International Journal of Innovative Research and Advanced Studies (IJIRAS)
Volume 10 Issue 5, May 2023

ISSN: 2394-4404

On The Modelling Of Covid-19 Cases In Nigeria

Akin-Awonira, Bukola .O

Salami, Taofeek Adeola

Adeyemi, Adekunle Mudasiru

Yusuf, Jamal Asamu

Arojo, Toyin Maria

Department of Mathematics/Statistics, Faculty of Pure and Applied Sciences,
Osun State College of Technology, Esa-Oke, Osun State, Nigeria

Abstract: ***Nigeria is one of the ten most affected countries in Africa with over 190,000 cases and over 2000 deaths to COVID-19 and this situation call for concern, as it is necessary to study the rate at which the country records new cases on a daily basis. This research paper models the series of COVID-19 in Nigeria using ARIMA analysis and it covers the data on total daily-confirmed cases of the coronavirus in Nigeria from February 27, 2020 to June 6, 2021 which was extracted from Coronavirus Daily Situation Report by Nigerian Centre for Disease Control. The study established comparison of three adopted models (Least Square Model and ARIMA model of two different order). The findings revealed that; (i) there is an upward trend in the total daily cases on COVID-19, (ii) the series was stationary after first differencing and (iii) the appropriate model among the adopted model is ARIMA (1,1,1). Finally, we deduced that the total daily covid-19 cases increasing regularly in Nigeria while findings showed that ARIMA (1,1,1) is a better choice for prediction.***

Keywords: Covid-19, ARIMA, ACF, R-Square, NCDC, Cases, Coronavirus, PACF.

I. INTRODUCTION

Coronavirus pandemic has engulfed the nations of the world for the first five months of 2020 and altered the pace, fabric, and nature of our lives. Coronavirus Disease 2019 also known as COVID-19 was first identified in Wuhan, China. This disease is a contagious respiratory and vascular disease (EBio Medicine, 2020). It mainly spread through the air when people are near each other long enough, primarily via small droplets or aerosols, as an infected person breathes, coughs, sneeze, sings, or speaks. It can be spread as early as two days before infected persons show symptoms. People remain infectious for up to ten days in moderate cases, and two weeks in severe cases.

The World Health Organization (WHO) officially declared the COVID-19 outbreak a Public Health Emergency of International Concern on January 30, 2020, and a global pandemic on March 11, 2020 (Cucinotta & Vanelli, 2020). Countries were urged by the World Health Organization to adopt strict social distancing and quarantine measures to avoid virus spread and to protect public health. Despite fragmented international efforts to contain the spread, it has spread all over the world. As at September, 2021, there are over 222 million confirmed cases of the virus with over 4.5 million death recorded (Worldometers, 2021).

In Africa, Egypt was the first country to record a case of coronavirus. During this first recorded case of the virus on the continent, there was tension that the virus could be more prevalent in African country due to poor health facilities available in almost all the countries on the continent. However, African countries have been hit relatively less by the virus compared to Europe, Asia and America (World Bank, 2020). Nigeria is one of the 10 most affected countries in Africa with over 190,000 cases and over 2000 deaths

(NCDC, 2021). The country first recorded its first case of coronavirus on 27th February, 2020 and adopted various measures to reduce the spread of the disease. Despite various measure that have been adopted, the country still record new cases and death daily (NCDC). It is therefore become necessary to keep track of the daily recorded cases of the disease in the country for proper intervention in cushioning the spread and its effect.

II. LITERATURE REVIEW

THE ORIGIN AND CLASSIFICATION OF VIRUS

The SARS-COV- 2 belongs to a family of viruses that may cause various symptoms such as pneumonia, fever, difficulty in breathing and pneumonitis (WMHC, 2020). The clinical manifestations include fever, cough, dyspnea, myalgia, fatigue, normal or decreased leukocyte counts and radiographic evidence of pneumonia. Organ dysfunction (such as shock, ARDS, acute cardiac injury and acute kidney injury) and death can occur in severe cases, yet severity seems to be associated with age, biological sex and comorbidities (Huang, 2020). Full-genome sequencing and phylogenic analysis indicated that SARS-Cov- 2 was a distinct clade from the beta-coronaviruses associated with humans, in addition to the six known coronaviruses that infect humans: HCoV-229E, HCoV-OC43, SARS-CoV, HCoV-NL63, HCoV-HKU1 and MERS-CoV (Zhu, 2019). It is closely similar to bat coronaviruses, with a homology of 85–96% to a bat SARS-like coronavirus (bat-SL- CoVZC45) at the whole-genome level (Zhou, 2020) and it has been postulated that bats may be the primary source, nevertheless no specific animal association has been identified. At this present moment, the origin of SARS Cov- 2 is yet to be completely ascertained. COVID-19 has been found to have higher levels of transmissibility and pandemic risk than the previous SARS-CoV, as the effective reproductive number (R) of COVID-19 (2.9) is estimated to be higher than the reported effective reproduction number (R) of SARS (1.77) at this early stage. Although, the average incubation period of COVID-19 was initially estimated to be 4.8±2.6, ranging from 2 to 11 days (Wu A, 2020), currently guidelines from health authorities state an average incubation period of 7 days, ranging from 2 to 14 days (Huang, 2020). Signs of Covid-19 infection overlap with other viral infections, which makes a clinical diagnosis very tricky. Diagnostic test based on detection of the viral sequence by real-time reverse-transcription (RT)-PCR is the gold standard confirmatory test. People with positive SARS-CoV- 2 RNA by respiratory tract specimens are probably an infectious source of COVID-19 (Jeong, 2014).

VIRUS TRANSMISSION MODE

According to World Health Organization (2020), transmission of SARS-CoV-2 can occur through direct, indirect, or close contact with infected people through infected secretions such as saliva and respiratory secretions or their respiratory droplets, which are expelled when an infected person coughs, sneezes, talks or sings. Respiratory droplets are >5-10 μm in diameter whereas droplets <5μm in diameter are referred to as droplet nuclei or aerosols. Respiratory droplet transmission can occur when a person is in close contact (within metre) with an infected person who has respiratory symptoms (e.g. coughing or sneezing) or who is talking or singing; in these circumstances, respiratory droplets that include virus can reach the mouth, nose or eyes of a susceptible person and can result in infection. Indirect contact transmission involving contact of a susceptible host with a contaminated object or surface (fomite transmission) may also be possible.

Airborne transmission: SARS-CoV-2 airborne transmission can occur during medical procedures that generate aerosols (“aerosol generating procedures”). WHO, together with the scientific community, has been actively discussing and evaluating whether SARS-CoV-2 may also spread through aerosols in the absence of aerosol generating procedures, particularly in indoor settings with poor ventilation (WHO, 2020).

The physics of exhaled air and flow physics have generated hypotheses about possible mechanisms of SARS-CoV-2 transmission through aerosols. These theories suggest that 1) a number of respiratory droplets generate microscopic aerosols (<5 μm) by evaporating, and 2) normal breathing and talking results in exhaled aerosols. Thus, a susceptible person could inhale aerosols, and could become infected if the aerosols contain the virus in sufficient quantity to cause infection within the recipient. However, the proportion of exhaled droplet nuclei or of respiratory droplets that evaporate to generate aerosols, and the infectious dose of viable SARS-CoV-2 required to cause infection in another person are not known, but it has been studied for other respiratory viruses (Gralton, 2013).

Fomite transmission: Respiratory secretions or droplets expelled by infected individuals can contaminate surfaces and objects, creating fomites (contaminated surfaces). Viable SARS-CoV-2 virus and/or RNA detected by RT-PCR can be found on those surfaces for periods ranging from hours to days, depending on the ambient environment (including temperature and humidity) and the type of surface, in particular at high concentration in health care facilities where COVID-19 patients were being treated. Therefore, transmission may also occur indirectly through touching surfaces in the immediate environment or objects contaminated with virus from an infected person (e.g. stethoscope or thermometer), followed by touching the mouth, nose, or eyes. Despite consistent evidence as to SARS-CoV-2 contamination of surfaces and the survival of the virus on certain surfaces, there are no specific reports which have directly demonstrated fomite transmission. People who come into contact with potentially infectious surfaces often also have close contact with the infectious person, making the distinction between respiratory droplet and fomite transmission difficult to discern. However, fomite transmission is considered a likely mode of transmission for SARS-CoV-2, given consistent findings about environmental contamination in the vicinity of infected cases and the fact that other coronaviruses and respiratory viruses can transmit this way (WH0, 2020).

ISSN: 2394-4404

PREVALENCE OF COVID-19 PANDEMIC IN NIGERIA

The first case of COVID-19 in Nigeria was confirmed on 27 February 2020. The case was a 44-year old Italian citizen who arrived Nigeria through the Murtala Mohammed International Airport, Lagos, on a flight via Milan, Italy. This index case led to the activation of COVID-19 Public Health Emergency Operation Centers (PHEOC) at national and sub-national levels, with associated active case finding via contact tracing. By 9 March 2020, 217 contacts were linked to this index case, out of which 136 (63.0%) were under follow-up, with one contact confirmed positive. The 14-day follow-up for contacts of the index case ended on 12 March 2020. During this period, two additional unlinked cases were reported in Nigeria. In addition, 42 suspected cases were identified across seven states in Nigeria namely the Federal Capital Territory (FCT), Edo, Kano, Lagos, Ogun, Rivers and Yobe (Ihekweazu, 2020).

The federal government of Nigeria enforced an initial 2-week lockdown on March 30, 2020, for three of 36 states (Lagos, Ogun, and Abuja) and April 13, extended it another 2 weeks. Since the confirmation of the first COVID-19 case in Nigeria, cases and deaths have risen steadily in the country, although the government implemented public health interventions – e.g. advocacy for physical distancing, complete and partial lockdown, and ban on large public gatherings including at churches and mosques – to contain or mitigate spread. By 6 June 2020, 35 (out of 36) states, plus the FCT, have reported at least one confirmed COVID-19 case. A descriptive analysis of the clinical characteristics, treatment modalities and outcomes of the first 32 COVID-19 patients admitted to Mainland Hospital in Lagos State, Nigeria, found that two-thirds of patients were male, and the mean age was 38.1 years. This early analysis however was insufficient to provide a national overview of COVID-19 epidemiology in Nigeria. The Nigeria Centre for Disease Control (NCDC) coordinates the public health response to COVID-19 in the country. Through NCDC's surveillance and laboratory network as well as coordination of state PHEOCs, epidemiological information on COVID-19 cases are captured into a real-time networked platform called Surveillance Outbreak Response Management and Analysis System (SORMAS). This forms the basis for the release of daily situation reports for COVID-19 on NCDC COVID-19 microsite (NCDC, 2020). By 6 June, thousands of individual records with laboratory diagnosis contained on SORMAS offered opportunities to expand and explore country-specific epidemiologic and clinical characteristics of COVID-19 from the onset of the outbreak. Despite the enforcement of lockdown and other preventive measures, new cases of the disease is still being recorded across the states in country. Nigeria is one of the worst affected countries in Africa by the COVID-19 pandemic, having reported 180,000+ infections and 2000+ deaths as of September, 2021 (NCDC, 2021). This number of cases is the highest among the West African country and tenth-highest on Africa continent (Statista, 2021).

Alzahrani, et al (2020) in their study used Autoregressive Integrated Moving Average (ARIMA) model to estimate the number of COVID-19 cases per day in Saudi Arabia. They implemented four different forecasting models. Autoregressive model, moving average and combined ARMA model was examined to determine the best model. It was found that the ARIMA model is superior to other models. The evaluation results indicate that if strict prudence and control measures are not taken to limit the spread of Saudi Arabia, Saudi Arabia's model will continue to evolve, possibly adding up to 7668 new cases per day, and 127,129 cases per day in just 1 month. Aslam M. (2020) proposed to use a more logical method of the Kalman channel in combination with the Autoregressive Integrated Moving Average (ARIMA) model in order to obtain a more accurate estimation of the prevalence, dynamic cases, active cases, and death cases of COVID-19 pandemic identification in Pakistan (Aslam 2020). A basic econometric model is proposed by Benvenuto, et al (2020) which may be very valuable for predicting the spread of COVID-2019. They carried out Autoregressive Integrated Moving Average (ARIMA) model prediction on Johns Hopkins epidemiological information to predict the epidemiological model of the prevalence and incidence of COVID-19. It was suggested in the study that, for further inspection or future consideration, the case definition and information classification must be maintained continuously.

Sahai, et al (2020) studied the timing information of the top five countries affected by COVID-19 to estimate the spread. From the online dataset, the time schedule information from February 15, 2020, to June 30, 2020, of all contaminated cases from five main countries (especially the USA, Brazil, India, Russia, and Spain) was collected. The detailed information of the ARIMA model was evaluated using Hannan and Rissanen calculations. The ARIMA model was used to process the test chart for the next 77days.The estimated values for the 18days in early July were compared, and within a satisfactory understanding range, the true information and graphic accuracy obtained by MAD and MAPE were found. The real track gauge info-graphic shows that although Russia and Spain have become the focus of the epidemic, the USA, Brazil, and India are still in exponential bending. The survey shows that India and Brazil will reach 1.38 million and 2.47millionmarks, respectively, while the USA will reach 4.29 million marks on July 31.

Ilie, et al (2020) in their study established an Autoregressive Integrated Moving Average (ARIMA) model used it to predict the epidemiological patterns of COVID-19 in Ukraine, Romania, the Republic of Moldova, Serbia, Bulgaria, Hungary, USA, Brazil, and India. In order to improve the accuracy, the main daily information of COVID-19 from March 10, 2020, to July 10, 2020, was collected from the official sites of the Romanian government (GOV.RO), the World Health Organization (WHO), and the European Center for Disease Control and Prevention (ECDC) website. ARIMA(1,1,0), ARIMA(3, 2, 2), ARIMA (3, 2, 2), ARIMA (3, 1, 1), ARIMA (1, 0, 3), ARIMA ((1, 2), 0), ARIMA (1, 1, 0), ARIMA (0, 2, 1), and ARIMA(0,2,0) models are selected as the best models, depending on their average absolute percentage error (MAPE) for Ukraine, Romania, Republic of Moldova, Serbia, Bulgaria, Hungary, the USA, Brazil, and India (4.70244, 1.40011, 2.67551, 2.16373, 2.98154, 2.11139, 3.21569, 4.10596, and 2.78051) respectively. This survey shows that the ARIMA model is reasonable for expectations in

current emergencies and provides ideas for the epidemiological stage of these regions.

AIM: This paper is basically designed to model the COVID-19 in Nigeria using ARIMA from February 27, 2020 to June 6, 2021.

The objectives of this paper are to:

- ✓ Draw the time plot of the COVID-19 cases recorded in Nigeria.
- ✓ Select the appropriate model for the data set.
- ✓ Test for the autocorrelation in the data.

III. METHODS OF ANALYSIS

Basically, data used in this study was secondary which was obtained from the daily reported cases of coronavirus by the Nigerian Centre for Disease Control (NCDC)

The basis of Box-Jenkins approach to modeling time series consists of three phases:

PHASE 1: MODEL IDENTIFICATION

Using plots of the data, autocorrelations, partial autocorrelations, and other information, a class of simple ARIMA models is selected. This amounts to estimating appropriate values for *p*, *d*, and *q*. Assuming for the moment that there is no seasonal variation, the objective of the model identification step is to select values of *d* and then *p* and *q* in the *ARIMA (p,d,q)* model.

STATIONARITY

The first step in model identification is to ensure the process is stationary. A stationary time series is one whose statistical properties such as mean, variance, autocorrelation etc. are all constant over time. A stationary time series is one whose properties do not depend on the time at which the series is observed. Thus, time series with trends, or with seasonality, are not stationary. Stationarity can be checked with Dickey-Fuller Test.

AUGMENTED DICKEY FULLER TEST

Augmented Dickey Fuller test (ADF Test) is a common statistical test used to test whether a given time series is stationary or not. ADF test is one of the most commonly used statistical test when it comes to analyzing the stationary of a series. The hypothesis of Augmented Dickey Fuller Test is:

H_0: $\theta = 0$(data is not stationary)

H_1: $\theta < 0$ (data is stationary)

When the null hypothesis is reject, this means the data doesn't need to be differenced but when null hypothesis is not reject, the data need to be differenced. Once the process is stationary, we fit the autoregressive and moving average components. To fit the model we use the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) with regard to the fundamental guide and examine various information criteria to select a better-fit-model.

ACF AND PACF PLOT

ACF and PACF Plot are often used in time series analysis and forecasting. These plots graphically summarize the strength of a relationship between present time series values and its lagged values. Auto-correlation Function (ACF) is a complete auto-correlation function which gives us values of auto-correlation of any series with its lagged values. A plot showing this values along with the confidence band and tada are known as ACF plot. ACF plot describes how well the present value of the series is related with its past values. A time series can have components like trend, seasonality, cyclic and residual. ACF considers all these components while finding correlations hence it's a 'complete auto-correlation plot'. Partial autocorrelation function (PACF), basically, instead of finding correlations of present with lags like ACF, it finds correlation of the residuals with the next lag value.

AKAIKE INFORMATION CRITERIA (AIC)

Akaike Information Criterion (AIC) is the most widely used information criteria for predictive modeling. It compares the quality of a set of statistical models to each other. AIC basic formula is defined as:

AIC: $2k - 2\ln(L)$

where:

K is the number of model parameters

ln (L) is a measure of model fit. The higher the number, the better the fit. This is usually obtained from statistical output.

PHASE 2: MODEL ESTIMATION

The estimation procedure involves using the model with p, d, q orders to fit the actual time series. The phis and thetas of the selected model are estimated using maximum likelihood techniques, back-casting, etc. The maximum likelihood equation is solved by nonlinear function maximization. Back-casting is used to obtain estimates of the initial residuals.

AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)

ARIMA models are, in theory, the most general class of models for forecasting a time series which can be made to be "stationary" by differencing (if necessary), perhaps in conjunction with nonlinear transformations such as logging or deflating (if necessary). A random variable that is a time series is stationary if its statistical properties are all constant over time. A stationary series has no trend, its variations around its mean have a constant amplitude, and it wiggles in a consistent fashion, i.e., its short-term random time patterns always look the same in a statistical sense. The latter condition means that its autocorrelations (correlations with its own prior deviations from the mean) remain constant over time, or equivalently, that its *power spectrum* remains constant over time. A random variable of this form can be viewed (as usual) as a combination of signal and noise, and the signal (if one is apparent) could be a pattern of fast or slow mean reversion, or sinusoidal oscillation, or rapid alternation in sign, and it could

also have a seasonal component. Autoregressive Integrated Moving Average (ARIMA) model of order p, q is given as:

$$\hat{y} = a + \emptyset y_{t-1} + \cdots + \emptyset_p y_{t-p} + \omega_t + \theta_1 \omega_{t-1} + \cdots + \theta_q \omega_{t-q} + \sum_{m=1}^{M} \beta_m X_{m,t}$$

MODEL ACCURACY

Mean Squared Error (MSE)

The mean squared error is a measure of how close a fitted line is to data points. It does this by taking the distances from the points to the regression line and squaring them. The square is done in order to remove any negative signs. A larger MSE means that the data values are disperse widely around its central moment (mean), and a smaller MSE means otherwise. A smaller MSE is preferred and desired choice as it shows that data values are dispersed closely to its central moment. The formula for computing MSE is given as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

$MSE = mean\ squared\ error$
$n = number\ of\ data\ points$
$Y_i = observed\ values$
$\hat{Y}_i = predicted\ values$

Root mean square error is the square root of the mean of the square of all of the error. A larger RMSE means that the data values are disperse widely around its central moment (mean), and a smaller MSE means otherwise. In this case, a smaller value is always a desire choice. RMSE formula is given as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$$

PHASE 3: DIAGNOSTIC CHECKING

Once a model has been fit, the final step is the diagnostic checking of the model. The checking is carried out by studying the autocorrelation plots of the residuals to see if further structure (large correlation values) can be found. If all the autocorrelations and partial autocorrelations are small, the model is considered adequate and forecasts are generated. If some of the autocorrelations are large, the values of *p* and/or *q* are adjusted and the model is re-estimated.

LJUNG-BOX Q-TEST

This test is a portmanteau test which assesses the null hypothesis that there is an absence of serial correlation in the residuals for a fixed number of lags L, against the alternative hypothesis that some autocorrelation coefficient $\rho(k), k=1,\ldots,L$, is non-zero.

The test statistic is;

$$Q = T(T+2)\sum k = 1L(\rho(k)2(T-k))$$

$Where$
$T = sample\ size$
$L = number\ of\ autocorrelation\ lags$
$\rho(k) = sample\ autocorrelation\ at\ lag$

IV. DISCUSSION OF RESULTS

Here the modeling of Cases of COVID-19 using Box-Jenkin fundamental approach of total daily cases of COVID-19 from February 27, 2020 to June 06, 2021 was used for the time series analysis.

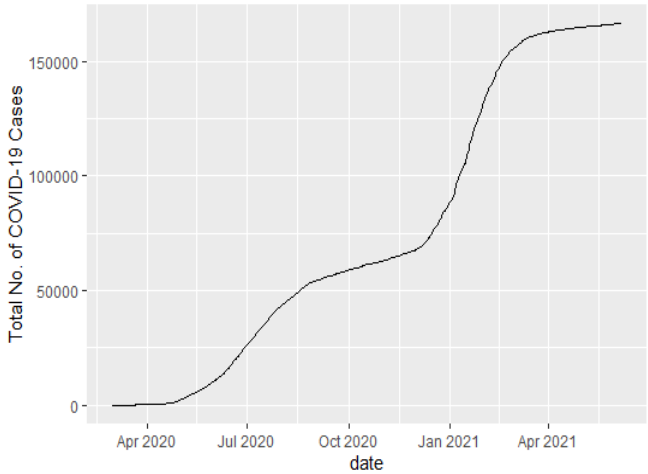

Figure 1: Time Series Plot of Total Daily cases of COVID-19 in Nigeria

Interpretation: There is an upward trend in number of total daily shown in Figure 1 above.

ESTIMATION OF LINEAR MODEL

Coefficients	Estimates	Std. error	t-value	p-value
Intercept	-25391.300	1271.612	-19.968	<0.0001
Trend	432.840	4.749	91.137	<0.0001

Table 1: Linear Model Coefficients

The above table indicated that the cumulative daily COVID-19 in Nigeria would continue to increase over subsequent time as the trend was significant (p-value<0.05). The intercept was negative but significant (p-value<0.05) which indicate that it will take some interval of days for increment to be significant. In other to examine if linear model can be improve upon, the Box & Jenkins approach to ARIMA modelling is employed and comparison were made on the bases of performance measure.

ESTIMATION OF ARIMA MODEL

By adopting this approach to modelling, the first step is to determine if the series is stationary. Consequently, by visual inspection of figure 1 above, the series can be assume to be non-stationary due to pattern of trend. Thus, there is need to make the series stationary by differencing. However, stationarity test is needed to be carried out to ascertain the stationarity status of the series. Hence, Augmented Dickey-Fuller test is used to test the stationarity of the series at level and first differencing.

STATIONARITY TEST

H_0: The series has a unit root (Non stationary) VS
H_1: There is no unit root (Stationary)

Series	Test Statistic	p-value
First Differencing	-439.91	0.01

Table 2: Augmented-Dickey Fuller Test

From table 1, the p-value of the series at level is greater than 0.05, hence we fail to reject H_0. This, implies series is non stationary at level. The p-value of the first order differencing of the series is 0.01 which is less than 0.05, this is in favor of the alternative hypothesis, and hence, the series is stationary after differencing.

MODEL IDENTIFICATION

In order to identify the best order of ARIMA model for the series, correlegram of ACF and PACF of the differenced series is plotted and inspected.

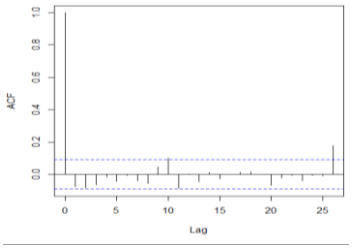

Figure 2: Plot of ACF

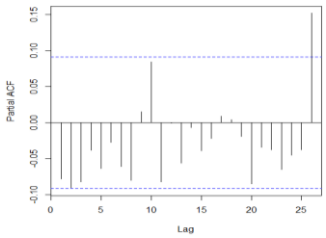

Figure 3: Plot of PACF

Checking the ACF plot above, it can be observed that there are no significant lag from lag 1 to lag 10 and also in the PACF plot there are no significant lag observed except at lag 25. However, in the PACF plot, the lag 2 can be seen to lie on the significance line. Based on this inspection two ARIMA models are proposed which are ARIMA (1,1,0) and ARIMA (1,1,1). The accuracy of these two proposed ARIMA models is further compared with the linear model using AIC, RMSE and MAE in order to select a better model that best fit the series.

MODEL COMPARISON

Model	AIC	RMSE	MAE
ARIMA (1,1,0)	4846.52	45.64065	9.572935
ARIMA (1,1,1)	4834.67	44.99815	12.38841
Linear Model	5755.062	120.2475	100.7301

Table 3: Model Accuracy

Table presents the value of AIC, RMSE and MAE. The value of the measures of accuracy indicates how close the fitted line is to the data point. A smaller value for each of the measure is a desired choice as it shows that data values are disperse closely to its central moment. From the table, it can be observed that the linear model has the highest value in each of the three measures, ARIMA (1,1,0) has the least MAE value while ARIMA (1,1,1) has the least AIC and RMSE value. This result indicates that, ARIMA (1,1,1) is better than the other two model. Thus, ARIMA (1,1,1) is been selected in the study as a better model.

MODEL DIAGNOSTIC

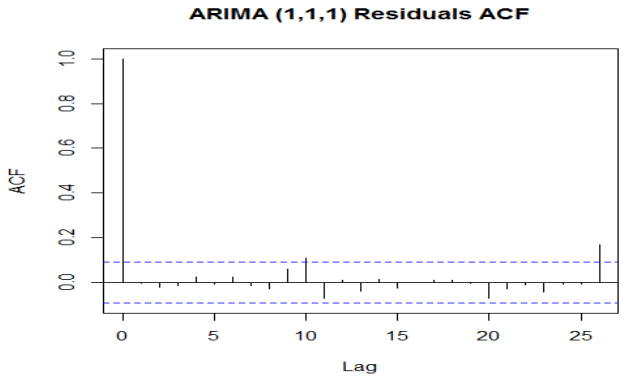

Figure 4: In-sample Residual ACF Plot

Fig. 4 shows the ACF plot of the in-sample forecast error. It can been seen that residuals from lag 1 through lag 9 are insignificant and only two lags were found to be significant afterward. This is an indication of no auto-correlation in the residual. Nevertheless, Box-Ljung test is further carried out to ascertain if the residuals of the model is statistically significant.

TEST FOR SERIAL AUTOCORRELATION

H_0: There is no auto-correlation
H_1: There is auto-correlation

Q*	Df	p-value
8.8503	8	0.3551

Table 4.5: Ljung-Box Test

The decision is to be made at 5% level of significance, thereby, a p-value less than 0.05 is in favor of the H_1 and if otherwise we don't have enough evidence to reject H_0. Since the p-value is greater than 0.05, we fail to reject H_0. This indicates that there is no residuals auto-correlation which indicates that the model good.

V. CONCLUSIONS

The study sought to determine the best model for predicting the trend of COVID-19 cases in Nigeria. The study made use of daily total cases of COVID-19 data obtained from NCDC for 2020-03-01 through 2021-06-06. The time plot of the series presented revealed that there is an upward trend in the recorded cases which indicates that the series is not stationary at level. Augmented Dickey-Fuller test presented in Table 1 shows that the series is stationary at first differencing when tested at 5% level of significance (p-value < 0.05). To determine the model that best predict the series, linear model was estimated and two ARIMA models. In the selection of the ARIMA model ACF and PACF plot were inspected, based on the inspection two ARIMA model were proposed which are, ARIMA (1,1,0) and ARIMA (1,1,1). Subsequently, the three models, linear, ARIMA (1,1,0) and ARIMA (1,1,1) were assessed for accuracy using AIC, RMSE and MAE. ARIMA (1,1,1) was found to have the least value in two of the measures examined, thus, ARIMA (1,1,1) was selected as the better model. Thereafter, serial correlation of the model residual was examined through ACF plot and Ljung-Box test.

It was found from the Ljung-Box test that, there is no serial correlation in the model residuals. This indicated that the ARIMA (1,1,1) is good and better prediction of covid-19 cases in Nigeria.

VI. RECOMMENDATIONS

For the purpose of the study and in line with Aslam, (2020), and others, we hereby recommend based on the findings that; (i) the government should carry out research on a more effective way to treat infected patients of covid-19, (ii) the measures of covid-19 should be strictly enforce by the Nigerian government, (iii) all Nigerian citizens should follow various measures of covid-19 and urge one and another to adhere to the rule in order to reduce the spread of the virus and (iv) the Nigeria Government should constantly give appropriate palliatives to citizen which are being affected by the pandemic.

ACKNOWLEDGEMENTS

The research of Aslam, (2020) and all the referees used in this paper, the paper reviewer of this research, were all acknowledged. Members of academic and Non Academic staff of Osun State College of Technology, Esa-Oke, and the Nigerian Centre for Disease Control body were all acknowledged.

REFERENCES

[1] Alzahrani, S.I., Aljamaan, I.A. & Al-Fakih EA (2020) Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. J Infect Public Health, 13(7):914–9.
[2] Aslam, M. (2020): Using the kalman filter with ARIMA for the covid-19 pandemic dataset of pakistan. Data in Brief, 31, 105854. https://doi.org/10.1016/j.dib.2020.
[3] Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S. & Ciccozzi M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. Data in brief. 31, 105340.
[4] Cucinotta, D.; Vanelli, M. (2020). WHO declares COVID-19 a pandemic Acta Biomed. 91, 157-160.
[5] EBio Medicine. (2020): COVID-19. doi:10.1016/j.ebiom.2020.102966, PMC 7438984.
[6] Gralton, J., Tovey, T.R., McLaws, M.L. & Rawlinson, W.D. (2013). Respiratory Virus RNA is detectable in airborne and droplet particles. J Med Virol, 85, 2151-9.
[7] Huang C, Wang Y, Li X, Ren L, Zhao J. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet, 395:497–506.
[8] Ihekweazu C., Elimian K.O. & Ochu C.L. (2020). Descriptive epidemiology of coronavirus disease 2019 in nigeria, 27 febuary - 6 june 2020. Epidemiol Infect, 148:e208 PMID: 32912370.
[9] Ilie, O.D., Cojocariu, R.O., Ciobica, A., Timofte, S.I., Mavroudis, I. & Doroftei B. (2020). Forecasting the spreading of COVID-19 across nine countries from Europe, Asia, and the American continents using the Arima models. Microorganisms, 8(8), 1158.
[10] Jeong J.H, Kim K.H, Jeong S.H, Park J.W, Lee S.M. (2014). Comparison of sputum and nasopharyngeal swabs for detection of respiratory viruses. J Med Virol., 86, 2122–2127.
[11] NCDC (2021). Covid-19 Nigeria. Retrieved 2020, September, 8 from https://covid19.ncdc.gov.ng
[12] NCDC. (2020). Covid-19 Nigeria. Retrieved from https://covid19.ncdc.gov.ng.
[13] Sahai, A.K., Rath, N., Sood, V. & Singh, M.P. (2020) ARIMA modelling &forecasting of COVID-19 in top five affected countries. Diabetes Metab Syndr Clin Res Rev., 14(5):1419–27.
[14] Statista (2021). Coronavirus cases in Africa as of august 8, 2021, by country. https://www.statista.com/statistics/1170463/coronavirus-cases-in-Africa
[15] WHO. (2020). Advice on the use of masks in the context of COVID-19. Interim. Guidance. Geneva (available at https://www.who.int/publications/i/item/advice-on-the-use-of-masks-in-the- community-during-home-care-and-in-healthcare-settings-in-the-context-of-the-novel-coronavirus-(2019-ncov)-outbreak).: World Health Organization.
[16] Worldometers. (2021): Covid-19 coronavirus pandemic. Retrieved from https://www.worldometers.info/coronavirus/
[17] World Bank (2020). The impact of covid-19 (coronavirus) on global poverty: why sub-Saharan Africa might be the region hardest hit. 2020. Retrieved from https://blogs.worldbank.org/opendata/impact-covid-19-coronavirus-global-poverty-why-sub-saharan-africa-might-be-region-hardest
[18] WMHC. (2020). Wuhan Municipal Health and Health Commission's Briefing on the Current Pneumonia Epidemic Situation in Our City. 2020. http:// www. wuhan. gov. cn/ front/ web/ show Detail/2019123108989.
[19] Wu A, Peng Y, Huang B, Ding X, Wang X. (2020). Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in china. Cell Host Microbe, 27, 325–328.
[20] Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature, 579, 270–273.