African Multidisciplinary
Journal of Sciences and Artificial Intelligence

ISSN : 3434-9871

Index: Harvard, Boston, Sydney University, Dimensions, Lens, ResearchGet Scilit, Semantic, Google Scholar, Base etc

https://doi.org/10.58578/AMJSAI.v1i2.4167

Effect of Box-Cox Transformation on a k-th Exponential Weighted Moving Average Processes for Time Series

Kenneth Kevin Uadiale & Emwinloghosa Kenneth Guobadia
Federal University of Wukari, Nigeria; Federal Medical Centre, Asaba, Nigeria
kengroove28@gmail.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Oct 20, 2024	Nov 5, 2024	Nov 17, 2024	Nov 22, 2024

Abstract

In the actual world, many time series are not stationary. The purpose of this research is to use the Box and Cox family of transformations to convert a nonstationary time series to a stationary time series in order to determine the influence of a transformation on the data. This is accomplished by setting particular values for the transformation parameter. The sample autocorrelation function (SACF) and the sample partial autocorrelation function (SPACF) were used to test for stationarity of the Box and Cox parameters. The ARIMA model is fitted to the transformed data using the techniques of Box-Jenkins, where the best ARIMA is selected among the competing ARIMA models using Akaike information corrected criterian (AICc) while the best k-th EWMA is selected among the competing models using some evaluation metrics such as root mean square error (RMSE) and mean absolute error (MAE). Finally, the optimal model is selected between ARIMA model and k-th EWMA using the RMSE and MAE. Our findings are that the transformed k-th EWMA models outperformed the classical ARIMA on the set of given data.

Keywords: Box-Cox transformation, k-th EWMA model, ARIMA model, SACF and SPACE

INTRODUCTION

In statistics, time series analysis is a crucial field that has been used to address numerous practical economic time series issues. The class of weighted methods (weighted moving averages, exponential smoothing methods, and simple moving averages) is a crucial component of time series modeling. These methods are commonly employed in forecasting, and their primary objective is to smooth out the time series' random continuous changes, Emwinloghosa Kenneth Guobadia & Kenneth Kevin Uadiale, (2024). Exponential smoothing has been quite helpful for many years. It was initially proposed by Holt (1957) to be applied to time series that are not seasonal and do not show any trend. This was then expanded upon by Holt (1958), who provided a method that addressed trend. Winters (1965) expanded the approach to account for seasonality. According to Box and Jenkins (1976), economic forecasting techniques for nonstationary time series which lack a natural mean invoke the class of exponentially weighted moving averages. In order to make nonstationary time series to become stationary time series, they advise using differencing. Alternatively, a class of Box-Cox transformations known as Box-Cox (1964) could be used. The impact of Box-Cox transformation on estimates of the genetic parameters for egg production qualities that do not adhere to parametric statistical analysis assumptions was examined by Unver et al. (2004). Shish and Tsokos (2008) and Tsokos (2010) presented the k-th weighted moving average and k-th exponential weighted moving average processes as weighted approaches for forecasting nonstationary time series, respectively. Safi and Dawoud (2013) considered this class of weighted algorithms. These weighted strategies can be used to turn the initial nonstationary series into a stationary series using a differencing filter, which can also be modeled using Box and Jenkins's (1976) ARIMA procedure. Safi and Dawoud (2013), Tsokos (2010), and Shish and Tsokos (2008) did not consider the usage of a differencing filter in conjunction with the Box-Cox transformation.

MATERIALS AND METHODS

The Nigeria Stock Group provided the data utilized in this study, which shows the daily value of the Nigeria Stock Market from March 13, 2017, until April 11, 2022.

The statistical methods employed for data analysis in this study are the ARIMA model introduced by Box Jenkins (1976) and the k-th EWMA model introduced by Shish and

Tsoko (2008). Initially, the Box-Cox transformation and differencing are used to convert the nonstationary data to stationary data. Additionally, Vijay et al. (2019) have taken this idea into consideration. According to Franses and De Bruin (2002), several of these Box-Cox factors are crucial for predicting. The most significant ones are $\lambda = 0$, $\lambda = 0.5$ and $\lambda = 1$. Therefore, the Box-Jenkins approaches for the ARIMA model and the Tsoko (2010) techniques for the k-th EWMA model are used to fit the models to the modified data. Using AICc, the least ARIMA model is chosen among the others as the best ARIMA model. Similarly, RMSE and MAE are used to choose the best k-th EWMA from the competing models. Using RMSE and MAE, the best model between the ARIMA and k-th EWMA models is finally chosen.

BOX-COX Family of Transformation

By applying the subsequent Box-Cox transformation, defined by

$$y_t = \begin{cases} log(x_t) - log(x_{t-1}), & (\lambda = 0) \\ x_t^{\lambda} - x_{t-1}^{\lambda}, & (\lambda \neq 0) \end{cases} \tag{1}$$

To get the original series back, we employed the following procedure

$$x_t = \begin{cases} x_{t-1} \exp(z_t), & \lambda = 0 \\ \left(y_t + x_{t-1}^{\lambda}\right)^{\frac{1}{\lambda}}, & \lambda \neq 0 \end{cases} \tag{2}$$

Autoregressive Integrated Moving Average model

Wei et al. (2006), constructed the well-known model of ARIMA (p.d.q) autoregressive integrated moving average as:

$$\phi_p(B) - (1 - B)^d x_t = x_t \theta_q(B) \varepsilon_t \tag{3}$$

Where d is the degree of differencing and $B^j x_t = x_{t-j}$

The Best Fit Procedure Criteria for Forecast Selection

After obtaining suitable models either directly or by observing the autocorrelation function (ACF) plot, which shows that the ARIMA model to be fitted to the series should have the smallest possible parameters, meaning that p and q should be less than or equal to (p, $q \leq$ 3), the next step is to choose the best model among them using information criteria like the Akaike information criterion corrected (AICc). The best model among the competing models is determined to be the ARIMA model with the lowest AICc. We obtain the AICc and AIC as

$$AICc = AIC + \frac{2k^2 + 2k}{n - k - 1}$$

where k is the number of parameters and n is the sample size.

The k-th Exponentially Weighted Moving Average Procedure.

For the k-th Exponentially weighted moving average based on the transformed time series we have

$$Z_t = \frac{1}{\sum_{j=0}^{n-1}(1-\alpha)^j}\sum_{j=0}^{k-1}(1-\alpha)^{k-j-1}x_{t-k+1+j} \quad (4)$$

where $z_t = f(x_t^{\lambda})$ defined in equation (1) and $t = k, k+1, \dots n$, which defined the smoothing factor of $\alpha = \frac{2}{k+1}$ regarding the basis of $\alpha = \frac{2}{k+1}$, since the smoothing factor is $0 < \alpha < 1$, $\alpha = \frac{2}{k+1}$ which implies that $0 < \alpha < 1$, and if the value of α is substituted, we then have that $0 < \frac{2}{k+1} < 1$, and in this case, it implies that $0 < 2 < 1 + k$ for $k > 1$ >1.

In equation (3) if $k = 1$, we have the original data as follow:

i $$\sum_{j=0}^{n-1}(1-\alpha)^j$$

reaches its maximum when $k = 3$, and as k increases, it gets closer and closer to 1

ii. As $k \rightarrow$n, the series $\{z_t\}$ becomes

$$z_t = \frac{1}{\sum_{j=0}^{n-1}(1-\alpha)^j}\sum_{j=0}^{n-1}(1-\alpha)^{n-j-1}x_{j+1}$$

iii. The EWMA weighs heavily on the most recent observation and decreases the weight exponentially as time increases.

We use the following example for $k = 3$ and $k = 5$ to demonstrate the times series smoothing method.

For $k = 3$

$$\sum_{j=0}^{3-1}(1-\alpha)^j = \sum_{j=0}^{2}\left(1 - {}^{1}/_{2}\right)^j$$

$$= \left({}^{1}/_{2}\right)^0 + \left({}^{1}/_{2}\right)^1 + \left({}^{1}/_{2}\right)^2 \quad (4)$$

$$Z_t = \frac{1}{{}^{7}/_{4}}\sum_{j=0}^{2}\left({}^{1}/_{2}\right)^{2-j} x_{i+j-2}$$

$$= \frac{4}{7}\left(\left({}^{1}/_{2}\right)^2 x_{t-2} + \left({}^{1}/_{2}\right)^1 x_{t-1} + \left({}^{1}/_{2}\right)^0 x_t\right) \quad (5)$$

$$Z_t = \frac{4}{7}\left(\frac{1}{4}x_{t-2} + {}^{1}/_{2}\, x_{t-1} + x_t\right) \quad (6)$$

In order to recover y_t , we make use of the following equation (2)

$$y_t = 7z_t - x_{t-2} - 2x_{t-1} \quad (7)$$

Where $x_t = f(z_t, x_t)$ defined in equation (1)

And if $k = 5$

$$Z_t = \frac{1}{\sum_{j=0}^{k-1}(1-\alpha)^j}\sum_{j=0}^{k-1}(1-\alpha)^{k-j-1} x_{t-k+1+j}$$

$$Z_t = \frac{1}{\sum_{j=0}^{5-1}(1-1/3)^j}\sum_{0}^{4}(1 - 1/3)^{5-0-1}x_{t-5+1+j} \quad (8)$$

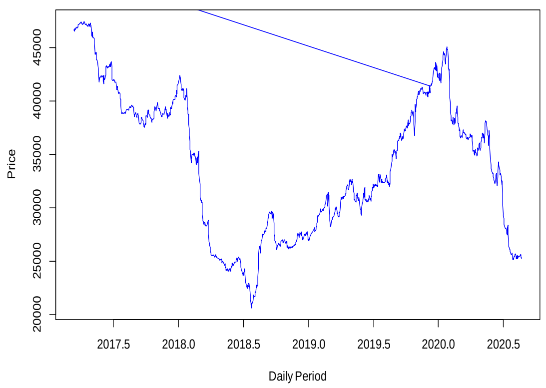
$$Z_t = \frac{81}{211}\left(\frac{32}{81}x_{t-4} + \frac{8}{27}x_{t-3} + \frac{4}{9}x_{t-2} + \frac{2}{3}x_{-1} + x_t\right) \quad (9)$$

Similarly, to recover 3.3, we have the following equation

$$y_t = 211z_t - 32x_{t-4} - 24x_{t-3} - 36x_{t-2} - 54x_{t-1} \quad (10)$$

Test for Stationarity

Visualizing a time plot or autocorrelation function (ACF) plot can indicate stationarity, while an Augmented Dickey-Fuller (ADF) test for unit root can confirm it. The ADF test does not directly assess stationarity, but rather the presence or lack of a unit root. If a series is not stationary, it may need to be transformed using differencing or power transformation.



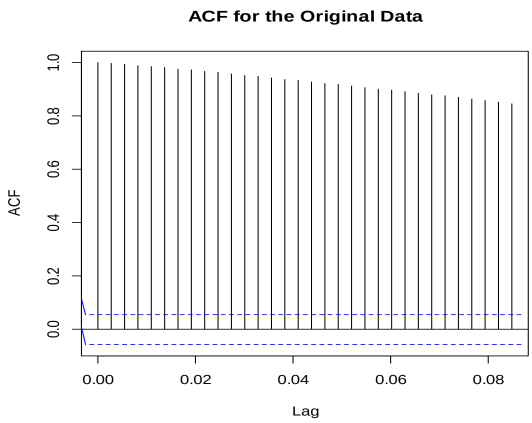


Figure A: The actual data time plot **Figure B: ACF plot for the original data**

The actual data time plot from 13 March 2017 to 11 April 2022

Figure A, it appears that the time series exhibits non-stationarity, as the mean and variance are not constant over time. The presence of an increasing and decreasing trend suggests that the series is not stationary in the mean, while the varying volatility indicates non-stationarity in the variance, it therefore require differencing. A sample test of the sample autocorrelation function (SACF) test is conducted for the time series data and the SACF is presented in figure B, shown slow decay as an evident that the time series data is non stationary since the SACF refuse to diminish quickly

RESULTS AND DISCUSSION

Table A to C provides the summary statistics for selecting the ARIMA model utilized in this study. When $\lambda = 0$, ARIMA (0,1,3) has the least Akaike Information Criterion Corrected (AICc = -12337.60) as a result, it was chosen as the best model among the competing models for table A, When $\lambda = 0.5$, ARIMA (0,1,3) has the least Akaike Information Criterion Corrected (AICc = 1082.44) as a result, it was chosen as the best model among the competing models for table B and When $\lambda = 1$, ARIMA (0,1,3) has the least Akaike Information Criterion Corrected (AICc = 15932.09) as a result, it was chosen as the best model among the competing models for table C.

Table A, B and C: ARIMA Model Selection Table

Table A: ARIMA Model selection when $\lambda = 0$		**Table B: ARIMA Model selection when $\lambda = 0.5$**		**Table C: ARIMA Model selection when $\lambda = 1$**	
ARIMA Model	**AICc**	**ARIMA Model**	**AICc**	**ARIMA Model**	**AICc**
ARIMA (0,1,1)	-12058.82	ARIMA (0,1,1)	1369.35	ARIMA (0,1,1)	16225.52
ARIMA (0,1,2)	-12090.34	ARIMA (0,1,2)	1327.94	ARIMA (0,1,2)	16175.42
ARIMA (0,1,3)	-12337.60	ARIMA (0,1,3)	1082.44	ARIMA (0,1,3)	15932.09
ARIMA (1,1,0)	-12042.12	ARIMA (1,1,0)	1385.98	ARIMA (1,1,0)	16241.48
ARIMA (2,1,0)	-12125.05	ARIMA (2,1,0)	1296.59	ARIMA (2,1,0)	16147.86
ARIMA (3,1,0)	-12233.73	ARIMA (3,1,0)	1191.40	ARIMA (3,1,0)	16044.97
ARIMA (1,1,1)	-12059.47	ARIMA (1,1,1)	1367.86	ARIMA (1,1,1)	16223.28
ARIMA (1,1,2)	-12292.79	ARIMA (1,1,2)	1129.64	ARIMA (1,1,2)	15981.35
ARIMA (2,1,1)	-12316.49	ARIMA (2,1,1)	1109.13	ARIMA (2,1,1)	15963.66

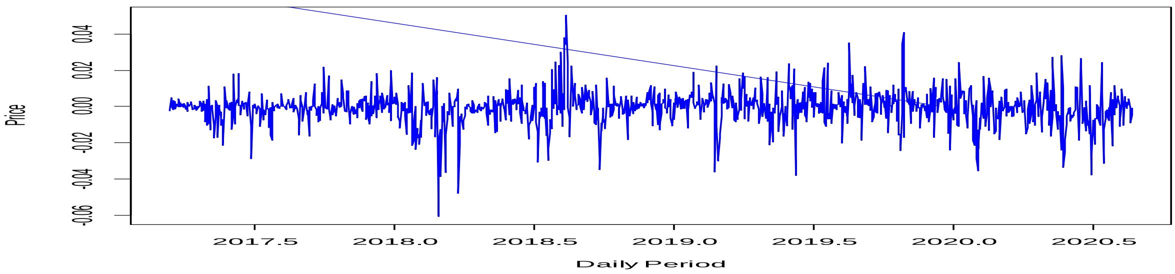


Figure C: Time series plot for first differenced data, when $\lambda = 0$

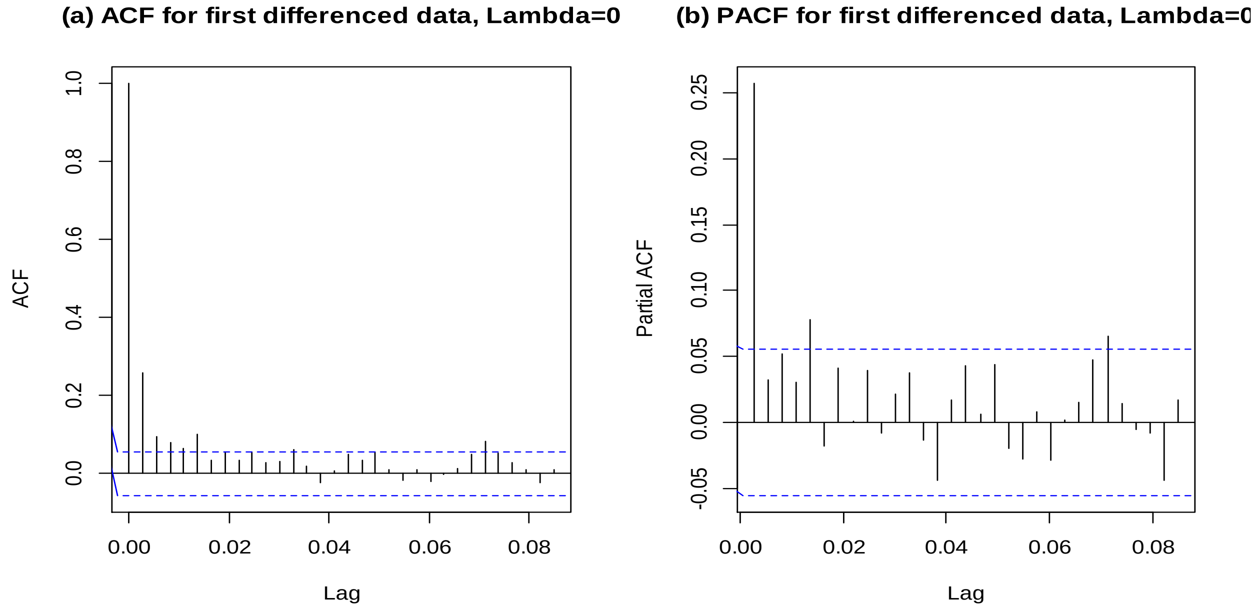


Figure C (i): (a) Shows the ACF plot for the first differenced data, $\lambda = 0$ and (b) shows the PACF plot for the first differenced data, $\lambda = 0$

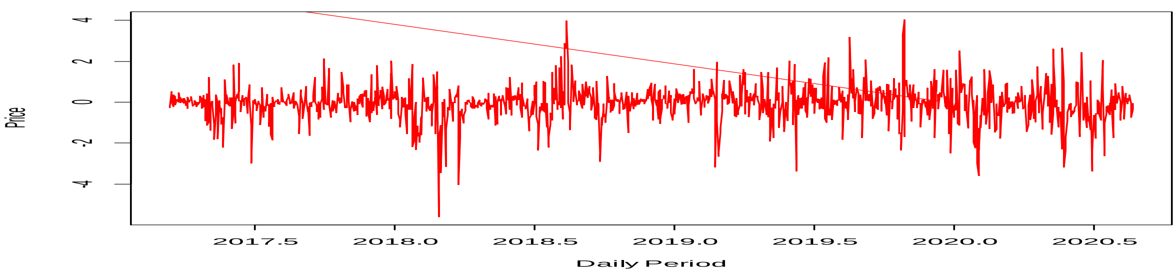


Figure D: Time series plot for first differenced data, when $\lambda = 0.5$

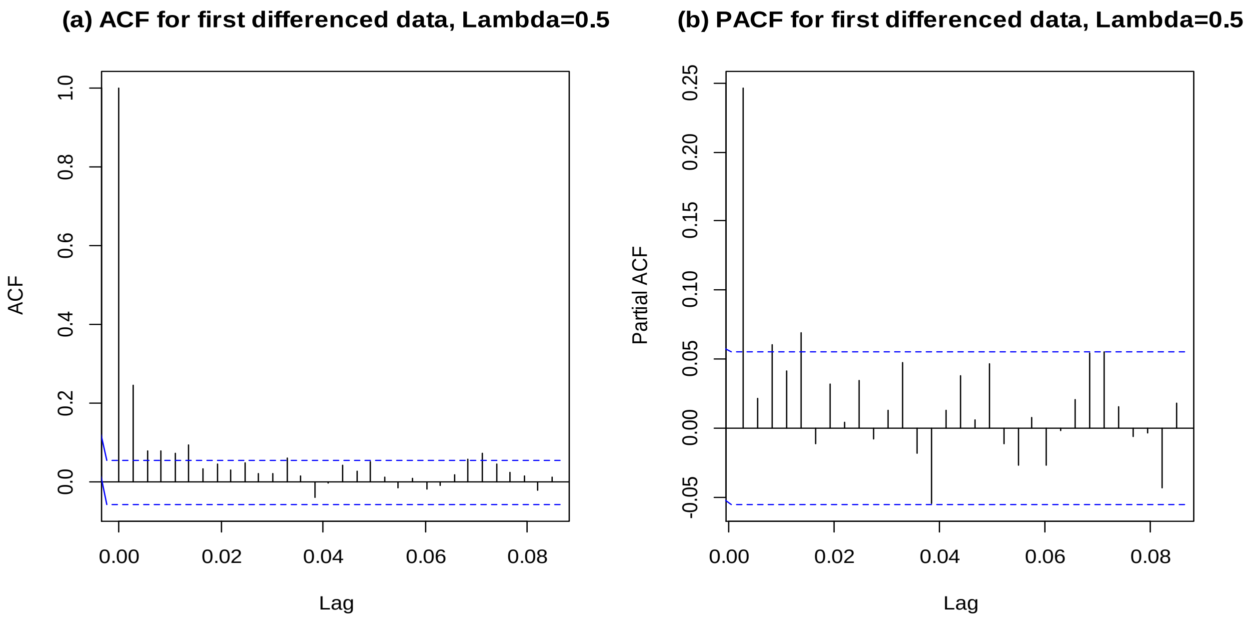


Figure D (i): (a) Shows the ACF plot for the first differenced data, $\lambda = 0.5$ and (b) shows the PACF plot for the first differenced data, $\lambda = 0.5$

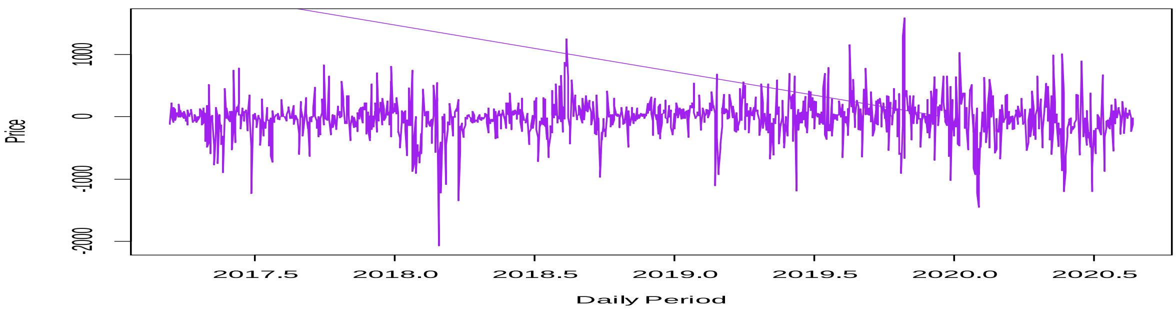


Figure F: Time series plot for first differenced data, when $\lambda = 1$

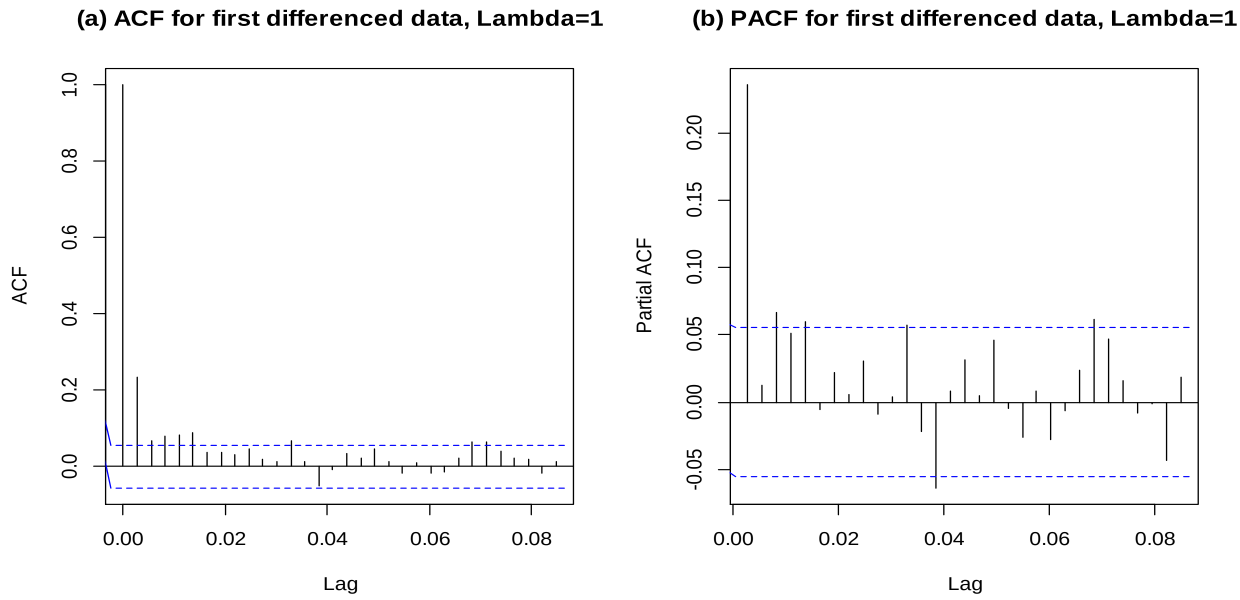


Figure F (i): (a) Shows the ACF plot for the first differenced data, $\lambda = 1$ and (b) shows the PACF plot for the first differenced data, $\lambda = 1$

Note that:

(i) **Figure C to F:** The pictorial display of the time plots for first order differencing data, showed stationary after applying the Box-Cox transformation when the value of the Box-Cox parameter $\lambda = 0$, 0.5 and 1.

(ii) **Figure C to F:** The ACF and PACF pictorial display of the time plots for first order differencing data with Box-Cox parameter $\lambda = 0$, 0.5 and 1 showed stationarity.

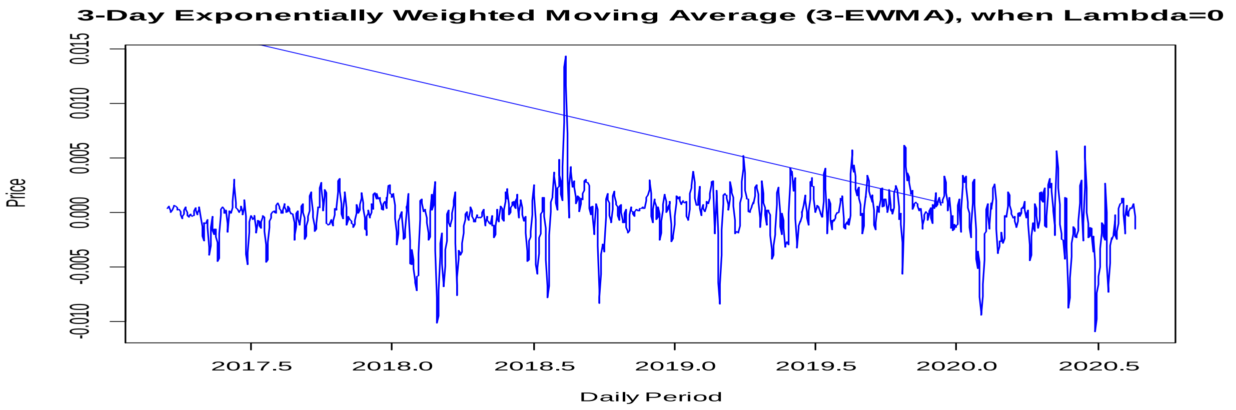


Figure G: Time plot for the Box-Cox transformation of 3- EWMA, when $\lambda = 0$

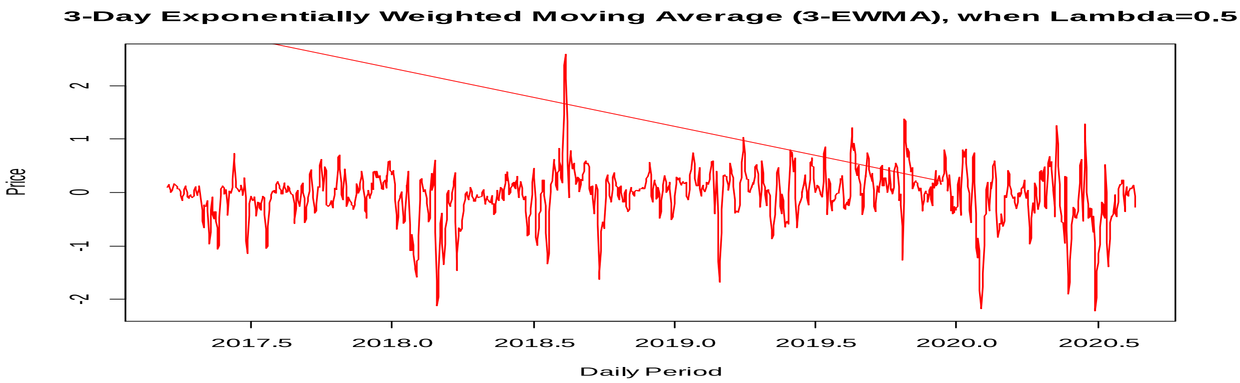


Figure H: Time plot for the Box-Cox transformation of 3- EWMA, when $\lambda = 0.5$

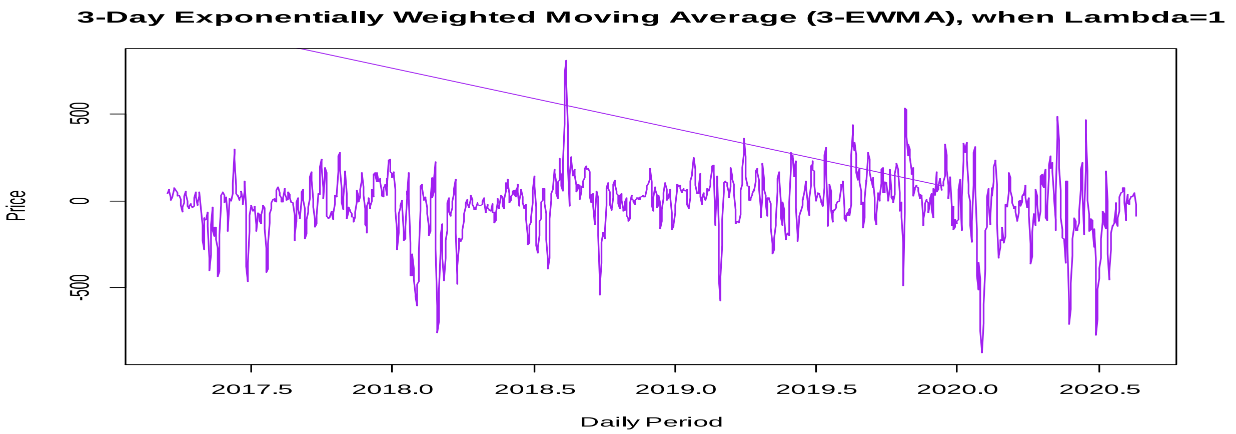


Figure I: Time plot for the Box-Cox transformation of 3- EWMA, when $\lambda = 1$

Tables 2 to 5 show the summary statistics for selecting the best model between the ARIMA model and the k-th EWMA model utilized in this study.

In table 2, shows that the 3-Day exponentially weighted moving average model with Box-Cox transformed data and $\lambda = 0$, has the lowest metrics (RMSE and MAE) compared to other models. However, the 3-EWMA model was chosen as the best forecasting model.

In table 3, shows that the 3-Day exponentially weighted moving average model with Box-Cox transformed data and $\lambda = 0.5$, has the lowest metrics (RMSE and MAE) compared to other models. However, the 3-EWMA model was chosen as the best forecasting model.

In table 4, shows that the 3-Day exponentially weighted moving average model with Box-Cox transformed data and $\lambda = 1$, has the lowest metrics (RMSE and MAE) compared to other models. However, the 3-EWMA model was chosen as the best forecasting model.

In table 5, shows that 3-EWMA with $\lambda = 0$, has the least metrics (RMSE and MAE). However, it is regarded as the finest linear forecast approach for predicting the daily Nigerian stock price.

In summary, the 3-EWMA model with $\lambda = 0$ performs well in predicting daily Nigerian stock prices, with the lowest RMSE and MAE values. Its categorization as the most accurate linear forecast approach demonstrates its effectiveness while also emphasizing the significance of considering aspects other than numerical accuracy when choosing forecasting models in financial situations.

Table 2: The comparison of the classical ARIMA and 3-Day exponentially weighted Moving Average, when $\lambda = 0$

Model	RMSE	MAE
Classical ARIMA	0.00896	0.00588
3-EWMA	0.00087	0.00058

Table 3: The comparison of the classical ARIMA and 3-Day exponentially weighted Moving Average $\lambda = 0.5$

Model	RMSE	MAE
Classical ARIMA	0.81411	0.53578
3-EWMA	0.18326	0.12081

Table 4: The comparison of the classical ARIMA and 3-Day exponentially weighted Moving Average $\lambda = 1$

Model	RMSE	MAE
Classical ARIMA	302.2399	198.3729
3-EWMA	67.76841	44.4749

Table 5: Comparing 3-EWMA with $\lambda = 0$, $\lambda = 0.5$, and $\lambda = 1$

Model	RMSE	MAE
3-EWMA when $\lambda = 0$	0.00087	0.00058
3-EWMA when $\lambda = 0.5$	0.18326	0.12081
3-EWMA when $\lambda = 1$	67.76841	44.4749

CONCLUSION

The purpose of this study is to compare a specific k-th EWMA with the classical ARIMA processes for modeling non-stationary time series data utilizing a class of Box-Cox Transformation. Its categorization as the most accurate linear forecast approach demonstrates its effectiveness while also emphasizing the significance of considering aspects other than numerical accuracy when choosing forecasting models in financial situations.

REFERENCES

Ammeh L.B., Guobadia E.K., Ugoh C.I. (2023), Exponential Smoothing State Space Innovation Model for Forecasting Export Commodity Price Index in Nigeria. African Journal of Economics and Sustainable Development 6(4), 74-84.DOI: 10.52589/AJESD-XQFAMZMY

Box G. E. P and Jenkins, G.M., (1976). "Time series analysis: „Forecasting and control," Holden-Day, San Francisco

Chatfield, C., Koehler, A.B., Keith Ord., J and Snyder, R.D. (2001). A New Look at Models for Exponential Smoothing. Journal of the Royal Statistical Society. Series D, Vol.50, No.2, pp.147-159

Christogonus Ifeanyichukwu Ugoh, Udochukwu Victor Echebiri, Gabriel Olawale Temisan, John Paul, Kenechukwu Iwuchukwu, Emwinloghosa Kenneth Guobadia (2022) On Forecasting Nigeria's GDP: A Comparative Performance of Regression with ARIMA Errors and ARIMA Method, International Journal of Mathematics and Statistics Studies, Vol.10, No.4, pp.48-64

Ekhosuehi, N. (2013). Effect of Box-Cox Transformation on k-th Moving Average processes for, Time series Forecasting Models. Journal of the Nigerian Statistical Association, Vol. 25 pp. 1-11.

Ekhosuehi N, Kenneth GE, Kevin UK. (2020). TheWeibull Length Biased Exponential Distribution: Statistical Properties and Applications. Journal of Statistical and Econometric Methods.9(4):15-30.

Emwinloghosa K.G., Pamela O.O., Paschal N.I., Eloho S.O., Agu C. (2023), Modeling and Forecasting Inflation in Nigeria: A Time Series Regression with ARIMA Method. African Journal of Economics and Sustainable Development 6(3), 42-53. DOI:10.52589/AJESD-HFYC2BNW

Emwinloghosa K.G., Pamela O.O., Christogonus I.U. (2023), Impact of Fiscal Policy on Inflation in Nigerian Economy. African Journal of Economics and Sustainable Development 6(4), 37-8.DOI: 10.52589/AJESD-E1RMOYKH

Guobadia Emwinloghosa Kenneth. (2021). Statistical Application of Regression techniques in Modeling Road Accidents in Edo State, Nigeria. Sch J Phys Math Stat, Jan 8(1): 14-18

Guobadia, E. K., & Uadiale, K. K. (2024). Effect of Box-Cox Transformation on a k-th Moving Average Processes for Time Series. African Multidisciplinary Journal of Sciences and Artificial Intelligence, 1(1), 655-668. https://doi.org/10.58578/amjsai.v1i1.3755

Guobadia Emwinloghosa Kenneth. Justification of the Nature of Fluctuations in Nigerian Bank Returns: An Empirical Analysis. Sch J Econ Bus Manag, 2021 Jan 8(1): 10-13.

Guobadia Emwinloghosa Kenneth & Ibeakuzie Precious Onyedikachi. (2021). Selected Economic Sector Contributionto Nigeria's Gross Domestic Product. Sch Bull, 7(3): 49-59.

Guobadia Emwinloghosa Kenneth & Ibeakuzi Precious Onyedikachi. (2021). Short Term Modeling of the NigerianNaira/United States Dollar Exchange Rate Using ARIMA Model. Sch J Phys Math Stat, Jan 8(1): 8-13.

Guobadia Emwinloghosa Kenneth. (2021). A Time Domain Approach to Modeling Nigeria's Gross Domestic Product. SchJ Phys Math Stat, Jan 8(1): 19-28.

Holt, C.C. (2004) Forecasting Seasonal and Trends by Exponentially Weighted Moving Averages. Int. J. Forecast. 2004, 20, 5–10.

Holt, C.C. (1957) Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages. ONR Memorandum, Vol. 52, Carnegie Institute of Technology, Pittsburgh. Available from the Engineering Library, University of Texas, Austin.

Ibeakuzie Precious Onyedikachi & Guobadia Emwinloghosa Kenneth (2021). Research Ethics Grasp and Enactment; A Casewith University of Benin. Sch Bull, 7(3): 60-71

Shih, S. and Tsokos, C.P. (2008). A weighted Moving Average process for Forecasting. Journal of Modern Applied Statistical Methods.

Tsokos, C.P. (2010), K-th Moving, Weighted and Exponential Moving Average for Time Series Forecasting Models, European Journal of pure and applied Mathematics. Vol.3, No.3 406-416

Winters, P.R. (1960), Forecasting Sales by Exponentially Weighted Moving Averages. Manag. Sci.,6, 324–342.

Ugoh C.I., Abode J.O., OnyiaC.T., Omoruyi P.O., Guobadia E.K. (2023), Evaluating the Determinants of Exchange Rates in Emerging Markets: Evidence from Nigeria and South Africa. African Journal of Economics and Sustainable Development 6(2), 49-63.DOI: 10.52589/AJESD-VB4NTHBE

Ugoh C.I., Igbinosa E.S. ,Akanno F.C., Omoruyi P.O.,Guobadia E.K. (2023),Investigating the Determinants of Inflation in Leading Economies in Africa: A Panel Data Analysis. African Journal of Economics and Sustainable Development6(2), 30-48. DOI:10.52589/AJESD-VGH59XLX

Unver, Y. Akbas, Y. and Oguz, I. 2004. Effect of Box-Cox transformation on genetic parameter estimation in layers. Turk Journal of Veterinary Science, 28, 249-255

Safi, S.K and Dawoud, I.A. (2013). Comparative study on forecasting accuracy among moving average models with simulation and PALTEL stock market data in Palestine. International