# Forecasting the Confirmed Cases of COVID-19 in Selected West African Countries Using ARIMA Model Technique

Musa Ganaka Kubi[1], Son-Allah Mallaka Philemon[2], Olope Ganiu Ibrahim[3]

[1]*Department of Mathematics and Statistics, Federal Polytechnic, Nasarawa, Nasarawa State, Nigeria*
[2]*Department of Management & Information Technology, ATBU, Bauchi, Nigeria*
[3]*Department of Mathematical Sciences, ATBU, Bauchi, Nigeria*

***Abstract:*** **COVID-19 is a disease caused by the novel coronavirus that was reported in China in 2019. The virus has infected more than a million people globally leading to hundred-thousands of deaths. Hence, forecasting the future confirmed cases to support prevention of the disease and aid in the healthcare service preparation is very important especially in developing countries. In order to support governments' effort in the prevention of COVID-19, we developed an ARIMA model which was used in forecasting future COVID-19 cases in selected West African Countries. The forecasting results from this study indicates an increase cases in the coming days. It is expected that the present prediction models will assist the government and medical personnel in the selected countries to be prepared for the upcoming conditions and have more readiness in healthcare systems.**

***Keywords:*** **COVID-19; Coronavirus; ARIMA; Forecasting; West Africa.**

## I. INTRODUCTION

A novel coronavirus that is related to the Middle East Respiratory Syndrome Virus (MERS-CoV) and Severe Acute Respiratory Syndrome virus (SARS-CoV) was reported in December, 2019 in Journal of Infection [1]. The virus which was first recognized in Wuhan, China is said to have been transmitted from animal (i.e. bat) to humans. The disease caused by this virus is named by the World Health Organization (WHO) as Corona Virus Disease 2019 (Henceforth, Covid-19). The highest world health body also declared the disease as Public Health Emergency of International Concern on January 30th, 2020 [2].

The virus spread from China to other countries of the world as seen in the reported cases in Asian countries, United Kingdom (UK), United States of America (USA), African countries and so on. In West Africa, the virus was first recorded in Nigeria and has spread to all the other 17 West African Countries within one month [3]. Although, the epidemic has not exploded as it was much anticipated in the region, it is no sign for laxity [4]. This is because, the healthcare system in the region are known to be fragile compared to other developed nations. Since the outbreak of the disease, a lot of economic activities were shut down by governments in the region in order to curtail the spread of the virus. Governments took various measures including a total or partial lockdown of the economy.

A search of the literature revealed that researchers have adopted various models to predict Covid-19 spread in countries such as China [2; 5-7], Italy [8-9], France [10], USA [11-13], India [14], and North-Korea [15]. However, studies related to countries in West Africa remains scanty. Furthermore, some of the studies did not use Autoregressive Integrated Moving Average (ARIMA) despite its importance and advantage over other forecasting models such as Wavelet Neutral Network and Support Vector Machine in forecasting natural disasters such as Covid-19 [16].

Forecasting the confirmed cases of COVID-19 will help governments in West Africa to make adequate preparation in their fight the virus in the region. As such, the aim of this study is to forecast the future confirmed cases of COVID-19 in West Africa. Five countries with the highest number of total confirmed cases as at 5th June, 2020 were selected for this study.

## II. METHODOLOGY

*Data Source*

Data for this study was collected from the daily confirmed cases of COVID-19 as announced by the health authorities of countries in West Africa. Five West African countries were selected based on countries with the highest number of confirmed cases in the region. To aid easy collection of the data, World Barometer and Centre for Disease Control daily confirmed cases figure covering a period of 99 days (from 28th February, 2020 to 5th June, 2020) was collected. A time series database was built in a Microsoft Excel sheet designed by the researchers for the five countries.

*ARIMA Models*

For forecasting a time series, ARIMA modeling is found out to be one of the best modeling techniques. Parameter ($p, d, q$) are always used in representing ARIMA models which are expressed as ARIMA ($p, d, q$). In this model, $p$ stands for the order of auto-regression, $d$ signifies the degree of trend difference while $q$ is the order of moving average. We applied an ARIMA model to the time series data of confirmed COVID-19 cases in the five selected West African Countries. Autocorrelation function (ACF) graph and partial autocorrelation (PACF) graph is used to find the initial number of ARIMA models. Two ARIMA models are proposed for each of the Country selected. An essential requirement in ARIMA model forecasting is the test for variance in normality and stationary which was tested in this study and there was no violation of the stationary requirement. We now proposed two models for each of the five selected countries which were checked for model fitting by observing their Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Normalized Bayesian Information Criteria (BIC) values to determine the finest model to forecast. The best model is the one which has the smallest value for all the measures. After fitting the model, its parameters are estimated followed by verification of the model. The best model was employed to forecast confirmed COVID-19 cases for the next 20 days, i.e. 6th June, 2020 to 25 June, 2020. The model for forecasting future confirmed COVID-19 cases is represented as,

$$ARIMA\ (p,d,q): Y_t = \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \ldots (1)$$

Here, $Y_t$ is the predicted number of confirmed COVID-19 cases at t$^{th}$ day, $\emptyset_1, \emptyset_2, \ldots, \emptyset_p$ $and$ $\theta_1, \theta_2, \ldots, \theta_q$ are parameters whereas $\varepsilon_t$ is the residual term for t$^{th}$ day. The pattern of future incidences of COVID-19 was estimated from the previous cases and a time series analysis is performed for this purpose. Time series forecasting refers to the employment of a model to forecast future data based on previously observed data [17]. In the present study, time series analysis is used to recognize the trends in confirmed COVID-19 cases in the five selected Countries in West Africa over the period covering 28 February 2020 to 5th June, 2020 and to predict future cases from 6th June, 2020 to 25th June, 2020. The level of statistical significance was set at 0.05. The SPSS version 23 was used for this study's data analysis.

## III. RESULTS

This study was aimed at developing a model to forecast COVID-19 confirmed cases in West African Countries. The results for measures of model fit for the two proposed ARIMA models in the selected Countries are presented in Table 1. A look at the RMSE, MAE, MAPE and Normalized BIC values suggests that ARIMA (0, 2, 1) model is the most accurate of all for forecasting future incidences of COVID-19 cases in Nigeria, Senegal and Mali. ARIMA (3,2,0) and ARIMA (1,2,2) model is the most accurate of all for forecasting future incidences of COVID-19 cases in Ghana, and Cote D' Ivoire respectively as they possesses the least value for all the measures of model fit.

Hence, the parameters are estimated for the ARIMA (0, 2, 1) model for Nigeria, Senegal and Mali and ARIMA (3,2,0) model for Ghana, while ARIMA (1,2,2) is for Cote D'Ivoire and the result are displayed in Table 2. It was observed that both the AR and MA components of the models have a p-value of 0.000, thus implying that the parameters are significant in the model for all the five selected Countries.

Table 1: Model Fit Measures

| Country | ARIMA models | RMSE | MAPE | MAE | NOR. BIC |
|---|---|---|---|---|---|
| Nigeria | ARIMA(0,2,2) | 55.676 | 138.047 | 35.175 | 8.228 |
| | ARIMA(0,2,1) | 56.014 | 108.451 | 34.019 | 8.193 |
| Ghana | ARIMA(1,2,1) | 141.533 | 32.587 | 81.111 | 10.094 |
| | ARIMA(3,2,0) | 122.135 | 25.704 | 77.166 | 9.846 |
| Senegal | ARIMA(0,2,1) | 18.865 | 10.428 | 11.504 | 6.016 |
| | ARIMA(0,2,2) | 18.949 | 10.482 | 11.528 | 6.072 |
| Cote D'Ivoire | ARIMA(0,2,2) | 26.711 | 22.406 | 18.904 | 6.812 |
| | ARIMA(1,2,2) | 26.530 | 21.796 | 18.537 | 6.792 |
| Mali | ARIMA(0,2,1) | 11.767 | 4.617 | 7.586 | 5.072 |
| | ARIMA(0,2,2) | 11.935 | 4.736 | 7.620 | 5.148 |

Table 2: Parameter Estimate of ARIMA Models

| Country | Type | Coeff. | SE Coeff. | T | P |
|---|---|---|---|---|---|
| Nigeria | Constant | 3.381 | 13079.091 | 0.000 | 0.999 |
| | MA(1) | 0.630 | 0.095 | 6.629 | 0.000 |
| Ghana | Constant | 2.918 | 21656.607 | 0.000 | 0.998 |
| | AR(1) | -1.016 | 0.085 | -11.907 | 0.000 |
| | AR(2) | -0.813 | 0.107 | -7.627 | 0.000 |
| | AR(3) | -0.574 | 0.086 | -6.715 | 0.000 |
| Senegal | Constant | 1.381 | 5224.034 | 0.000 | 0.989 |
| | MA(1) | 0.540 | 0.097 | 5.569 | 0.000 |
| Cote D'Ivoire | Constant | 1.742 | 6806.933 | 0.000 | 0.959 |
| | AR(1) | -0.970 | 0.266 | -3.642 | 0.000 |
| | MA(1) | -0.403 | 00.277 | -1.455 | 0.149 |
| | MA(2) | 0.572 | 0.157 | 3.645 | 0.000 |
| Mali | Constant | 0.247 | 1845.832 | 0.000 | 0.979 |
| | MA(1) | 0.762 | 0.103 | 7.425 | 0.000 |

Hence, the workable model obtained after the substitution of estimated parameters for the five selected Countries are represented below:

For Nigerian, the COVID-19 ARIMA model is presented as follows:

$$Y_t = 3.381 + 0.630\varepsilon_{t-1} + \varepsilon_t \quad \ldots\ldots\ldots\ldots..(2)$$

For Ghana, the COVID-19 ARIMA model is presented as follows:

$$Y_t = 2.918 - 1.016Y_{t-1} - 0.813Y_{t-2} - 0.574 + \varepsilon_t \ldots\ldots..(3)$$

For Senegal, the COVID-19 ARIMA model is presented as follows:

$$Y_t = 0.540\varepsilon_{t-1} + \varepsilon_t \quad \ldots..(4)$$

For Cote D'Ivoire, the COVID-19 ARIMA model is presented as follows:

$$Y_t = 1.742 - 0.970Y_{t-1} - 0.403\varepsilon_{t-1} + 0.572 + \varepsilon_t \; \ldots.(5)$$

For Mali, the COVID-19 ARIMA model is presented as follows:

$$Y_t = 0.247 + 0.762\varepsilon_{t-1} + \varepsilon_t \; \ldots \ldots (6)$$

Table 3: Cumulative figures for forecasted confirmed COVID-19 cases and their lower and upper limits for 20 days (6th June, 2020 to 25th June, 2020)

| Country/ Date | Nigeria | | | Ghana | | | Senegal | | | Cote D'Ivoire | | | Mali | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Forecast | LL | UL | Forecast | LL | UP | Forecast | LL | UP | Forecast | LL | UP | Forecast | LL | UP |
| 06/06/2020 | 12197 | 11988.7 | 12407.9 | 9400.21 | 9160.53 | 9639.89 | 4263.85 | 4226.66 | 4301.04 | 3537.34 | 3487.84 | 3586.84 | 1514.68 | 1493.08 | 1536.28 |
| 07/06/2020 | 12558 | 12206.4 | 12912.6 | 9652.76 | 9313.81 | 9991.72 | 4372.70 | 4309.39 | 4436.00 | 3632.01 | 3550.56 | 3713.45 | 1545.80 | 1513.10 | 1578.49 |
| 08/06/2020 | 12923 | 12421.8 | 13431.8 | 9930.55 | 9488.83 | 10372.27 | 4481.55 | 4390.63 | 4572.47 | 3736.58 | 3631.78 | 3841.38 | 1588.27 | 1545.56 | 1630.99 |
| 09/06/2020 | 13294 | 12630.1 | 13970.6 | 10173.18 | 9620.63 | 10725.73 | 4590.40 | 4469.76 | 4711.04 | 3836.10 | 3706.13 | 3966.06 | 1627.64 | 1575.19 | 1680.09 |
| 10/06/2020 | 13671 | 12830.4 | 14530.4 | 10397.69 | 9674.54 | 11120.84 | 4699.25 | 4546.72 | 4851.77 | 3940.94 | 3787.74 | 4094.13 | 1662.69 | 1600.52 | 1724.86 |
| 11/06/2020 | 14053 | 13022.4 | 15111.9 | 10631.70 | 9761.40 | 11501.99 | 4808.10 | 4621.57 | 4994.62 | 4043.90 | 3866.24 | 4221.56 | 1707.00 | 1635.00 | 1779.01 |
| 12/06/2020 | 14440 | 13206.1 | 15715.5 | 10868.59 | 9843.53 | 11893.65 | 4916.95 | 4694.40 | 5139.49 | 4149.98 | 3948.22 | 4351.73 | 1738.34 | 1656.33 | 1820.36 |
| 13/06/2020 | 14834 | 13381.5 | 16341.3 | 11084.96 | 9898.27 | 12271.65 | 5025.80 | 4765.28 | 5286.31 | 4255.71 | 4028.95 | 4482.46 | 1774.96 | 1678.35 | 1871.58 |
| 14/06/2020 | 15233 | 13548.9 | 16989.6 | 11294.20 | 9924.98 | 12663.42 | 5134.64 | 4834.30 | 5434.99 | 4363.49 | 4111.54 | 4615.45 | 1811.58 | 1700.29 | 1922.87 |
| 15/06/2020 | 15638 | 13708.4 | 17660.8 | 11505.32 | 9960.99 | 13049.65 | 5243.49 | 4901.53 | 5585.46 | 4471.67 | 4193.77 | 4749.57 | 1848.20 | 1722.06 | 1974.34 |
| 16/06/2020 | 16049 | 13860.1 | 18354.7 | 11712.50 | 9986.17 | 13438.82 | 5352.34 | 4967.04 | 5737.65 | 4581.39 | 4277.12 | 4885.66 | 1884.82 | 1743.58 | 2026.06 |
| 17/06/2020 | 16465 | 14004.2 | 19072.2 | 11907.72 | 9994.72 | 13820.72 | 5461.19 | 5030.90 | 5891.49 | 4691.85 | 4360.54 | 5023.16 | 1921.44 | 1764.82 | 2078.06 |
| 18/06/2020 | 16888 | 14140.9 | 19813.0 | 12098.84 | 9990.64 | 14207.05 | 5570.04 | 5093.15 | 6046.94 | 4803.61 | 4444.75 | 5162.47 | 1958.06 | 1785.75 | 2130.37 |
| 19/06/2020 | 17316 | 14270.2 | 20578.5 | 12287.98 | 9986.30 | 14589.67 | 5678.89 | 5153.85 | 6203.93 | 4916.29 | 4529.24 | 5303.33 | 1994.68 | 1806.35 | 2183.00 |
| 20/06/2020 | 17751 | 14392.3 | 21366.0 | 12471.87 | 9971.85 | 14971.89 | 5787.74 | 5213.05 | 6362.43 | 5030.14 | 4614.37 | 5445.91 | 2031.30 | 1826.63 | 2235.97 |
| 21/06/2020 | 18192 | 14507.4 | 22178.7 | 12648.24 | 9946.96 | 15349.53 | 5896.59 | 5270.79 | 6522.39 | 5144.99 | 4699.89 | 5590.09 | 2067.92 | 1846.57 | 2289.27 |
| 22/06/2020 | 18639 | 14615.4 | 23016.0 | 12821.09 | 9914.53 | 15727.66 | 6005.44 | 5327.11 | 6683.77 | 5260.97 | 4785.98 | 5735.96 | 2104.54 | 1866.17 | 2342.91 |
| 23/06/2020 | 19091 | 14716.5 | 23878.0 | 12990.46 | 9878.43 | 16102.50 | 6114.29 | 5382.05 | 6846.53 | 5377.98 | 4872.52 | 5883.43 | 2141.16 | 1885.43 | 2396.89 |
| 24/06/2020 | 19551 | 14810.9 | 24765.1 | 13154.89 | 9834.29 | 16475.49 | 6223.14 | 5435.63 | 7010.64 | 5496.08 | 4959.61 | 6032.55 | 2177.78 | 1904.35 | 2451.20 |
| 25/06/2020 | 20016 | 14898.6 | 25677.6 | 13314.00 | 9783.06 | 16844.93 | 6331.99 | 5487.90 | 7176.07 | 5615.24 | 5047.19 | 6183.29 | 2214.40 | 1922.94 | 2505.85 |

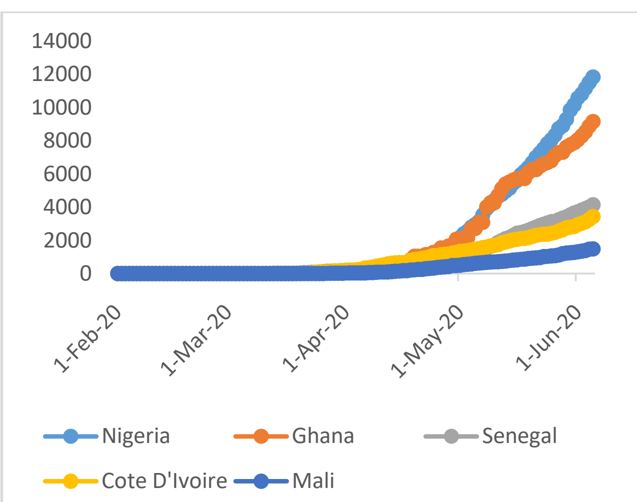


Figure 1. A comparative plot of confirmed COVID-19 cases in Nigeria, Ghana, Senegal, Cote D'Ivoire and Mali from 28th February, 2020 to 5th June, 2020

The fitted models equation 2-6 are used to forecast confirmed COVID-19 cases in Nigeria, Ghana, Senegal, Cote D'Ivoire and Mali for the next 20 days, that is, 6th June, 2020 to 25th June, 2020. The forecasted COVID-19 cumulative cases are presented in Table 3 with 95% confidence interval (CI). According to the forecast, the number of confirmed COVID-19 cases is expected to increase considerably in the coming 20 days.

The trend for the number of confirmed cases with respect to time due to COVID-19 infections in the five selected countries in West Africa depicted in Figure 1. According to the plot, Nigeria is the most infected while Mali is the least infected of the five selected countries in West Africa. The increase is highly suspected by the researchers to be associated with the attitude of the people. There is the belief among people in West Africa that COVID-19 is not real and so do not adhere to rules and regulation put in place by health experts. People still engage in social gathering and they interact with infected people that are still untraceable. Thus, these people may increase the transmissions of the virus and lead to higher number of infected figures. Also, negligence on the side of the few people who did not follow the suggested 14 days isolation after returning from abroad or state with higher cases might be consider as another reasons. Social media to some extent, is also contributing towards some cases owing to the fake information being spread through the platform. Hence, it is very important to control such communications as they result in people moving out of their places due to wrong information obtained from the platform. Other reasons associated with the increase in the case, is the belief that the lack of trust in the government agencies confirmed COVID-19 results with many equating the disease to malaria, the believe that the disease affects only the rich and those who travels overseas. All these circumstances can end up in transmissions.

## V. CONCLUSION AND RECOMMENDATIONS

Coronavirus disease (COVID-19) declared as pandemic by WHO and is currently a major global threat. In order to support the prevention of the disease and aid in the healthcare service planning and preparation, we have conducted this study to examine the best model for the prediction of confirmed COVID-19 infection cases and to use that model for forecasting future COVID-19 infection cases in five selected countries in West Africa. Based on the model forecast, with all other things being equal, the confirmed cases are expected to rise greatly in the coming days. This means that unless restrictive measures are put in place, West African countries should expect a rising number of confirmed cases of COVID-19. This selected Countries can still control the situation if the prevention measures such as quarantine and city sanitization are strictly followed. This study has several implications to practice especially the government. Specifically, government can use the result of this study to design measures that will curtail the uprising trend of COVID-19 in the region. The upward trend of confirmed COVID-19 cases can be prevented when appropriate prentive measures are taken by the people. This implies that people should be adequately sensitized about COVID-19 and preventive practices. The prediction models come up in this study will help the government and medical workforce to be prepared for the upcoming situations and have more readiness in healthcare systems. The African Healthcare system is known to be weak and a laxity in taking stringent measures to curtail the disease will spell result into greater loss of lives to the virus.

## REFERENCES


[1]. Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PloS one*, *15*(3), e0230405.

[2]. Fang, Y., Nie, Y., & Penny, M. (2020). Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. *Journal of medical virology*, *92*(6), 645-659.

[3]. OECD (2020). COVID-19 and West Africa in Numbers. Organization for Economic Cooperation and Development. Retrieved June 17, 2020 from http://www.oecd.org/swac/coronavirus-west-africa/

[4]. Okonta, C. (2020). COVID-19 in West Africa: Let's Prepare for Long Distance Run. Africa Centre for Disease Control. Retrieved June 16, 2020 from: https://www.msf.org/covid-19-west-africa-preparing-long-distance-run

[5]. Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J. M., ... & Chowell, G. (2020). Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infectious Disease Modelling*, *5*, 256-263.

[6]. Li, Q., Feng, W., & Quan, Y. H. (2020). Trend and forecasting of the COVID-19 outbreak in China. *Journal of Infection*, *80*(4), 469-496.

[7]. Liu, Z., Magal, P., Seydi, O., & Webb, G. (2020). Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data. *arXiv preprint arXiv:2002.12298*.

[8]. Grasselli, G., Pesenti, A., & Cecconi, M. (2020). Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: early experience and forecast during an emergency response. *Jama*, *323*(16), 1545-1546.

[9]. Russo, L., Anastassopoulou, C., Tsakris, A., Bifulco, G. N., Campana, E. F., Toraldo, G., & Siettos, C. (2020). Tracing DAY-ZERO and Forecasting the Fade out of the COVID-19 Outbreak in Lombardy. *Italy: a compartmental modelling and numerical optimization approach. medRxiv*.

[10]. Massonnaud, C., Roux, J., & Crépey, P. (2020). COVID-19: Forecasting short term hospital needs in France. *medRxiv*.

[11]. Liu, P., Beeler, P., & Chakrabarty, R. K. (2020). COVID-19 Progression Timeline and Effectiveness of Response-to-Spread Interventions across the United States. *medRxiv*.

[12]. Lover, A. A., & McAndrew, T. (2020). Sentinel Event Surveillance to Estimate Total SARS-CoV-2 Infections, United States. *medRxiv*.

[13]. Wise, T., Zbozinek, T. D., Michelini, G., & Hagan, C. C. (2020). Changes in risk perception and protective behavior during the first week of the COVID-19 pandemic in the United States.

[14]. Gupta, R., & Pal, S. K. (2020). Trend Analysis and Forecasting of COVID-19 outbreak in India. *medRxiv*.

[15]. Kim, S. K. (2020). AAEDM: Theoretical Dynamic Epidemic Diffusion Model and Covid-19 Korea Pandemic Cases. *medRxiv*.

[16]. Zhang, Y., Yang, H., Cui, H., & Chen, Q. (2019). Comparison of the Ability of ARIMA, WNN and SVM Models for Drought Forecasting in the Sanjiang Plain, China. *Natural Resources Research*, 1-18.

[17]. Imdadullah, M. (2014). Time series analysis. *Basic statistics and data analysis*.