Original Research Article

Modelling and Forecasting of COVID-19 New Cases in the Top 10 Infected African Countries, 2020

Abstract

Introduction: The total COVID-19 cases reached 3,581,783 with 102,201 total deaths and 3,214,512 total recovered in the top 10 infected African countries on April 28, 2021 at 10:30 am. This study aimed to model and forecast COVID-19 new cases in the top 10 infected African countries from February 14 to September 06, 2020. **Methods:** The COVID-19 dailly new cases datas were taken from our World in data COVID-19 database and from Worldometer information for the top 10 infected African countries from February 14 to September 06, 2020. And, different time serious models were used to model and forecast COVID-19 new cases for the next times. **Results:** Monthly prevalence of COVID-19 cumulative cases was declined in South Africa, Cote d′Ivoire, Egypt, Ghana, Cameron, Nigeria, and Algeria by 31%, 26%, 22%, 20%, 14%, 12%, and 4% from July to August, respectively. But, it was raised in Ethiopia, Morocco, and Kenya by 41%, 38%, and 1% from July to August, respectively. In the time series analysis; the Algeria, Egypt, and South Africa COVID-19 new cases data have fitted the ARIMA (0,1,0), ARIMA (0,1,0), and ARIMA (0,1,14) models, respectively. The Cameroon, Côte d'Ivoire, Ghana and Nigeria data have fitted the simple exponential smoothing models. Ethiopia, Kenya, and Morocco data have followed the Damped trend, Holt, and Brown exponential smoothing models, respectively. **Conclusion:** The findings of the study may be 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 before the virus is distributed every where and attacked more and more.

Keywords: COVID-19 new cases, Time Series Models, African Countries

Introduction

An outbreak of coronavirus disease 2019 (COVID-19), now a pandemic, caused by 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 (Gilbert et al., 2020; WHO, 2020).

Globally on the date of April 28, 2021 at 10:30 am, as the Worldometer report showed that the total number of COVID-19 cases, total deaths, and total recoveries were 149,366,430, 3,149,673 with death rate of 2.11%, and 127,047,644 with recovery rate of 85.1%, respectively (Worldometer, 2021).

In Africa region on this date, the top 10 COVID-19 infected African countries were South Africa (1st), Morocco (2nd), Tunisia (3rd), Ethiopia (4th), Egypt (5th), Libya (6th), Nigeria (7th), Kenya (8th), Algeria (9th), and Ghana (10th). These countries shared 1,577,200 (44%), 509,972 (14.2%), 303,584 (8.5%), 254,044 (7.1%), 224,517 (6.3%), 176,254 (4.9%), 164,912 (4.6%), 157,492 (4.6%), 121,344 (3.4%), and 92,464 (2.6%) total cumulative cases of COVID-19 from Africa region, respectively (Worldometer, 2021).

The pandemic rate was highly increased as compare with the virus distribution on September 06, 2020. The increament was due to the number of daily laboratory tests were increased in each country, and the distribution of the virus was increased rapidly as Worldometer reports shown (Worldometer, 2020).

Thus, the African region has been described as one of the most 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 volume of business people travel to the region. Since the epicenter is now in Europe and Ameria, due to the close tie between Africa and countries, African countries face even bigger threat (Gilbert et al., 2020).

It is essential to create a reliable and suitable predictive model that can help governments and other stakeholders to control the further spread of Novel coronavirus. Time series forecasting models is the statistical technique which gives a decent predictions and has been widely applied for trend of infectious disease in quick time (Ceylan, 2020; Dehesh, Mardani-Fard, & Dehesh, 2020; Grasselli, Pesenti, & Cecconi, 2020; Liu, Beeler, & Chakrabarty, 2020; Petropoulos, Makridakis, & Stylianou, 2020; Shi & Fang, 2020; Takele, 2020; Wu, Leung, & Leung, 2020; Yonar, Yonar, Tekindal, & Tekindal, 2020; Zhao et al., 2020). In this study, with the view to predicting coronavirus (COVID-19) prevalence in the top ten infected african countries, using ARIMA modeling strategy, which would be a useful guidance for timely prevention and control measure to be effectively planned in advance.

Thus, the aim of this study was to model and forecast the number of COVID-19 new cases in the top 10 infected African countries, which should be valuable in informing the official and public in the preparedness against COVID-19 spread, modeling and forecasting the trend, and highlighting the importance of sustaining strict measures in order to curtail the spread of the virus.

Methods

Data

The top ten infected African countries COVID-19 new cases data was downloaded in our World in data COVID-19 database and Worldometer information from February 14 to September 06, 2020. It is available in https://github.com/owid/covid-19-data.

Time Serious 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 recoreded 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 serious models were used to forceste COVID-19 new cases for the next times (Armstrong, 2001; Brockwell & Davis, 2016; Li et al., 2012; Tseng & Shih, 2020).

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:

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

Where ϕ_p are the parameter values for an autoregressive operator, α_q are the error term coefficients, θ_q are the parameter values for moving average operator, and Y_t is the time series of the original series difference at the degreed.

Exponential Smoothing Methods

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:

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

Where $\theta = 1 - \alpha$ and B represents the backshift operator such that $B^r X_t = x_{t-r}$ 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.

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

Where α and γ 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.

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

Damped trend model: It is well established for an accurate forecasting method. The new stated damped trend model is written as follow:

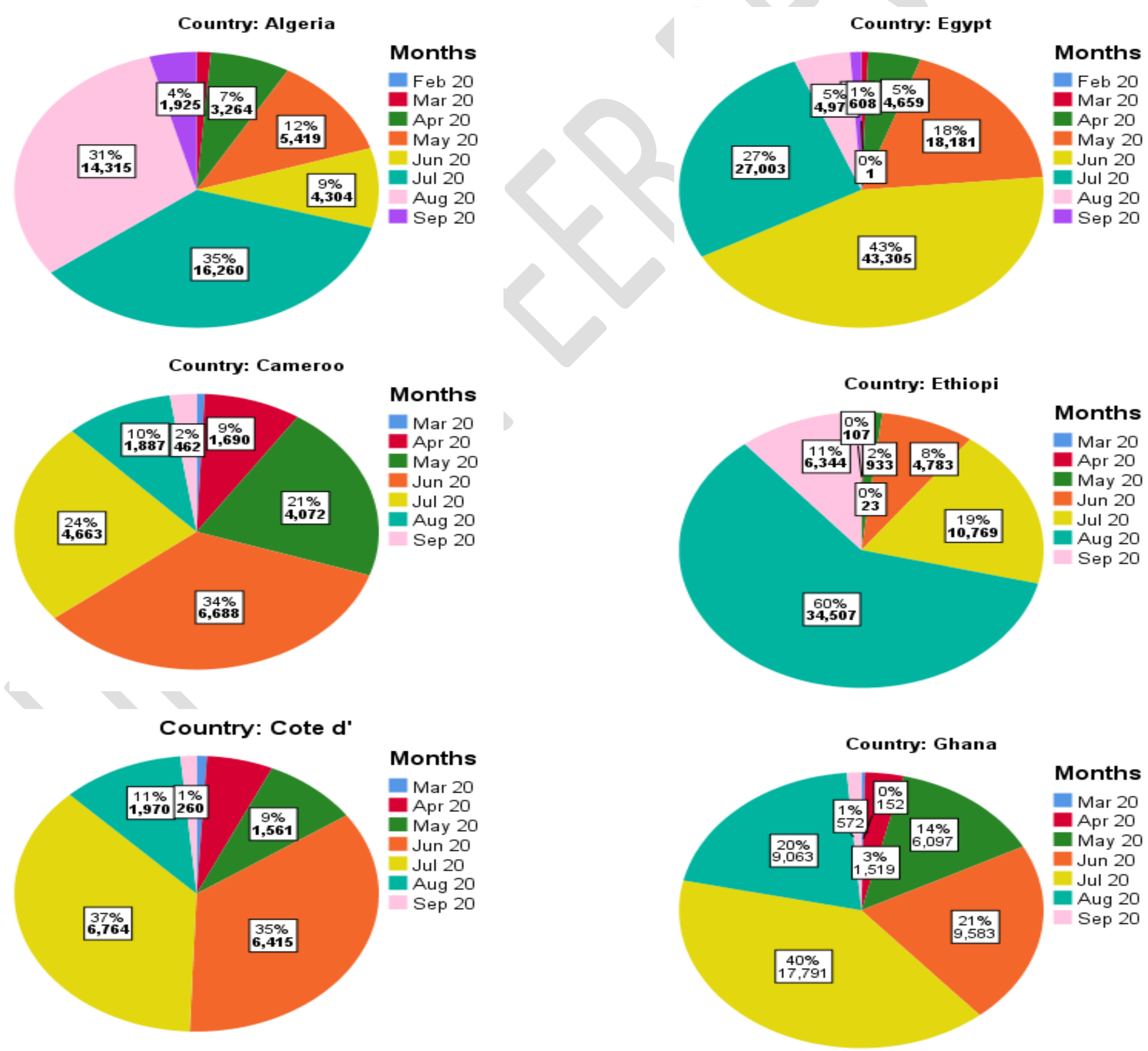
$$\mathbf{Y_t = l_{t-1} + A_t b_{t-1} + \varepsilon_t} \quad (5)$$

Where Y_t is the observed series, l_t is its level and b_t is the gradient of its linear trend. This model has a single source of error, ε_t.

Results

Monthly Prevalence of COVID-19 Cumulative Cases

Monthly prevalence of COVID-19 cumulative cases declined in South Africa, Cote d′Ivoire, Egypt, Ghana, Cameron, Nigeria, and Algeria by 31%, 26%, 22%, 20%, 14%, 12%, and 4% from July to August, respectively. However, it rose in Ethiopia, Morocco, and Kenya by 41%, 38%, and 1% from July to August, respectively. Specifically in Ethiopia and Morocco, 60% and 54% of the cumulative cases have been recorded in August month only, respectively. This indicates that more laboratory tests were conducted in the two countries. The two highly infected countries (South Africa and Egypt) have recorded only 22% and 5% of the countries cases in August month, respectively (Figure 1).



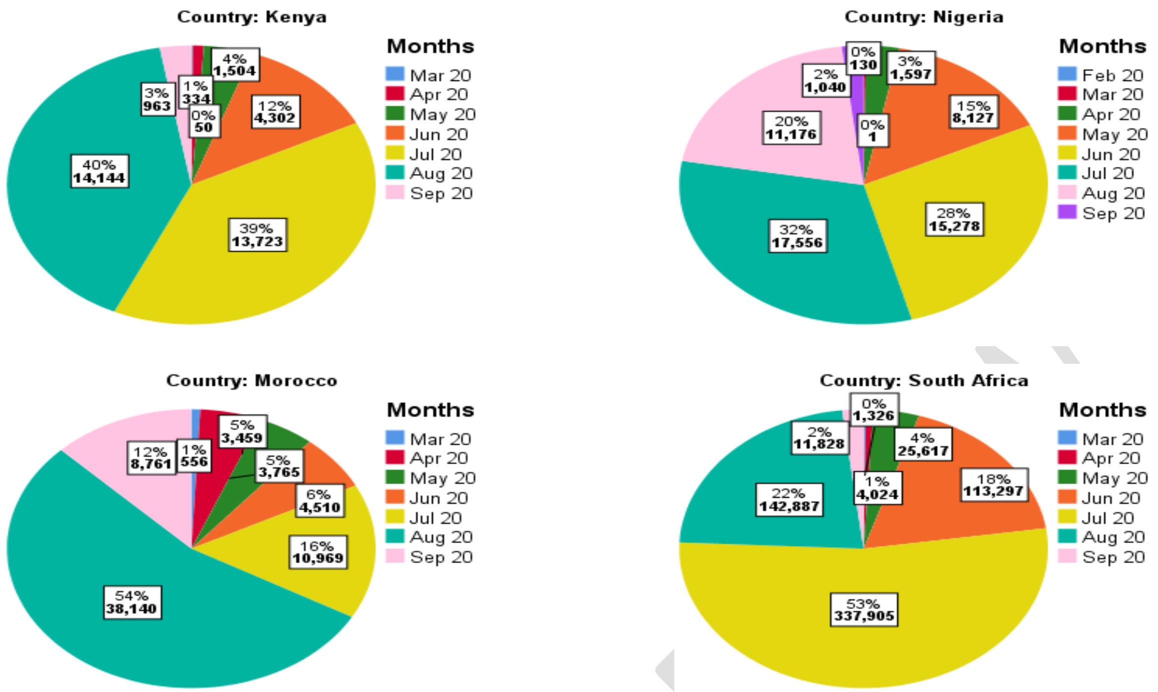


Figure 1. The Pie charts for COVID-19 cumulative new cases by months of the countries from February 14 to September 06, 2020.

Time Series Models Summaries

An appropriate model selection is based on normalised BIC with the smallest value. The Algeria, Egypt, and South Africa COVID-19 new cases data fitted the ARIMA (0,1,0), ARIMA (0,1,0), and ARIMA (0,1,14) models, respectively. The Cameroon, Cote d'Ivoire, Ghana, and Nigeria data fitted the same model type (Simple exponential smoothing model). And, the Ethiopia, Kenya, and Morocco data followed the Damped trend, Holt, and Brown exponential smoothing models, respectively. All of these fitted models have relatively the smallest normalized BIC, root mean square error, and mean absolute percentage error values with the highest stationary R-squared and R-squared values. The ARIMA and exponential smoothing models summaries of the countries' COVID-19 new cases were presented in Table 1. The stationary of the residuals was examined and the ACF and PACF graphics of the series for countries were shown in Figure 2.

Table 1. The ARIMA and exponential smoothing models summaries for COVID-19 new cases for the countries from February 14 to September 06, 2020

Country	Model Type	Model Fit statistics					Ljung-Box Q(18)	
		Stationary R^2	R^2	RMSE	MAPE	Normalized BIC	DF	Sig.
Algeria	ARIMA (0,1,0)	0.638	0.994	15.655	14.669	5.788	18	0.576
Cameroon	Simple	0.438	0.292	183.52	191.19	10.472	17	0.193
Cote d'Ivoire	Simple	0.272	0.679	66.815	56.825	8.435	17	0.000
Egypt	ARIMA (0,1,0)	0.564	0.977	82.233	30.387	9.081	18	0.601
Ethiopia	Damped Trend	0.387	0.899	158.06	53.397	10.222	15	0.033
Ghana	Simple	0.333	0.578	196.80	53.78	10.601	17	0.412
Kenya	Holt	0.787	0.861	84.317	57.394	8.93	16	0.000
Morocco	Brown	0.725	0.721	310.65	184.81	11.507	17	0.005
Nigeria	Simple	0.304	0.84	87.487	28.188	8.973	17	0.553
South A	ARIMA (0,1,14)	0.518	0.974	680.08	91.046	13.303	14	0.464

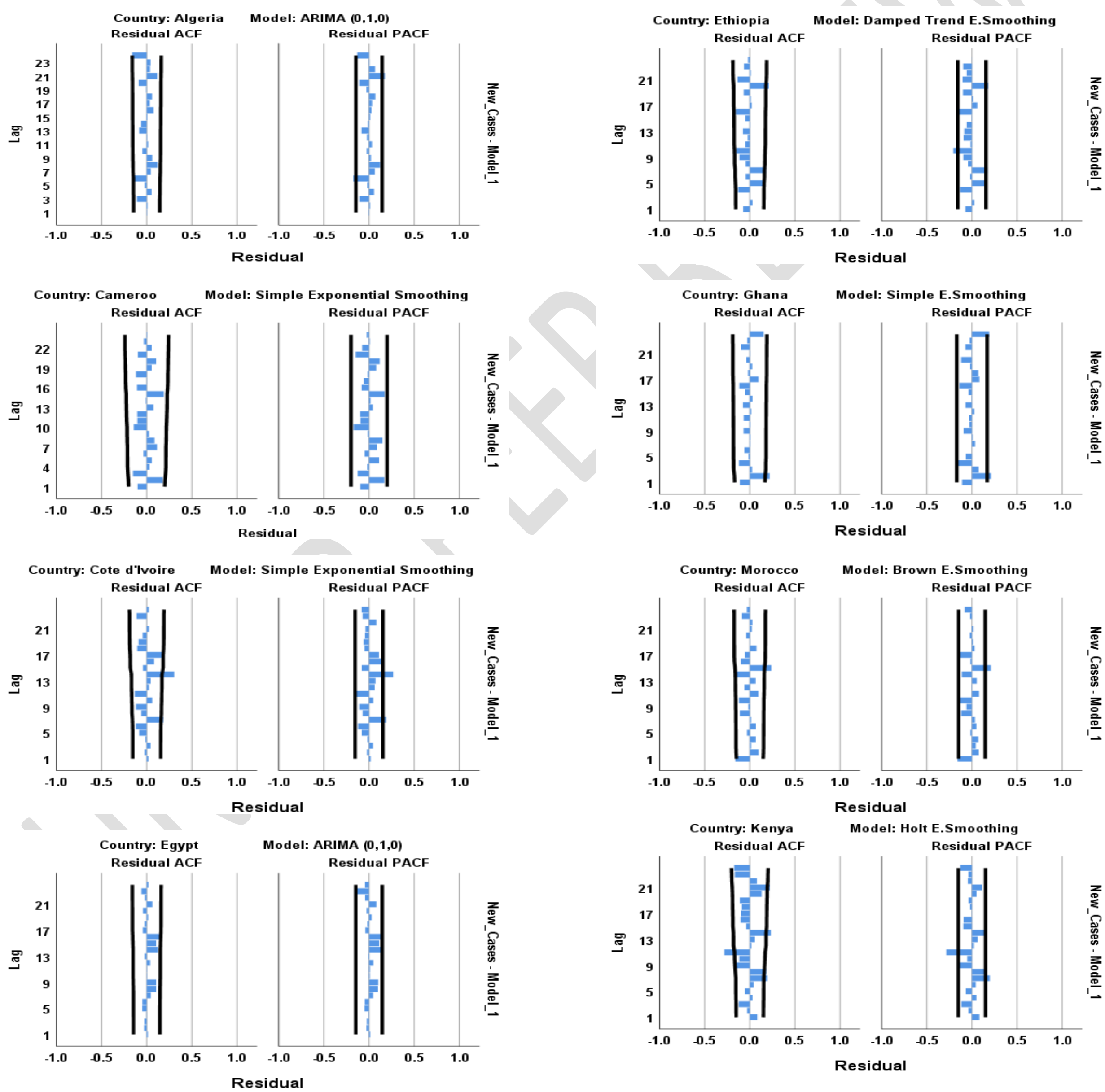
Forecasting and Trends of COVID-19 New Cases

In this study, the trends of COVID19 new cases declined only in Algeria and Ethiopia from September 7 to October 6, 2020. It was constant in Cameroon, Cote d'Ivoire, Ghana, and Nigeria. But, the trends was raised slightly in Egypt, Kenya, Morocco, and South Africa. And, from the highest to lowest forecasted COVID-19 new infections were 2,807 in Morocco , 2,444 in South Africa, 789 in Ethiopia, 285 in Kenya, 224 in Egypt, 172, 168 in Nigeria, 96 in Ghana, 59 in Cote d'Ivoire , and 25 in Algeria. Thus, Morocco, South Africa, and Ethiopia had the highest forecasted COVID-19 new infections on 6 October, 2020. The forecasted values and the trend curves of COVID-19 new infections results were presented in Table 2 and Figure 3.

Table 2. Forecast and trends of COVID-19 new cases for the next month for the countries for the countries from February 14 to September 06, 2020.

Country	Model Type		Forecasted Values and Trends Status			
		From Date	Values	Trends	LCL	UCL
Algeria	ARIMA (0,1,0)	Sep 7 to Oct 6	289 – 25	Decline	258	197
Cameroon	Simple	Sep 5 to Oct 4	172	Constant	0	726
Cote d'Ivr.	Simple	Sep 5 to Oct 4	59	Constant	0	356

Egypt	ARIMA (0,1,0)	Sep 7 to Oct 6	133 – 224	Increase	63	933
Ethiopia	Damped Trend	Sep 7 to Oct 6	948 – 789	Decline	636	2,141
Ghana	Simple	Sep 6 to Oct 5	96	Constant	0	1,022
Kenya	Holt	Sep 7 to Oct 6	169-285	Increase	3-0	1,222
Morocco	Brown	Sep 7 to Oct 6	1,900-2,807	Increase	1,287	3,858
Nigeria	Simple	Sep 7 to Oct 6	168	Constant	0	543
South Afr.	ARIMA (0,1,14)	Sep 7 to Oct 6	2,056-2,444	Increase	1,258	9,229



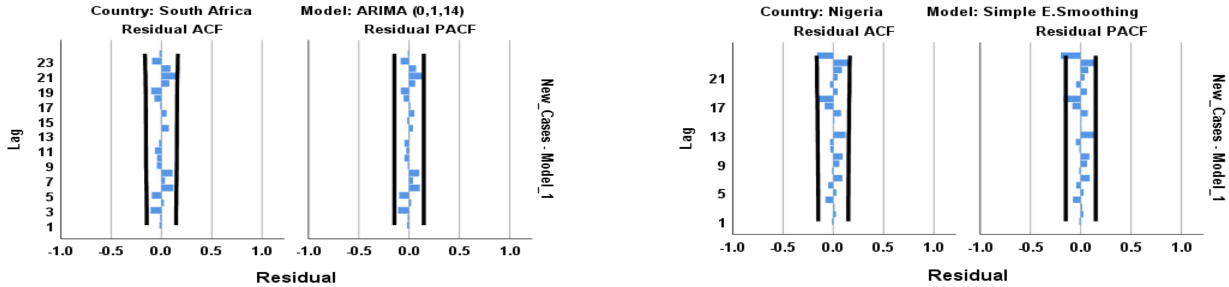
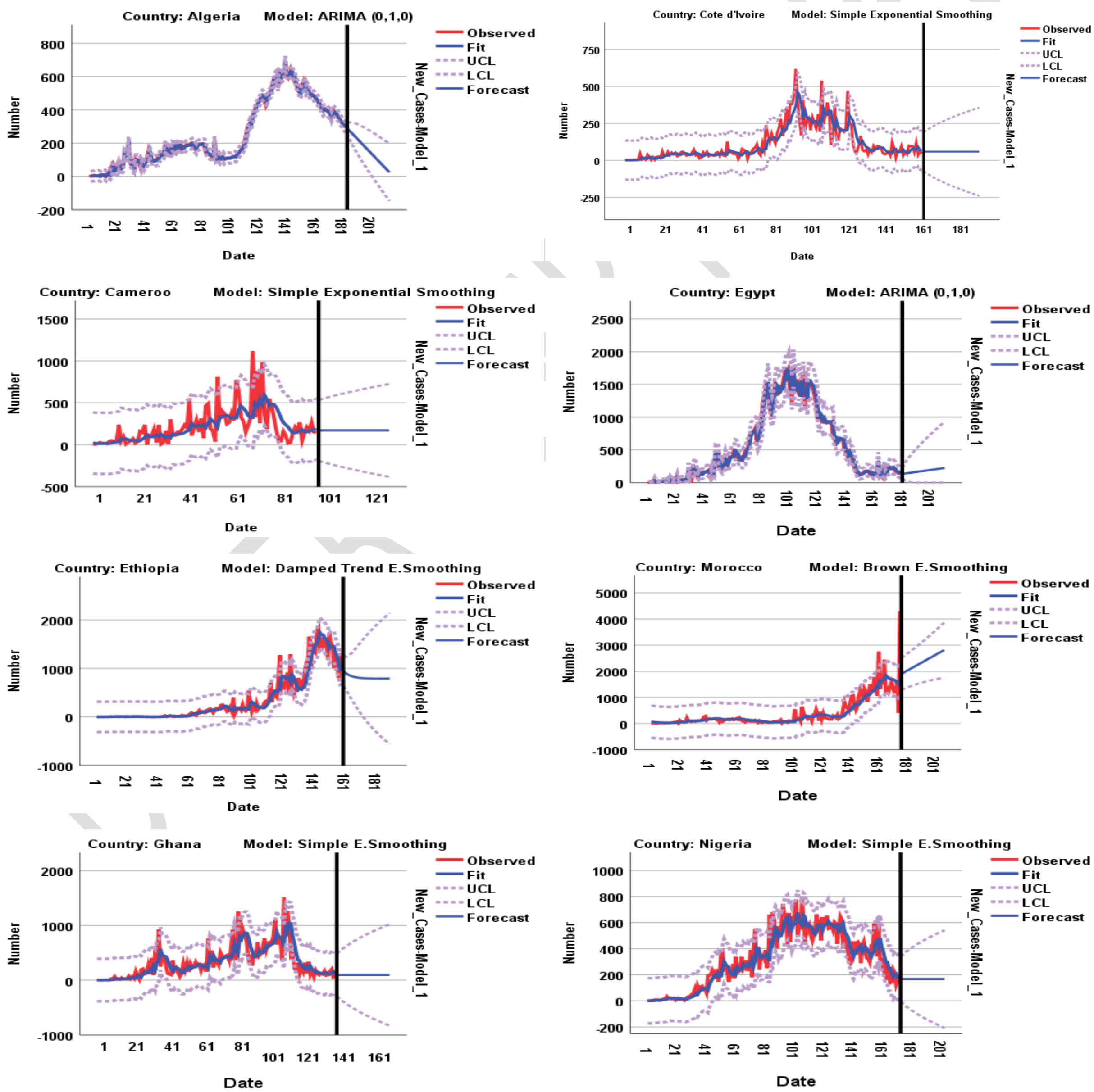


Figure 2. The graphs of ACF and PACF residuals



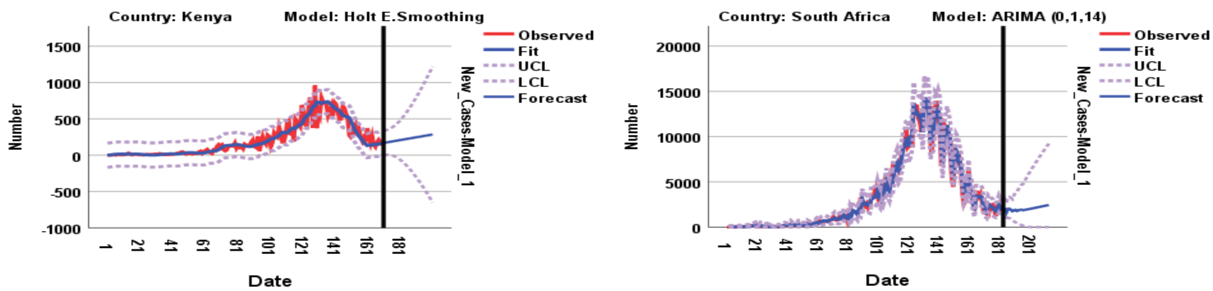


Figure 3.The fitting models and forecasts graphs of COVID-19 new cases for the countries from February 14 to September 06, 2020.

Discussions

The time series models for Algeria, Egypt, and South Africa COVID-19 new cases data fitted the ARIMA (0,1,0), ARIMA (0,1,0), and ARIMA (0,1,14), respectively. The Cameroon, Cote d'Ivoire, Ghana, and Nigeria data fitted the same model type (Simple exponential smoothing model). And, Ethiopia, Kenya, and Morocco data followed the Damped trend, Holt, and Brown exponential smoothing models, respectively.

A similar study was conducted using the four most affected African countries COVID-19 cases data in 2020. The study found 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 SA, 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 (Lukman et al., 2020).

A study used stochastic modelling for predicting COVID-19 prevalence in East Africa Countries in 2020. In the study, the optimal models for Ethiopia, Djibouti, Somalia and Sudan, were ARIMA (1,2,1), ARIMA (2,1,1) with drift,ARIMA (1,2,2), ARIMA (0,2,1) respectivel. In the views of worst-case and average-case scenarios, cumulated infection rate of COVID-19 were predicted by the end of October. And, rate of infection will expect to increase and magnify in Sudan and Ethiopia in the coming four month (Takele, 2020).

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. And, the number of COVID 19 epidemic cases data were modeled and forecasted that Japan (Holt Model), Germany (ARIMA (1,4,0) and France (ARIMA (0,1,3) were provided statistically significant. The UK (Holt Model), Canada (Holt Model), Italy (Holt Model), and Turkey (ARIMA (1,4,0) were not statistically significant (Yonar et al., 2020).

Likassa (2020) had shown 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 (Likassa, Xain, Tang, & Gobebo, 2021).

Achoki et al. (2020) found that the 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 2,581366 and 254,403 cumulative cases, respectively. In the Western Africa sub-region, cumulative cases of infection were dominated by Cote d'Ivoire 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 5778830 cases in May to 8,044,927 cases by the end of July (Achoki et al., 2020).

Study using African COVID-19 cases showed that the 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 the African region (Musa et al., 2020).

An 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 were 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 (Zhao et al., 2020).

A similar study by Ceylan (2020) 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 study shown that ARIMA models are suitable for predicting the prevalence of COVID-19 in the 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 (Ceylan, 2020).

Another similar study 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 (Dehesh et al., 2020).

Conclusions and Recommendations

Conclusions

Monthly prevalence of COVID-19 cumulative cases was declined in South Africa, Cote d′Ivoire, Egypt, Ghana, Cameron, Nigeria, and Algeria by 31%, 26%, 22%, 20%, 14%, 12%, and 4% from July to August, respectively. But, it was raised in Ethiopia, Morocco, and Kenya by 41%, 38%, and 1% from July to August, respectively. In the time series analysis; the Algeria, Egypt, and South Africa COVID-19 new cases data have fitted the ARIMA (0,1,0), ARIMA (0,1,0), and ARIMA (0,1,14) models, respectively. The Cameroon, Côte d’Ivoire, Ghana and Nigeria data have fitted the simple exponential smoothing models. Ethiopia, Kenya, and Morocco data have followed the Damped trend, Holt, and Brown exponential smoothing models, respectively. 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.

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 travelers 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 before the virus is distributed every where and attacked more and more. While these are necessary actions, the appropriate public health and social measures must be instituted on the ground. These measures include again, but are not limited to, early detection of suspect cases and tracing contacts to confirmed cases, case management, risk communication, and keeping up with infection prevention and control guidelines. In future studies, more data and healthier evaluations can be made as a matter of course. However, since this study provides information about the levels that the number of cases can reach if the course of the current situation cannot be intervened, it can guide countries to take the necessary measures again. And, the researcher recommended that cumulative cases of COVID-19 will be conducted.

COMPETING INTERESTS DISCLAIMER:

Authors have declared that no competing interests exist. The products used for this research are commonly and predominantly use products in our area of research and country. There is absolutely no conflict of interest between the authors and producers of the products because we do not intend to use these products as an avenue for any litigation but for the advancement of knowledge. Also, the research was not funded by the producing company rather it was funded by personal efforts of the authors.

References

Achoki, T., Alam, U., Were, L., Gebremedhin, T., Senkubuge, F., Lesego, A., . . . Kinfu, Y. (2020). COVID-19 pandemic in the African continent: forecasts of cumulative cases, new infections, and mortality. *MedRxiv*.

Armstrong, J. S. (2001). *Principles of forecasting: a handbook for researchers and practitioners* (Vol. 30): Springer.

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting*: Springer.

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

Dehesh, T., Mardani-Fard, H., & Dehesh, P. (2020). Forecasting of covid-19 confirmed cases in different countries with arima models. *MedRxiv*.

Gilbert, M., Pullano, G., Pinotti, F., Valdano, E., Poletto, C., Boëlle, P.-Y., . . . Altmann, M. (2020). Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. *The Lancet, 395*(10227), 871-877.

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.

Li, Q., Guo, N.-N., Han, Z.-Y., Zhang, Y.-B., Qi, S.-X., Xu, Y.-G., . . . Liu, Y.-Y. (2012). 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*, *87*(2), 364.

Likassa, H. T., Xain, W., Tang, X., & Gobebo, G. (2021). Predictive models on COVID 19: What Africans should do? *Infectious Disease Modelling*, *6*, 302-312.

Liu, P., Beeler, P., & Chakrabarty, R. K. (2020). COVID-19 progression timeline and effectiveness of response-to-spread interventions across the United States. *MedRxiv*.

Lukman, A. F., Rauf, R. I., Abiodun, O., Oludoun, O., Ayinde, K., & Ogundokun, R. O. (2020). COVID-19 prevalence estimation: Four most affected African countries. *Infectious Disease Modelling, 5*, 827-838.

Musa, S. S., Zhao, S., Wang, M. H., Habib, A. G., Mustapha, U. T., & He, D. (2020). Estimation of exponential growth rate and basic reproduction number of the coronavirus disease 2019 (COVID-19) in Africa. *Infectious diseases of poverty*, *9*(1), 1-6.

Petropoulos, F., Makridakis, S., & Stylianou, N. (2020). COVID-19: Forecasting confirmed cases and deaths with a simple time series model. *International journal of forecasting*.

Shi, Z., & Fang, Y. (2020). Temporal relationship between outbound traffic from Wuhan and the 2019 coronavirus disease (COVID-19) incidence in China. *MedRxiv*.

Takele, R. (2020). Stochastic modelling for predicting COVID-19 prevalence in East Africa Countries. *Infectious Disease Modelling*, *5*, 598-607.

Tseng, Y.-J., & Shih, Y.-L. (2020). Developing epidemic forecasting models to assist disease surveillance for influenza with electronic health records. *International Journal of Computers and Applications, 42*(6), 616-621.

WHO. (2020). World Health Organization coronavirus disease (COVID-19) dashboard. *World Health Organization*.

Worldometer. (2020). *CORONAVIRUS PANDEMIC REPORT ON SEPTEMBER 06, 2020.* Retrieved from

Worldometer. (2021). *COVID-19 CORONAVIRUS PANDEMIC REPORT ON APRIL 28, 2021*. Retrieved from

Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, *395*(10225), 689-697.

Yonar, H., Yonar, A., Tekindal, M. A., & Tekindal, M. (2020). 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*, *4*(2), 160-165.

Zhao, Z., Li, X., Liu, F., Zhu, G., Ma, C., & Wang, L. (2020). 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*, *729*, 138959.