SOME SMOOTHING TOOLS FOR CONFIRMED CORONAVIRUS CASES IN NIGERIA

Supported by tetfund

J. N. Onyeka-Ubaka[1]*, R. K. Ogundeji[1] and O. Abass[2]
[1]Department of Mathematics, University of Lagos, Nigeria
[2]Department of Mathematics and Computer, Bells University, Ota, Nigeria
*Corresponding author: jonyeka-ubaka@unilag.edu.ng, rogundeji@unilag.edu.ng

Received: September 23, 2020 **Accepted:** November 13, 2020

Abstract: This paper evaluated and compared the performance of a family of smoothing models such as autoregressive integrated moving average (ARIMA) models andHolt's exponential smoothing methods: Additive and multiplicative; to forecast the daily confirmed coronavirus (COVID-19) cases in Nigeria for the sampled period. The predictive capabilities were compared in terms of forecast accuracy measures, Akaike information criterion and Schwarz Bayesian information criterion based on the validated data set. The Holt's linear exponential smoothing model with parameter α and β was found to have best described the data having the lowest ranked error statistics in an out of sample performance among other exponential smoothing models while the autoregressive integrated moving average (ARIMA (0, 1, 1)) with the smallest AIC was selected as the best. The forecast values of the two selected models show that the COVID-19 pandemic will live with us for a long term. The forecast results imply that the government and citizenry will and must adhere to preventive measures while going about their normal businesses.

Keywords: COVID-19, ARIMA models, Holt's linear exponential smoothing model, AIC

Introduction

A virus-caused infectious disease called Coronavirus (COVID-19) was declared by the World Health Organization (WHO) on January 30, 2020 as a global pandemic that require public health emergency of international scope. Coronavirus was first recorded in Nigeria on February 29, 2020. As of July 6, 2020, COVID-19 has spread to over 250 countries and territories causing over 11.4 million infections with 534,780 deaths (Johns Hopkins University, 2020). Nigeria has recorded a total of 28,711 confirmed cases of COVID-19 with 645 deaths and 11,665 discharged spreading across 35 states and the FCT (NCDC, 2020). In every epidemic, some individuals become sick, and some may die, whereas others may recover from illness and still others show no signs or symptoms of disease, while nonetheless carrying it and being potential sources of infection. The COVID-19 virus does not move except if it is moved by human beings. This leads the WHO, in conjunction with Nigeria Centre for Disease Control and Ministry of Health to come with the approaches to stop or minimize the spread of COVID-19. Some of the approaches are wash your hands with clean water soap for 20 seconds; use hand sanitizers where water are not available; wear face masks on public places; maintain social and physical distancing of about 1 – 2 metres; sneeze or cough into tissue papers and dispose accordingly and stay at home. Despite the regular orientation on preventive measures, the lockdown of international and inter-state movements, no gathering of more than 20 people, closure of schools and worship centres, viewing centres and markets, the surge of COVID-19 infection kept increasing in exponential manner. This necessitates for adequate planning.

Planning for future events is an integral aspect of operating any business. Planning allows actions to be taken that will meet lead time requirements and create a competitive operation. The process of planning, however, assumes that forecasts of the future are readily available. Government and health sector, in particular, must anticipate future surge for Coronavirus cases, demand for health products or services and plan to provide facilities and resources necessary to meet that demand. Forecasting is the first step in planning. It is one of the most important tasks, as many other organizational decisions are based on a forecast of the future. The quality of these decisions can only be as good as the quality of the forecast upon which they are based. This paper, therefore, evaluated and compared the performance of a family of smoothing models such as simple exponential smoothing, Holt's linear/exponential trend and Holt's damped methods: additive and multiplicative; and family of autoregressive integrated moving average (ARIMA) models to forecast the daily Coronavirus cases in Nigeria.

The Model

Exponential smoothing forecasting methods use constants that assign weights to current demand and previous forecasts to arrive at new forecasts. Their values influence the responsiveness of forecasts to actual COVID-19 cases and hence influence forecast error. In Simple Exponential Smoothing, observed case (X_t) is level with only random variations around some average. The forecast X_{t+m} for the upcoming period is the estimate of average level S_t at the end of period *t*.

$$\mathrm{X}_{t+m} = S_t + \alpha(\mathrm{X}_t - S_t) = \alpha \mathrm{X}_t + (1-\alpha)S_{t-1} \quad (1)$$

Where: α, the smoothing constant, is between 0 and 1. The new estimate of level may be seen as a weighted average of X_t, the most recent information of average level, and S_t the previous estimate of that level. Small values of α imply that the revision of the old forecast, in light of the new COVID-19 case is small; the new forecast is not very different from the previous one. The method requires an initial forecast S_t which has to be either assumed or estimated.

In Exponential Smoothing with Trend Adjustment (Double Exponential Smoothing), the time series exhibits a trend; in addition to the level component, the trend (slope) has to be estimated. The forecast, including trend for the upcoming period $t+1$, is given by;

$$\mathrm{X}_{t+1} = S_t + \mathrm{T}_t \quad (2)$$

Here, S_t is the estimate of level made at the end of period *t* and is given by;

$$S_t = \alpha \mathrm{X}_t + (1-\alpha)S_t \quad (3)$$

T_t is the estimate of trend at the end of period *t* and is given by;

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \qquad (4)$$

Where: β is also a smoothing constant between 0 and 1 and plays a role similar to that of α. Again, small values of α and β imply that consecutive estimates of level and trend components do not differ much from each other. Any revision in the light of the new case is small. This method requires estimation of the initial level component S_1 and the initial trend component T_1 to start off the series of forecasts. Smoothing constants are key to successful forecasting with exponential smoothing, but there are no consistent guidelines in the forecasting literature on how they should be selected. There arevarious variety of notation existing in the literature but for continuity, the paper adapts Hyndman's *et al*. (2002) taxonomy, as extended by Taylor (2003) in describing the methods. Each method is denoted by one or two letters for the trend (row heading) and one letter for seasonality (column heading). Method N-N denotes no trend with no seasonality, or simple exponential smoothing (Brown, 1959). The other non-seasonal methods are Holt's (1957) additive trend (A-N), Gardner and McKenzie's (1985) damped additive trend (DA-N), Pegels' (1969) multiplicative trend (M-N), and Taylor's (2003) damped multiplicative trend (DM-N). The parameters in the trend methods can be constrained using discounted least squares (DLS) to produce special cases often called Brown's methods (Gardner, 2005). All seasonal methods are formulated by extending the methods in Winters (1960). Note that the forecast equations for the seasonal methods are valid only for a forecast horizon (*m*) less than or equal to the length of the seasonal cycle (*p*).

The standard exponential smoothing as captured by Gardner (2005) are presented in Table 1.

When a combined observation occurs, in the N-N equation we replace X_t with X_t / k, where *k* is the number of periods combined, and we replace α with the expression $1-(1-\alpha)^k$. This adjustment assumes that the data are spread evenly over the combined periods. Each exponential smoothing method in Table 1 is equivalent to one or more stochastic models. The possibilities include regression, ARIMA, and state-space models.All linear exponential smoothing methods have equivalent ARIMA models. The easiest way to see the non-seasonal models is through the DA-N method, which contains at least six ARIMA models as special cases (Gardner and McKenzie, 1988). If $0 < \phi < 1$, the DA-N method is equivalent to the ARIMA (1, 1, 2) model, which can be written as:

$$(1-B)(1-\phi B)X_t = [1-(1+\phi-\alpha-\phi\alpha\gamma)B - \phi(\alpha-1)B^2]e_t \qquad (5)$$

We obtain an ARIMA (1, 1, 1) model by setting $\alpha = 1$. With $\alpha = \gamma = 1$, the model is ARIMA (1, 1, 0). When $\phi = 1$, we have a linear trend (A-N) and the model is ARIMA (0, 2, 2):

$$(1-B)^2 X_t = [1-(2-\alpha-\alpha\gamma)B - (\alpha-1)B^2]e_t \qquad (6)$$

When $\phi = 0$, we have simple smoothing (N-N) and the equivalent ARIMA (0, 1, 1) model:

$$(1-B)X_t = [1-(1-\alpha)]e_t \qquad (7)$$

The ARIMA (0, 1, 0) random walk model can be obtained from (7) by choosing $\alpha = 1$. ARIMA-equivalent seasonal models for the linear exponential smoothing methods exist, although most are so complex that it is unlikely they would ever be identified through the Box-Jenkins methodology. That is, each of the linear exponential smoothing models with additive errors has an ARIMA equivalent. However, the linear models with multiplicative errors and the nonlinear models are beyond the scope of the ARIMA class.

For the A-A method, an analytical variance expression was derived by Yar and Chatfield (1990), who assumed only that one-step-ahead errors are uncorrelated. But for this to be true, the equivalent ARIMA model must be optimal. The first class includes linear models with additive errors and ARIMA equivalents, corresponding to the N-N, A-N, DA-N, N-A, A-A, and DA-A methods. Simple smoothing (N-N) is certainly the most robust forecasting method and has performed well in many types of series not generated by the equivalent ARIMA (0, 1, 1) process. Such series include the very common first-order autoregressive processes and a number of lower-order ARIMA processes (Cogger, 1973; Tiao and Xu, 1993). Bossons (1966) showed that simple smoothing is generally insensitive to specification error, especially when the miss-specification arises from an incorrect belief in the stationarity of the generating process. Related work by Hyndman (2001) shows that ARIMAmodel selection errors can inflate MSEs compared to simple smoothing. Hyndman simulated time series from an ARIMA (0, 1, 1) process and fitted a restricted set of ARIMA models of order (0, 1, 1), (1, 1, 0), and (1, 1, 1), each with and without a constant term. The best model was selected using Akaike's information criterion (AIC) (Akaike, 1970) and Bayesian information criterion (BIC) (Schwarz, 1978). The ARIMA forecast mean square errors (MSEs) were significantly larger than those of simple smoothing due to incorrect model selections, a problem that became worse when the errors were non-normal. Simple smoothing has done especially well in forecasting aggregated economic series with relatively low sampling frequencies. Rosanna and Seater (1995) show that such series not generated by an ARIMA (0, 1, 1) process often can be approximated by an ARIMA (0, 1, 1) process. This finding has been misinterpreted by some researchers in that the series were sums of averages over time of data generated more frequently than the reporting interval. The effects of averaging and temporal aggregation were to destroy information about the generating process, producing series for which the ARIMA (0, 1, 1) process was merely an artifact. Much the same problem can occur in company-level data.For example, simple exponential smoothing was a very competitive method in Schnaars' (1986) study of annual unit sales series for a variety of products. Satchell and Timmerman (1995) derive an explicit formula for weights when the time series has a finite history and give a different explanation for the performance of simple smoothing in economic time series. They found that exponentially declining weights are surprisingly robust as long as the ratio of the variance of the random walk process to the variance of the noise component is not exceptionally small.Simple smoothing was shown to be equivalent to a random walk with noise model, assuming that the process began an infinite number of periods ago (Muth, 1960).

Table 1: Standard exponential smoothing methods (Gardner, 2005)

	Seasonality		
Trend	**N** **None**	**A** **Additive**	**M** **Multiplicative**
N **None**	$S_t = \alpha X_t + (1-\alpha)S_{t-1}$ $\hat{X}_t(m) = S_t$	$S_t = \alpha(X_t - I_{t-p}) + (1-\alpha)S_{t-1}$ $I_t = \delta(X_t - S_t) + (1-\delta)I_{t-p}$ $\hat{X}_t(m) = S_t + I_{t-p+m}$	$S_t = \alpha(X_t \vert I_{t-p}) + (1-\alpha)S_{t-1}$ $I_t = \delta(X_t \vert S_t) + (1-\delta)I_{t-p}$ $\hat{X}_t(m) = S_t I_{t-p+m}$
	$S_t = S_{t-1} + \alpha e_t$ $\hat{X}_t(m) = S_t$	$S_t = S_{t-1} + \alpha e_t$ $I_t = I_{t-p} + \delta(1-\alpha)e_t$ $\hat{X}_t(m) = S_t + I_{t-p+m}$	$S_t = S_{t-1} + \alpha e_t \vert I_{t-p}$ $I_t = I_{t-p} + \delta(1-\alpha)e_t \vert S_t$ $\hat{X}_t(m) = S_t I_{t-p+m}$
A **Additive**	$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + T_{t-1})$ $T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1}$ $\hat{X}_t(m) = S_t + mT_t$	$S_t = \alpha(X_t - I_{t-p}) + (1-\alpha)(S_{t-1} + T_{t-1})$ $T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1}$ $I_t = \delta(X_t - S_t) + (1-\delta)I_{t-p}$ $\hat{X}_t(m) = S_t + mT_t + I_{t-p+m}$	$S_t = \alpha(X_t \vert I_{t-p}) + (1-\alpha)(S_{t-1} + T_{t-1})$ $T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1}$ $I_t = \delta(X_t \vert S_t) + (1-\delta)I_{t-p}$ $\hat{X}_t(m) = (S_t + mT_t)I_{t-p+m}$
	$S_t = S_{t-1} + T_{t-1} + \alpha e_t$ $T_t = T_{t-1} + \alpha\gamma e_t$ $\hat{X}_t(m) = S_t + mT_t$	$S_t = S_{t-1} + T_{t-1} + \alpha e_t$ $T_t = T_{t-1} + \alpha\gamma e_t$ $I_t = I_{t-p} + \delta(1-\alpha)e_t$ $\hat{X}_t(m) = S_t + mT_t + I_{t-p+m}$	$S_t = S_{t-1} + T_{t-1} + \alpha e_t \vert I_{t-p}$ $T_t = T_{t-1} + \alpha\gamma e_t \vert I_{t-p}$ $I_t = I_{t-p} + \delta(1-\alpha)e_t \vert S_t$ $\hat{X}_t(m) = (S_t + mT_t)I_{t-p+m}$
DA **Damped** **Additive**	$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + \phi T_{t-1})$ $T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\phi T_{t-1}$ $\hat{X}_t(m) = S_t + \sum_{i=1}^{m} \phi^i T_t$	$S_t = \alpha(X_t - I_{t-p}) + (1-\alpha)(S_{t-1} + \phi T_{t-1})$ $T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\phi T_{t-1}$ $I_t = \delta(X_t - S_t) + (1-\delta)I_{t-p}$ $\hat{X}_t(m) = S_t + \sum_{i=1}^{m} \phi^i T_t + I_{t-p+m}$	$S_t = \alpha(X_t \vert I_{t-p}) + (1-\alpha)(S_{t-1} + \phi T_{t-1})$ $T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\phi T_{t-1}$ $I_t = \delta(X_t \vert S_t) + (1-\delta)I_{t-p}$ $\hat{X}_t(m) = (S_t + \sum_{i=1}^{m} \phi^i T_t)I_{t-p+m}$
	$S_t = \alpha S_{t-1} + \phi T_{t-1} + \alpha e_t$ $T_t = \phi T_{t-1} + \alpha\gamma e_t$ $\hat{X}_t(m) = S_t + \sum_{i=1}^{m} \phi^i T_t$	$S_t = S_{t-1} + \phi T_{t-1} + \alpha e_t$ $T_t = \phi T_{t-1} + \alpha\gamma e_t$ $I_t = I_{t-p} + \delta(1-\alpha)e_t$ $\hat{X}_t(m) = S_t + \sum_{i=1}^{m} \phi^i T_t + I_{t-p+m}$	$S_t = S_{t-1} + \phi T_{t-1} + \alpha e_t \vert I_{t-p}$ $T_t = \phi T_{t-1} + \alpha\gamma e_t \vert I_{t-p}$ $I_t = I_{t-p} + \delta(1-\alpha)e_t \vert S_t$ $\hat{X}_t(m) = (S_t + \sum_{i=1}^{m} \phi^i T_t)I_{t-p+m}$
M **Multiplicative**	$S_t = \alpha X_t + (1-\alpha)(S_{t-1}R_{t-1})$ $R_t = \gamma(S_t \vert S_{t-1}) + (1-\gamma)R_{t-1}$ $\hat{X}_t(m) = S_t R_t^m$	$S_t = \alpha(X_t - I_{t-p}) + (1-\alpha)S_{t-1}R_{t-1}$ $R_t = \gamma(S_t \vert S_{t-1}) + (1-\gamma)R_{t-1}$ $I_t = \delta(X_t - S_t) + (1-\delta)I_{t-p}$ $\hat{X}_t(m) = S_t R_t^m + I_{t-p+m}$	$S_t = \alpha(X_t \vert I_{t-p}) + (1-\alpha)S_{t-1}R_{t-1}$ $R_t = \gamma(S_t \vert S_{t-1}) + (1-\gamma)R_{t-1}$ $I_t = \delta(X_t \vert S_t) + (1-\delta)I_{t-p}$ $\hat{X}_t(m) = (S_t R_t^m)I_{t-p+m}$
	$S_t = S_{t-1}R_{t-1} + \alpha e_t$ $R_t = R_{t-1} + \alpha\gamma e_t \vert S_{t-1}$ $\hat{X}_t(m) = S_t R_t^m$	$S_t = S_{t-1}R_{t-1} + \alpha e_t$ $R_t = R_{t-1} + \alpha\gamma e_t \vert S_{t-1}$ $I_t = I_{t-p} + \delta(1-\alpha)e_t$ $\hat{X}_t(m) = S_t R_t^m + I_{t-p+m}$	$S_t = S_{t-1}R_{t-1} + \alpha e_t \vert I_{t-p}$ $R_t = R_{t-1} + (\alpha\gamma e_t \vert S_{t-1})I_{t-p}$ $I_t = I_{t-p} + \delta(1-\alpha)e_t \vert S_t$ $\hat{X}_t(m) = (S_t R_t^m)I_{t-p+m}$
DM **Damped** **Multiplicative**	$S_t = \alpha X_t + (1-\alpha)(S_{t-1}R_{t-1}^{\phi})$ $R_t = \gamma(S_t \vert S_{t-1}) + (1-\gamma)R_{t-1}^{\phi}$ $\hat{X}_t(m) = S_t R_t^{\sum_{i=1}^{m} \phi^i}$	$S_t = \alpha(X_t - I_{t-p}) + (1-\alpha)S_{t-1}R_{t-1}^{\phi}$ $R_t = \gamma(S_t \vert S_{t-1}) + (1-\gamma)R_{t-1}^{\phi}$ $I_t = \delta(X_t - S_t) + (1-\delta)I_{t-p}$ $\hat{X}_t(m) = S_t R_t^{\sum_{i=1}^{m} \phi^i} + I_{t-p+m}$	$S_t = \alpha(X_t \vert I_{t-p}) + (1-\alpha)S_{t-1}R_{t-1}^{\phi}$ $R_t = \gamma(S_t \vert S_{t-1}) + (1-\gamma)R_{t-1}^{\phi}$ $I_t = \delta(X_t \vert S_t) + (1-\delta)I_{t-p}$ $\hat{X}_t(m) = (S_t R_t^{\sum_{i=1}^{m} \phi^i})I_{t-p+m}$
	$S_t = S_{t-1}R_{t-1}^{\phi} + \alpha e_t$ $R_t = R_{t-1}^{\phi} + \alpha\gamma e_t \vert S_{t-1}$ $\hat{X}_t(m) = S_t R_t^{\sum_{i=1}^{m} \phi^i}$	$S_t = S_{t-1}R_{t-1}^{\phi} + \alpha e_t$ $R_t = R_{t-1}^{\phi} + \alpha\gamma e_t \vert S_{t-1}$ $I_t = I_{t-p} + \delta(1-\alpha)e_t$ $\hat{X}_t(m) = S_t R_t^{\sum_{i=1}^{m} \phi^i} + I_{t-p+m}$	$S_t = S_{t-1}R_{t-1}^{\phi} + \alpha e_t \vert I_{t-p}$ $R_t = R_{t-1}^{\phi} + (\alpha\gamma e_t \vert S_{t-1})I_{t-p}$ $I_t = I_{t-p} + \delta(1-\alpha)e_t \vert S_t$ $\hat{X}_t(m) = (S_t R_t^{\sum_{i=1}^{m} \phi^i})I_{t-p+m}$

Where: α is the smoothing parameter for the level of the series; γ, smoothing parameter for the trend; δ, smoothing parameter for seasonal indices; ϕ, autoregressive of damping parameter; S_t, smoothed level of the series, computed after X_t is observed; also, the expected value of the data at the end of the period *t* in some models. T_t, smoothed additive trend at the end of period *t*; R_t, smoothed multiplicative trend at the end of period *t*; I_t, smoothed seasonal index at the end of period *t*. It can be additive or multiplicative. X_t, observed value of the time series in period *t*; *m*, number of periods in the forecast lead-time; *p*, number of periods in the seasonal cycle; $\hat{\mathrm{X}}_t(m)$, forecast for m periods ahead from origin *t* and e_t is one-step-ahead forecast error, $e_t = \mathrm{X}_t - \hat{\mathrm{X}}_{t-1}(1)$. Note that $e_t(m)$ should be used for other forecast origins.

Methodology

Holt's two-parameter model, also known as linear exponential smoothing, is a popular smoothing model for forecasting data with trend. Holt's model has three separate equations that work together to generate a final forecast. The first is a basic smoothing equation that directly adjusts the last smoothed value for last period's trend. The trend itself is updated over time through the second equation, where the trend is expressed as the difference between the last two smoothed values. Finally, the third equation is used to generate the final forecast. Holt's model uses two parameters, one for the overall smoothing and the other for the trend smoothing equation. The method is also called double exponential smoothing or trend-enhanced exponential smoothing.

The paper adapts method-selection procedures using time series characteristics proposed by Gardnerand McKenzie (1988), Shah (1997), and Meade (2000). The aim is not to improve accuracy but to avoid fitting a damped trend when simpler methodsserve just as well. The method selection rules are summarized in Table 2:

Table 2: Method selection rules

CASE	Series yielding minimum variance	Method
A	X_t	N-N
B	$(1-\mathrm{B})\mathrm{X}_t$	DA-N
C	$(1-\mathrm{B})^2\mathrm{X}_t$	A-N
D	$(1-\mathrm{B}^p)\mathrm{X}_t$	N-M
E	$(1-\mathrm{B})(1-\mathrm{B}^p)\mathrm{X}_t$	DA-M
F	$(1-\mathrm{B}^2)(1-\mathrm{B}^p)\mathrm{X}_t$	A-M

In Case A, the N-N method is suggested if differencing serves only to increase variance and trend or seasonal pattern are not allowed. In Case B, the DA-N method is recommended because it is equivalent to an ARIMA process with a difference of order 1 as can be seen in our empirical data analysis of Section 4. Although the N-N method is also equivalent to an ARIMA process with a difference of order 1, the DA-N method is suggested for reasons of robustness. In Case C, the A-N method is justified by its equivalence to an ARIMA process with a difference of order 2. A multiplicative trend is another possibility in Case C; although Gardner and McKenzie (1988) argue that such trends are dangerous in automatic forecasting systems. In Cases D, E, and F, a seasonal method is called for because a seasonal difference reduces variance. Forecast accuracy was slightly better than the DA-N method applied to all nonseasonal series, with the DA-M method applied to all seasonal series. We also estimate discriminant scores from standard statistics such as autocorrelations and coefficients of skewness and kurtosis. The sample accuracy results are combined with discriminant scores to determine the best method for the daily recorded confirmed coronavirus cases in Nigeria within the sampled period.

The DA-N method can be used to forecast multiplicative trends with the autoregressive or damping parameter ϕ restricted to the range $1<\phi<2$, a method sometimes called “generalized Holt.” This collects special versions of the standard Holt-Winters methods to cope with missing or irregular observations, irregular update intervals, planned discontinuities, series containing a fixed drift, and series containing two or more seasonal cycles. We can also simplify the A-A method by merging the level and seasonal components, and adapt several methods to multivariate series. If the time between updates of the N-N method is irregular, the data for several periods may be reported as a combined observation. Obviously, the smoothing parameter should be increased to give more weight to combined observations. Due to the fundamental importance of time forecasting in many practical situation, proper care should be taken while selecting a particular model, to estimate forecast accuracy and to compare different models. The following indicators measure the forecasting results:

Root mean squared error (RMSE) = $\sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$;

Mean absolute error (MAE) = $\frac{1}{n}\sum_{t=1}^{n} |e_t|$;

Mean percentage error (MPE) = $\frac{1}{n}\sum_{t=1}^{n} \frac{e_t}{\mathrm{X}_t} \times 100$ and

Mean absolute percentage error (MAPE) = $\frac{1}{n}\sum_{t=1}^{n} \left|\frac{e_t}{\mathrm{X}_t}\right| \times 100$.

The criteria included the Akaike information criteria (AIC) and Bayesian information criteria (BIC) that penalize the likelihood of the data by a function of the number of parameters in the model. This is completely automatic expert system that selects from a different set of candidate methods: the N-N and DA-N methods, classical decomposition, and a combination of all candidates.

Results and Discussion

The paper adapts Box-Jenkins strategyto analyze the daily recorded confirmed COVID-19 cases in Nigeria from February 29, 2020 to July 6, 2020 (NCDC, 2020). First, we plot the series as presented in Fig. 1 and look for trend, seasonal variation, outliers, and changes in structure that may be slow or sudden and may indicate that exponential smoothing is not appropriate in the first place. The plot displays a non-stationary upward trending behavior in the recorded daily coronavirus diseases within the sampled period. This non-stationarity is further depicted by the slow decreasing nature of autocorrelation function (ACF) and partial autocorrelation function (PACF) plotted in Figs. 2 and 3.

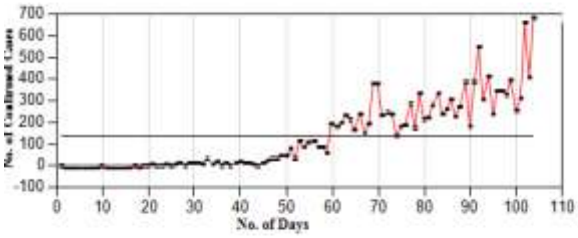

Fig. 1: Time plot of the number of confirmed cases in Nigeria

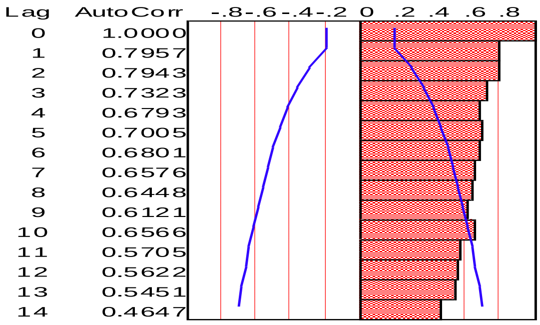

Fig. 2: ACF of original series

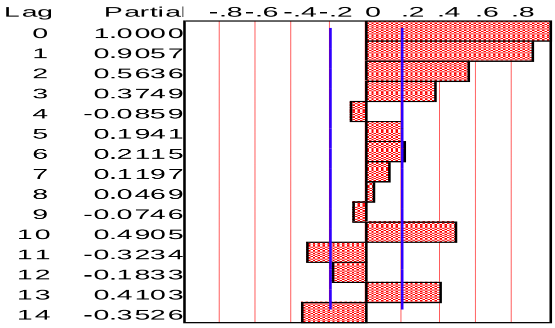

Fig. 3: PACF of original series

We also examine for any outliers, consider making adjustments, and then decide on the form of the trend and seasonal variation. At this point, we consider the possibility of transforming the data, either to stabilize the variance or to make the seasonal effect additive. The original data of daily recorded confirmed COVID-19 cases were transformed by taking the first difference and plotted in Fig. 4. The plot shows an estimated mean level of zero and an approximately constant variance. The autocorrelation function of the differenced series plotted in Fig. 5 decreases fast after lag one indicating a moving average model of order one (MA(1)). The partial autocorrelation function in Fig. 6 also decreases fast after lag one indicating an autoregressive model of order one (AR(1)). The PACF also spikes on lag 9 and beyond pointing that autoregressive moving average (ARMA) model should be considered as well as competing models.

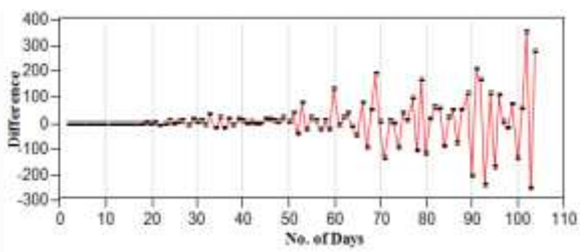

Fig. 4: 1st difference of the number of confirmed cases in Nigeria [(1-B)^1]

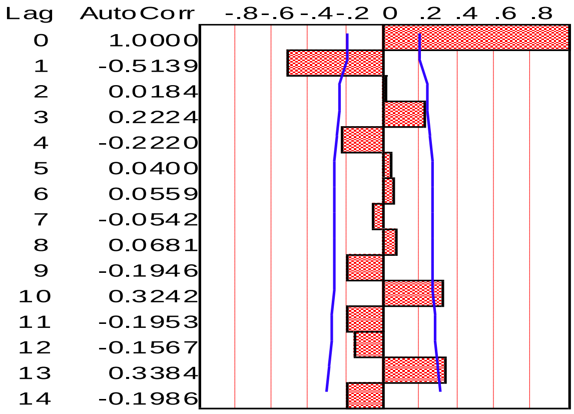

Fig. 5: ACF of Difference (1-B)^1

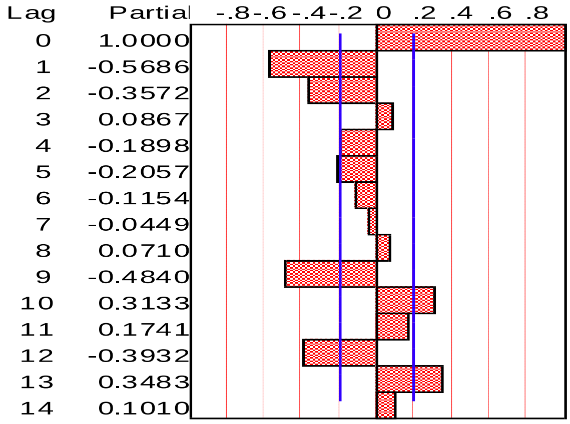

Fig. 6: PACF of Difference (1-B)^1

Table 3: Model comparison

Model	DF	Variance	AIC	SBC	RSquare	–2LogLH
ARI(1, 1)	101	5135.7907	884.03088	889.30033	0.797	878.39221
IMA(1, 1)	101	4600.3442	872.6903	877.95976	0.818	867.49541
ARIMA(1, 1, 1)	100	4613.1156	874.97585	882.88004	0.819	866.69847
Linear (Holt) Exponential Smoothing	100	4667.2096	865.7283	870.97824	0.815	865.65497

Given Table 3 evidences, all four models appear to be correctly specified. We, however, choose ARIMA (0, 1, 1) model among the ARIMA models and linear (Holt) exponential smoothing models asmodels for forecasting due to lower values parameter diagnostics anda higher probability of the residuals being serially uncorrelated.

The models are estimated are presented in Table 4. In the linear non-seasonal methods, the parameters are always invertible if they are chosen from the usual [0, 1] interval. Invertible parameters create a model in which each forecast can be written as a linear combination of all past observations, with the absolute value of the weight on each observation less than one, and with recent observations weighted more heavily than older ones. Non-invertibility usually occurs when one or more parameters fall near boundaries, or when trend and/or seasonal parameters are greater than the level parameter.

Next, we fit an appropriate method, produce forecasts, and check the adequacy of the method by examining the one-step-ahead forecast errors, particularly their autocorrelation function. The findings may lead to a different method or a modification of the selected method. For a sample of reasonable size, it would be useful to have results for this strategy as a validation of the automatic method selection procedures discussed above. It does not appear that any of the automatic procedures have been validated in such a manner.

The accuracy of the models: N-N, A-A, DA-N and DA-M using root mean square error (RMSE), mean absolute error (MAE), mean percentage error (MAE), and mean absolute percentage error (MAPE) refer to as the magnitude of the error rate (errors) of an estimate, the smaller the value of these test statistics, the better the forecasts. The results of the evaluation of the forecasting models displayed in Table 5. The model accuracy analysis showed that Holt's with exponential trend model has the lowest ranked error rate. For Simple Exponential Smoothing model (N-N), the estimated parameter of α is approximately one, as the series is clearly trending over time. From the analysis, it is evident that rising trend of COVID-19 cases in Nigeria is not affected by seasonality. This is true as the Holt's linear exponential trend model is the most accurate method among the exponential smoothing models to describe the data according to the RMSE (2.1756), MAE (1.2891), and MAPE (5.5472), while Holt's damped multiplicative is most accurate according to MPE (- 1.3549). In resolving this type of conflicting results, we selected the least ranked value of the error test statistics, which is the Holt's linear exponential smoothingmodel.

For the A-A method, assuming only that one-step-ahead errors are uncorrelated,an analytical variance of linear exponential smoothingmodel equivalent to ARIMA model must be optimal; the width of the multiplicative prediction intervals depends on the time origin and can change with seasonal peaks and troughs. Hyndman *et al*. (2005b) is an extremely valuable reference because it contains all known results for variances and prediction intervals around point forecasts. The models are divided into three classes. The first class includes linear modelswith additive errors and ARIMA equivalents, corresponding to the N-N, A-N, DA-N, N-A, A-A, and DA-A methods. The second class includes the same models, but now the errors are assumed to be multiplicative to enable the variance to change with the level and trend of the time series. In the third class, including the N-M, A-M, and DA-M methods, the variance changes with level, trend, and the multiplicative seasonal pattern (Gardner, 2005). In fitting additive seasonal models, it is alarming that some combinations of [0, 1] parameters fall within the ARIMA invertible region, yet the weights on past data diverge.

Table 4: Model parameter estimates

Model	Term	Lag	Estimate	Std Error	t Ratio	Prob>|t|
ARIMA(1, 1, 0)	AR1	1	-0.5628767	0.0863048	-6.52	<.0001
	Intercept	0	5.69417229	4.4897797	1.27	0.2076
	Constant Estimate		8.89928912			
ARIMA(0, 1, 1)	MA1	1	0.74945248	0.0766922	9.77	<.0001
	Intercept	0	4.92275746	1.7230526	2.86	0.0052
	Constant Estimate		4.92275746			
ARIMA(1, 1, 1)	AR1	1	-0.1463363	0.1718897	-0.85	0.3966
	MA1	1	0.65580404	0.1673657	3.92	0.0002
	Intercept	0	5.09477048	2.0511706	2.48	0.0147
	Constant Estimate		5.8403204			
Linear (Holt) Exponential Smoothing	Level Smoothing Weight		0.2431455	0.0787781	3.09	0.0026
	Trend Smoothing Weight		0.051954	0.051954	1.28	0.2042

Table 5: Error estimation for different forecast models

Exponential Smoothing	Estimated errors				Rank
	RMSE	MAE	MPE	MAPE	
Simple Exponential Smoothing (N-N) $\alpha = 0.9788$	2.3459	1. 5637	3.3218	6.0015	16
Holt's Exponential Additive (A-A) $\alpha = 0.8716$ $\beta = 0.0010$	2.1756	1.2891	-0,1849	5.4872	5
Holt's Damped Additive (DA-N) $\alpha = 0.8975$ $\phi = 0.96$ $\beta = 0.0513$	2.3567	1. 4891	1.2893	5.511	10
Holt's Damped Multiplicative (DA-M) $\alpha = 0.8904$ $\phi = 0.96$ $\beta = 0.0010$	2.1801	1.5137	-1.3549	6.1290	15

Table 6: 60-day forecast of Coronavirus cases in Nigeria

Time	ARIMA (0, 1, 1)			Linear (Holt) Exponential Smoothing Model		
	Forecast Value	95% CI Lower	95% CI Upper	Forecast Value	95% CI Lower	95% CI Upper
105	536.79683	396.337127	677.256534	503.137245	369.238427	637.036062
106	626.864722	473.571967	780.157476	513.490255	375.280838	651.699672
107	585.066895	398.751313	771.382476	523.843265	381.037408	666.649122
108	617.493206	414.391304	820.595109	534.196276	386.515453	681.877098
109	608.14048	383.9397	832.341261	544.549286	391.722906	697.375665
110	622.304201	381.809981	862.798421	554.902296	396.668084	713.136508
111	623.231062	365.925177	880.536947	565.255306	401.359491	729.151121
112	631.608643	359.356119	903.861166	575.608317	405.805663	745.41097
113	635.792387	348.929365	922.655409	585.961327	410.015036	761.907617
114	642.336744	341.810576	942.862912	596.314337	413.995858	778.632816
115	647.552367	333.82939	961.275344	606.667347	417.75611	795.578585
116	653.515903	327.198804	979.833003	617.020357	421.303464	812.737251
117	659.058457	320.577847	997.539067	627.373368	424.645249	830.101487
118	664.837972	314.636571	1015.03937	637.726378	427.788432	847.664324
119	670.484107	308.930501	1032.03771	648.079388	430.739614	865.419162
120	676.205318	303.651371	1048.75927	658.432398	433.50503	883.359767
121	681.884271	298.642257	1065.12628	668.785409	436.090557	901.48026
122	687.58701	293.948857	1081.22516	679.138419	438.501725	919.775112
123	693.27636	289.508652	1097.04407	689.491429	440.743733	938.239125
124	698.973247	285.324517	1112.62198	699.844439	442.821465	956.867414
125	704.665891	281.366425	1127.96536	710.19745	444.739504	975.655395
126	710.360923	277.626062	1143.09578	720.55046	446.502159	994.598761
127	716.054612	274.085649	1158.02357	730.90347	448.11347	1013.69347
128	721.749056	270.735064	1172.76305	741.25648	449.577238	1032.93572
129	727.443075	267.56189	1187.32426	751.609491	450.897032	1052.32195
130	733.137334	264.556739	1201.71793	761.962501	452.07621	1071.84879
131	738.831458	261.710036	1215.95288	772.315511	453.117933	1091.51309
132	744.525657	259.013636	1230.03768	782.668521	454.02518	1111.31186
133	750.219814	256.459753	1243.97988	793.021531	454.800756	1131.24231
134	755.913995	254.04143	1257.78656	803.374542	455.447314	1151.30177
135	761.608163	251.752157	1271.46417	813.727552	455.967357	1171.48775
136	767.302338	249.585986	1285.01869	824.080562	456.363254	1191.79787
137	772.996508	247.537381	1298.45564	834.433572	456.637248	1212.2299
138	778.690681	245.601225	1311.78014	844.786583	456.791465	1232.7817
139	784.384853	243.772753	1324.99695	855.139593	456.827922	1253.45126
140	790.079026	242.047533	1338.11052	865.492603	456.748538	1274.23667
141	795.773198	240.421426	1351.12497	875.845613	456.555135	1295.13609
142	801.46737	238.890562	1364.04418	886.198624	456.249449	1316.1478
143	807.161543	237.451319	1376.87177	896.551634	455.833135	1337.27013
144	812.855715	236.100296	1389.61113	906.904644	455.307772	1358.50152
145	818.549887	234.834299	1402.26547	917.257654	454.674869	1379.84044
146	824.244059	233.650323	1414.8378	927.610665	453.935868	1401.28546
147	829.938232	232.545534	1427.33093	937.963675	453.09215	1422.8352
148	835.632404	231.517258	1439.74755	948.316685	452.145038	1444.48833
149	841.326576	230.562969	1452.09018	958.669695	451.095801	1466.24359
150	847.020749	229.680277	1464.36122	969.022705	449.945657	1488.09975
151	852.714921	228.866917	1476.56292	979.375716	448.695776	1510.05566
152	858.409093	228.120741	1488.69744	989.728726	447.347285	1532.11017
153	864.103265	227.439712	1500.76682	1000.08174	445.901267	1554.26221
154	869.797438	226.82189	1512.77299	1010.43475	444.358766	1576.51073
155	875.49161	226.265433	1524.71779	1020.78776	442.72079	1598.85472
156	881.185782	225.768585	1536.60298	1031.14077	440.988311	1621.29322
157	886.879955	225.329672	1548.43024	1041.49378	439.162266	1643.82529
158	892.574127	224.947097	1560.20116	1051.84679	437.243563	1666.45001
159	898.268299	224.619336	1571.91726	1062.1998	435.23308	1689.16652
160	903.962471	224.344932	1583.58001	1072.55281	433.131666	1711.97395
161	909.656644	224.122492	1595.1908	1082.90582	430.940144	1734.87149
162	915.350816	223.95068	1606.75095	1093.25883	428.659311	1757.85835
163	921.044988	223.82822	1618.26176	1103.61184	426.289941	1780.93374
164	926.739161	223.753886	1629.72443	1113.96485	423.832785	1804.09691

Table 6 list a 60-day period of confirmed COVID-19 cases in Nigeria forecast values and the 95% confidence interval estimates of ARIMA (0, 1, 1) and linear (Holt) exponential smoothing models respectively. Estimate of confirmed COVID-19 cases from day to day are close to one another. Confidence intervals for both forecast values have widths of 0.10 or 0.16 in all days showing remarkable precision of the forecast. The forecast values of the two selected models show that the COVID-19 pandemic will live with us for a long term. For a stationary series and model, the forecasts of future values will eventually converge to the mean and then stay there. For the purpose of the study, the applied ARIMA models remain the most suitable statistical tool since the data they are applied on are not volatile as obtainable with high frequency data such as financial data. These stylized volatile data are captured with autoregressive conditional heteroskedasticity (ARCH) models proposed by Engle (1982). The ARIMA models (being a crucial forecasting tool) are equally adopted in identifying parameter orders in generalized autoregressive conditional heteroskedasticity GARCH (p, q) models used in empirical applications of financial data (Bollerslev, 1986). The application of exponential smoothing to volatility forecasting is very different to the usual exponential smoothing applications. With financial returns, the mean is often assumed to be zero or a small constant value, and attention turns to predicting the variance. In the additive seasonal methods, it is not necessary to renormalize the seasonal indices if forecast accuracy is the only concern, but this is rarely the case in practice when repetitive forecasts are made over time. Forecasting methods require regular maintenance, a job that is easier to accomplish when the method components can be interpreted without bias.

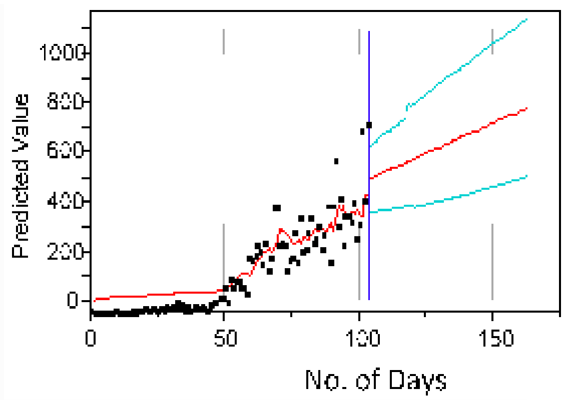

Fig. 6: ARIMA (0, 1, 1) forecast of COVID-19 cases

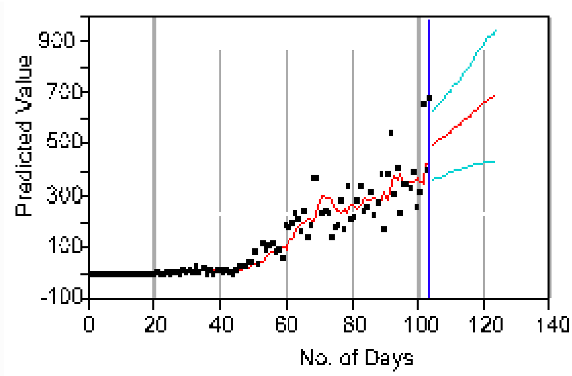

Fig. 7: Linear (Holt) exponential smoothing model forecast of COVID-19 cases

In Figs. 6 and 7, we observe that at short forecast horizons, the residual random noise term dominates while as the horizon increases, the importance of noise term is superseded by sampling error.

Conclusion

The ARIMA models are appropriate for large sample greater than 42 data points and the selected data are not stochastically volatile while exponential smoothing methods are used in the situations for which they are appropriate (simple exponential smoothing where there is no underlying trend and double exponential smoothing where there is an underlying linear trend), along with good starting forecasts, the best smoothing constants tend to be very small, if not zero. Significantly large smoothing constants signal the presence of either trend (simple exponential smoothing) or changes in trend (double exponential smoothing). This is a strong argument for the use of adaptive smoothing methods – methods that monitor forecast errors continuously and change the smoothing constants to keep them within predetermined limits – and more coverage of them in forecasting classes.

The empirical results of optimal smoothing constants seem to suggest the following: When the initial forecast is good, α values will very often be zero. In fact, small non-zero values of α are indicative of local trends. Larger non-zero values of α are indicative of sustained trends which might be better accounted for with a technique, like double exponential smoothing; Large values of the smoothing constants are certainly possible and should not be rejected without detailed examination of the underlying series or of the quality of the initial forecasts used; When there is a linear trend in the data, the performance of double exponential smoothing depends on the initial estimates of the level and trend components. Where these are good, α and β will be very small. This is true of small as well as large series. Larger values might be indicative of poor initial estimates of level and trend. The impact of poor initial forecasts is felt less on longer series than on smaller ones. The values of α and β decrease with series length; When there is a nonlinear trend in the data, the results are mixed and not easily generalizable. The best values of α and β depend on the particular kind of nonlinearity involved. The best approach is to graph the time series and pick appropriate starting values before finding the optimal values of α and β.

Conflict of Interest

Authors have declared that there is no conflict of interest reported in this work.

References

Akaike H 1970. Statistical predictor identification. *Annals of the Institute of Statistical Mathematics*, 22: 203-217.

Bollerslev T 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31: 307-327.

Bossons J 1966. The effects of parameter misspecification and non-stationary on the applicability of adaptive forecasts. *Management Science*, 12: 659-669.

Brown RG 1959. Statistical Forecasting for Inventory Control. New York: McGraw-Hill.

Cogger KO 1973. Specification analysis. *J. Am. Stat. Assoc.*, 68: 899-905.

Engle RF 1892. Autoregressive conditional heteroskedasticity with estimates of United Kingdom inflation. *Econometrica*, 50: 987-1007.

Gardner ES Jr. 2005. *Exponential Smoothing: The State of the art-Part II.* Bauer College of Business, University of Houston, Houston.

Gardner ES Jr. & McKenzie E 1988. Model identification in exponential smoothing, *J. Operational Res. Soc.*, 39: 863-867.

Johns Hopkins University 2020. World Confirmed Coronavirus (COVID-19) Cases: July, 2020.

Holt CC 2004. Forecasting seasonals and trends by exponentially weighted moving averages. *Int. J. Forecasting*, 20: 5-10.

Hyndman RJ, Koehler AB, Snyder RD & Grose S 2002. A state space framework for automatic forecasting using exponential smoothing methods. *Int. J. Forecasting*, 18: 439-454.

Meade N 2000. Evidence for the selection of forecasting methods. *Journal of Forecasting*, 19: 515-535.

Muth JF 1960. Optimal properties of exponentially weighted forecasts. *J. Am. Stat. Assoc.*, 55: 299-306.

NCDC 2020. Nigeria Confirmed Coronavirus (COVID-19) Cases. Nigeria Centre for Disease Control (NCDC), July, 2020.

Pegels C 1969. Exponential forecasting: Some new variations. *Management Science*, 15: 311-315.

Rosanna, R.J., &Seater, J.J. (1995). Temporal aggregation and economic time series. *J. Bus. and Econ. Stat.*, 13: 441-451.

Satchell S & Timmermann A 1995. On the optimality of adaptive expectations: Muth revisited. *Int. J. Forecasting*, 11: 407-416.

Schnaars SP 1986. A comparison of extrapolation models on yearly sales forecasts. *Int. J. Forecasting*, 2: 71-85.

Schwarz G 1978. Estimating the dimension of a model. *The Annals of Statistics*, 6: 461-464.

Shah C 1997. Model selection in univariate time series forecasting using discriminant analysis. *Int. J. Forecasting*, 13: 489-500.

Taylor JW 2004a. Volatility forecasting with smooth transition exponential smoothing. *Int. J. Forecasting*, 20, 273-286.

Tiao GC & Xu D 1993. Robustness of maximum likelihood estimates for multi-step predictions: The exponential smoothing case. *Biometrika*, 80: 623-641.

Winters PR 1960. Forecasting sales by exponentially weighted moving averages. *Management Science*, 6: 324-342.

Yar M & Chatfield C 1990. Prediction intervals for the Holt-Winters forecasting procedure. *Int. J. Forecasting*, 6: 127-137.