



# Time Series Models for COVID-19 New Cases in Top Seven Infected African Countries

Alemayehu Siffir Argawu <sup>a\*</sup>

<sup>a</sup> *Department of Statistics, Ambo University, Ambo, Oromia Regional State, Ethiopia.*

## **Author's contribution**

*The sole author designed, analyzed, interpreted and prepared the manuscript.*

## **Article Information**

DOI: 10.9734/JPRI/2021/v33i60B34703

## **Open Peer Review History:**

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: <https://www.sdiarticle5.com/review-history/78043>

**Original Research Article**

**Received 15 October 2021**  
**Accepted 20 December 2021**  
**Published 23 December 2021**

## **ABSTRACT**

**Introduction:** In Africa region on the date of December 09, 2021 at 14:46 GMT, the total cumulative cases of COVID-19 was 8,889,437 with total deaths and total recoveries of 224,731 (2.5% of death rate) and 8,185,382 (92% of recovery rate) respectively. Thus, this study aimed modelling and forecasting of COVID-19 new cases in top seven infected African countries using time series models.

**Methods:** The top seven infected African countries COVID-19 new cases dataset was taken from our World COVID-19 dataset. The study period was from February 14 to September 06, 2020. Different time series models were used for modelling and forecasting of COVID-19 new cases data. Models comparisons were done by normalized BIC, root mean square error (RMSE), mean absolute percentage error (MAPE), and R-squared values.

**Results:** The COVID-19 new cases data of Algeria, Egypt, Ethiopia, Morocco, and South Africa were fitted by ARIMA (0,1,0), ARIMA (0,1,0), Damped trend, Brown, and ARIMA (0,1,14) models in the study period, respectively. Whereas Ghana, and Nigeria COVID-19 new cases data were followed by simple exponential smoothing models. The 95% confidence levels for lowest to highest forecasted COVID-19 new cases were 258 to 197 with decreasing trend in Algeria, 63 to 933 with increasing trend in Egypt, 636 to 2,141 with increasing trend in Ethiopia, 0 to 1,022 with constant trend in Ghana, 1,900 to 2,807 with increasing trend in Morocco, 0 to 543 with constant trend in Nigeria, and 2,056 to 2,444 with increasing trend in South Africa for the next one month (from September 7 to October 6, 2020).

**Conclusion:** The findings of the study used for preparedness planning against further spread of the COVID-19 epidemic in African countries. The author recommends that as many countries continue to relax restrictions on movement and mass gatherings, and more are opening their air

spaces, and the countries' other public and private sectors are reopening and then strong appropriate public health and social measures must be instituted on the ground again and again before the virus is distributed and attacked more and more peoples in the region. And, the researcher recommended that risk factors of COVID-19 new cases should be conducted for next time in Africa countries.

*Keywords: COVID-19 new cases; time series models; African countries.*

## 1. INTRODUCTION

### 1.1 Background of the Study

The COVID-19 was first identified on 31 December 2019 in the city of Wuhan, which is the capital of Hubei Province in China [1]. The World Health Organization (WHO) on March 11, 2020, has declared the novel coronavirus (COVID-19) outbreak a global pandemic [2]. It caused the severe acute respiratory syndrome coronavirus 2 (SARSCoV-2) has hit the world severely on early of 2020. Many countries are facing a rapid increasing trend of confirmed cases. The case-fatality-rate varies wildly from country to country [3,4].

Globally on the date of December 09, 2021 at 14:46 GMT, the total cumulative cases of COVID-19 was 268,326,045 with total deaths and total recoveries of 5,298,933 (2% of death rate) and 241,501,470 (90% of recovery rate) respectively as Worldometer report shown [5].

In Africa region on this date, the total cumulative cases of COVID-19 was 8,889,437 with total deaths and total recoveries of 224,731 (2.5% of death rate) and 8,185,382 (92% of recovery rate)

### 1.2 Statement of Problem

It is essential to create a reliable and suitable predictive model that can help governments and stakeholders to control the further spread of COVID-19. Time series forecasting models are the statistical techniques which give decent predictions and have been widely applied for trends of infectious disease as different researchers used in their studies [8-17]. Thus, the researcher aimed to conduct study on September 06, 2020 by considering the top seven infected African countries since COVID-19 new cases were alarmingly increased in Ethiopia and other African countries in previous two months (July to August). On this date, top seven infected African countries were South Africa (1<sup>st</sup>) by 636,884, Egypt (2<sup>nd</sup>) by 99,712, Morocco (3<sup>rd</sup>)

respectively as this report shown. In the region, there was some rates increment on both sides. On this date, the top seven infected African countries were South Africa (1<sup>st</sup>) by 3,071,064, Morocco (2<sup>nd</sup>) by 950,946, Tunisia (3<sup>rd</sup>) by 718,866, Libya (4<sup>th</sup>) by 376,873, Ethiopia (5<sup>th</sup>) by 372,588, Egypt (6<sup>th</sup>) by 365,831, and Kenya (7<sup>th</sup>) by 255,652 total cumulative cases of COVID-19. South Africa is still leading by huge difference. And, these reports were presented on Fig. 1.

On this date, the COVID-19 cases and rates were highly increased as compared with the study period from February 14 to September 06, 2020. The increments may be due to the increasing of daily laboratory tests in the country [6,7]. And, it's nature of rapid distribution and less protectiveness of individual in the countries.

The African region was described as one of the vulnerable with the COVID-19 infection in the initial phase, due to the fact that Africa is important commercial partner of China and as a result, large volumes of business people travel to the region. Since the epicentre is now in Europe and America, due to the close tie between Africa and countries, African countries face even bigger threat [4].

by 70,160, Ethiopia (4<sup>th</sup>) by 57,466, Nigeria (5<sup>th</sup>) by 54,905, Algeria (6<sup>th</sup>) by 46,071, and Ghana (7<sup>th</sup>) by 44,777 total cumulative cases of COVID-19, respectively [18]. In the region, it was 1,297,434 total cases. And, South Africa covered almost 50% of regional COVID-19 total cases. The report was charted on Fig. 2.

### 1.3 Aim of the Study

This study aimed modelling and forecasting of COVID-19 new cases in top seven infected African countries using time series models. Modelling and forecasting COVID-19 new cases would be a useful guidance for timely prevention and control measure to be effectively planned in advance, and it is also useful for sustaining strict measures in order to curtail spread of the virus.

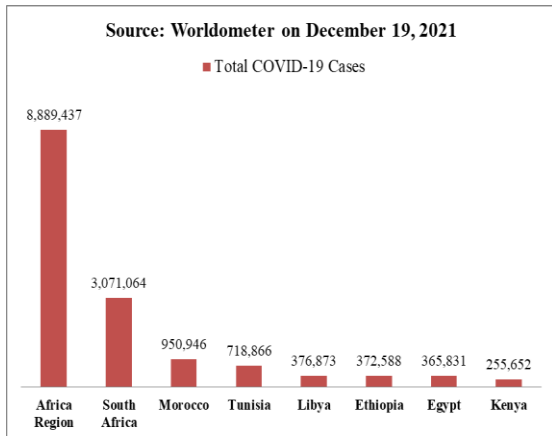


Fig. 1. Charts of total COVID-19 cases for top seven infected African country in Dec. 2021



Fig. 2. Charts of total COVID-19 cases for top seven infected African country in Sep. 2020

## 2. MATERIALS AND METHODS

### 2.1 Study Period

### 2.3.2 Exponential smoothing models

There are four types of non-seasonal exponential smoothing models. These are Simple, Holt's linear trend, Brown's linear trend, and Damped trend models.

**Simple Model:** It is used for forecasting a time series when there is no trend or seasonal pattern. The simple exponential smoothing model is given by the model equation:

$$(1 - B)Y_t = (1 - \theta B)a_t \quad (2)$$

Where:

$\theta = 1 - \alpha$  and B represents the backshift operator such that

The study period was from February 14 to September 06, 2020.

### 2.2 Source of the Data

The top seven infected African countries COVID-19 new cases dataset was downloaded from our World COVID-19 dataset from February 14 to September 06, 2020. It is available in <https://github.com/owid/covid-19-data>.

### 2.3 Time Series Models

A time series is a set of observations  $x_t$ , each one being recorded at a specific time  $t$ . Discrete-time time series are recorded when observations are made at fixed time intervals. Continuous-time time series are obtained when observations are recorded continuously over some time interval. Then, different time series models were used to forecast COVID-19 new cases for the next times [19-22].

#### 2.3.1 ARIMA models

ARIMA model becomes AR ( $p$ ), MA ( $q$ ), or ARMA ( $p, q$ ) if the time series is stationary. The expression of ARIMA ( $p, d, q$ ) model can be defined as follows:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \alpha_1 - \theta_1 \alpha_{t-1} - \alpha_2 - \theta_2 \alpha_{t-2} - \dots - \alpha_q - \theta_q \alpha_{t-q} \quad (1)$$

Where  $\phi_p$  are the parameter values for an autoregressive operator,  $\alpha_q$  are the error term coefficients,  $\theta_q$  are the parameter values for moving average operator, and  $Y_t$  is the time series of the original series difference at the degree.

$$B^r X_t = x_{t-r} \text{ for any given time series } x_t.$$

**Holt's Linear Trend Model:** This model is appropriate for a series with a linear trend and no seasonality. Its relevant smoothing parameters are level and trend, and, in this model, they are not constrained by each other's values. The estimates are made using the equations below.

$$Y'_t = \alpha Y_t + (1 - \alpha)(Y'_{t-1} + B_{t-1}) \quad (3)$$

Where  $\alpha$  and  $\gamma$  are the smoothing constants in the range of [0, 1].

**Brown’s linear trend model:** In this model, the parameters are assumed that the level and trend are equal. In this method, estimates are made using the equations below.

$$Y'_t = \alpha Y_t + (1 - \alpha)(Y'_{t-1}) \quad (4)$$

**Damped Trend Model:** It is well established for an accurate forecasting method, and it's new stated damped trend model is written as follow:

$$Y_t = I_{t-1} + A_t b_{t-1} + \varepsilon_t \quad (5)$$

Where  $Y_t$  is the observed series,  $I_t$  is its level and  $b_t$  is the gradient of its linear trend. This model has a single source of error,  $\varepsilon_t$ .

And, models comparisons were done by normalized BIC, root mean square error (RMSE), mean absolute percentage error (MAPE), and R-squared values.

### 3. RESULTS

#### 3.1 Fitted Time Series Models

The COVID-19 new cases data of Algeria, Egypt, Ethiopia, Morocco, and South Africa were fitted

by ARIMA (0,1,0), ARIMA (0,1,0), Damped trend, Brown, and ARIMA (0,1,14) models in the study period, respectively. Whereas Ghana, and Nigeria COVID-19 new cases data were followed by simple exponential smoothing models. All the fitted models had relatively the smallest normalized BIC, root mean square error (RMSE), mean absolute percentage error (MAPE) values, and with the highest R-squared values. The result was presented in Table 1.

#### 3.2 Forecasting and Trends of COVID-19 New Cases

The 95% confidence levels for lowest to highest forecasted COVID-19 new cases were 258 to 197 with decreasing trend in Algeria, 63 to 933 with increasing trend in Egypt, 636 to 2,141 with increasing trend in Ethiopia, 0 to 1,022 with constant trend in Ghana, 1,900 to 2,807 with increasing trend in Morocco, 0 to 543 with constant trend in Nigeria, and 2,056 to 2,444 with increasing trend in South Africa for the next one month (from Sep 7 to Oct 6, 2020). The results were presented in Table 2 and Fig. 3.

**Table 1. Fitted models summaries for COVID-19 new cases for top seven infected African countries from February 14 to September 06, 2020**

Country	Model	R <sup>2</sup>	RMSE	MAPE	Normalized BIC
Algeria	ARIMA (0,1,0)	0.994	15.655	14.669	5.788
Egypt	ARIMA (0,1,0)	0.977	82.233	30.387	9.081
Ethiopia	Damped Trend	0.899	158.06	53.397	10.222
Ghana	Simple	0.578	196.80	53.78	10.601
Morocco	Brown	0.721	310.65	184.81	11.507
Nigeria	Simple	0.84	87.487	28.188	8.973
South Africa	ARIMA (0,1,14)	0.974	680.08	91.046	13.303

**Table 2. Forecast and trends of COVID-19 new cases for the top seven African countries from 5<sup>th</sup> of September to 6<sup>th</sup> of October 2020**

Country	Model	Trend	95% Forecasted values (LCL, UCL)	
Algeria	ARIMA (0,1,0)	Decreasing	258	197
Egypt	ARIMA (0,1,0)	Increasing	63	933
Ethiopia	Damped Trend	Increasing	636	2,141
Ghana	Simple	Constant	0	1,022
Morocco	Brown	Increasing	1,287	3,858
Nigeria	Simple	Constant	0	543
South Africa	ARIMA (0,1,14)	Increasing	1,258	9,229

*LCL and UCL are the lower and upper confidence levels, respectively*

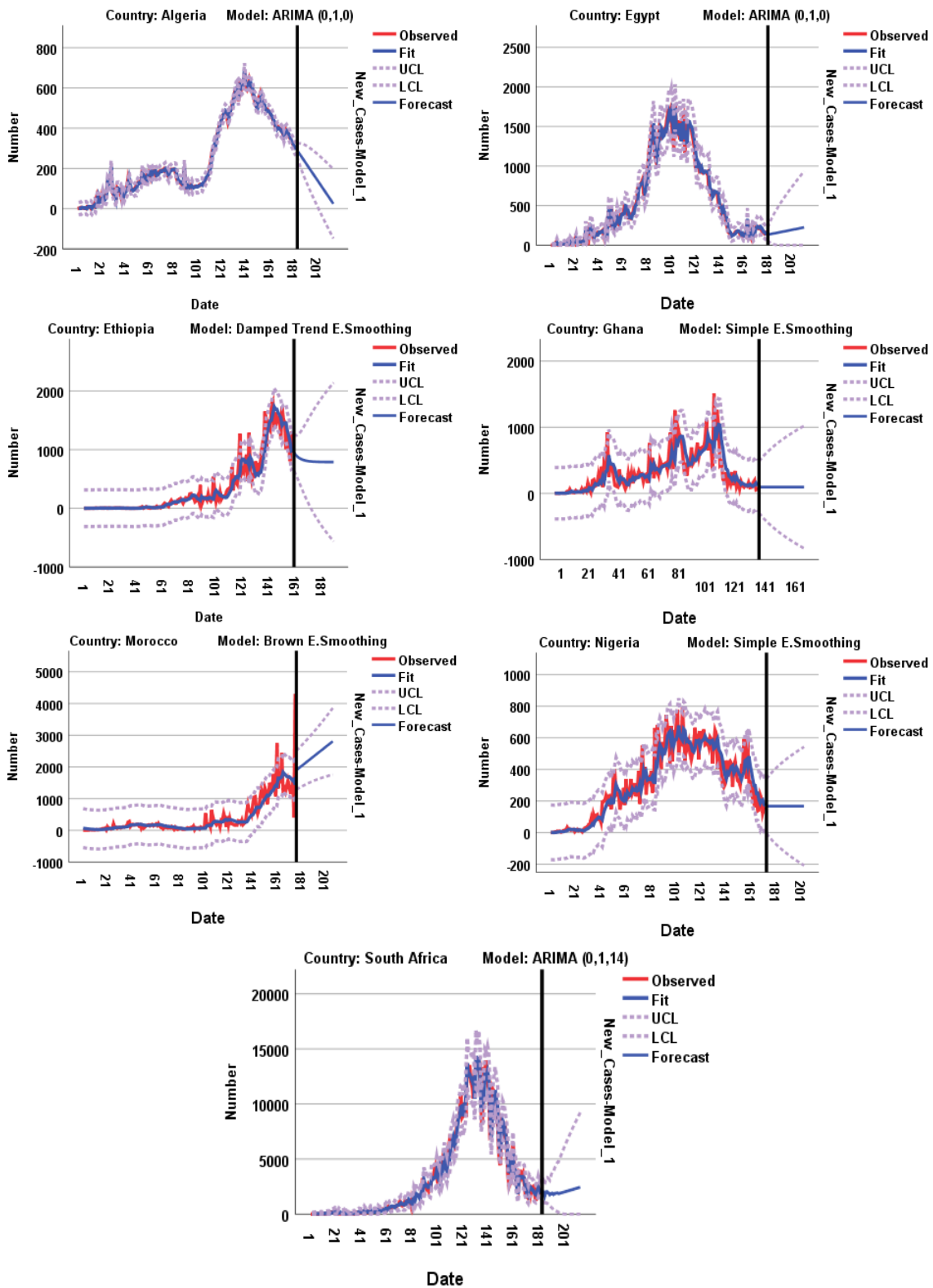


Fig. 3. Graphs of forecasted and trends of COVID-19 new cases data for the top seven African countries from February 14 to September 06, 2020

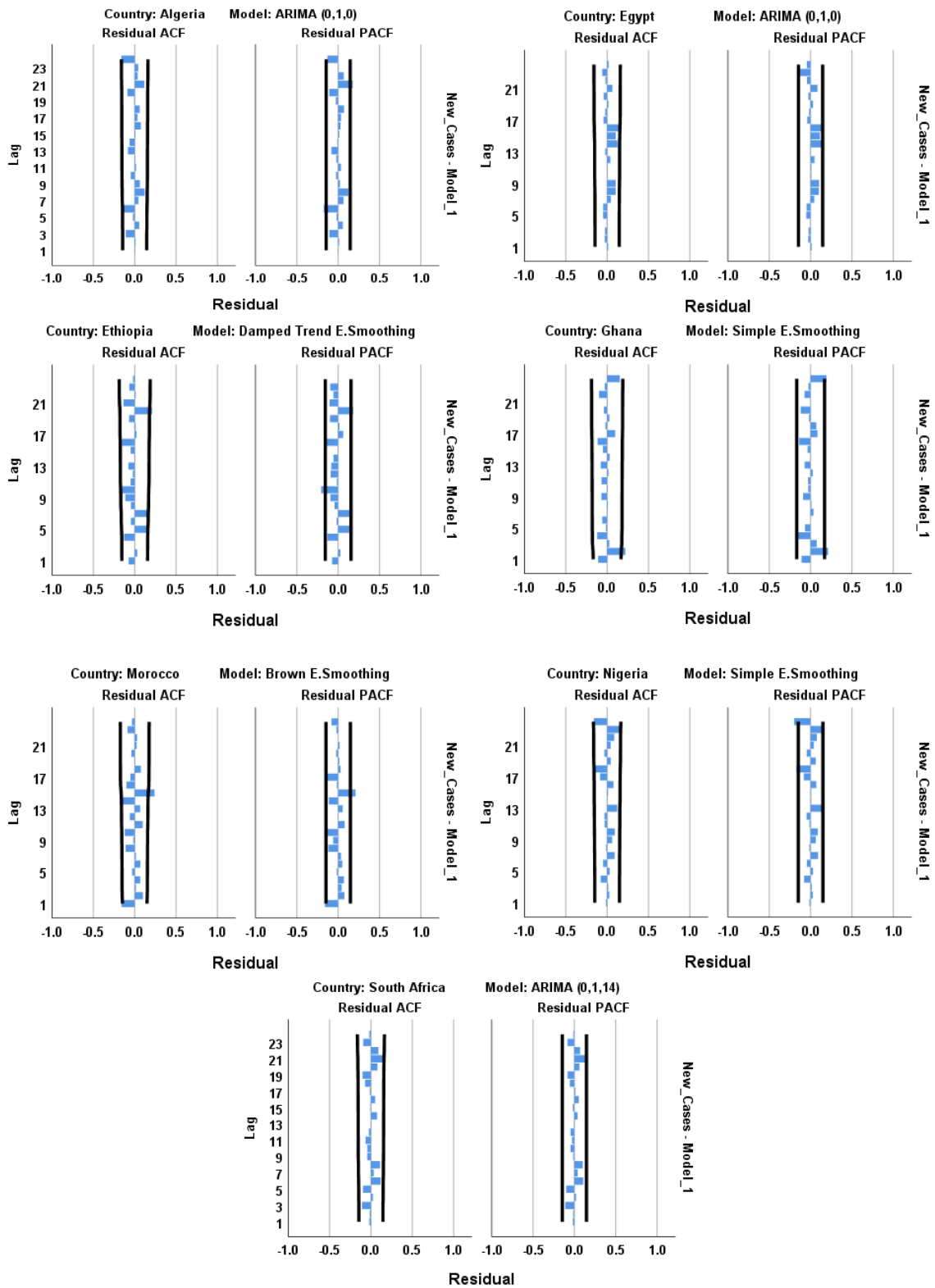


Fig. 4. Graphs of auto-correlation and partial auto-correlation functions residuals

### 3.3 Residuals Stationary Tests

The residuals stationary tests were examined and checked by auto-correlation function and partial auto-correlation function graphs. Both were presented in Fig. 4.

## 4. DISCUSSION

The principal purpose of this research was modelling and forecasting of COVID-19 new cases in top seven infected African countries using time series models. As the results revealed that The COVID-19 new cases data of Egypt, Ethiopia, Morocco, and South Africa were fitted by ARIMA (0,1,0), Damped trend, Brown, and ARIMA (0,1,14) models in the study period, respectively. Whereas Libya, Nigeria, and Tunisia COVID-19 new cases data were followed by similar models of simple exponential smoothing. This result is in line with various studies as stated below.

Argawu et al. [6] found that the COVID-19 new cases was positively correlated with the number of days, daily laboratory tests, new cases of males, new cases of females, new cases from Addis Ababa city, and new cases from foreign natives. In the multiple linear regression model, COVID-19 new cases was significantly predicted by the number of days at 5%, the number of daily laboratory tests at 10%, and the number of new cases from Addis Ababa city at 1% levels of significance [6].

Dehesh et al. [8] used parameters for ARIMA were (2,1,0) for Mainland China, ARIMA (2,2,2) for Italy, ARIMA(1,0,0) for South Korea, ARIMA (2,3,0) for Iran, and ARIMA(3,1,0) for Thailand. Mainland China and Thailand had almost a stable trend. The trend of South Korea was decreasing and will become stable in near future. Iran and Italy had unstable trends [8].

Takele [13] applied autoregressive integrated moving average (ARIMA) modeling approach for projecting coronavirus (COVID-19) prevalence patterns in East Africa Countries, mainly Ethiopia, Djibouti, Sudan and Somalia. The finding showed that the optimal models for Ethiopia, Djibouti, Somalia and Sudan were ARIMA (1, 2, 1), ARIMA (2, 1, 1), ARIMA (1, 2, 2), and ARIMA (0, 2, 1) respectively. In the views of worst-case and average-case scenarios, cumulated infection rate of COVID-19 were predicted by the end of October, 2020. Rate of

infection was expected to increase and magnify in Sudan and Ethiopia from Nov. to Feb. 2020/21 [13].

Earlier study in the African region indicated that the epidemic was controlled in late April with strict control of scenario one, manifested by the circumstance in South Africa and Senegal. Under moderate control of scenario two, the number of infected peoples was increase by 1.43–1.55 times of that in scenario one, the date of the epidemic being controlled was delayed by about 10 days, and Algeria, Nigeria, and Kenya were following this situation. In the third scenario of weak control, the epidemic was controlled by late May, and the total number of infected cases was double that in scenario two, and Egypt was in line with this prediction [14].

Study in the selected G8 European countries (Germany, United Kingdom, France, Italy, Russian, Canada, Japan, and Turkey) for the number of COVID 19 epidemic cases data was fitted the cubic regression models with the curve estimations. The number of COVID 19 epidemic cases data were modelled and forecasted that Japan (Holt Model), Germany (ARIMA (1, 4, 0) and France (ARIMA (0, 1, 3) were provided statistically significant. UK (Holt Model), Canada (Holt Model), Italy (Holt Model), and Turkey (ARIMA (1, 4, 0) were not statistically significant [15].

Ceylan [17] found that the ARIMA (0, 2, 1), ARIMA (1, 2, 0), and ARIMA (0, 2, 1) models with the lowest MAPE values (4.7520, 5.8486, and 5.6335) were selected as the best models for Italy, Spain, and France, respectively. The ARIMA models are suitable for predicting the prevalence of COVID-19 in future. The results of the analysis can shed light on understanding the trends of the outbreak and give an idea of the epidemiological stage of these regions [17].

Parallel study shown that the ARIMA (0, 2, 3), ARIMA (0, 1, 1), ARIMA (3, 1, 0) and ARIMA (0, 1, 2) models were chosen as the best models for South Africa, Nigeria, and Ghana and Egypt, respectively. Forecasting was made based on the best models. It is noteworthy to claim that the ARIMA models are appropriate for predicting the prevalence of COVID-19. The researchers noticed a form of exponential growth in the trend of this virus in Africa in the days to come [23].

Likassa et al. [24] also revealed that the spatial and temporal pattern of this novel virus was varying, spreading and covering the entire world within a brief time. In the study, the fitting effect of the cubic model ( $R^2=99.6\%$ ) was the best outperforming compared to the other six families of exponentials [24].

Similar study by Achoki et al. [25] found that spatial pattern of cumulative COVID-19 cases in Morocco was the leading contributor to the burden of COVID-19 in Northern African on June 30, 2020. Morocco had forecasted 4,459,877 cumulative cases of COVID-19 and this was almost double the estimated number for Algeria, a country with the next highest burden, 2,804,674 by the end of June 2020. In Southern Africa, South Africa and Swaziland are the leading contributors to the pandemic. By the end of June 2020, the countries were expected to have 2,581,366 and 254,403 cumulative cases, respectively. In the Western Africa sub-region, cumulative cases of infection were dominated by Libya and Ghana, despite Nigeria having a larger population than both countries combined. And the numbers of new COVID-19 infections were expected to increase from 2,453,700 cases in April to 577,883 cases in May to 8,044,927 cases by the end of July [25].

Another similar study using African COVID-19 cases showed that estimated exponential growth rate was 0.22 per day, and the basic reproduction number ( $R_0$ ) was 2.37 based on the assumption that the exponential growth starting from March 1, 2020. With an  $R_0$  at 2.37, the researchers quantified the instantaneous transmissibility of the outbreak by the time-varying effective reproductive number to show the potential of COVID-19 to spread across African region [26].

## 5. CONCLUSION

The aim of this investigation was modelling and forecasting of COVID-19 new cases in top seven infected African countries using time series models. The study was based on secondary data obtained from our World COVID-19 dataset. The trends of COVID-19 new cases were increased in Egypt, Ethiopia, Morocco, and South Africa from September 5 to October 6, 2020. But, it was constant in Tunisia, Nigeria, and Libya. And, the measures taken by countries such as the individual attitudes of the societies towards the specified measures and the number of virus tests to be performed are factors that may affect the

number of cases. Since this study was conducted with the current measures, the forecasts obtained may differ from the number of cases that occur in the future. Thus, the study findings should be useful in preparedness planning against the further spread of the COVID-19 epidemic in Africa.

## 6. RECOMMENDATIONS

The author recommend that as many countries continue to relax restrictions on movement and mass gatherings, and more are opening up their airspaces to international travellers with easing of quarantine measures for returning residents and visitors, and the countries' different public and private sectors (like Schools, Universities, Stadiums, and others) are reopening, then strong appropriate public health and social measures must be instituted on the ground again and again before the virus is distributed and attacked more and more peoples in the region. And, the researcher recommended that risk factors of COVID-19 new cases should be conducted for next time in Africa countries.

## CONSENT AND ETHICAL APPROVAL

It is not applicable.

## ACKNOWLEDGEMENT

The author thanked our World COVID-19 dataset organization for freely released the COVID-19 cases dataset in their web page. It is available in <https://github.com/owid/covid-19-data>. And, I acknowledged my organization Ambo University for free and fast internet access.

## COMPETING INTERESTS

Author has declared that no competing interests exist.

## REFERENCES

1. Wu F, et al. Pei 465 YY et al: A new coronavirus associated with human respiratory disease in 466. China. *Nature*. 2020;579(7798):265-269.
2. WHO, WHO Director-General's opening remarks at the media briefing on COVID-19-11 March 2020. Geneva, Switzerland; 2020.
3. WHO, World Health Organization coronavirus disease (COVID-19)



- dashboard. World Health Organization; 2020.
4. Gilbert M, et al., Preparedness and vulnerability of African countries against importations of COVID-19: A modelling study. *The Lancet*. 2020;395(10227):871-877.
  5. Worldometer, <COVID Live Update\_268,326,045 Cases and 5,298,933 Deaths from the Coronavirus - Worldometer on December 09, 2021.pdf>; 2021.
  6. Argawu AS, et al. Prediction of COVID-19 New Cases Using Multiple Linear Regression Model Based on May to June 2020 Data in Ethiopia. *Journal of Pharmaceutical Research International*. 2021;33(51A):54-63.
  7. Weber GM, et al. International comparisons of harmonized laboratory value trajectories to predict severe covid-19: Leveraging the 4ce collaborative across 342 hospitals and 6 countries: A retrospective cohort study. *medRxiv*. 2021;2020:12. 16.20247684.
  8. Dehesh T, Mardani-Fard H, Dehesh P. Forecasting of covid-19 confirmed cases in different countries with arima models. *MedRxiv*; 2020.
  9. Grasselli G, Pesenti A, Cecconi M. Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: Early experience and forecast during an emergency response. *JAMA*. 2020;323(16):1545-1546.
  10. Liu P, Beeler P, Chakrabarty RK. COVID-19 progression timeline and effectiveness of response-to-spread interventions across the United States. *medRxiv*; 2020.
  11. Shi Z, Fang Y. Temporal relationship between outbound traffic from Wuhan and the 2019 coronavirus disease (COVID-19) incidence in China. *MedRxiv*; 2020.
  12. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *The Lancet*. 2020;395(10225):689-697.
  13. Takele R. Stochastic modelling for predicting COVID-19 prevalence in East Africa Countries. *Infectious Disease Modelling*. 2020;5:598-607.
  14. Zhao Z, et al. Prediction of the COVID-19 spread in African countries and implications for prevention and control: A case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya. *Science of the Total Environment*. 2020;729: 138959.
  15. Yonar H, et al. Modeling and Forecasting for the number of cases of the COVID-19 pandemic with the Curve Estimation Models, the Box-Jenkins and Exponential Smoothing Methods. *EJMO*. 2020;4(2):160-165.
  16. Petropoulos F, Makridakis S, Stylianou N. COVID-19: Forecasting confirmed cases and deaths with a simple time series model. *International Journal of Forecasting*; 2020.
  17. Ceylan Z. Estimation of COVID-19 prevalence in Italy, Spain, and France. *Science of The Total Environment*. 2020;729:138817.
  18. Worldometer, <Coronavirus Update (Live): 27,087,151 Cases and 884,110 Deaths from COVID-19 Virus Pandemic - Worldometer on September 06, 2020. pdf>. 2020.
  19. Armstrong JS. Principles of forecasting: A handbook for researchers and practitioners. Springer. 2001;30.
  20. Brockwell PJ, Davis RA. Introduction to time series and forecasting. Springer; 2016.
  21. Li Q, et al. Application of an autoregressive integrated moving average model for predicting the incidence of hemorrhagic fever with renal syndrome. *The American Journal of Tropical Medicine and Hygiene*. 2012;87(2):364.
  22. Tseng YJ, Shih YL. Developing epidemic forecasting models to assist disease surveillance for influenza with electronic health records. *International Journal of Computers and Applications*. 2020;42(6): 616-621.
  23. Lukman AF, et al. COVID-19 prevalence estimation: Four top seven affected African countries. *Infectious Disease Modelling*. 2020;5:827-838.
  24. Likassa HT, et al. Predictive models on COVID 19: What Africans should do? *Infectious Disease Modelling*. 2021;6:302-312.
  25. Achoki T, et al. COVID-19 pandemic in the African continent: forecasts of cumulative cases, new infections, and mortality. *MedRxiv*; 2020.

26. Musa SS, et al. Estimation of exponential growth rate and basic reproduction number of the coronavirus disease 2019 (COVID-19) in Africa. *Infectious Diseases of Poverty*, 2020;9(1):1-6.

---

© 2021 Argawu; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Peer-review history:*  
*The peer review history for this paper can be accessed here:*  
<https://www.sdiarticle5.com/review-history/78043>