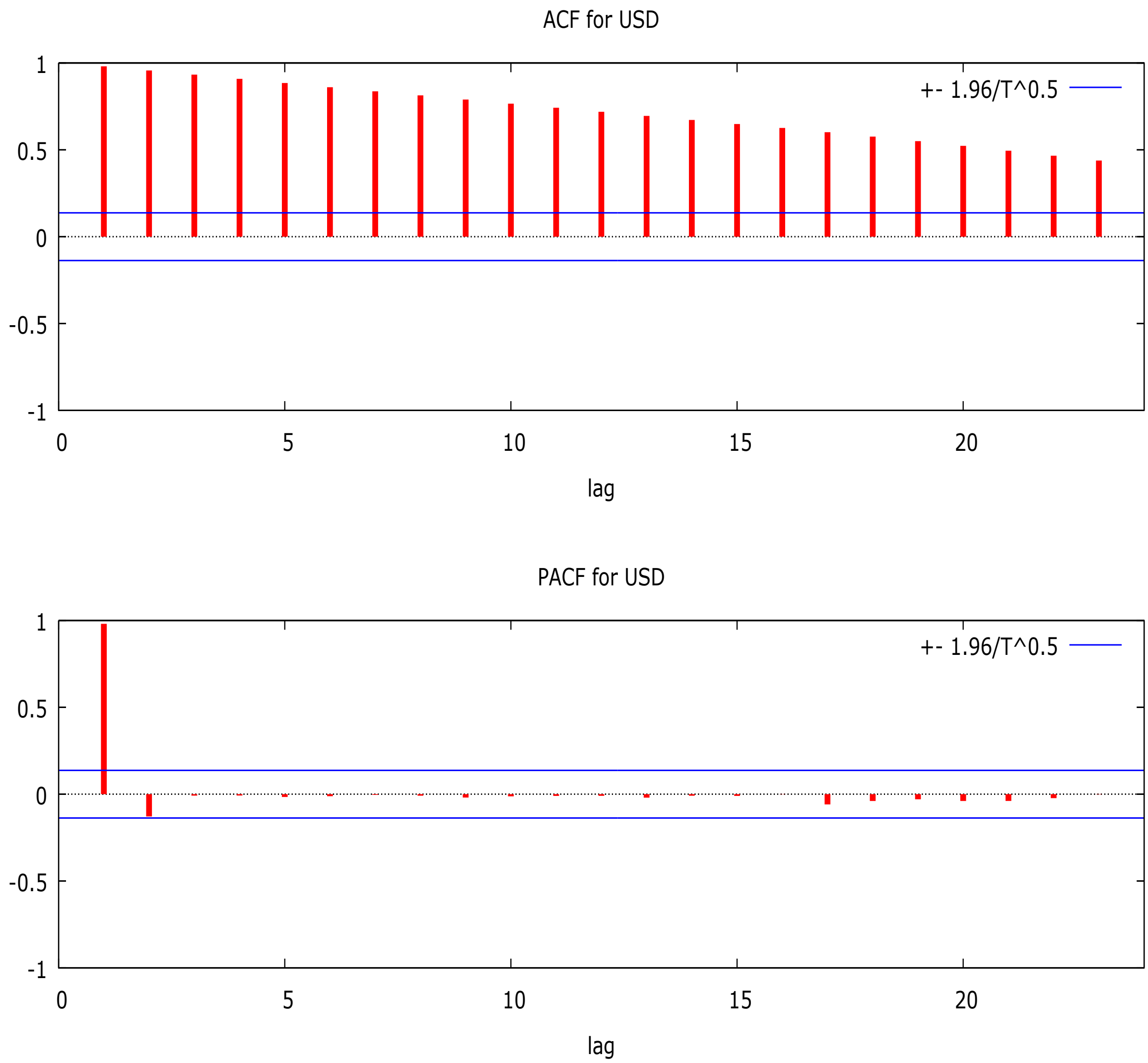

Figure 2. ACF and PACF of monthly exchange rate of Naira to Dollar from January 2002 to December 2018

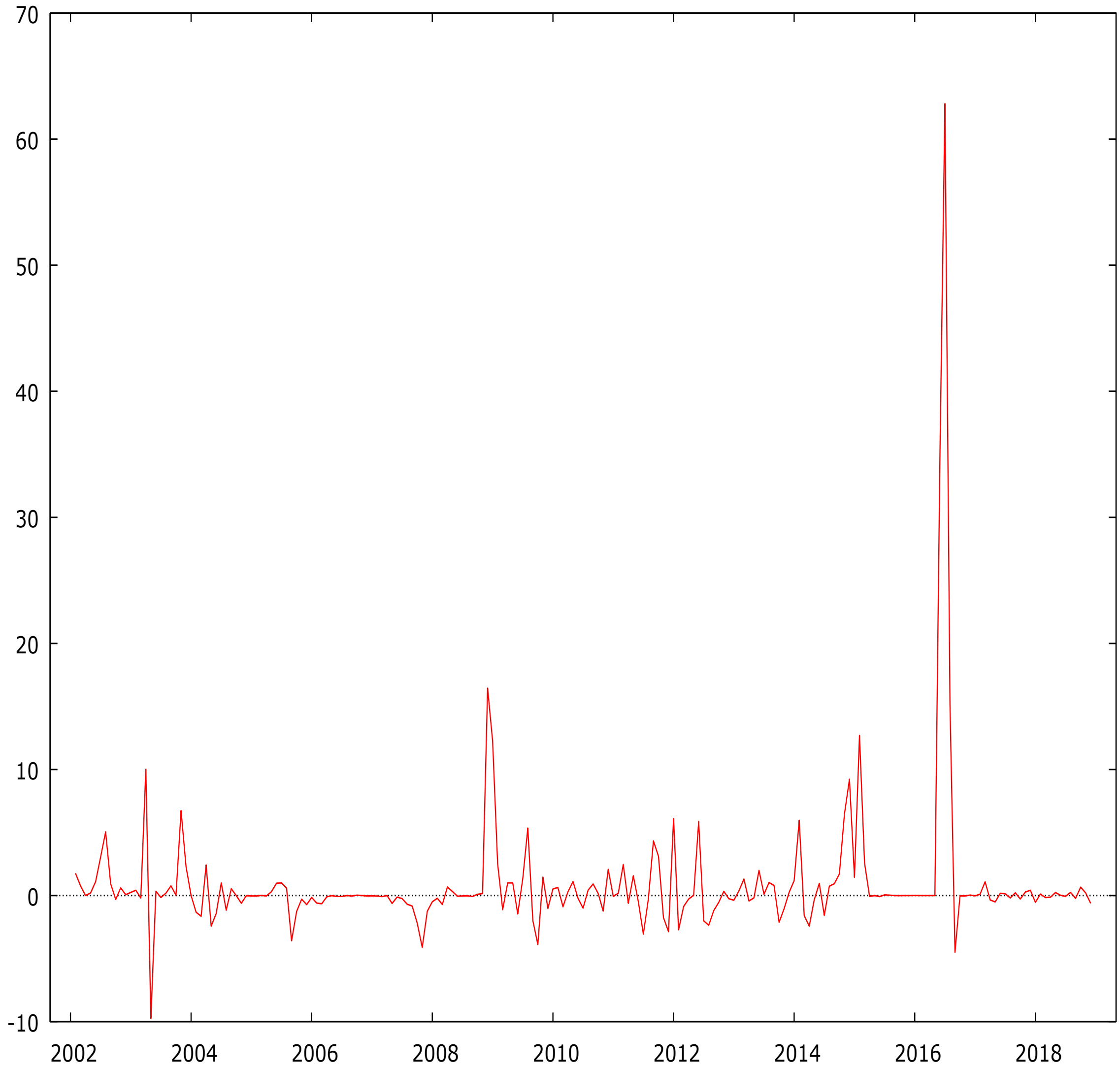

Figure 3. Time plot of the first difference of monthly exchange rate of Naira to Dollar from January 2002 to December 201.

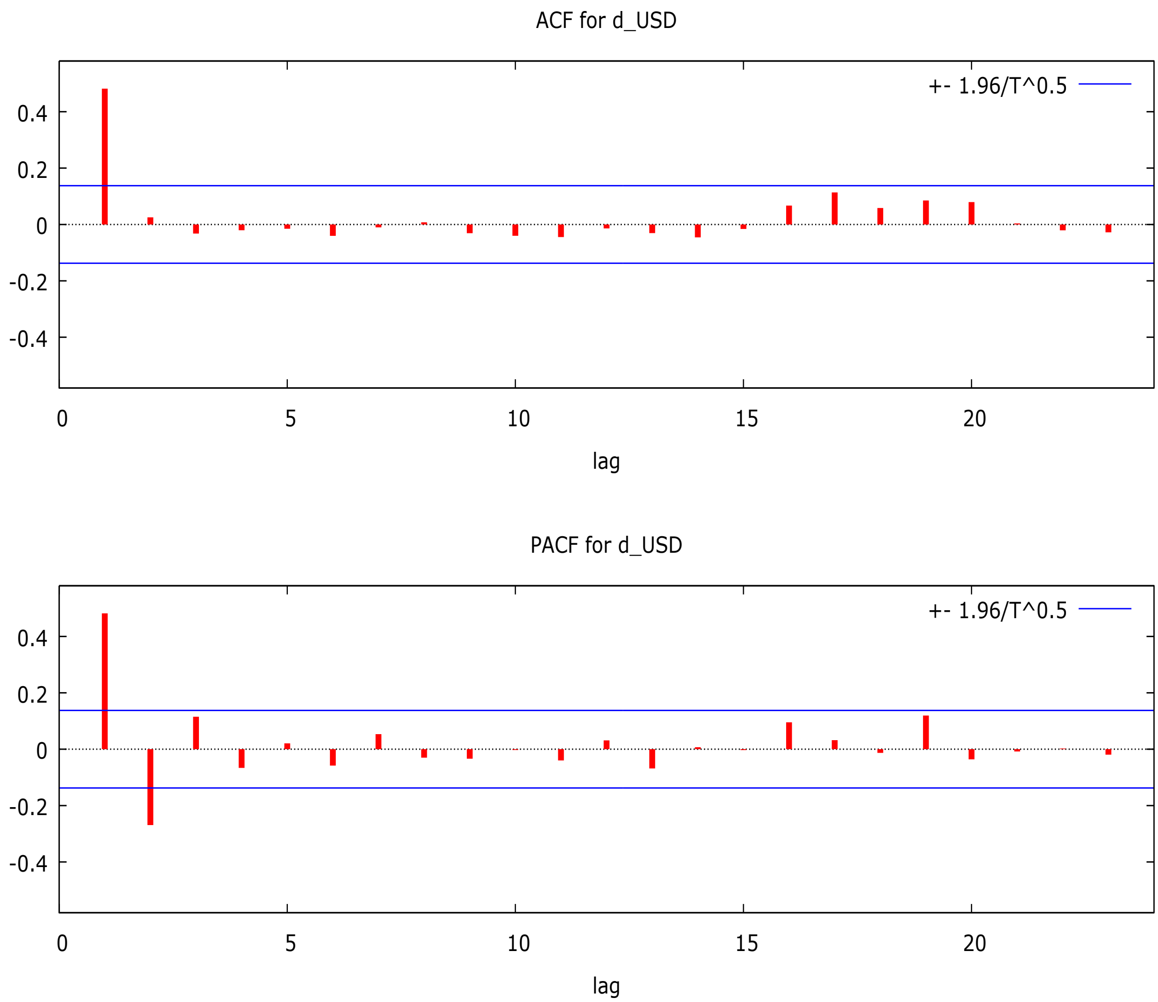

Figure 4. ACF and PACF of the first difference of monthly exchange rate of Naira to Dollar from January 2002 to December 2018

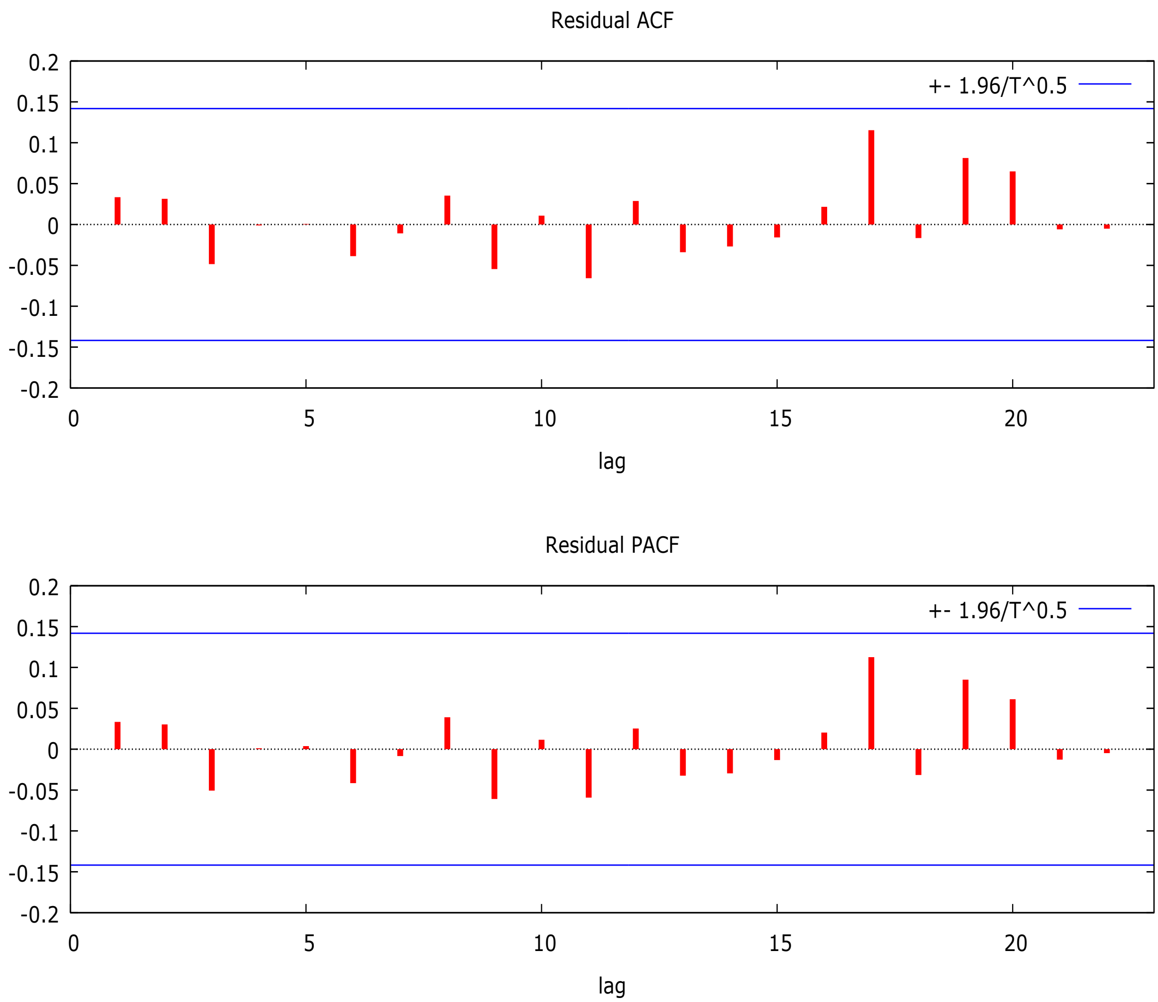

Figure 5. ACF and PACF of residuals of ARIMA (0, 1, 1)

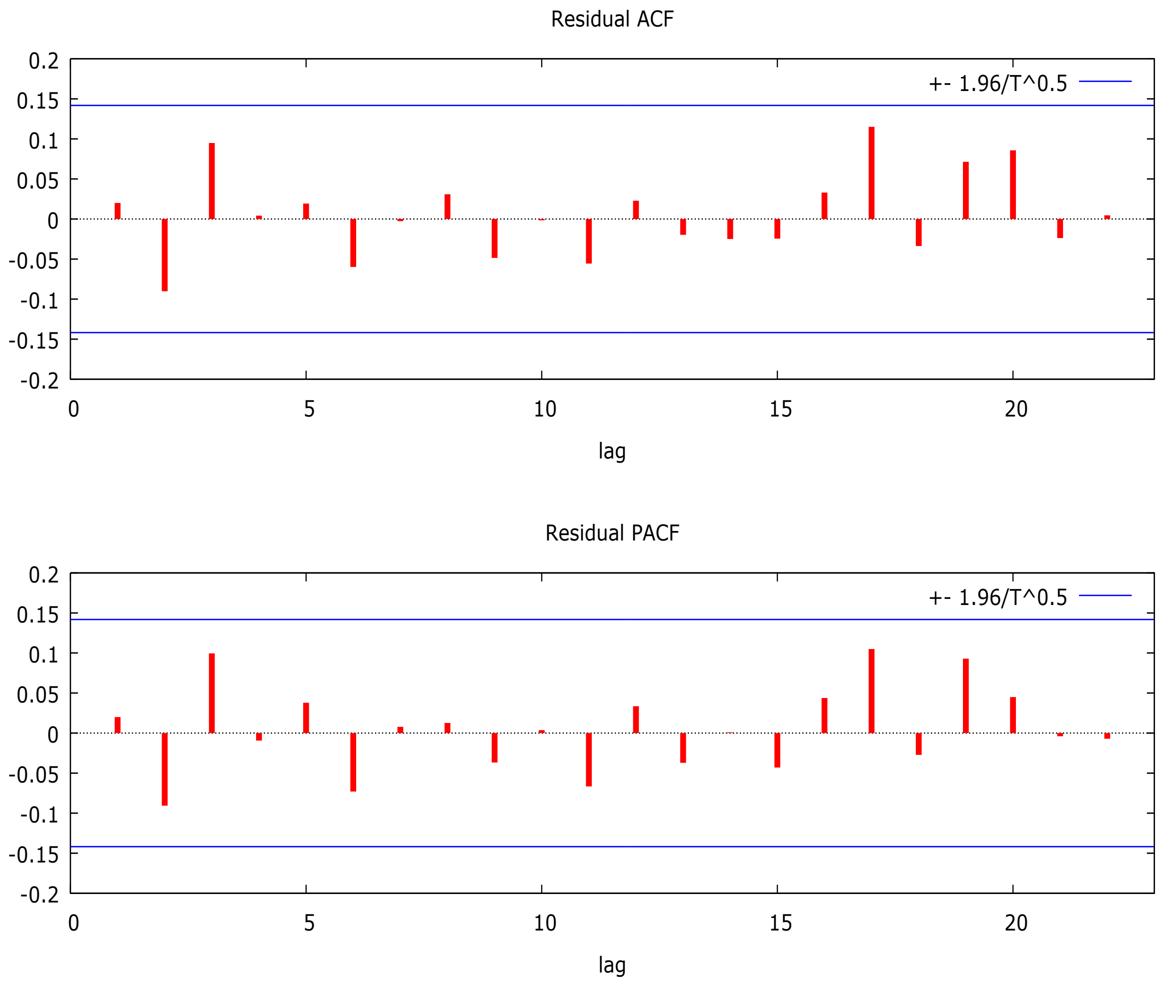

Figure 6. ACF and PACF of residuals of ARIMA (2, 1, 0)

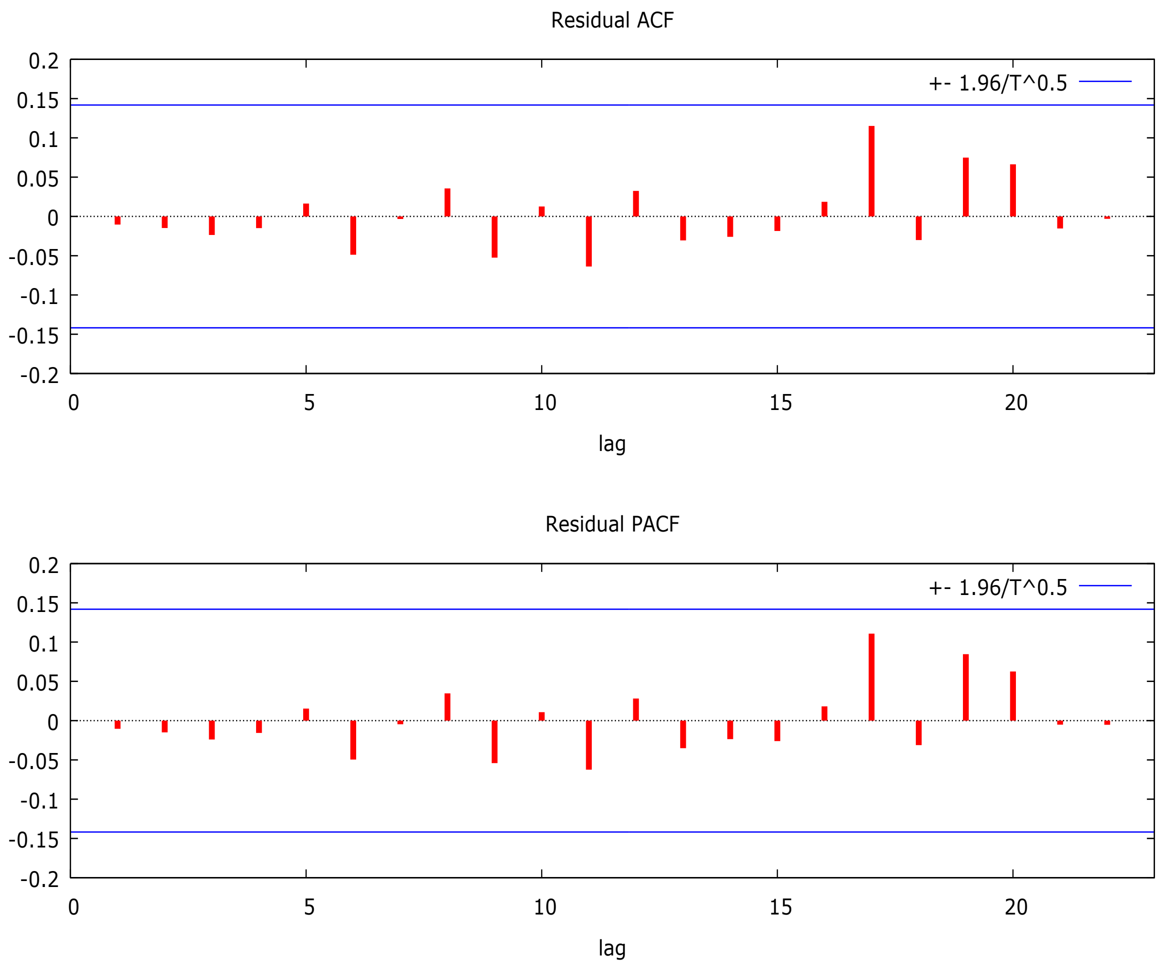

Figure 7. ACF and PACF of residuals of ARIMA (1, 1, 2)

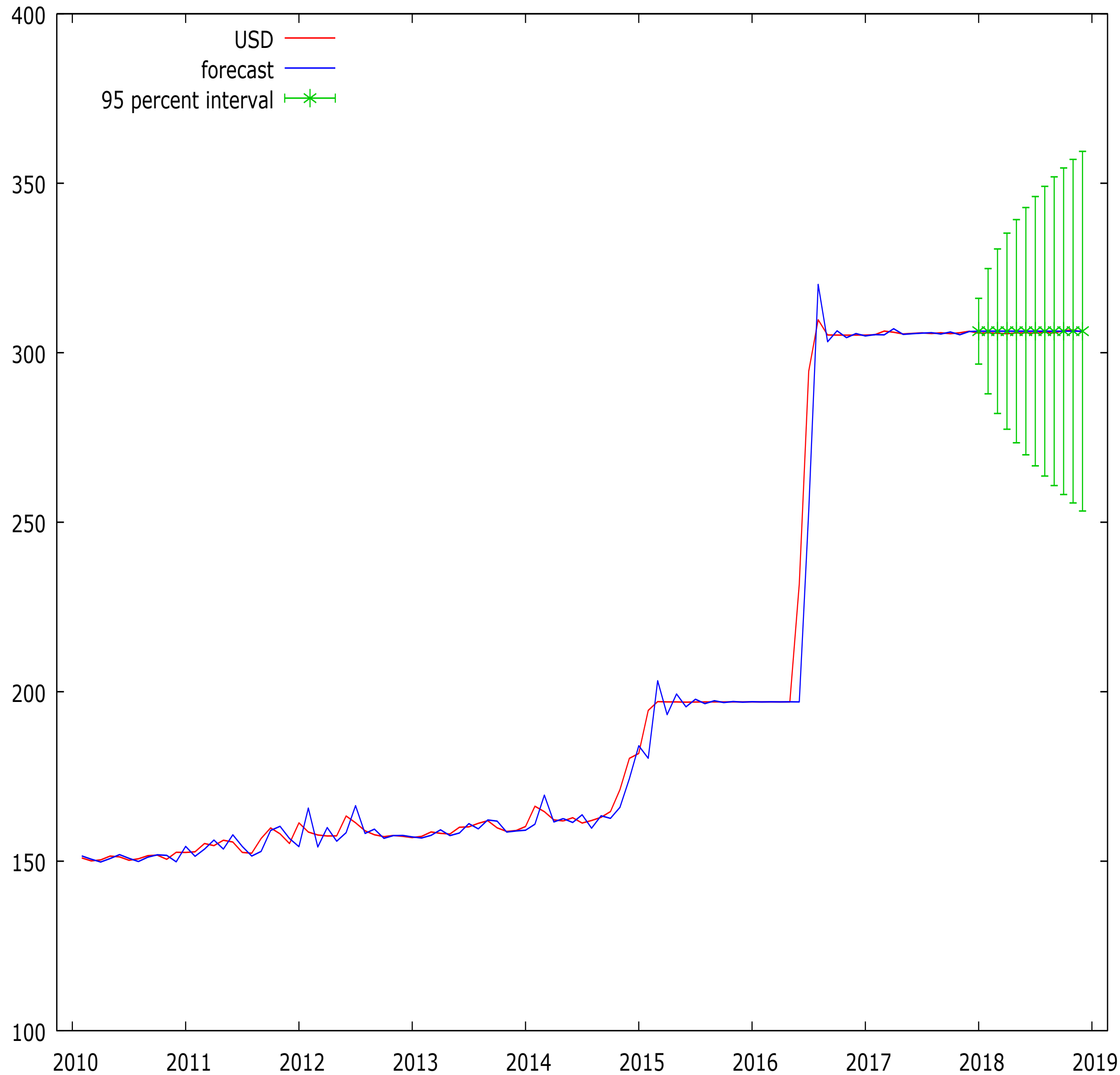

Figure 8. Plot of ARIMA (0, 1, 1) out-of-sample actual against predicted values from 2018:01 to 2018:12.

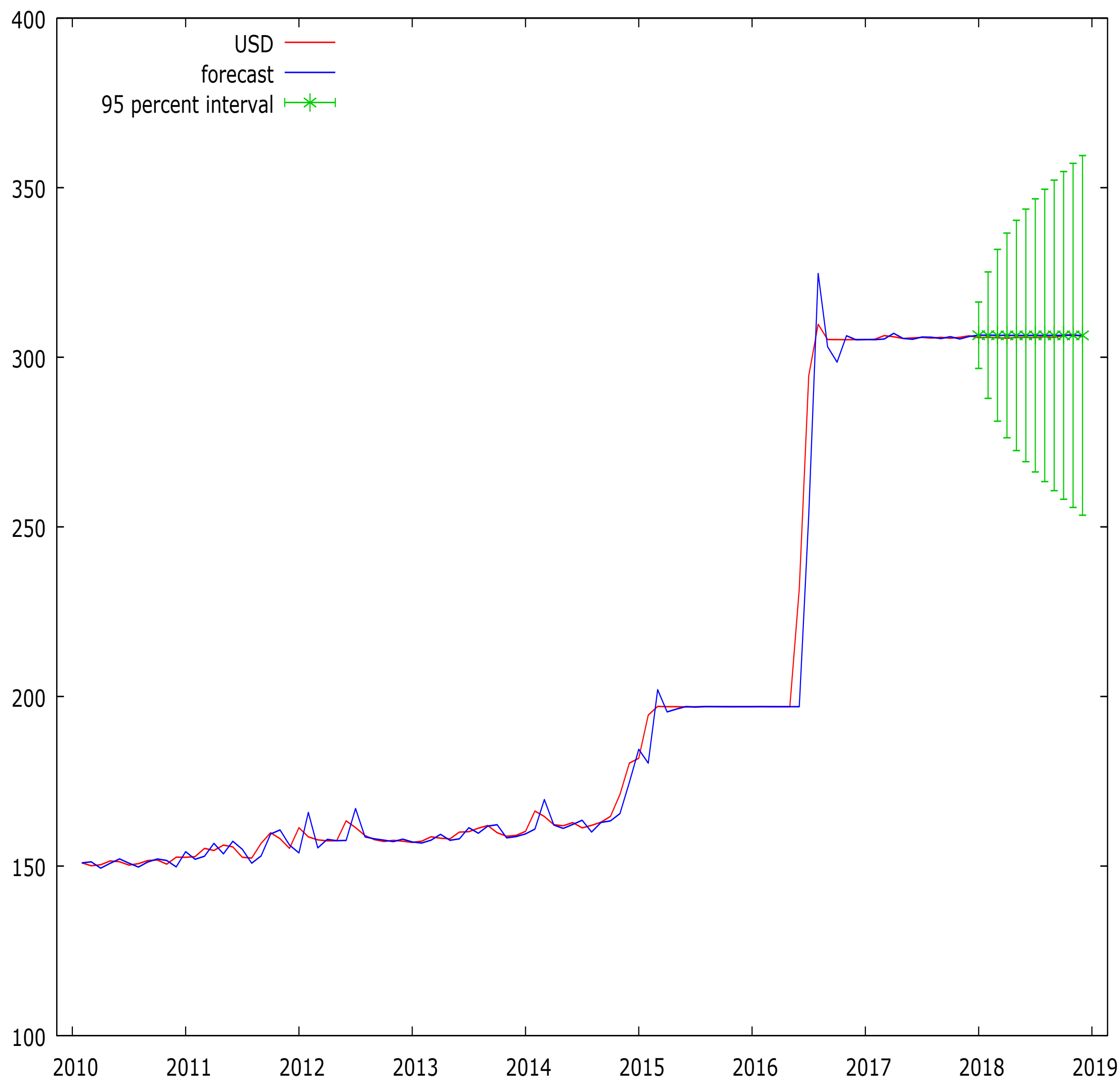

Figure 9. Plot of ARIMA (2, 1, 0) out-of-sample actual against predicted values from 2018:01 to 2018:12

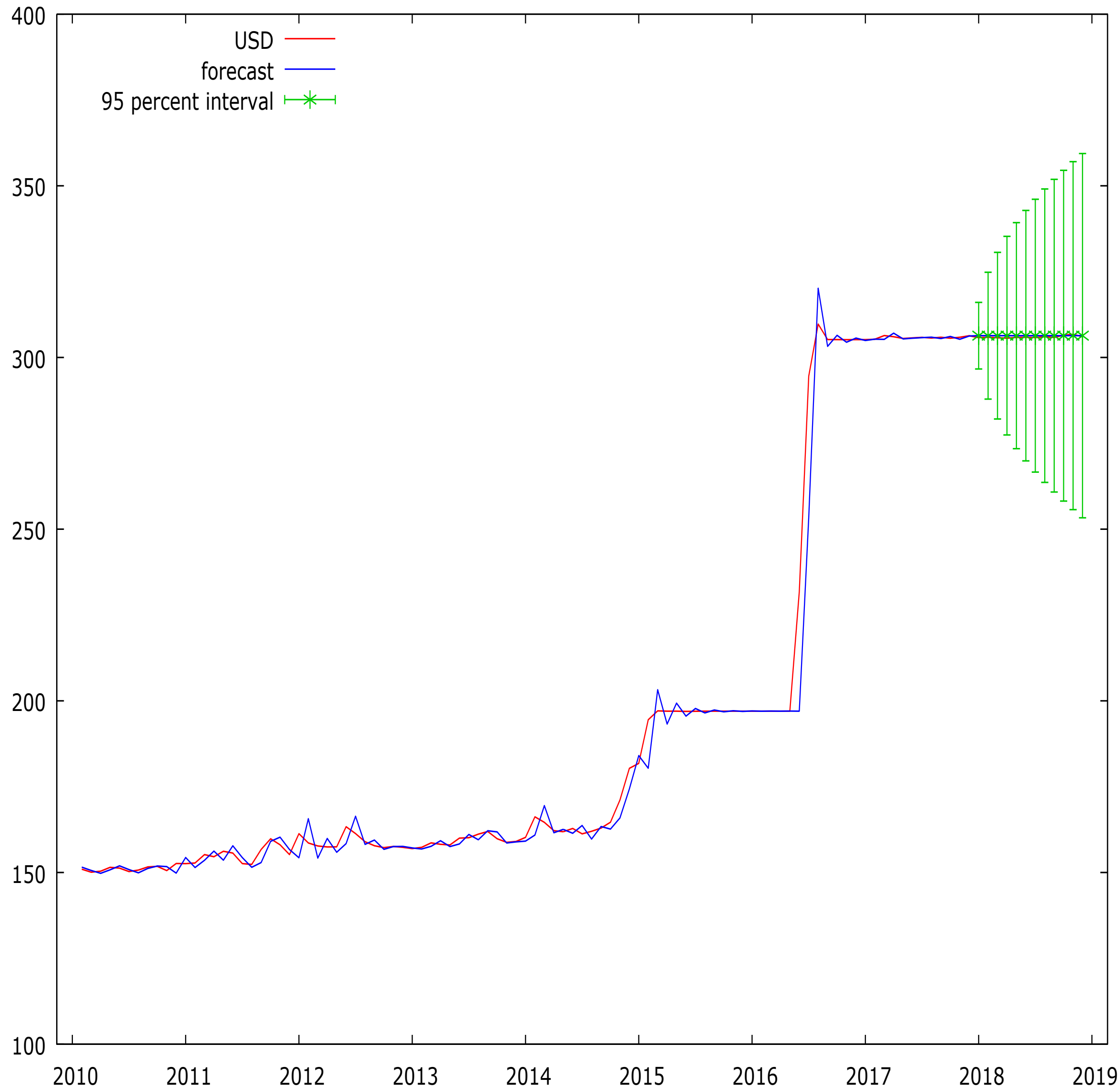

Figure 10. Plot of ARIMA (1, 1, 2) out-of-sample actual against predicted values from 2018:01 to 2018:12