Transactions of the Nigerian Association of Mathematical Physics
***Volume 12, (July – Sept., 2020 Issue*), pp97 – 100**
© Trans. of NAMP

TREND ANALYSIS AND PREDICTION FOR DAILY NEW CONFIRMED COVID-19 CASES IN NIGERIA

S. Suleiman[1] and A. B. Zoramawa[2]

Department of statistics, Usmanu Danfodiyo University, Sokoto

Abstract

The current outbreak of the COVID-19 pandemic across the world, and with an alarming increasing number of cases in Nigeria, government at all levels need to take urgent action to halt the transmission of the deadly virus. Nigeria was first Sub Saharan country to record a case of COVID-19 on 27th February, 2020 after an Italian citizen tested positive for the virus. Two weeks after the index case, COVID-19 cases began relatively increase day by day. In this research, trend of the COVID 19 new cases in Nigeria were analyzed. Linear, Quadratic, Exponential growth and S-curve trend models were employed to predict COVID 19 new cases in Nigeria. The results show COVID 19 new cases in Nigeria are time dependent and models predict it very well in relation time. Forecast evaluation indices indicate that COVID 19 new cases in Nigeria is best predicted by quadratic trend model.

Keywords: COVID 19, Pandemic, time series, trend, prediction

1. Introduction

A time series in general is supposed to be affected by four main components which can be separated from the observed data. These components are Trend, Seasonal, Cyclical and Irregular components. The general tendency of a time series to increase, decrease or stagnate over a long period of time is termed as secular Trend or simply Trend [1, 2].

In time series analysis, it is vital to estimate the trend function because it provides information about the underlying properties of the time series Fuller [3].

The increasing trend curve of Coronavirus confirmed cases against time since December 2019 when it was first reported in Wuhan city of China, has been of serious concern to researchers worldwide.

The Coronavirus also known as (COVID-19) caused by severe acute respiratory syndrome coronavirus 2(SARS-CoV-2), was declared a global pandemic on 11th March, 2020 by World Health Organization (WHO) [4]. According to the situation report of 8th July, 2020 there were **11,669,259** confirmed cases and **539,906** deaths globally [5]. The disease has spread to over two hundred countries in the world and is likely to reach all the countries of the world, thus, understanding its trend movement becomes priority to researchers. Nigeria had 30,249 confirmed cases and 684 deaths according to NCDC as at 8th July, 2020. Actually, this was the Nigerian situation report of 8th July, 2020 [6].

The ability to identify the speed by which the pandemic is spreading is critical in the fight against the virus. Governments can be guided in public health planning and policy-shaping in trying to address the consequences of the pandemic when the rate at which virus spread at any given point in time is determined [7, 8, 9]

In Nigeria, Covid-19 confirmed cases showed an initial gradual rise from the first reported case on the 27th February, 2020up to the second week in April when number of COVID-19 began to relatively sharper increase and this trend has continued till date [10]

Recently COVID 19 infections were modeled using both linear and non -linear models. Novel Coronavirus incidences were forecasted by [11] using Nonlinear Autoregressive Artificial Neural Network (NAR). They found that training error in the network prediction which represents its regression and performance is 2.65 % in cases forecast and 3.22 % in deaths forecast for global prediction. COVID -19 Prevalence in Italy, Spain and France were estimated by [12]. In his study, Autoregressive integrated moving average (ARIMA) models were developed to predict the epidemiological trend of COVID-19 prevalence of Italy, Spain and France. Several models ARIMA models were formulated with different parameters but ARIMA (0,2,1) ARIMA (1,2,0) and ARIMA (0,2,1) were found to have lowest MAPE and were selected as the best models for Italy, Spain and France respectively. According [13] mathematical models based on various factors and

Correspondence Author: Suleiman S., Email: macostamkd@gmail.com, Tel: +2347038132534

analyses are subject to potential bias and hence proposed an application of ARIMA model prediction on the Johns Hopkins epidemiological data to predict the epidemiological trend of the prevalence and incidence of COVID 19.Linear regression, Multilayer perceptron and Vector auto regression methods were employed by [14] to forecast COVID-19 pandemic in India. They used COVID-19 Kaggle data to anticipate the epidemiological example of the ailment and pace of COVID-2019 cases in India. Anticipated the potential patterns of COVID-19 effects in India dependent on data gathered from Kaggle.

Researchers usually tend to favour linear time series without subjecting the data to different functions and determining the most the approriate model fitting the data. In this research, non-linear trends such as quadratic and exponentials trends were considered in addition to the linear trend.

2. Methodology

Daily new confirmed Covid-19 cases data for this research were obtained from the official COVID-19 site (https://covid19.ncdc.gov.ng/) of Nigeria Centre for Diseases Control (NCDC). It covers the period from February 27, 2020, when the country recorded her first case until July 8, 2020. Since the research considered thenewcovid 19 data, the number of cases at a time t depends on previous time t-1. In this research, the number of new infected cases of Covid 19 cases in Nigeria were forecasted using four different trend analysis models. These four different forecasting models were linear trend model, Quadratic trend model, Exponential growth model and S-curve model which were used to find the best fitted model for daily new cases of covid 19 in Nigeria. These forecasting models as also used by [15,16, 17,18] to forecast different phenomena were given below:

Linear model:

$$Y_t = \beta_0 + (\beta_1 t) + \varepsilon_t \quad (1)$$

Quadratic model:

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t \quad (2)$$

Exponential growth model:

$$Y_t = \beta_0 e^{\beta_0 t} + \varepsilon_t \quad (3)$$

S Curve Model (Pearl-Reed logistic trend model):

$$Y_t = \frac{10^a}{\beta_0 + \beta_1 \times \beta_2^t} + \varepsilon_t \quad (4)$$

Where $Y_t$is the number of the daily new confirmed cases of Covid-19 in Nigeria; t is the trend which determines the tendency of time series data to increase or decrease over time; β_0, β_1 and β_2: Parameters of the model .

Model selection criterion and evaluation indices.

The differences between the value obtained NCDC (observed value) and the predicted values from the four models were compared to determine the accuracy of the four forecasting models used in this study. The mean absolute error (MAE), mean absolute percentage error (MAPE), and the root mean square error (RMSE) were chosenfor the evaluation because they are widely used in combining and selecting forecasts for measuring bias and accuracy of models [19].

$$MAE = n^{-1} \sum_{t=1}^{n} |y_t - \hat{y}_t| \quad (5)$$

$$MAPE = n^{-1} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{\hat{y}_t} \right| * 100 \quad (6)$$

$$RMSE = (n^{-1} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2)^{1/2} \quad (7)$$

Smaller values of all these indices determine a good fitted model with minimum forecasting errors [15,20, 21].

Table 1: Linear Trend Model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-134.8847	16.76690	-8.044706	0.0000
@TREND	5.490849	0.219593	25.00467	0.0000
R-squared	0.826773	Mean dependent var		227.5113
Adjusted R-squared	0.825451	S.D. dependent var		232.7201
S.E. of regression	97.22836	Akaike info criterion		12.00693
Sum squared resid	1238389.	Schwarz criterion		12.05039
Log likelihood	-796.4605	Hannan-Quinn criter.		12.02459
F-statistic	625.2335	Durbin-Watson stat		1.107877
Prob(F-statistic)	0.000000			

Table 2: Quadratic Trend Model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-28.30520	21.58537	-1.311315	0.1921
@TREND	0.609343	0.755653	0.806380	0.4215
@TREND^2	0.036981	0.005540	6.674967	0.0000

R-squared	0.870989	Mean dependent var	227.5113
Adjusted R-squared	0.869004	S.D. dependent var	232.7201
S.E. of regression	84.22911	Akaike info criterion	11.72726
Sum squared resid	922290.6	Schwarz criterion	11.79245
Log likelihood	-776.8625	Hannan-Quinn criter.	11.75375
F-statistic	438.8339	Durbin-Watson stat	1.488656
Prob(F-statistic)	0.000000		

Table 3: Growth Curve Trend Model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.057908	0.176355	0.328361	0.7432
@TREND	0.059164	0.002310	25.61553	0.0000

R-squared	0.833578	Mean dependent var	3.962730
Adjusted R-squared	0.832308	S.D. dependent var	2.497306
S.E. of regression	1.022653	Akaike info criterion	2.897602
Sum squared resid	137.0024	Schwarz criterion	2.941066
Log likelihood	-190.6905	Hannan-Quinn criter.	2.915264
F-statistic	656.1556	Durbin-Watson stat	0.746281
Prob(F-statistic)	0.000000		

Table 4: S-Curve model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
- 0.00173717596518925/(0.059163968631232 7+0.036981105440536*0*@TREND^2)	-7748.478	687.2608	-11.27444	0.0000

R-squared	0.762976	Mean dependent var	227.5113
Adjusted R-squared	0.760612	S.D. dependent var	232.7201
S.E. of regression	232.7201	Akaike info criterion	13.74504
Sum squared resid	7148942.	Schwarz criterion	13.76677
Log likelihood	-913.0452	Hannan-Quinn criter.	13.75387
Durbin-Watson stat	0.097263		

Table 5: Models evaluation indices

Forecasting models	Evaluation indices		
	RMSE	MAE	MAPE
Linear trend model	96.49455	77.33719	2758.183
Quadratic trend model	83.27374	59.72137	770.7524
Exponential growth model	452.5286	205.4350	8142.1129
S- curve model	231.8436	197.4862	7614.939

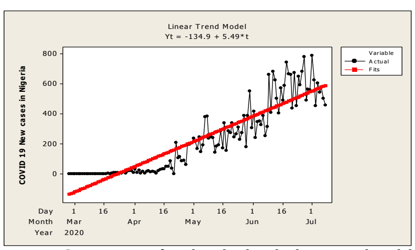

Figure1: Comparisons of predicted values by linear trend model with observed values

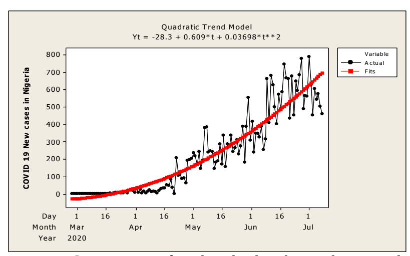

Figure 2: Comparisons of predicted values by quadratic trend model with observed values

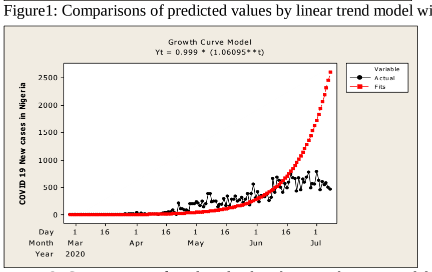

Figure 3: Comparisons of predicted values by growth curve model with observed values

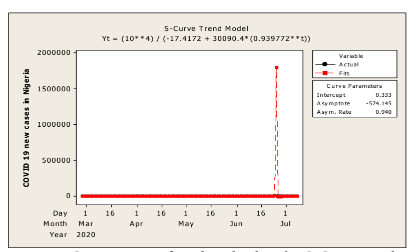

Figure 4: Comparisons of predicted values by S-Curve trend model with observed value

3. Results

The parameters estimation for each model was first done to ascertain their significant values. Tables 1, 2, 3 and 4 respectively presented results for linear, quadratic, exponential and S curve models. The results show high significant trend values for linear, quadratic and exponential (growth curve) and S-curve trend models since their respective p-values are less than 0.1%. High R- squared values indicate that time predicts new COVID-19 cases in Nigeria very well. The performance of the model is determined by the closeness of prediction values for test data to the observed values. Table 5 presented evaluation indices for the models. The results in table 5 indicated that quadratic trend model gives better predictions of the COVID 19 new cases in Nigeria as it gives smaller values for RMSE, MAE and MAPE measures. Figure 1,2,3 and 4 compare the observed values with the predicted values for the models, it can be observed from the figures that quadratic trend model has the best fitted values. In figure 2, the predicted values by the quadric trend model relate closely with the observed i.e. the two lines move closely in the same direction indicating best fit.

4. Conclusion

Based on the analysis carried out and findings made, the result showed there was a general increasing trend for COVID 19 new cases in Nigeria. The four forecasting models namely linear, quadratic, exponential growth and S-curve trend models revealed that COVID 19 new cases in Nigeria were time dependent. The results further revealed that the models can effectively and efficiently forecast COVID 19 new cases in Nigeria. COVID 19 new cases in Nigeria were best predicted by quadratic trend model as it gives smaller values of RMSE, MSE and MAPE performance metrics. The quadratic trend model has therefore minimized the forecasting error and is chosen as the model the data fits.

Reference

[1] Balogun, O. S., Ogunleye, O. M. (2015). On the Time Series Modelling of Crude Oil Exportation in Nigeria, *European Journal of Academic Essays* 2(3): 1-11.

[2] Agatamudi L. R., Bruk S. And Tafere T. (2018).Time Series Analysis Of International Tourist Arrival To Ethioia 2006-2015, A Statistical Analysis, *International Journal of Business Management and Economic Review*,1(6), 117-141

[3] Fuller W. A. (1996), *Introduction to Statistical Time Series*, 2nd Ed., New York, John Wiley and Sons, pp. 475 – 484

[4] Francesco D.G., Damiano P., Claudia M., Mario A., Vincenzo R., Nicola V., and Lee S., (2020) Coronavirus Diseases (COVID-19) Current Status and Future Perspectives: A Narrative Review, *International Journal of Environmental Research and Public Health*

[5] WHO (2020): Coronavirus disease 2019 (COVID-19) situation report, available at https://www.who.int/emergencies/diseases/novel-coronavirus-2019, Accesses July, 8th, 2020

[6] NCDC (2020): http://covid19.ncdc.gov.ng, Accessed July, 8th 2020

[7] Vasilis P., Pantelis L., and Sotiris K. (2020). COVID-19: A Comparison of Time Series Methods to Forecast Percentage of Active Cases per Population, *Appl. Sci*., 10, 3880; doi:10.3390/app10113880

[8] Preiser, W.; Van Zyl, G.; Dramowski, A. (2020). COVID-19: Getting ahead of the epidemic curve by early implementation of social distancing. *S. Afr. Med* J., 110, 1. [CrossRef]

[9] Klompas, M.(2020). Coronavirus Disease 2019 (COVID-19): Protecting hospitals from the invisible. *Ann. Intern. Med*., 172, 619–620. [CrossRef]

[10] Odukoya O., Adeleke I., Jim C., Isikekpei B., Obiodunukwe C., Lesi F., Osibogun A.O., (2020). Evolutionary trends of the COVID-19 epidemic and effectiveness of government interventions in Nigeria: A data-driven analysis, medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20110098

[11] Nouby M. G., Muhammad A. A. and Abd El-Aziz A. A.(2020). Novel Coronavirus Forecasting Model using Nonlinear Autoregressive Artificial Neural Network. *International journal of Advanced Science and Technology*, 29(5),1831-1849.

[12] Ceylan Z., (2020) Estimation of COVID 19 prevalence in Italy, Spain and France. *Science of the Total Environment*, 729, 138817

[13] Domenico B., Marta G., Lazzaro V., Silvia A., and Massimo C. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset, *Data in brief* 29 (2020) 105340

[14] Sujath R., Chatterjee J.M. and Hassanien A. E (2020). A machine learning forecasting model for COVID-19 pandemic in India, *Stochastic Environmental Research and Risk Assessment*, 34:959–972 https://doi.org/10.1007/s00477-020-01827-8

[15] Saleem A., Nasir J., , Muhammad Z. A. and Saleem Z., (2018). Exponential Growth Model for Forecasting of Area and Production of Potato Crop in Pakistan, Pakistan Journal of Agricultural Research, 31(1): 24-28. DOI http://dx.doi.org/10.17582/journal.pjar/2018/31.1.24.28

[16] Imoh U.M., Emmanuel A.A., (2014). Modeling and Forecasting Trend Function of a Discrete - Time Stochastic Process, *Am. J. Sci. Ind. Res*., 5(6): 195-202

[17] Musa B., (2017). Comparison of various models on cancer rate and forecasting, *J. Appl. Sci. Environ. Manage*. 21 (5) 957-959

[18] Reindolf A., Enoch A., Isaac O. and Edmund I. O. (2018). Time series analysis of malaria in Kumasi: Using ARIMA models to forecast future incidence, Cogent Social Sciences, 4:1, 1461544, https://doi.org/10.1080/23311886.2018.1461544

[19] Chai T. and Draxler R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature, *Geosci. Model Dev.,* 7, 1247–1250, doi:10.5194/gmd-7-1247-2014

[20] Yorucu V.(2003).The Analysis of Forecasting Performance by Using Time Series Data for Two Mediterranean Islands. *Review of Social, Economic & Business Studies*, 2, 175-196

[21] Karim R. and Akhter M. (2010). Forecasting of wheat production in Bangladesh. Bangla desh *J. Agric. Res*. 35(1): 17-28. https://doi. org/10.3329/bjar.v35i1.5863