

Time Series Filters Applied to Improve Box-Jenkins Forecasting Approach

J. N. Onyeka-Ubaka^{1*}, R. K. Ogundeji² and R. U. Okafor³

^{1,2,3}*Department of Mathematics, Faculty of Science, University of Lagos,
Akoka, Nigeria*

(Received: 06 August 2020; accepted: 21 December 2020)

Abstract. Data obtained from observations collected sequentially over time are extremely common in business, agriculture, manufacturing, biological science, meteorology and virtually endless. Box-Jenkins methodology have been severally used to obtain forecasts through the process of identification, estimation and diagnostics checks. This study employed the use of time series filters to improve or obtain more quality forecasts by the modification of the Box-Jenkins forecasting approach. This study used the Nigerian Naira versus the US Dollar exchange rates data (4th January, 2010 to 17th September, 2019) to demonstrate the tool of time series filters to improve or obtain more quality forecasts. The study identified and reviewed two band-pass filters (Baxter-King Filter and Christiano-Fitzgerald Filter) designed for use in a wide range of economic applications. The classical forecasting technique of Box and Jenkins approach based on ARIMA model was also used to forecast the exchange rates. The results show that the use of Baxter-King filter ARIMA performed better than the Christiano-Fitzgerald filter and the classical Box-Jenkins forecasting ARIMA models. The forecast results indicate that the Nigerian economy within the sampled period did not conform to the theory behind ARIMA models for forecasting future exchange rates.

Keywords: ARIMA models, Box-Jenkins approach, Baxter-King filter, Christiano-Fitzgerald filter, exchange rates, forecasting, time series decomposition.

Published by: Department of Statistics, University of Benin,
Nigeria

1. Introduction

Data obtained from observations collected sequentially over time are extremely common in many areas of endeavour. In business, examples are weekly interest

*Corresponding author. Email: jonyeka-ubaka@unilag.edu.ng

rates, daily closing stock prices, monthly price indices, yearly sales figures, and so forth. In meteorology, examples include daily high and low temperatures, annual precipitation and drought indices, and hourly wind speeds. In agriculture, record annual figures for crop and livestock production, soil erosion, and export sales. In the biological sciences, observations are made of the electrical activity of the heart at milli-second intervals. In ecology, records of the abundance of an animal species (Cryer and Chan, 2008). The general ARIMA model for non-stationary series introduced by Box and Jenkins (1976) follows some practical steps which include identification, estimation, diagnostic checking and forecasting in capturing the empirical regularities of real life situations. Fitting non-stationary time series to statistical models can be challenging. As part of the identification stage of Box and Jenkins methodology, smoothing can be applied using time series filters. Smoothing is usually done to help us better see patterns or trends in time series. The term filter is sometimes used to describe a smoothing procedure. A band-pass filter will remove both high frequencies and low frequencies of a time series and leave only frequencies in a band in the middle (Nilsson and Gyomai, 2011). Band-pass filters tend to make even noise-look periodic, or at least quasi-periodic. Thus, filtering time or space series is such that certain frequencies or wave-numbers are removed and some are retained. That is, a method of accentuating certain frequencies and removing others.

The Christiano-Fitzgerald (CF) random walk filter is a band-pass filter that was built on the same principles as the Baxter and King (BK) filter. These filters formulate the de-trending and smoothing problem in the frequency domain. Given a continuous and/or infinitely long time series the frequency filtering could be an exact procedure. However, the granularity and finiteness of real life time series do not allow for perfect frequency filtering. Both the BK and CF filters approximate the ideal infinite band-pass filter. Box, et al. (2008) assume that the time series can be split into a sum of some deterministic components: seasonal oscillation, linear trend and a stochastic error term by the means of wavelet transform. That is, one can decompose a time series into a linear combination of different frequencies. Pre-filtering is a common practice for empirical analyses in economics to extract relevant information from economic data as accurately as possible. Particularly in macroeconomics, it attracts a great concern to identify trends and cyclical components in time series. For example, a trend component is used as an approximation to the potential output. Common studies in this regards includes Baxter and King (1999), Hodrick and Prescott (1999), Mills (1999), Box, et al. (2008) and Florian (2011). In contrast, cyclical components may approximate the business cycle, and are used to investigate the statistical validity of the real business cycle models, as done by Chatfield (1978). Also, seasonal effects may blur relations between economic variables, as such seasonally-adjusted series might be relevant for economic analyses. Ekhoehi and Ogbonmwan (2014) determine the exchange rate by solving an optimal control problem wherein the objective is to maximize the

balance-of-payments and the domestic interest rate are under statistical control.

A standard modelling approach for data sets with outliers is to describe extreme observations by means of heavy-tailed distributions. There is a vast literature on modelling continuous-valued time series with heavy-tailed distributions. Amongst are: Gorgi (2020) who discussed Beta-negative binomial autoregressions for modelling integer-valued time series with extreme observations. Chen et al. (2019) have discussed the estimation of integer-valued autoregressive models and binomial autoregressive models that are contaminated with innovational and additive outliers. Hall (2001, 2003) derived the extremal index of integer-valued moving averages. Scotto et al. (2018) studied the properties of a first-order max-integer-valued autoregressive model. Zhu et al. (2015) considered influence analysis of outliers for integer-valued generalized autoregressive conditional heteroscedastic (INGARCH) models. Importantly, Qian et al. (2020) gave their contributions on the first-order integer-valued autoregressive model. A widely used approach is the classical forecasting autoregressive integrated moving average (ARIMA) model which captures inter-temporal linear dependence in the data itself as well as in the error term. Existing trends are treated by modelling not the data but the differences which is then called ARIMA model (see Box, et al. (2008)).

This paper identified and reviewed two band-pass filters applied to improve the Box-Jenkins methodology. The identified band-pass time series filters that are useful for smoothing and extracting trend and cyclical components of a time series are Baxter-King Filter and Christiano-Fitzgerald Filter. The aim of this study is to apply the time series filters separately to recent economic data in the identification stage of Box-Jenkins methodology with a view to obtain optimum forecasts in comparison to the classical Box-Jenkins methodology. To check for the optimization of the forecasts generated using the respective time series filters, diagnostic tools of the in-sample error measures are used. The Box and Jenkins approach (ARIMA) model was used separately to forecast the Nigeria Naira versus United States Dollar exchange rates. Then the Baxter-King Filter and Christiano-Fitzgerald Filter were used to filter the data before applying the Box and Jenkins approach (ARIMA) model. The filters decomposed the data into trend and cycles. The filtered data using BK and CF, and the raw data were modelled by the ARIMA models to compare their model performances.

2. Methodology

2.1 *Filtration and decomposition of time series*

Time series decomposition is a mathematical procedure which transforms a time series into multiple different time series. Time series decomposition is a pattern-based technique (Nau, 2014). The four main components of time series data are trend, seasonality, cyclic and irregularity. Using band-pass filters, certain accen-

tuating frequencies or wave-numbers are removed and some are retained in the process of filtering and decomposition. Thus, both filtering and decomposition are geared towards improving forecasts estimation.

Given a time series $\{x_t\}_{t=1}^T$, x_t that may be decomposed into:

$$x_t = y_t + \bar{x} \quad (1)$$

where y_t is the isolated component from x_t . y_t oscillates between pi and pu (perceived usefulness), $2 \leq pi < pu < \infty$. The component y_t is assumed to have power only in the frequencies in the interval. In situations where x_t data are of infinite length,

$$y_t = B(L)x_t \quad (2)$$

where $B(L)$ is the lag operator defined as:

$$B(L) = \sum_{j=-\infty}^{\infty} B_j L^j \quad (3)$$

$$L^k x_t = x_{t-k}. \quad (4)$$

The ideal band-pass filter weights are given by

$$B_j = \frac{\sin(jb) - \sin(ja)}{\pi j}, \quad (5)$$

where $a = 2\pi/pu$ and $b = 2\pi/pi$. The finite sample approximation to the ideal band-pass filter uses the alternative filter

$$y_t = \hat{B}(L)x_t = \sum_{j=-\infty}^{\infty} \hat{B}_{t,j} x_{t+j} \quad (6)$$

where $\hat{B}_{t,j}$ is a solution to $\hat{B}_{t,j} = \arg \min \{(y_t - \hat{y}_t)\}$.

2.2 Baxter-King Filter

The Baxter-King filter is intended to explicitly deal with the periodicity of the business cycle. The band-pass filters designed for use in a wide range of economic applications. This filter 'detrends' the data, in the sense that it will render stationary time series that are integrated of order two or less, or that contain deterministic time trends (Baxter and King, 1999). By applying their band-pass filter to a series, they produce a new series that does not contain fluctuations at higher or lower than those of the business cycle. The Baxter-King filter is

a finite data approximation to the ideal band-pass filter with following moving average weights:

$$y_t = \hat{B}(L)x_t = \sum_{j=-\infty}^n \hat{B}_{t,j}x_{t+j} = \hat{B}_0x_t + \sum_{j=1}^n \hat{B}_t(x_{t-j} + x_{t+j}), \quad (7)$$

$$\hat{B}_j = B_j - \frac{1}{2n+1} \sum_{j=-n}^n \hat{B}_j. \quad (8)$$

2.3 Christiano-Fitzgerald Filter (CF)

The Christiano-Fitzgerald filter is a generalization of Baxter-King filter and can thus also be seen as weighted moving average. However, the CF filter is asymmetric about t as well as using the entire series. The Christiano-Fitzgerald filter computes cyclical and trend components of the time series using several band-pass approximation strategies to the ideal band-pass filter and minimizes the mean squared error defined in the equation:

$$\hat{B}_{t,j} = \arg \min E \{ (y_t - \hat{y}_t)^2 \}. \quad (9)$$

In the R package according to Shumway and Stoffer (2006), the default is $\theta = 1$ (i.i.d. series). If $\theta = (\theta_1, \theta_2, \dots)$, then the series is assumed to be

$$\hat{x}_t = \mu + 1_{\text{root}}x_{t-1} + \theta_1e_t + \theta_2e_{t-1} + \dots, \quad (10)$$

where $1_{\text{root}} = 1$ if the option `root = 1` and $1_{\text{root}} = 0$ if the option `root = 0`, e_t is a white noise. If drift in time series is evident (i.e. if `drift = TRUE`), the drift adjusted series is obtained as

$$\tilde{x}_t = x_t - t \left(\frac{x_T - x_1}{T - 1} \right), \quad t = 0, 1, 2, \dots, T - 1, \quad (11)$$

where $\tilde{x}_t$ is the undrifted series.

2.4 The Box and Jenkins approach

The approach starts with the assumption that the process that generated the time series can be approximated using an ARMA model if it is stationary or an ARIMA model if it is non-stationary. The process is a stochastic model building and an iterative approach that consists of the following steps make an ARIMA forecast model are model identification, estimation, model diagnostics and forecasting.

- (i) **Model Identification:** This is to identify the order of differencing (d) and choosing plausible values of p and q to identify the parameters of an ARMA model for the data. Differencing was adopted for the identification of p and q using the autocorrelation function (ACF) and partial autocorrelation function (PACF). If the data are stationary, the ACF should decay to zero fairly rapidly. Otherwise, differencing the series a second time may be necessary for a seasonal component. A random walk without drift is a process where the dependent variable can be estimated on one lagged period of itself plus an error term, assumed to be white noise, known as a random shock. The mean is constant over time in a random walk without drift, however, the variance is increasing indefinitely over time, making it a nonstationary stochastic process.

$$\text{Random walk without drift: } X_t = X_{t-1} + u_t. \quad (12)$$

Similarly, a random walk with drift is a process where the variable is dependent on its own lagged values and a random shock. However, the model that may be used to estimate a random walk with drift includes an intercept known as the drift parameter, denoted by δ . This parameter indicates if the time series is trending upwards or downwards, depending on whether δ is positive or negative. A random walk with drift is a non-stationary stochastic process since the mean and variance are increasing over time (Gujarati and Porter, 2008).

$$\text{Random walk without drift: } X_t = \delta + X_{t-1} + u_t. \quad (13)$$

The general $ARIMA(p, d, q)$ model introduced by Box and Jenkins (1976) and Box et al. (2008) includes autoregressive and moving average parameters, and explicitly involves differencing in their formulation. The general form of the $ARIMA(p, d, q)$ process is of the form

$$\phi_p(B)(1 - B)^d X_t = \theta_0 + \theta_q(B)\epsilon_t. \quad (14)$$

where the autoregressive (AR) operator $\phi_p = (1 - \phi_1 B - \dots - \phi_p B)$ are invertible while the moving average (MA) operator $\theta_p = (1 - \theta_1 B - \dots - \theta_p B)$ are stationary. Both AR and MA share no common factors.

Exploratory results which include descriptive statistics and time plots, among others, is used to determine the patterns, trends, cycles and seasonality to check if the series is stationary or not. A further test called Augmented Dickey Fuller test (ADF) is also used to determine whether the series is stationary. The ADF tests the null hypothesis that the series is non-stationary. If the P-value of the ADF test is above 0.05, it means that the data is non-stationary. Hence, the need to difference the series until we obtain a stationary time series. The differencing is called the integration

point in ARIMA which is the d parameter of the $ARIMA(p, d, q)$ model. Once a stationary time series is achieved, the next step is to select the appropriate ARIMA model finding the most appropriate values for p and q in the $ARIMA(p, d, q)$ model.

Another method to test for stationarity is by computing the autocorrelation function, also known as the ACF. The autocorrelation function is the ratio between the covariance at a specific lag, generally expressed as lag k , to the variance. At lag k , ρ_k denotes the ACF and is defined as follows; $\rho_k = \gamma_k / \gamma_0$, where γ_k is the covariance at lag k and γ_0 is the variance. The ACF can be plotted by using a "correlogram". In the correlogram, if all or most of the lags are statistically insignificant, there is no specific pattern, constant variance, and the autocorrelations at various lags hovers around zero, the time series could be regarded as stationary. This means that a time series is most likely stationary if the ACF correlogram resembles a white noise process.

- (ii) **Parameter Estimation:** Estimation involves using numerical methods to minimize a loss or error term. It involves the estimation of the parameters of the models identified. In this paper, least squares method was adopted to obtain optimal values. The idea of least squares estimation is to find the parameters that minimize the sum of squared errors. Hence, the sum of squared errors is given as

$$S = \sum_{t=1}^n a_t^2 = \sum_{t=1}^n \left(\sum_{j=1}^q \theta_j a_{t-j} + y_t - \delta - \sum_{i=1}^p \phi_i y_{t-i} \right)^2. \quad (15)$$

- (iii) **Diagnostic Checks and Performances Evaluation:** The adequacy of a model and its forecasting performance is evaluated using information criteria and some diagnostics statistics. Akaike's information criterion (AIC) (1970) and Bayesian Information Criterion (BIC) (1978) are information criteria that measures the goodness of fit of a model by assigning an informational cost to the number of parameters to estimate. Generally, a good model is obtained using the minimized values of the AIC and BIC. Also forecasted results from the models can be tested using any of the following statistics (error measures):

- (a) The Root mean square error (RMSE) depends on the scale of the dependent variable. It should be used as relative measure to compare forecasts for the same series across different models. The smaller the error, the better the forecasting ability of that model according to the

RMSE criterion.

$$RMSE = \sqrt{\frac{1}{n+1} \sum_{t=s}^{s+n} (\hat{y}_{t-1}(1) - y_t)^2}. \quad (16)$$

One problem associated with the use of the RMSE or similar measures is the fact that the forecast error variance vary across time. It can vary because of nonlinearities in the model and because of variation in exogenous variables (if included in the model).

(b) Mean square error (MSE) measures the average squared difference between the forecast and actual values. The MSE mean squared forecast error can be decomposed as

$$\sum_{t=1}^n (\hat{y}_{t-1}(1) - y_t)^2 / n = (\bar{\hat{y}} - \bar{y})^2 + (s_{\hat{y}} - s_y)^2 + 2(1 - r)s_{\hat{y}}s_y. \quad (17)$$

where $\bar{\hat{y}}$ and $\bar{y}$ are the means of the forecast variable and of the true variable, respectively. $s_{\hat{y}}$ and s_y are the standard deviations of the forecast variable and of the true variable, respectively. r is the correlation between $\hat{y}$ and the true variable y . The proportions are defined as follows: The bias proportion

$$\frac{(\bar{\hat{y}} - \bar{y})^2}{\sum_{t=1}^n (\hat{y}_{t-1}(1) - y_t)^2 / n}$$

measures how far the mean of the forecast is from the mean of the actual series. The variance proportion

$$\frac{(s_{\hat{y}} - s_y)^2}{\sum_{t=1}^n (\hat{y}_{t-1}(1) - y_t)^2 / n}$$

measures how far the variance of the forecast is from the variance of the actual series. The covariance proportion

$$\frac{2(1 - r)s_{\hat{y}}s_y}{\sum_{t=1}^n (\hat{y}_{t-1}(1) - y_t)^2 / n}$$

measures the remaining unsystematic forecasting errors. If the forecast is 'good', the bias and the variance proportions should be small so that most of the bias should be concentrated on the covariance proportions.

(c) The Mean absolute error (MAE) is dependent on the scale of

the dependent variable but it is less sensitive to large deviations than the usual squared loss.

$$MAE = \frac{1}{n+1} \sum_{t=s}^{s+n} |\hat{y}_{t-1}(1) - y_t|. \quad (18)$$

One problem with the MAE is that the relative size of the error is not always obvious. Sometimes it is hard to tell a big error from a small error. To deal with this problem, we can find the mean absolute error in percentage terms. Both RMSE and MAE are relative and scale dependent i.e. should be used to compare forecasts of the same time series across different forecasting models. The smaller the RMSE and MAE, the better the forecasting performance of the model. The mean percentage error (MPE) is the computed average of percentage errors by which forecasts of a model differ from actual values of the quantity being forecast.

(d) The mean absolute percentage error (MAPE) is one of the most widely used measures of forecast accuracy, due to its advantages of scale-independency and interpretability. MAPE measures the (absolute) size of each error in percentage terms, then averages all percentages. The MAPE is scale independent given as

$$MAPE = \frac{1}{n+1} \sum_{t=s}^{s+n} \left| \frac{\hat{y}_{t-1}(1) - y_t}{y_t} \right|. \quad (19)$$

MAPE allows us to compare forecasts of different series in different scales but was criticized for the problem of asymmetry and instability when the original value is small. MAPE as accuracy measure is affected by the four problems: (i) Equal errors above the actual value result in a greater absolute percentage error; (ii) Large percentage errors occur when the value of the original series is small; (iii) Outliers may distort the comparisons in empirical studies; (iv) MAPEs cannot be compared directly with naive models such as random walk.

(e) Theil's U statistics or Theil's coefficient of inequality is another criterion that measures forecast accuracy. There are two Theil's coefficients labeled as Theil U_1 and Theil U_2 coefficient.

$$U_1 = \frac{\sqrt{\frac{1}{n+1} \sum_{t=s}^{s+n} (\hat{y}_{t-1}(1) - y_t)^2}}{\sqrt{\frac{1}{n+1} \sum_{t=s}^{s+n} \hat{y}_{t-1}^2(1)} + \sqrt{\frac{1}{n+1} \sum_{t=s}^{s+n} y_t^2}}. \quad (20)$$

<http://www.bjs-uniben.org/>

$$U_2 = \frac{\sqrt{\frac{1}{n+1} \sum_{t=s}^{s+n} (\hat{y}_{t-1}(1) - y_t)^2}}{\sqrt{\frac{1}{n+1} \sum_{t=s}^{s+n} y_t^2}}. \quad (21)$$

Values closer to 0 for both, Theil's U_1 and U_2 criteria, indicate better forecasting performance of the evaluated models; if Theil's U_1 and U_2 are equal to zero then the forecast is perfect. Theil U_1 is bounded between 0 and 1, whereas Theil U_2 is not bounded.

Model verification can further be achieved by computing the residuals from the fitted model to check that they are consistent with white noise. The errors from an ideal model would resemble white noise that is, a Gaussian distribution with a mean of zero and a symmetrical variance.

- (iv) **Forecasting:** Forecasting is one of the main objectives of time series analysis. The forecasting errors are often included in the measurements and two measurements that are based on the relative forecasting errors are the mean percentage error and the mean absolute percentage error (Philip and Hugo, 2020). A forecast “error” is the difference between an observed value and its forecast, therefore we use the criterion of minimum mean squared errors to produce point forecasts. Like other statistical forecasts, there are two main sources of uncertainty involved. The first uncertainty is concerned with “future” variables and the second is the uncertainty about the model used. In wavelet-based forecasting, wavelets can be used to improve the forecasting accuracy of the classical forecasting models like the ARMA and its extension ARIMA models. (See Brockwell and Devis, 2002; Sue, 2009; Schluter and Deusche, 2010; Fat and Eva, 2011; Idrees et al. 2019).

3. Results and Discussion

A 10-year historical data recorded daily from 4th January 2010 to 17th September 2019 of US Dollars to Nigeria Naira exchange rates data were collected from Central Bank of Nigeria website (www.cbn.gov.ng). Its attributes include buying price, central price and the selling price and contains a total of 2394 observations. The central price of the exchange rate was used for this paper. Importantly, exchange rate influences decisions taken by the participants of the foreign exchange market, namely investors, importers, exporters, bankers, financial institutions, business, tourists and policy makers both in the developing and developed world as well. Given the US Dollars to Nigerian Naira exchange rates daily data (4th January, 2010 to 17th September, 2019), exploratory results of the descriptive statistics and time plot are shown in Table 1 and Figure 1(a), respectively. The minimum and maximum values for which naira was traded

for dollar under the period considered are N147.1 and N324.5, respectively. The mean exchange rate is N210.5. The sharp drop in the time plot seems to suggest that while falling oil prices (2013-2014) would support activity and reduce inflation globally, Nigeria being an oil-exporting country, is under stress as falling oil-related revenues put fiscal balances under pressure and exchange rates depreciate on deteriorating growth prospects. Again, macroeconomic fundamentals and financial exposure of countries are found to have played a key role in the transmission process of US shocks: in particular Nigeria with low foreign exchange reserves, weak current account positions and high direct financial exposure vis-à-vis the United States have experienced substantially larger currency depreciations during the crisis overall. The time plot also shows an additive time series since the random fluctuations in the data are roughly constant in size over time. Thus, indicating that the data set is non-stationary. A stochastic process is weakly stationary if it has constant mean and variance and the covariance is time invariant, i.e. the statistics do not change over time (Gujarati and Porter 2008). The data set being an additive time series, seasonal fluctuations or the variations around trend does not vary much with time as shown in Figure 1(b) where the decomposition varies from trend, seasonal and random. The observed part of the decomposed data further shows that the data is non-stationary.

Table 1: Summary statistics of US Dollars to Nigerian Naira exchange rates

	Min.	Q1	Q2	Mean	Q3	Max.
Dollars to Naira Exchange Rates	147.1	155.2	156.6	210.5	305.1	324.5

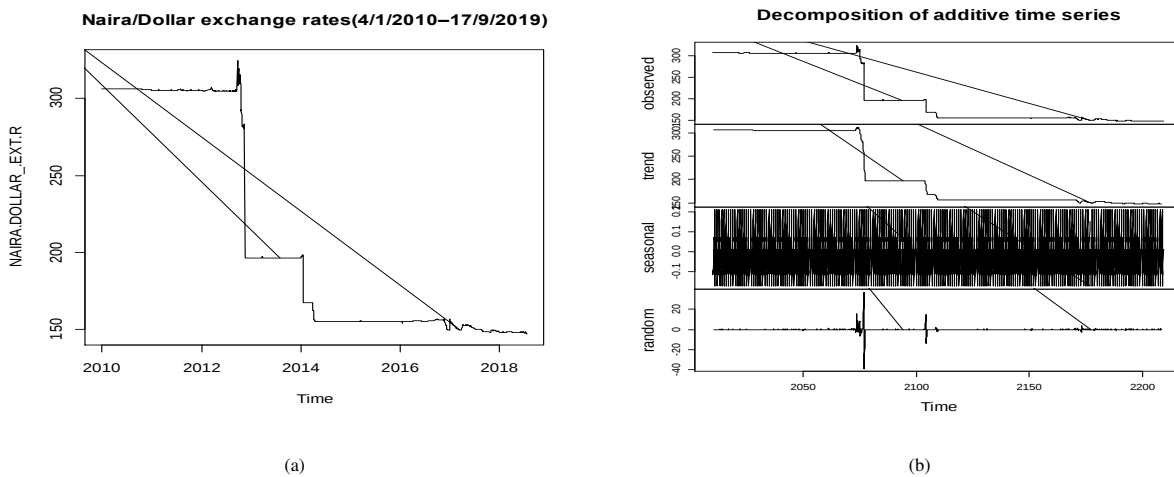


Figure 1: (a) Time plot of Naira/USD exchange rate. (b) Plot of decomposed Naira/Dollar exchange rate.

The Augmented Dickey-Fuller test based on the following test of hypotheses (H_0 : Non-stationary series against H_1 : Stationary series) was used where if the p-value of the ADF test is above 0.05, it means that the data is non-stationary.

ADF test results indicate that the data gives a p-value of 0.8925 which supports the null hypothesis in affirmation that the data is not stationary (Dickey-Fuller = -1.258, lag order = 13, p-value = 0.8925). The plot of ACF (Figure 2) is

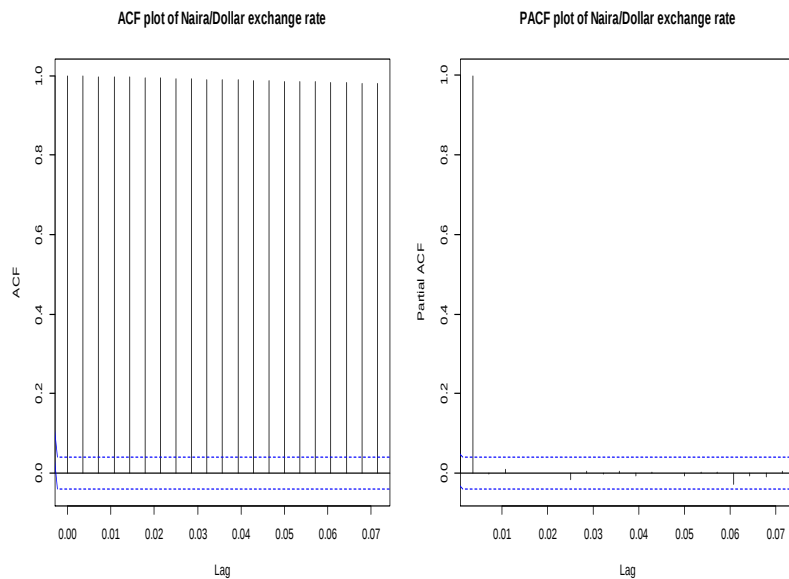


Figure 2: Plot of ACF and PACF of non-stationary data

decreasing or decaying very slowly and remains well above the significance range (dotted blue line). This is an indication that the series is a non-stationary time series.

After the first differencing of the series, the ADF test gave results: Dickey-Fuller = -22.33, lag order = 13, p-value = 0.01. The p-value of 0.01 supports the acceptance of the alternative hypothesis indicating that the data set is now stationary. Also the ACF and PACF confirm that the series is now stationary as they deceased fast and spike after some lags indicating that ARMA model should be tried as well (see Figure 3).

Using least squares method, the values of parameters were estimated for possible ARIMA (p, d, q) models with AIC to determine the best model among the models (Table 2 and 3).

Table 2: Box-Jenkins ARIMA models

Model	AIC
ARIMA(2,1,2) with drift	9972.049
ARIMA(0,1,0) with drift	9963.407
ARIMA(1,1,0) with drift	9966.329
ARIMA(0,1,1) with drift	9965.326
ARIMA(0,1,1) with drift	9965.326
ARIMA(0,1,0)	9964.222
ARIMA(1,1,1) with drift	9968.327

The ARIMA (0, 1, 0) with drift which is a random walk was found to be suitable for modelling the data. These results are consistent with the works of Nor et al. (2013). The optimal out-of-sample forecast performance of the models is

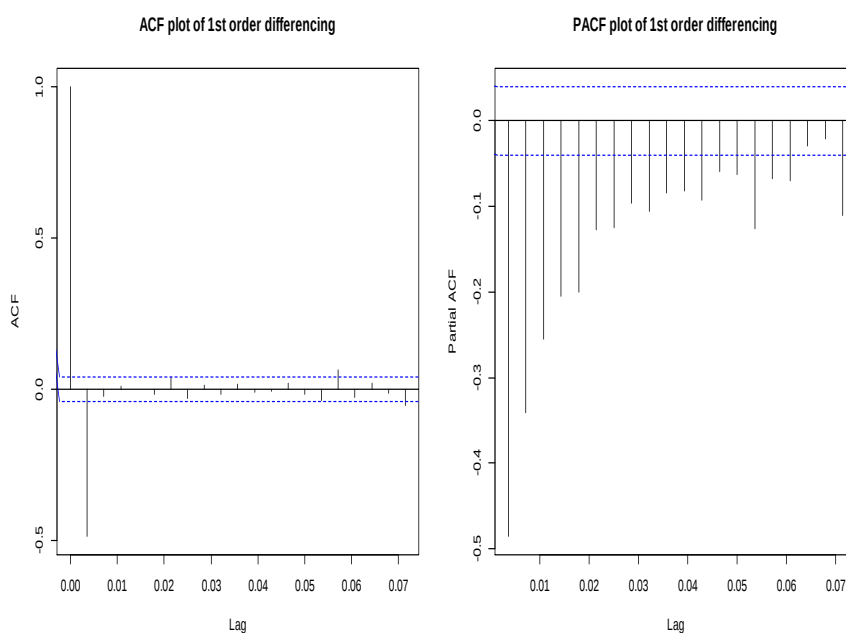


Figure 3: ACF and PACF plots of differenced data

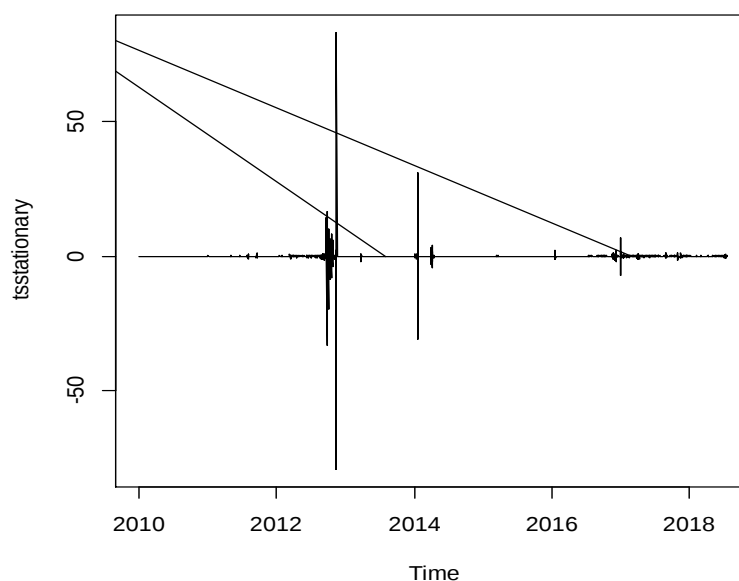


Figure 4: Stationary time plot of Naira/USD exchange rate (4/1/2010 – 17/9/2019)

measured in terms of the errors statistics of the selected ARIMA models fitted to the data set and presented on Table 4. All the models appeared to be good but the random walk model outperforms them as indicated by the Theil's U coefficient in the case of the Naira/USD exchange rate. This is apparently consistent with Meese and Rogoff (1983) who showed the superiority of the random walk model in out-of-sample exchange rate forecast. However, the Theil's U evaluation statistic shows that the ARIMA (0, 1, 0) model outperforms other ARIMA models and is therefore chosen as the best model. The ARIMA (0, 1, 0) with

Table 3: Summary statistics of US Dollars to Nigerian Naira exchange rates

ARIMA(0,1,0) with drift	
Coefficients	
Drift	-0.0666

s.e. 0.0397, σ^2 estimated as 3.769, log-likelihood = -4980.37,
AIC = 9963.407, AIC=9964.74, BIC = 9976.29.

Table 4: Summary statistics of US Dollars to Nigerian Naira exchange rates

Models	AIC	Theil's U	ME	RMSE	MAE	MPE	MAPE
ARIMA(2,1,2)	9972.049	1.1843	0.3537	1.9511	1.2481	2.8077	9.5822
ARIMA(2,1,0)	9964.222	1.1854	0.5532	1.9849	1.2556	2.8094	9.6710
ARIMA(0,1,2)	9970.135	1.1886	0.6943	2.0357	1.3492	2.8175	9.8727
ARIMA(1,1,0)	9966.329	1.0867	0.5796	1.8596	1.2857	2.7681	9.6971
ARIMA(0,1,1)	9965.326	1.0893	0.7354	2.0348	1.3598	2.7413	9.8946
ARIMA(0,1,0)	9963.407	1	1.0864	1.9405	1.2212	2.5223	9.2543
ARIMA(1,1,1)	9968.327	1.0852	0.3675	1.9412	1.2341	2.6512	9.7461

drift was used to make a 226-point ahead forecast of Naira/USD exchange rate, that is, between September 18, 2019 – April 30, 2020. The time axis label of the time periods (in years) of 0, 500, 1000, 1500, 2000 and 2500 represent 2010, 2012, 2014, 2016, 2018 and the forecast 2500 represents 2020, respectively as earlier seen in Figure 1. The forecast was at 80% and 95% confidence interval. The forecast results for the model is as plotted in Figure 5. The Dark gray part

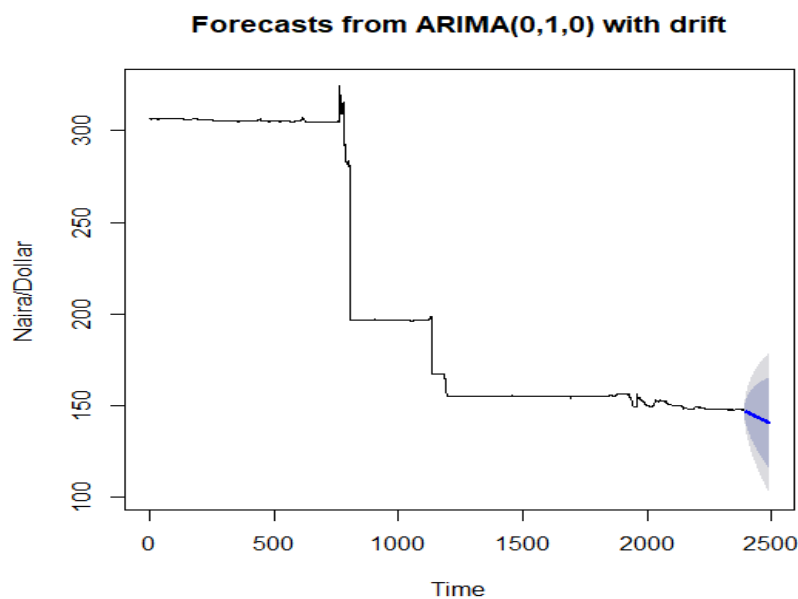


Figure 5: Plot of Box-Jenkins ARIMA with forecasts

of the plot is the 80% confidence interval, the light gray part is the 95% confidence interval, and the blue line is the actual prediction. The forecast errors are stored as “residuals” were noted for further analysis to check the pattern in the residuals of the chosen model. The ACF, PACF and Q-Q plot of the residuals for the ARIMA (0, 1, 0) model are given in Figure 6. As seen from the ACF and PACF plots for the ARIMA (0, 1, 0) residuals, almost all the correlations are below the threshold limit. This means that the residuals are behaving like white

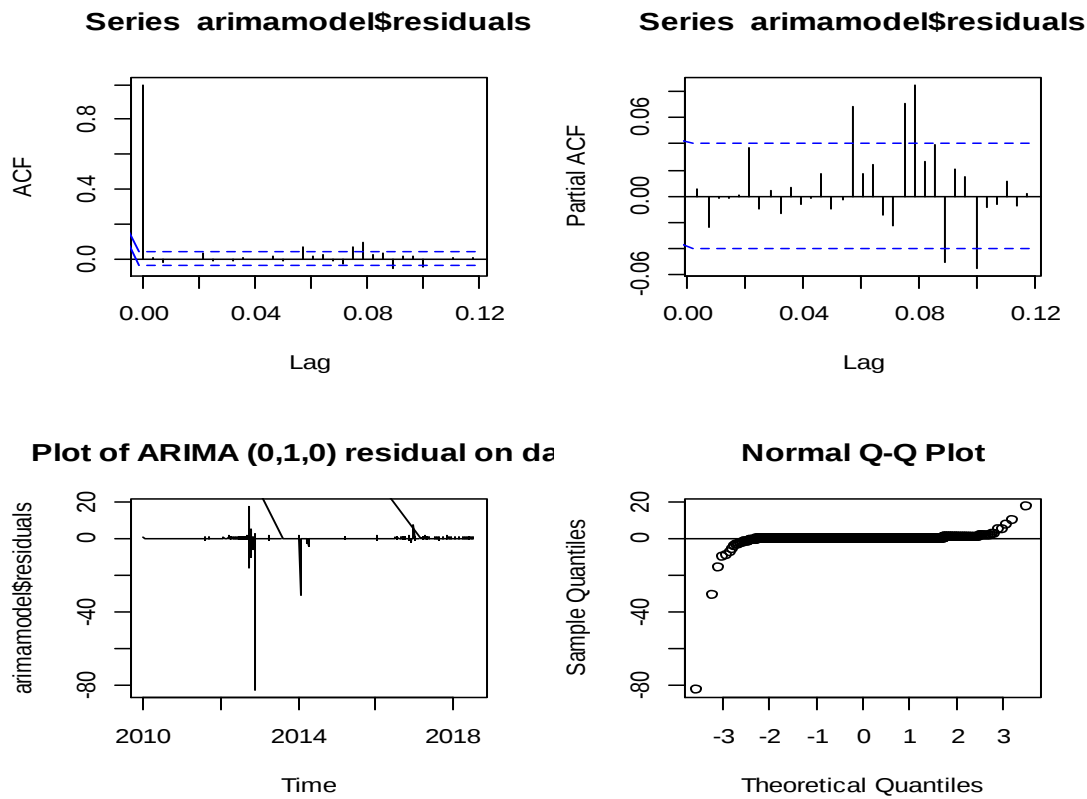


Figure 6: ACF, PACF, time plot and Q-Q-plot of ARIMA (0, 1, 0) residuals

noise and the residuals are distributed normally as testified by Q-Q plot, Figure 6.

There are 4 forecasts generated by a random walk (ARIMA (0, 1, 0)), an ARIMA (1, 1, 0), an ARIMA (0, 1, 1), and a forecast generated by a model selected based on BIC criterion. Table 5 presents the change of the values produced by the accuracy measures using the last 20 points. The first two accuracy measures-MAPE and RMSE-produced very different values after the transformation since they are not scale transformation invariant. The last accuracy measure- Theil's inequality coefficient-had some minor changes. This suggested that the first two measures are generally not good for cross-series comparison of forecasting procedures since a linear transformation of the original series may change the ranking of the forecasts. Hence, in terms of forecasting performance, the results in Table 5 indicate that ARIMA (0, 1, 0) is the best among the ARIMA models considered, but both the original and new series of this model are as good as the naïve model. Analyzing the data with smoothing

Table 5: Comparison of different criteria after transformation $y^{new} = 10y^{original} - 10\text{mean}(y^{original})$

Forecast Series	ARIMA Original	(0, 1, 0) New	ARIMA Original	(1, 1, 0) New	ARIMA Original	(0, 1, 1) New	ARIMA Original	BIC New
MAPE	0098	0327	0094	0331	0091	0321	0028	0406
RMSE	0618	6394	0625	6401	0610	6328	0756	7899
Theil's U_1	1	1	1.029	1.029	1.092	1.092	1665	1574

techniques, the Baxter-King filter (BKF) and Christiano-Fitzgerald Filter (CFF) results of trends and cyclic components are presented in Figures 7 and 8. The summary statistics are in Tables 6 and 7 for the BKF and CFF respectively.

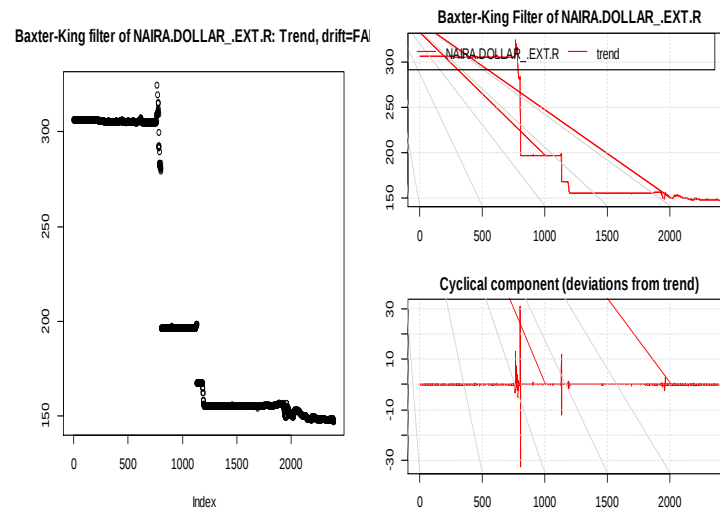


Figure 7: BK filter of Naira/ Dollar exchange rates of trend and cyclic components

Table 6: Baxter-King filter decomposition of series into trend and cycle

Summary Statistics	Naira/USD Ex-Rate	Trend	Cycle
Minimum	147.1	147.4	32.74084
1st Quarter	155.2	155.2	-0.00947
Median.	156.6	156.4	0.00000
Mean.	210.5	210.4	0.00003
3rd Quarter	305.1	305.1	0.00930
Maximum	324.5	312.7	31.06650

Table 7: Baxter-King filter decomposition of series into trend and cycle

Summary Statistics	Naira/USD Ex-Rate	Trend	Cycle
Minimum	147.1	64.79	-66.36
1st Quarter	155.2	133.81	-40.10
Median.	156.6	205.93	31.15
Mean.	210.5	189.75	20.73
3rd Quarter	305.1	243.91	66.20
Maximum	324.5	320.77	140.27

3.1 Modelling the BK filtered data with ARIMA model

The BK filter decomposed the data into trend and cycles. Each of the two components was modelled with ARIMA. The trend component of the filtration was modelled with ARIMA (4, 1, 1) with zero mean, variance estimated as 0.01603, log likelihood = 1548.47, AIC = 8082.93, BIC = 8145.49. While cyclic component of decomposed data was modelled with ARIMA (0,0,2) with zero mean, variance estimated as 1.166, log likelihood = -3570.36, AIC = 7146.71, BIC =

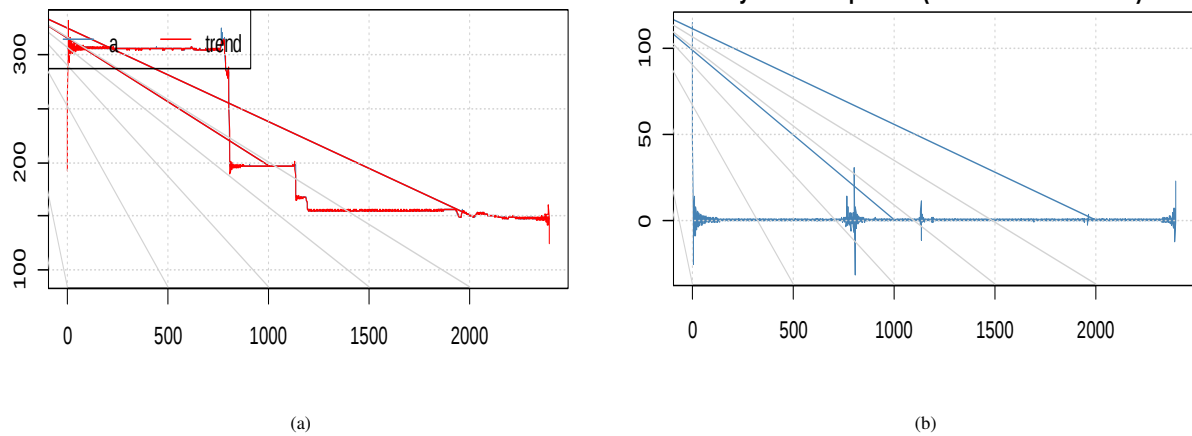


Figure 8: CF filter of Naira/Dollars rate trend and cyclic component, respectively

7164.05 (see Table 8 and 9). The model performance for the trend and cyclic components based on BK Filter are presented on Tables 10 and 11 with the value of MAPE having the value of infinity. However, we note that MAPE has the significant disadvantage that it produces infinite or undefined values for zero or close-to-zero actual values. This may be addressed using the mean arctangent absolute percentage error (MAAPE) which is a slope as an angle, while MAPE is a slope as a ratio by considering a triangle with adjacent and opposite sides that are equal to an actual value and the difference between the actual and forecast values, respectively. MAAPE inherently preserves the philosophy of MAPE, overcoming the problem of division by zero by using bounded influences for outliers in a fundamental manner through considering the ratio as an angle instead of a slope.

Table 8: ARIMA modelling of BKF trend component

Model	AIC
ARIMA(2,1,2) with drift	-2988.597
ARIMA(0,1,0) with drift	5711.006
ARIMA(1,1,0) with drift	1366.945
ARIMA(0,1,1) with drift	2659.154
ARIMA(0,1,0) without drift	5725.488
ARIMA(2,1,1) with drift	-2823.007
ARIMA(3,1,1) with drift	3074.304
ARIMA(4,1,2) with drift	8062.994
ARIMA(4,1,1) with drift	8082.932
ARIMA(4,1,0) with drift	-2979.854
ARIMA(5,1,1) with drift	Infinity
ARIMA(3,1,0) with drift	-2807.641

3.2 Modelling of Christiano-Fitzgerald filtered data with ARIMA model

The CF filter also decomposed the data into trend and cycles. Each of the two components was modelled with ARIMA. Based on the AIC, ARIMA (0, 1, 0),

Table 9: ARIMA modelling of BKF cycle component

Model	AIC
ARIMA(2,0,2) with non-zero mean	Infinity
ARIMA(0,0,0) with non-zero mean	7635.3
ARIMA(1,0,0) with non-zero mean	7633.853
ARIMA(0,0,1) with non-zero mean	7627.067
ARIMA(0,0,0) with zero mean	7633.3
ARIMA(1,0,1) with non-zero mean	7593.053
ARIMA(2,0,1) with non-zero mean	Infinity
ARIMA(1,0,2) with non-zero mean	Infinity
ARIMA(0,0,2) with non-zero mean	7146.388
ARIMA(0,0,3) with non-zero mean	Infinity
ARIMA(1,0,3) with non-zero mean	Infinity
ARIMA(0,0,2) with zero mean	7144.388
ARIMA(0,0,1) with zero mean	7625.067
ARIMA(1,0,2) with zero mean	Infinity

Table 10: Model performance for the trend component based on BK filter

Measures	ME	RMSE	MAE	MPE	MAPE	MASE
Statistics	0.0001332925	0.1264074	0.01932013	0.000509902	0.00908376	0.1971003

Table 11: Model performance for cyclic component based on BK filter

Measures	ME	RMSE	MAE	MPE	MAPE	MASE
Statistics	0.0002217488	1.079319	0.1502145	0.17902473	Infinity	0.8880269

with variance estimated as 5.405, log likelihood = -5412.08, AIC = 8826.15, BIC = 8931.93 which is a random walk was found suitable for the trend component of the Christiano-Fitzgerald filtered data. While ARIMA(2, 0, 0) which is AR(2) model with variance estimated as 5.641, log likelihood = -5465.34, AIC = 7938.68, BIC = 7961.8 was found suitable for the cyclic component of the Christiano-Fitzgerald filter (see Tables 12-15).

Table 12: ARIMA modelling of CF trend component

Model	AIC
ARIMA(2,1,2) with drift	Infinity
ARIMA(0,1,0) with drift	8825.7
ARIMA(1,1,0) with drift	6580.228
ARIMA(0,1,1) with drift	9558.548
ARIMA(0,1,0) without drift	8824.47
ARIMA(2,1,0) with drift	2215.268
ARIMA(5,1,0) with drift	Infinity
ARIMA(3,1,0) with drift	-2276.204
ARIMA(4,1,0) with drift	-6821.881
ARIMA(5,1,0) without drift	Infinity
ARIMA(5,1,0) with drift	-11410.95
ARIMA(4,1,0) with drift	Infinity

The Baxter-King Filter (BKF) and Christiano-Fitzgerald filter (CFF) compared on the basis of the Mean Error (ME), Mean Square Error (MSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), and Mean Absolute Percentage Error (MAPE) were presented on Table 16.

The results of all the statistics presented in Table 16 showed that the Baxter-King filter performed better than the Christiano-Fitzgerald filter as evinced from the smaller error measures.

Table 13: ARIMA modelling of CF cycle component

Model	AIC
ARIMA(2,0,2) with non-zero mean	Infinity
ARIMA(0,0,0) with non-zero mean	1291.35
ARIMA(1,0,0) with non-zero mean	12023.57
ARIMA(0,0,1) with non-zero mean	12061.23
ARIMA(0,0,0) with zero mean	12490.53
ARIMA(2,0,0) with non-zero mean	7938.68
ARIMA(3,0,0) with non-zero mean	Infinity

Table 14: Model performance for the trend component based on CF filter

Measures	ME	RMSE	MAE	MPE	MAPE	MASE
Statistics	-0.04156251	2.324321	0.4637462	0.03748878	0.2255418	0.9997547

Table 15: Model performance for the cyclic component based on CF filter

Measures	ME	RMSE	MAE	MPE	MAPE	MASE
Statistics	0.002631386	2.373621	0.3391235	21.61788	152.4224	0.6435376

Table 16: Comparison of BK and CF filtered data ARIMA models

In-sample error measures	Baxter-King filter	Christiano-Fitzgerald filter
ME	2.838e-05	2.073e+01
MSE	1.432e+00	3.950e+03
MAE	1.311e-01	5.506e+01
MPE	-4.457e-05	7.137e-02
MAPE	5.944e-04	2.708e-01

Based on the statistics in Table 17, the BK filtered data modelled with ARIMA model gives the best model followed by the CF filter and then the raw data modelled with Box-Jenkins ARIMA. Thus, filtering the data actually improved the performance of the Box-Jenkins ARIMA model. The results also show that BK filtration of the data improved the performance of the Box-Jenkins ARIMA model more than the CF filter.

4. Conclusion

This paper has demonstrated the effectiveness of two identified time series filters to improve the Box-Jenkins approach to time series analysis and forecasting. The results confirm the statements of Tran and Uyen (2016), Tsay (2016), and Jason (2017) that wavelet-based processes improve the forecasting quality. The Naira/Dollar exchange rate data decomposed into cycles and trend with the aid of Baxter-King and the Christiano-Fitzgerald filters as smoothing techniques that use wavelet-based methods of forecasting give optimal forecasts in terms of the error measures than the general ARIMA. The filtered data using BK generated ARIMA (4, 1, 1) model while filtered data using CF generated ARIMA (0, 1, 0) model similar to ARIMA (0, 1, 0) model generated by the classical Box-Jenkins approach. In comparison and based on assessment, the result showed that the BK filtration of the data improved the performance of the model more than the CF filter and the classical Box-Jenkins ARIMA model.

Table 17: Comparison of performance by ARIMA models

Statistical Measures	Box-Jenkins ARIMA (0,1,0)	BK (trend) ARIMA (4,1,1)	BK (cyclic) ARIMA (0,0,2)	CF (trend) ARIMA (0,1,0)	CF (cyclic) ARIMA (2,0,0)
RMSE	1.94051	0.1264074	1.079319	2.324321	2.373621
MAE	0.192118	0.01932013	0.1502145	0.4637462	0.3391235
MPE	0.000289	0.000509902	0.17902473	-0.03748878	21.61788
MAPE	0.092543	0.00908376	Infinity	0.2255418	152.4224
AIC	9964.73	8082.93	7146.71	8826.15	7938.68
BIC	9976.29	8145.49	7164.05	8931.93	7961.8

Thus, for an improved time series analysis and forecasting, it is imperative to include an appropriate time series filter in the classical Box-Jenkins approach.

It should be worthy to note that the forecast results in this paper indicate that the Nigeria economy within the period under review did not conform to the theory behind ARIMA model for forecasting future exchange rates. This may be attributed or due to the currency devaluation by the Central Bank of Nigeria within the sample period that can make the exchange rate to follow an exponential function of time. The paper, therefore, suggests that the government should adopt either a flexible exchange rate policy over a long period or a fixed exchange rate regime as opposed to the demand and supply schedules of Iyoha (2004) to improve on the macroeconomic policy framework for exchange rate determination in Nigeria. It is therefore imperative that further explorative research be carried out to determine the empirical ways to minimize or eradicate the fluctuation of the Nigeria Naira.

Acknowledgement

The authors are grateful to the anonymous reviewers/referees for many helpful constructive criticisms and suggestions which significantly improve the paper.

References

- Akaike, H. (1970). Statistical predictor identification. *Annals of the Institute of Statistical Mathematics*, **22**: 203-217.
- Baxter, M and King, R.G. (1999). Measuring Business Cycles: Approximate band-pass Filters. *The Review of Economics and Statistics*, **81**(4): 575 - 593.
- Box, G. and Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control* (Revised Ed.). Holden- Day, Oakland, California.
- Box, G., Jenkins, G., and Reinsel, G. (2008). *Time Series Analysis: Forecasting and Control* (4th Ed.). Wiley, Oxford.
- Brockwell, P.J. and Davis, R.J. (2002). *Introduction to Time Series and Forecasting*. (2nd Edition) Springer.
- Chen, H., Li, Q. and Zhu, F. (2019). Binomial AR (1) processes with innovational outliers. *Commun. Statist. Theory Meth.* doi 10.1080/03610926.2019.1635704.
- Chatfield, C. (1978). The Holt-Winters forecasting procedure. *Journal of Royal Statistical Society: Series C*, **27**(3): 264-279.
- Cryer J. D. and Chan K. S. (2008). *Time Series Analysis with Applications in R*. Springer Texts in Statistics, Springer Science + Business Media, LLC.

- Ekhosuehi, V. U. and Ogbonmwan, S. M. (2014). Determination of the optimal exchange rate via control of the domestic interest rate in Nigeria. *Operations Research and Decisions*, **24**(1): 23-36. doi: 10.5277/140102.
- Fat, C. M. and Eva, D. (2011). Exchange rates forecasting: Exponential smoothing techniques and ARIMA models. *RePEc:ora:journal*, **1**(1): 499 - 508.
- Florian, P. (2011). *Univariate Time Series*. University of Lausanne, Ecole des HEC Department of Mathematics (IMEA-Nice). Lecture5: Box – Jenkins methodology, 3 - 32.
- Gorgi, P. (2020). Beta-negative binomial auto-regressions for modelling integer-valued time series with extreme observations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. <https://doi.org/10.1111/rssb.12394>.
- Gujarati, N.D. and Porter, D.C. (2008). *Basic Econometrics*. 5th Edition, New York: McGraw-Hill.
- Hall, A. (2001). Extremes of integer-valued moving average models with regularly varying tails. *Extremes*, **4**: 219- 239.
- Hall, A. (2003). Extremes of integer-valued moving average models with exponential type tails. *Extremes*, **6**: 361- 379.
- Hodrick, R.J. and Prescott, E.C. (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, **29**(1): 1-16.
- Idrees, S. M., Afshar Alam, M., Agarwal, P. (2019). A prediction approach for stock market volatility based on time series data. *IEEE Access*, **7**: 17287-17298. doi: 10.1109/ACCESS.2019.2895252.
- Iyoha, M. A. (2004). *Macroeconomics: Theory and Policy*. Mindex Publishing, Benin City.
- Jason, B. (2017). A gentle introduction to exponential smoothing for time series forecasting in Python and a gentle introduction to the Box-Jenkins method of time series. <https://www.machinelearningmastery.com>
- Meese, R.A. and Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, **14**(1-2): 3-24. doi.org/10.1016/0022-1996(83)90017-x
- Mills, T. C. (1999). *The Econometric Modelling of Financial Time Series* (2nd Ed.). Cambridge University Press..
- Nau, R. (2014). *Mathematical Structure of ARIMA Models*. Fuqua School of Business, Duke University. <https://www.People.duke.edu/mall/forecasting.htm>
- Nilsson, R. and Gyomai, G. (2011). Cycle Extraction: A Comparison of the Phase-Average Trend Method, the Hodrick-Prescott and Christiano-Fitzgerald Filters. *OECD Statistics working paper*, 2011(04): 4 - 8.
- Nor, H., Miswan P., Yean, P. and Maizah, H. A. (2013). Parameter estimation for Malaysian gold prices Modelling and Forecasting. *Int. Journal of Math. Analysis*, **7**(22): 1059 - 1068.
- Philip, J. and Hugo, L. (2020). *ARIMA Modeling Forecasting Indices on the Stockholm Stock Exchange*. Bachelor Thesis-Finance, Karlstad Business School Karlstad University SE-651 88 Karlstad.
- Qian, L., Li, Q. and Zhu, F. (2020). Modelling heavy-tailedness in count time series. *Appl. Math. Modelling*, **82**: 766–784.
- Shumway, R.H. and Stoffer, D.S. (2006). *Time Series Analysis and Its Applications with R Examples* (2nd Ed.). Springer.
- Schluter, S. and Deuschle, C. (2010). *Using Wavelets for Time Series Forecasting: Does it pay off?* University of Erlangen-Nuremberg, Institute for Economics, 2-10.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**: 461-464.
- Scotto, M. G., Weiß, C. H., Möller, T. A. and Gouveia, S. (2018). The max-INAR (1) model for count processes. *TEST*, **27**: 850– 870.
- Sue, N. (2009). Financial

- Planning Using Excel: Forecasting, Planning and Budgeting Techniques (2nd Ed.). Elsevier. <https://www.Sciencedirect.com>
- Tran, M. and Uyen, N. (2016). Forecasting foreign exchange rate by using ARIMA model: A case of VND/USD exchange rate. *Research Journal of Finance and Accounting*, **7**(12): 38 - 44.
- Tsay, R. (2016). Time Series Analysis for Forecasting and Model Building. Graduate School of Business, University of Chicago. BUS 41910, 1-13.
- Vipul, M. (2017). Forecasting USD to INR foreign exchange rate using time series analysis techniques like Holt-Winters simple exponential smoothing, ARIMA and Neural Networks. doi: 10.13140/RG.2.2.23465.08800.
- Zhu, F., Shi, L. and Liu, S. (2015). Influence diagnostics in log-linear integer-valued GARCH models. *Adv. Statist. Anal.*, **99**: 311-335.