*Journal of Computing Research and Innovation (JCRINN) Vol. 7 No. 1 (2022) (pp1-16)*
*https://jcrinn.com : eISSN: 2600-8793 doi: 10.24191/jcrinn.v7i1.258*

# Data Driven Mathematical Models for Forecast of Covid-19 Disease in Nigeria

**Adedire, O.[1,2]*, Sadiku, Y.[2], Adedire, O.O.[3] , Collina, K.[4] , Oladejo, A.O.[5] , Sikiru, G.K.[6]**
*[1] Department of Mathematics, University of Jos, Plateaus State Nigeria*
*[2,4,5,6]Federal College of Forestry, Jos, Plateau State, Nigeria*
*[3]Federal College of Forestry Mechanization, Afaka , Kaduna State, Nigeria*

*Corresponding author: * dharenss@gmail.com*
Received Date: 21 October 2021
Accepted Date: 10 December 2021
Published Date: 30 March 2022

## HIGHLIGHTS

- Two mathematical models are derived from two statistical Autoregressive Integrated Moving Average (ARIMA) models.
- The derived data-driven models provide an alternative approach for forecasting both confirmed daily COVID-19 incidence and total active daily infectious COVID-19 disease in Nigeria.
- The models compete favourably with existing data hence they offer promising nature for forecast and adequate predictions of COVID-19 in Nigeria.

## ABSTRACT

*In this research, two mathematical models are proposed for investigation of laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases using data obtained from Nigeria Centre for Disease Control. Due to the observed patterns in the raw data, the Autoregressive Integrated Moving Average (ARIMA) method is used on the data which covers a period of 521 days (27 February, 2020- 1st August 2021). While diagnostic check of ARIMA (11,1,0) indicate Ljung-BoxQ(18) statistics value of 12.544 with p-value of 0.084, diagnostic check of ARIMA(1, 1, 1) indicate Ljung Box Q(18) statistics value of 22.420 with p-value of 0.130. Furthermore, stationary R- squared values are 0.803 and 0.858 at 95% confidence bound for ARIMA (11, 1, 0) and ARIMA (1, 1, 1) respectively which are indicative of good models. Results from ARIMA (11, 1, 0) forecast show a slightly moderate upward trend in confirmed daily COVID-19 incidence in Nigeria and results from ARIMA (1, 1, 1) indicate significant upward trend in total active daily infectious COVID-19 cases in the Nigerian population. Therefore, the developed models can be adopted by the presidential task force and other agencies in the health sector regarding future vaccination towards prevention of the spread of COVID-19 disease in Nigeria provided that the present general prevailing conditions of disease spread remain fairly the same.*

***Keywords:*** *ARIMA, model, disease, coronavirus, COVID-19, forecast, Nigeria.*

## INTRODUCTION

Coronavirus disease 2019 (COVID - 19) is a fatal disease acquired from acute respiratory syndrome coronavirus 2(SARS-COV 2) which usually targets the human respiratory system (Adhikari et al., 2020). The disease was first discovered in Hubei province in China (Rothana & Byrareddy, 2020). SARS-COV 2 has about two weeks incubation period in the human population consequently allowing infected asymptomatic individuals to travel around the world while unknowingly transmitting the disease to people around them (Lauer et al., 2020; Azooz , 2020). The traveller who became the first index case in Nigeria was an Italian citizen whose diagnosis of COVID- 19 was confirmed after a laboratory test on the 27th of February, 2020 having arrived Nigeria on the 24th February, 2020 aboard Turkish airline from Milan, Italy (NCDC, 2021). Since the confirmation of the first index case, many preventive measures have been introduced by the Federal Government of Nigeria and other stakeholders to mitigate the effect of the spread of the virus. Some of such control measures include formation of Presidential Task Force for the control of the virus in the country (Agbakwuru*, 2020;* Daka*,* 2020), postponement of National Sport Festival *(*Eludini, 2020), suspension of national orientation exercise of National Youth Service Corps for 2020 batch A stream-one indefinitely (Awojulugbe, 2020), placement of travel ban on some countries with high spread of the virus (Ogundele, 2020) and closure of schools (Olaleye, 2020) and restrictions on social gatherings by some state governments (Asishana, 2020; Shobiye, 2020).

Vaccinations of citizens started on 5th March 2021 as part of government efforts to control effects of the spread of COVID-19. About 2,534,205 individuals have been vaccinated with the first dose of COVID -19 vaccines while 1,404,740 have taken the second dose. However, despite the control measures put in place, some stakeholders announced the commencement of the second wave of the disease spread as well as discovery of a new coronavirus variant in Nigeria (Adebowale, 2020). As of August 1, 2021, Nigeria had reported 174,305 laboratories confirmed cumulative COVID -19 cases with 2,149 deaths, 7,161 active daily infectious COVID-19 cases and 165,005 discharged cases after 2,558,969 laboratory sample tests have been conducted nationwide (NCDC, 2021). Some researchers have developed mathematical models to predict the transmission dynamics of COVID-19 in Nigeria. While some studies are centred on Nigeria (Adewole et al., 2021; Daniel, 2020; Irany et al., 2020; Iboi et al., 2020; Olaniyi et al., 2020), other studies are centred on some states within Nigeria with focus on nonlinear deterministic and stochastic models (Adedire and Ndam,2021; Misau et al., 2020; Okuonghae and Omame, 2020). Some other studies have also been conducted on Nigeria outside the scope of deterministic and stochastic models (Amzat et al., 2020; Habib et al., 2021; Ajisegiri, et al., 2020; Agusi et al., 2020; Omaka-Amari et al., 2020; Obi-Ani et al., 2020).

Other studies in literature have used ARIMA models to investigate behaviour of certain time series data. Such studies include prediction of next-day electricity prices using ARIMA model (Contreras et al., 2003), investigation of China's public fiscal revenue with ARIMA model (Liu and Wang, 2015) and Day-ahead electricity price prediction with wavelet transform and ARIMA models ( Conejo et al ., 2005). Another research relative to application of ARlMA models include the study of Prybutok et al. (2000) where they compared ARIMA and regression models with neural networks for prediction of Houston's daily concentration. While the work of Ho et al. (2002) considered a study which comparatively investigate neural network and Box-Jenkins ARIMA modelling in time series prediction, the study of Melard (1984) considered a fast algorithm for the exact likelihood of autoregressive moving average models and that of Simon (2007), Jadevicius and Huston (2015), Abu and Rosbi (2017) used ARIMA models in their studies. From the foregoing research, none of the studies to date in literature has considered the Autoregressive Integrated Moving Average (ARIMA) method on Nigeria for modelling forecasts of laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases.

The motivation for this study centres on the need to have alternative means of generating data for the forecast of laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases in Nigeria. Considering the high cost of carrying out contact tracing, laboratory tests and quarantine of infected individuals, the models will be beneficial as tools to forecast future values of laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases in Nigeria. Another benefit that could be derived from the model is the possibility of using the models as checks to any confirmed cases which could suggest further study relative to control measures whenever sharp deviations from the forecast of the proposed models are observed in Nigeria.

In this study, we develop mathematical models using ARIMA methods for daily forecast of laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases in Nigeria. The remaining part of this study is organised as follows: Methodology is presented in section 2 while section 3 centres on Results and Discussion. Conclusion comes up in section 4.

## METHODOLOGY

### Study Area

The study area for this study is the Federal Republic of Nigeria, a country in West Africa with a coast along the Atlantic Ocean on the Gulf of Guinea. It consists of 36 states and Federal Capital Territory (FCT) and has average range of high temperature between 25°C – 28°C all year round with Jos Plateau State having average temperature of 22°C. It is the most populated country in Africa with an estimated population of about 193,392,517 as at 2016 population census estimate (National Bureau of Statistics, 2017).

There are many groups which contribute to the rich culture of Nigeria but disagreements among such ethnic groups often constitute major challenges to nation building and administration of vaccines. Nigeria is located in the tropics of Africa with land area of 923,768 square kilometres and has latitude of 9.0820°N and a longitude of 8.6753°E showing that it is both in the northern and the eastern hemispheres of the world map.

### Epidemic Data

The data used for this study was obtained from the records of the Nigeria Centre for Disease Control (NCDC). Part of data made available to the public consists of laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases in Nigeria. In the present study, data on reported laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases in Nigeria are considered for the period between 27$^{th}$ February, 2020 and 1$^{st}$ August, 2021 representing 521 days of active daily data points adequate for time series analysis.

### Model Specification

For the time series secondary data on reported daily COVID-19 disease, the Autoregressive Integrated Moving Average is adopted (ARIMA). With this specification, the time series data are described by their lagged values and stochastic error terms. ARIMA is one of the parametric approaches which uses linear relationships to describe current behaviour with their past values. It should be noted that the ARIMA method is a classical method that is covered by common texts on time series analysis (Brockwell and Davies, 2003; Tsay 2005; Wei, 2005).

Consequently, this study does not intend to duplicate here well-documented method but to give a practical representation of its application in science for derivation of simple mathematical models capable of predicting COVID-19 data for Nigeria. Of course, the approach could be used on time series data for effective planning towards vaccination of the citizens for disease eradication and national development.

### Definition 1

Let $\boldsymbol{A} = \boldsymbol{Y}_{\boldsymbol{t-1}}, \boldsymbol{Y}_{\boldsymbol{t-2}}, \dots, \boldsymbol{Y}_{\boldsymbol{t-n}}$ be sequence of some functions $\boldsymbol{Y}_{\boldsymbol{t-i}}$, $[\boldsymbol{t}, \boldsymbol{i} \in \boldsymbol{N}]$ and let $\boldsymbol{B} = \sum_{i=1}^{n} \boldsymbol{\varphi}_{\boldsymbol{i}} \boldsymbol{Y}_{\boldsymbol{t-i}}$ be a series of the sequence $\boldsymbol{A}$ for $\boldsymbol{\varphi} \boldsymbol{\epsilon} \boldsymbol{Z}$, then a property displayed by A of adjacent items not being independent of each other is the autocorrelation of B.

### Definition 2

Let $\boldsymbol{p}, \boldsymbol{d}, \boldsymbol{q} \boldsymbol{\epsilon} \{\boldsymbol{0}, \boldsymbol{Z}^{+}\}$be order of autoregression, degree of difference and order of moving average respectively. ARIMA (p, d, q) is an autoregressive (AR) integrated (I) moving average (MA) method for fitting time series data to better comprehend the data and predict future points in the series (Box et al., 2003; Rojas and Pomares, 2016).

### Definition 3

Let p represent the order of the autoregressive process for both laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases in Nigeria for non-seasonal components, d represents the order of stationarity and q stands for moving average process of the reported cases where $\boldsymbol{p}, \boldsymbol{d}, \boldsymbol{q} \boldsymbol{\epsilon} \{\boldsymbol{0}, \boldsymbol{Z}^{+}\}$. Choose

$$\text{ARIMA } (p, d, q) \tag{1}$$

and let $\boldsymbol{Y}_{\boldsymbol{t}} \boldsymbol{\epsilon} \boldsymbol{R}$ be time series secondary data on reported daily COVID-19 for an integer index t, then the generalised form of ARIMA (p, d, q) for a series $Y_t$ with seasonal component can be found in (Milenkovic et al, 2020) with more parameters, but without seasonal component could be written as

$$\emptyset_p (B) \Delta\ Y_t = \theta_q (B) \varepsilon_t \tag{2}$$

where $\Delta = (1 - B)^d$ is a non seasonal differencing operator, $\varepsilon_t$ is the residual or error terms and B is a backshift operator such that

$$\emptyset_p(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p \tag{3}$$

$$\theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q \tag{4}$$

Substitution (3) and (4) into (2) with $\Delta = (1 - B)^d$ gives

$$(1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p)\,(1 - B)^d\ Y_t = (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q)\,\varepsilon_t \tag{5}$$

which in compact form becomes

$$(1-\sum_{i=1}^{p} \emptyset_i B^i)\,(1-B)^d\; Y_t = (1+\sum_{i=1}^{q} \theta_i B^i)\;\varepsilon_t$$

(6)

**Definition 4**

Let $y_i$ , $i = 1,2,\dots,N$ be observed data, $N$ is the number of observations, $\theta$ is the number of predictors for some model M, then Bayesian Information criterion defined in terms of Mean Squared Error (MSE) is given by (Yaffee, 2000)

$$BIC = N\ ln\ (MSE) + (\theta+1)\ ln\ (N) \tag{7}$$

and

$$Normalised\ \ BIC = N\ ln\ (MSE) + (\theta+1)\ \frac{ln\ (N)}{N} \tag{8}$$

where $\theta+1$ is the number of total parameters including coefficients and intercept of the model M.

**Definition 5**

Let $N$ be number of data points or sample size, $y_i$ , $(i = 1,2,\dots,N)$ be actual measurements at some point $i$ and $\underline{y_i}$ be corresponding prediction for $y_i$ , then the Root Mean Square Error (RMSE) is given by (Hyndman, 2006; Pontius et al., 2008)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left|y_i - \underline{y_i}\right|^2}{N}} \tag{9}$$

the Mean Absolute Error (MAE) is given by (Willmot, 2005)

$$MAE = \frac{\sum_{i=1}^{N} \left|y_i - \underline{y_i}\right|}{N} \tag{10}$$

and the Mean Absolute Percentage Error (MAPE) is given by (Hyndman, 2006; Kim and Heeyoung, 2016; Makridakis, 1993)

$$MAPE = \frac{100\,\%}{N} \sum_{i=1}^{N} \left|\frac{y_i - \underline{y_i}}{y_i}\right| \tag{11}$$

**Definition 6**

Let $n$ be sample size, N time lags under consideration for sample autocorrelation $r$ at lag L. Then the Ljung- Box test statistics Q is defined by Ljung and Box (1978)

$$Q(N) = n(\,n+2\,)\sum_{L=1}^{N} \frac{r^2}{n\ -\ L}$$

(12)

*Journal of Computing Research and Innovation (JCRINN) Vol. 7 No. 1 (2022) (pp1-16)*
https://jcrinn.com : *eISSN: 2600-8793 doi: 10.24191/jcrinn.v7i1.258*

for a null hypothesis $H_o$ which states that data are independently distributed and an alternate hypothesis $H_1$ which states that data are not independently distributed.

## FINDINGS AND DISCUSSIONS

The null hypothesis tested in this study is that there are no significant errors in the use of ARIMA models investigated for the data on laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases in Nigeria. However, the justification for the choice of ARIMA (11, 1, 0) and ARIMA (1, 1, 1) is centred on diagnostic checks performed on the raw data whose results are shown in Figures and Tables displayed in subsequent paragraphs. It can be seen from Table 1 that p values (0.084) and (0.130) for ARIMA (11, 1, 0) and ARIMA (1, 1, 1) respectively are greater than $\propto = 0.05$ level of significance. This indicates that there is no statistical evidence to reject the null hypothesis. Hence, there is no significant errors in the use of ARIMA (11, 1, 0) and ARIMA (1, 1, 1) models respectively on laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases in Nigeria which justifies their use in this study.

The residual autocorrelation function (ACF) and partial autocorrelation function (PACF) are shown for ARIMA (11, 1, 0) and ARIMA (1, 1, 1)

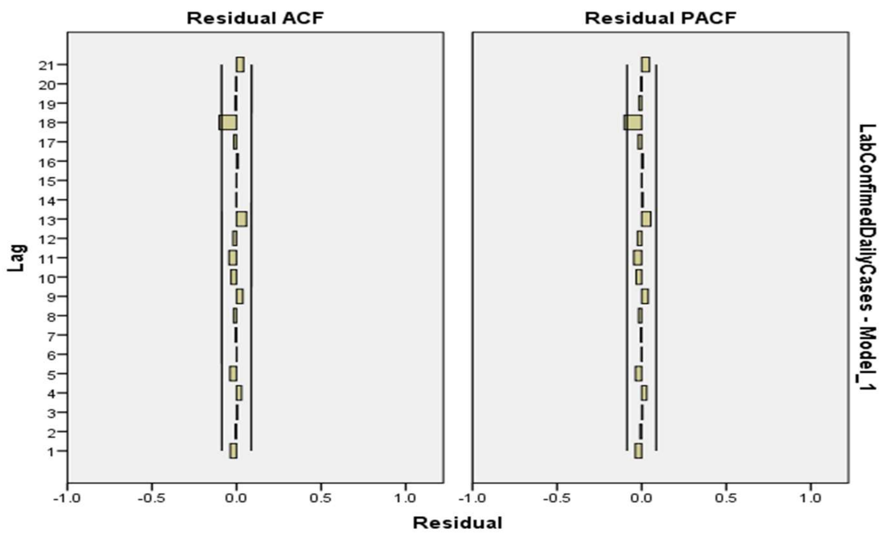


**Figure 1:** Residual ACF and PACF for laboratory confirmed daily COVID-19 disease incidence in Nigeria (27th Feb 2020 - 1st Aug 2021) based on ARIMA (11, 1, 0) at 95% confidence limit.

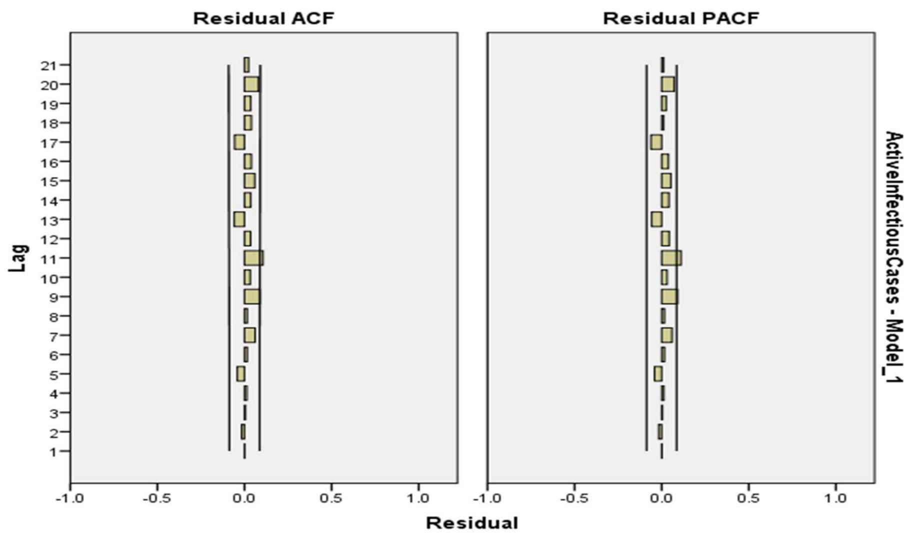


**Figure 2:** Residual ACF and PACF for total active daily infectious COVID-19 cases in Nigeria (27th Feb 2020 - 1st Aug 2021) based on ARIMA (1, 1, 1) at 95% confidence limit.

While residual ACF and PACF for laboratory confirmed daily COVID-19 disease incidence are shown in Figure 1, total active daily infectious COVID-19 cases in Nigeria have their ACF and PACF represented in Figure 2. The results from Figures 1 and 2 indicate differenced time series for d = 1 in ARIMA which has the structure of the form ARIMA (p, 1, q).

From Figure 1, ACF and PACF of the differenced time series shows only one very significant contribution of residual at eighteenth lag at 95% confidence limit. This means that there are not many significant autocorrelations except at lag 18 for laboratory confirmed cases of COVID-19 in Nigeria. Also from Figure 2, the only significant autocorrelations exist at lags 9, 11 and 20 for ACF and at lags 9 and 11 only for PACF at 95% confidence limit indicated by vertical lines. Since many autocorrelations do not exist in Figures 1 and 2, then the existing autocorrelations would be shown in Table 1 by Ljung Box statistics that they are not significant enough to cause inaccurate results and might have existed due to some random errors synonymous with stochastic processes.

It should be noted that substitution of $p = 0, 1, 2, ..., 10$, $d = 1$ and $q = 0$ into eq.(1) produced set of ARIMA methods whose diagnostics check do not satisfy assumptions of good fitted model for laboratory confirmed daily COVID-19 disease incidence. However, substitution of $p = 11$, $d = 1$ and $q = 0$ into eq. (1) gives ARIMA( 11, 1, 0) whose diagnostics checks give the Ljung- Box Q(18) statistics $= 12.544$, and p-value $> 0.05$ which showed that the residuals are white noise hence there is no significant autocorrelation in the differenced series of laboratory confirmed daily COVID-19 disease incidence as indicated in Table 1 at 95% confidence limits represented by vertical lines shown in Figure 1.

**Table 1: Ljung-Box statistics of the ARIMA (11, 1, 0) and ARIMA (1, 1, 1) models of COVID-19 disease in Nigeria.**

| Model No | Ljung-Box(18) | | |
|---|---|---|---|
| | Statistics | DF | p-value |
| **ARIMA (11, 1, 0)** | 12.544 | 7 | 0.084 |
| **ARIMA (1, 1, 1)** | 22.420 | 16 | 0.130 |

For total active daily infectious COVID-19 cases in Nigeria, substitution of p = 1, d = 1 and q = 1 into eq. (1) gives ARIMA (1, 1, 1) whose diagnostics checks give Ljung- Box Q (18) = 22.420 and p-value > 0.05 which indicate well fitted model. Hence ARIMA (1, 1, 1) meets the standard assumption that the residuals of the model are independent, and no significant autocorrelations exist in the model. This means that any existing autocorrelations in the model are merely caused by random errors.

It should be noted further that for any given number of lags N that Ljung-Box Q (18) helps to test the hypothesis that the residuals from the ARIMA models have no autocorrelation as similarly stated in Definition 6 in the context of data being independently distributed. This suggests in simple terms that acceptance of null hypothesis indicates that the model performs well and does not show lack of fit.

The choice of p = 11, d = 1 and q = 0 produced ARIMA (11, 1, 0) whose analysis for the data gives normalised Bayesian Information Criterion (BIC) of 9.129, Root Mean Square Error (RMSE) of 78.195, Mean Absolutely Error (MAE) of 53.250, Mean Absolute Percentage Error (MAPE) of 40.328 based on the formulae displayed in Definitions 4 and 5.

**Table 2: Fit statistics of ARIMA (11, 1, 0) and ARIMA (1, 1, 1) models of COVID-19 disease in Nigeria.**

| Model | Stationary R-squared | R-Squared | RMSE | MAE | MAPE | normalised (BIC) |
|---|---|---|---|---|---|---|
| **ARIMA (11, 1, 0)** | 0.803 | 0.963 | 78.195 | 53.250 | 40.328 | 9.129 |
| **ARIMA (1, 1, 1)** | 0.858 | 0.999 | 250.021 | 152.108 | 15.517 | 11.381 |

Also for p = 1, d = 1 and q = 1, the fit statistics for ARIMA (1, 1, 1) has normalized Bayesian Information Criterion (BIC) of 11.381, Root Mean Square Error (RMSE) of 250.021, Mean Absolute Error (MAE) of 152.108, Mean Absolute Percentage Error (MAPE) of 15.517 as shown in Table 2.

While Table 2 indicates that 80.3% of variance (stationary R- squared) of the laboratory confirmed daily COVID-19 disease incidence was covered by ARIMA (11, 1, 0) model, it also shows that 85.8% of variance (stationary R- squared) of the total active daily infectious COVID-19 cases in Nigeria was covered by ARIMA (1, 1, 1). The total active daily infectious COVID-19 cases are the total population of people that are capable of transmitting the COVID- 19 disease to other people around them.

Results indicated in Table 3 are for the coefficients of ARIMA (11, 1, 0) model which govern equation for predicting the laboratory confirmed daily COVID-19 disease incidence in Nigeria.

**Table 3: Coefficients of ARIMA (11, 1, 0) model for the governing equation of laboratory confirmed daily COVID-19 cases in Nigeria.**

| ARIMA(11, 1, 0) | Coefficients | SE | t | p-value |
|---|---|---|---|---|
| **Constant** | 0.198 | 2.291 | 0.086 | 0.931 |
| **AR lag 1** | -0.513 | 0.039 | -12.979 | 0.000 |
| **lag 2** | -0.452 | 0.042 | -10.641 | 0.000 |
| **lag 3** | -0.243 | 0.047 | -5.204 | 0.000 |
| **lag 4** | -0.271 | 0.045 | -6.014 | 0.000 |
| **lag 5** | -0.074 | 0.044 | -1.693 | 0.091 |
| **lag 6** | -0.085 | 0.044 | -1.930 | 0.054 |
| **lag 7** | 0.276 | 0.045 | 6.104 | 0.000 |
| **lag 8** | 0.440 | 0.047 | 9.414 | 0.000 |
| **lag 9** | 0.194 | 0.049 | 3.986 | 0.000 |
| **lag 10** | 0.108 | 0.042 | 2.554 | 0.011 |
| **lag 11** | 0.107 | 0.039 | 2.759 | 0.006 |

Analogously, Table 4 shows the coefficients of ARIMA (1, 1, 1) which govern the equation predicting the total active daily infectious COVID-19 cases in the Nigerian population.

**Table 4: Coefficients of ARIMA (1, 1, 1) model for the governing equation of total active daily infectious COVID-19 cases in Nigeria.**

| ARIMA(1, 1, 1) | Coefficients | SE | t | p-value |
|---|---|---|---|---|
| **AR lag 1** | 0.167 | 0.081 | 2.055 | 0.042 |
| **MA lag 1** | 0.918 | 0.126 | 7.281 | 0.000 |

Comparison of the fitted models with actual raw data of laboratory confirmed daily COVID-19 disease incidence is shown in Figure 3.

*Journal of Computing Research and Innovation (JCRINN) Vol. 7 No. 1 (2022) (pp1-16)*
*https://jcrinn.com : eISSN: 2600-8793 doi: 10.24191/jcrinn.v7i1.258*

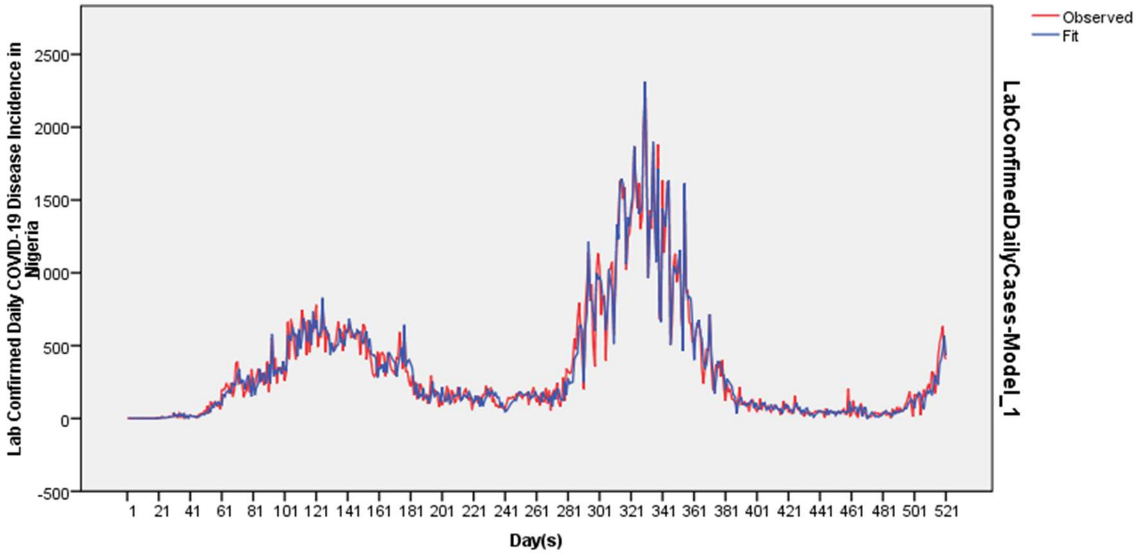


**Figure 3:** Comparison of ARIMA model fit with observed values of lab confirmed daily COVID-19 disease incidence (521 days) based on ARIMA (11, 1, 0).

Comparison of the fitted models with total active daily infectious COVID-19 cases is displayed in Figure 4.

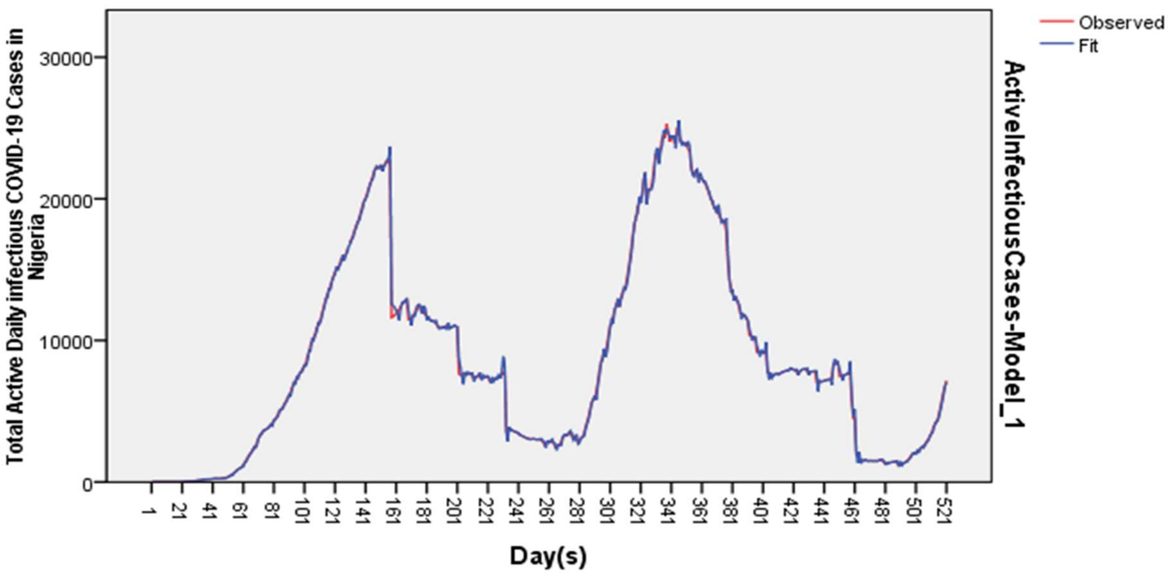


**Figure 4:** Comparison of ARIMA model fit with observed value of total active daily infectious COVID-19 cases (521 days) based on ARIMA (1, 1, 1)

Results from Figures 3 and 4 show similarity between the raw data and the fitted model over a long range of time. This gives assurance that the derived models would meaningfully give adequate estimates when used to predict the future occurrences.

*Journal of Computing Research and Innovation (JCRINN) Vol. 7 No. 1 (2022) (pp1-16)*
*https://jcrinn.com : eISSN: 2600-8793 doi: 10.24191/jcrinn.v7i1.258*

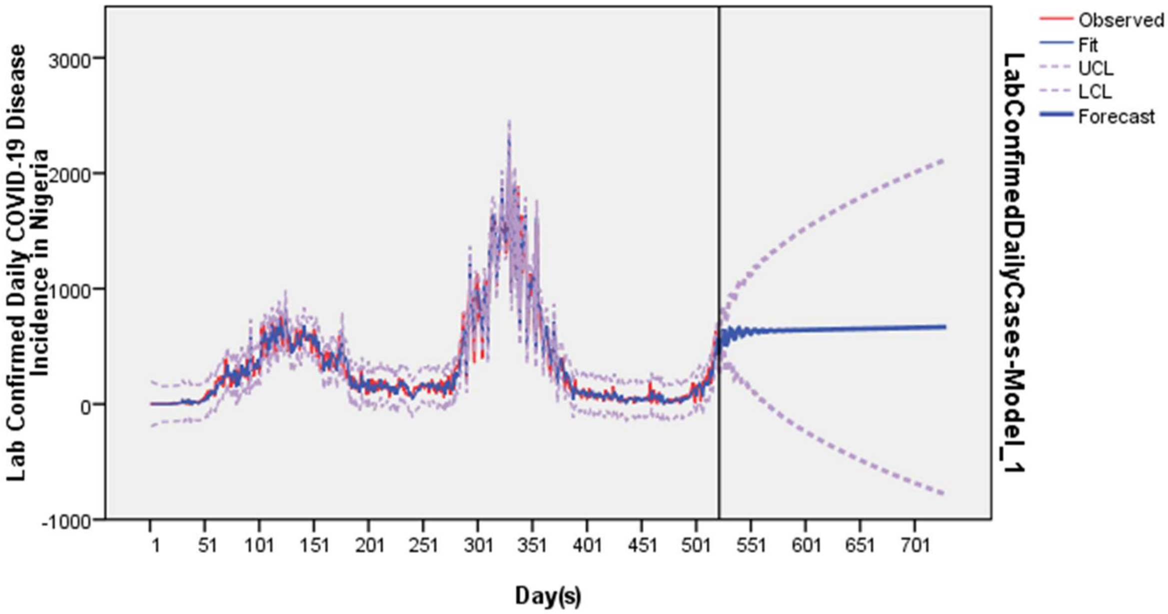


**Figure 5:** ARIMA (11, 1, 0) model chart of confirmed daily COVID-19 disease incidence (521 days) with forecast (522days - 730 days) in Nigeria.

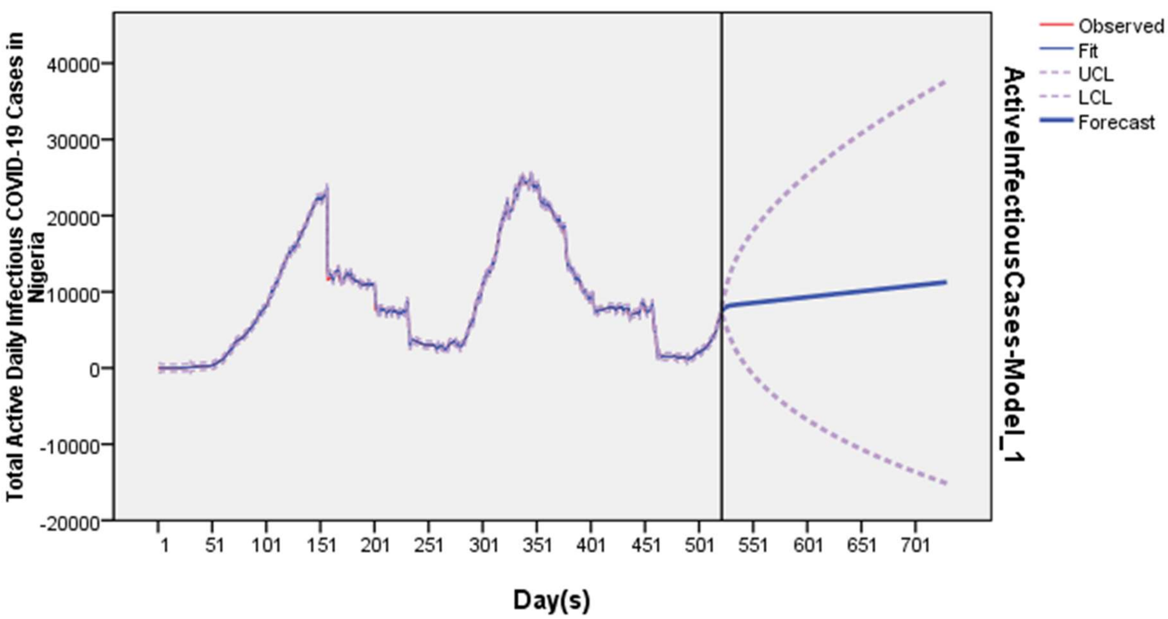


**Figure 6:** ARIMA (1, 1, 1) model chart of total active daily infectious COVID-19 cases (521 days) with forecast (522days - 730 days) in Nigeria

Figures 5 and 6 show a 209 days forecast of laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases in Nigeria.

Substitution of equations (2.3) and (2.4) into (2.2) for ARIMA (11, 1, 0) and use of coefficients listed in Table 3 with the aid of Maple software gives the following mathematical model equation

$$Y_t = 0.198 + 0.487\, Y_{t-1} + 0.061\, Y_{t-2} + 0.209 Y_{t-3} - 0.028\, Y_{t-4} + 0.197\, Y_{t-5} - 0.011\, Y_{t-6} + 0.361\, Y_{t-7} + 0.164\, Y_{t-8} - 0.246\, Y_{t-9} - 0.086\, Y_{t-10} - 0.001\, Y_{t-11} - 0.107\, Y_{t-12} + \varepsilon_t \quad (7)$$

for the forecast of laboratory confirmed daily COVID-19 disease incidence in Nigeria.

Analogously for ARIMA (1,1,1), substitution of (3) and (4) into (2) ) and use of coefficients listed in Table 4 with the aid of Maple software gives the following mathematical model equation

$$Y_t = 1.167\, Y_{t-1} - 0.167\, Y_{t-2} + \varepsilon_t + 0.918\, \varepsilon_t \quad (8)$$

for total active daily infectious COVID-19 cases in Nigeria.

## CONCLUSION AND RECOMMENDATIONS

In this paper, two mathematical models are derived from ARIMA (11, 1, 0) and ARIMA(1, 1, 1) methods with targets on forecast of laboratory confirmed daily COVID-19 disease incidence and total active daily infectious COVID-19 cases in Nigeria respectively for data which covered a period of 521 days. Various checks on the derived models indicate that they are appropriate for modelling generation of COVID-19 data.

Results from a mathematical model obtained from ARIMA (11, 1, 0) method indicate an upward trend in laboratory confirmed daily COVID-19 disease incidence in the Nigerian population. Also, a significant upward trend is observed in total active daily infectious COVID-19 cases in Nigeria from a mathematical model obtained from ARIMA (1, 1, 1). Thus, forecasting using these models is recommended as valuable decision support tools for the presidential task force and other stakeholders in the health sector regarding future vaccination towards prevention of the spread of COVID -19 disease in Nigeria.

## ACKNOWLEDGEMENT

The authors appreciate the reviewers for their contributions towards improving the quality of this research.

## CONFLICT OF INTEREST DISCLOSURE

All authors declare that they have no conflicts of interest to disclose.

## REFERENCES

Abu, B.N., and Rosbi, S. (2017). Data clustering using autoregressive integrated moving average (ARIMA) model for Islamic country currency: An econometrics method for Islamic financial engineering. The international Journal of Engineering and Sciences, 6 (6), 22-31.

Adebowale, N. (2021). "Updated: COVID-19 variant, causing anxiety in UK, found in Nigeria – Official". The Premium Times. Retrieved January 29, 2021, from https://www.premiumtimesng.com/news/top-news/438652-covid-19-variant-causing-anxiety-in-uk-found-in-nigeria-official.html

Adedire, O., Ndam, J. N. (2021). A model of dual latency compartments for the transmission dynamics of COVID-19 in Oyo State, Nigeria. Journal of Engineering and Applied Science Letters, 4(1), 1-13, doi:10.30538/psrp-easl2021.0056.

Adewole, M.O., Onifade, A. A.,·Abdullah, F.A., Kasali, F·, Ismail, A. I. M.(2021). Modeling the Dynamics of COVID-19 in Nigeria. Int.J.Appl.Comput.Math, 7:67, 2-25, https://doi.org/10.1007/s40819-021-01014-5.

Adhikari, S. P., Meng, S., Wu, Y. J., Mao, Y. P., Ye, R. X., Wang, Q. Z., Sun, C., Sylvia, S., Rozelle, S., Raat, H., and Zhou, H. (2020). Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review, Infectious Diseases of Poverty 9:29

Agbakwuru, J. (2020). "Buhari sets up 12 member Task Force to control Coronavirus". Vanguard Newspaper. Retrieved March 11, 2020, from https://www.vanguardngr.com/2020/03/buhari-sets-up-12-member-task-force-to-control-coronavirus

Agusi, E.R., Ijoma, S.I., Nnochin, C.S.,, Njoku-Achu, N.O., Nwosuh, C.I., Meseko, C.A.(2020). The COVID-19 pandemic and social distancing in Nigeria: ignorance or defiance. Pan Africa Medical Journal. 2020;35(2):52. ISSN: 1937-8688, doi: 10.11604/pamj.2020.35.2.23649

Ajisegiri, W.S, Odusanya, O.O & Joshi, R. (2020). COVID-19 Outbreak Situation in Nigeria and the Need for Effective Engagement of Community Health Workers for Epidemic Response. Global Biosecurity, 2020; 1(4).

Amzat, J., Aminu, K., Kolo, V.I., Akinyele, A.A., Ogundairo, J.A., Danjibo, M.C. (2020). Coronavirus outbreak in Nigeria: Burden and socio-medical response during the first 100 days. International Journal of Infectious Diseases 98 (2020), 218–224,. https://doi.org/10.1016/j.ijid.2020.06.067

Asishana, J. (2020). "BREAKING: Niger declares lockdown over COVID-19". The Nation Newspaper. Retrieved 23 March 2020, *https://thenationonlineng.net/breaking-niger-declares-lockdown-over-covid-19*

Awojulugbe, O. (2020). "NYSC shuts orientation camps over coronavirus". The Cable. Retrieved 18 March 2020. *https://www.thecable.ng/breaking-nysc-shuts-orientation-camps-over-coronavirus-fears*.

Azooz, A. A. (2020). Empirical parameterization of COVID-19 virus pandemic, Journal of Multidisciplinary Modeling and Optimization 3(1), 12-26.

Box, G.E.P., Jenkins G.M. and Reinsel, G.C. (2020). Time series analysis: Forecasting and control, John Wiley and Sons, New York, 2013.

Brockwell, P.J. and Davies, R.A. (2003). Introduction to time series and forecasting, Springer, New York, 2003.

Daka, F. T. (10 March 2020). "Buhari names task force on coronavirus". The Guardian Newspaper (Nigeria). Retrieved 11 March 2020. https://guardian.ng/news/buhari-names-task-force-on-coronavirus.

Daniel, D.O. (2020). Mathematical model for the transmission of COVID-19 with nonlinear forces of infection and the need for prevention measure in Nigeria. Journal Infectious Disease Epidemiology, 6(5) : 158, 1-12, doi.org/10.23937/2474-3658/1510158.

Eludini, T. (2020). "Nigeria postpones National Sports Festival over coronavirus". Premium Times. Retrieved 18 March 2020, https://www.premiumtimesng.com/news/headlines/382312-just-in-nigeria-postpones-national-sports-festival-over-coronavirus.html.

Habib, M.A., Dayyab, F.M., Iliyasu, G., Habib, A.G. (2021). Knowledge, attitude and practice survey of COVID-19 pandemic in Northern Nigeria. PLoS ONE 16(1) : e0245176, https://doi.org/10.1371/journal.pone.0245176.

Hyndman, R.J. and Koehler, A.B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4): 679–688. doi: 10.1016/j.ijforecast.2006.03.001.

Iboi, E.A., Sharomi, O., Ngonghala., C.N., and Gumel, A.B. (2020). Mathematical modelling and analysis of COVID-19 pandemic in Nigeria. Mathematical Biosciences and Engineering, 17(6): 7192–7220. DOI: 10.3934/mbe.2020369

Irany, F.A., Akwafuo, S.E., Abah, T., Mikler, A.R., (2020). Estimating the Transmission Risk of COVID-19 in Nigeria: A Mathematical Modelling Approach. J Health Care and Research. 05;1(3):135-43

Jadevicius A., and Huston, S. (2015). ARIMA modelling of Lithuanian house price index. International Journal of Housing Markets and Analysis, 8(1), 135-147.

Kim, S., and Heeyoung, K. (2016). A new metric of absolute percentage error for intermittent demand forecast. International Journal of Forecasting, 32 (3): 669-679

Lauer, S. A., Grantz, K. H. Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., Azman, A. S., Reich, N. G., Lessler, J. (2020). The incubation period of Coronavirus Disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application, Annals of Internal Medicine 172(9), 577-582

Ljung, G. M., Box, G. E. P. (1978). On a measure of a lack of fit in time series models. Biometrika, 297-303, doi: 10.1093 /biomet/65.2.297

Makridakis, S. (1993). Accuracy measures: theoretical and practical concerns. International Journal Forecasting, 9 (4): 527-529, doi:10. 10.1016/0169-2070(93) 90079-3.

Milenkovic, M., Svadlenka, L., Melichar, V., Bojović, N., Avramovic, Z. (2018). SARIMA modelling approach for railway passenger flow forecasting. Transport, 2018, 33(5): 1113–1120, ISSN 1648-4142 / eISSN 1648-3480, doi:10.3846/16484142.2016.1139623.

Misau, Y.A., Nansak, N., Maigoro, A., Malami, S., Mogere, D., Mbaruk, S., Mohammed, R., Lawal, S., Usman, S.U. (2020). Mathematical modelling of COVID-19 transmission and control strategies in the population of Bauchi State, Nigeria. Annals of African Medical Research 2020; 3:120 doi:10.4081/aamr.2020.120

National Bureau of Statistics (2017). Demographic statistics bulletin, Federal Republic of Nigeria. Vol 3, Retrieved September 01, 2020, from www.nigerianstat.gov.ng.

Nigeria Centre for Disease Control [NCDC], (2021). An update of COVID-19 outbreak in Nigeria. Retrieved April, 22, 2021 from https://ncdc.gov.ng/diseases/sitreps/?cat=14&name=An%20update%20of%20COVID-19%20outbreak%20in%20Nigeria

Obi-Ani, N.A., Anikwenze, C. & Isiani, M.C. (2020) Social media and the Covid-19 pandemic: Observations from Nigeria. Cogent Arts & Humanities, 7:1, 1799483, doi: 10.1080/23311983.2020.1799483

Ogundele, K. (2020). "UPDATED: FG places travel ban on China, Italy, US, UK, nine others". The Punch Newspaper. Retrieved March 18, 2020, from *https://punchng.com/breaking-fg-places-travel-ban-on-china-italy-us-uk-others*

Okuonghae, D., and Omame, A. (2020). Analysis of a mathematical model for COVID-19 population dynamics in Lagos, Nigeria. Chaos, Solitons & Fractals, 139(6),1-19.

Olaleye, A. (2020). "[UPDATED] Coronavirus: FG orders closure of varsities, schools nationwide". The Punch Newspaper. Retrieved 19 March 2020, https://punchng.com/breaking-coronavirus-fg-orders-closure-of-varsities-schools-nationwide

Olaniyi, S., Obabiyi, O.S., Okosun, K. O., Oladipo, A. T., Adewale, S. O. (2020). Mathematical modelling and optimal cost-effective control of COVID-19 transmission dynamics, European Physical Journal Plus.135, 135:938, 1-20, https://doi.org/10.1140/epjp/s13360-020-00954-z.

Omaka-Amari, L.N., Aleke, C.O., Obande-Ogbuinya, N.E., Ngwakwe, P.C., Nwankwo, O. and Afoke, E.N.(2020). Coronavirus (COVID-19) Pandemic in Nigeria: Preventive and Control Challenges within the First Two Months of Outbreak. African Journal of Reproductive Health June 2020 (Special Edition on COVID-19); 24 (2):87, doi: 10.29063/ajrh2020/v24i2s.13.

Pontius, R., Thontteh, O. and Chen, H. (2008). Components of information for multiple resolution comparison between maps that share are real variable. Environmental Ecological Statistics , 15(2): 111-142, doi: 10.1007/s10651 -007-0043 - y.

Rojas, I., Pomares, H. (2016). Time series analysis and forecasting, Springer, New York.

Rothana, H. A., Byrareddy S.N. (2020). The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. J Autoimmun 2020;109:102433.

Shobiye, H. (2020). "COVID-19: Lagos bans gatherings of over 20". Vanguard Newspaper. Retrieved 21 March 2020, https://www.vanguardngr.com/2020/03/covid-19-lagos-bans-gatherings-of-over-20

Simon, S. K. (2007). A comparison of the forecasting ability of ARIMA models. Journal of Property and Investment and Finance, 25 (3) 223-240.

Tsay, R.S. (2005). Analysis of financial time series, John Wiley and Sons, New York.

Wei, W.W.S. (2005) . Time series analysis: Univariate and multivariate methods, Addison Wesley, New York, 2005.

Willmott, C.J., Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error ((RMSE) in assessing average model performance. Climate Research. 30: 79-82, doi: 10.3354/cr030079.

Yaffee, R.A., McGee, M. (2000). An introduction to time series analysis and forecasting: with applications of SAS and SPSS, Academic Press.