Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Modeling Ivory Coast COVID-19 cases: Identification of a high-performance model for utilization

Ugochinyere Ihuoma Nwosu, Chukwudi Paul Obite*

Department of Statistics, Federal University of Technology Owerri, Nigeria

ARTICLE INFO

Keywords:
- Pandemic disease
- COVID-19 cases
- Econometric models
- ARIMA
- Early warning

ABSTRACT

This study modelled the reported daily cumulative confirmed, discharged and death Coronavirus disease 2019 (COVID-19) cases using six econometric models in simple, quadratic, cubic and quartic forms and an autoregressive integrated moving average (ARIMA) model. The models were compared employing R-squared and Root Mean Square Error (RMSE). The best model was used to forecast confirmed, discharged and death COVID-19 cases for October 2020 to February 2021. The predicted number of confirmed and death COVID-19 cases are alarming. Good planning and innovative approaches are required to prevent the forecasted alarming infection and death in Ivory Coast. The applications of findings of this study will ensure that the COVID-19 does not crush the Ivory Coast’s health, economic, social and political systems.

Introduction

The COVID-19 is a respiratory pandemic disease which has recently threatened and frightened the world with huge economic losses and disruption of global calendar as part of the consequences. The urgent need to offer guidance for policy and measurement using updated data to save Ivory Coast’s economy motivated the present study. With increased research and funding, new pathological conditions have been characterized to help people to understand the disease. According to Ayinde et al. [1] and Mayo Clinic [2], the symptoms of the disease are fever, dry cough and tiredness, difficulty in breathing or shortness of breath, chest pain, pressure and loss of speech or movement, headache, pains, debility, sore throat, diarrhea, conjunctivitis, loss of taste or smell, a rash on the skin and discoloration of fingers or toes. Recent findings have shown that type 1 diabetes, high blood pressure, asthma, liver disease, chronic lung diseases, brain and nervous system disorders, weakened immune system from bone marrow transplant, human immunodeficiency virus and some medications may increase the risk of serious COVID-19 illness [2]. The first outbreak of the novel disease was recorded in Wuhan city, Hubei Province, China in late December 2019 and was precisely reported to the World Health Organization (WHO) on December 31, 2019 [3]. At short duration of 30 days, the infection rate was alarming in China and have spread to other countries of the world including Thailand, Japan, Republic of Korea, Vietnam, Germany, United States and Singapore [1,4,5]. The WHO pronounced COVID-19 a disease of international concern on January 31, 2020 [6] and a pandemic on March 11, 2020 because of the alarming spread and severity in the world [7]. According to WHO global situation reports of September 19, 2020 (10:35 am CEST) there have been 30, 295, 744 confirmed cases of COVID-19, 947,933 deaths and more than 19.6 million patients have recovered. The COVID-19 has now been reported in at least 188 countries (with different degrees of spread, death and recovery), culminating in global calamity [8].

In Ivory Coast, the first COVID-19 confirmed case was reported on March 11, 2020 by the Ministry of Public Health and Hygiene of Ivory Coast when a forty-five-year-old Ivorian man who returned from Italy the week prior and presented himself to public health authorities after developing symptoms [9]. This made Ivory Coast the eighth country in sub-Saharan Africa and the twelfth country in Africa to have a confirmed case of COVID-19. This occurred after many suspected cases (including an Ivorian student who came back from China) were negative after tests were carried out between January and March 2020 [10]. According to Bone [9], due to the confirmed case, health authorities immediately began to trace contacts the man had before testing positive. His wife tested positive for COVID-19 the next day, making the total number of cases, two. Few days later (March 14, 2020), authorities published three additional confirmed cases and they were all Ivorian citizens, and two had traveled abroad (to Italy and France) in recent weeks. The third case was a health worker who had not travelled out of Ivory Coast in recent times, raising fears that community transmission was going on in the
country. This unexpected development alarmed citizens and forced many concerned groups into action. Result of this concern was the chronological pronouncement of lockdown policies by the Government of the country. This became necessary to ensure that COVID-19 does not overwhelm the country’s health systems, economics, politics and stability.

The lockdown policies strengthened surveillance in the country, controlled borders, restricted internal travels and movement, shutdown schools, restaurants, bars, nightclubs, cinemas and other places of entertainment, prohibited gatherings of more than 50 people and enforced use of hand sanitizers and wearing of face masks. The plot of Ivory Coast confirmed, recovered and death cases for the first 173 days is shown in Fig. 1 and Ivory Coast proportion of active, discharged and death cases for the first 173 days is shown in Fig. 2.

It was believed that social distance and sanitary measures would limit the spread of COVID-19 [9,10]. Despite these restrictions, the virus has continued to infect people, spread and caused varying degrees of health and economic problems in the country. Following increased pressure from citizens and an understanding of the disease phenomenon, recently, some of the lockdown policies were relaxed.

The COVID-19 is a novel disease that has not been entirely understood. There is a need to monitor/forecast important parameters of the pandemic. At the early stage of the pandemic, the unavailability of data made prognosis a difficult task [1]. In the light of data availability and urgency to identify efficient models to predict the epidemic and adopt measures that limit the spread, the present investigation sought to model the reported daily cumulative confirmed, discharged and death COVID-19 cases using twenty-five statistical model estimators to identify a high-performance/best model for utilization in forecast making.

Ndolane [16] proposed a stochastic model to predict the confirmed COVID-19 cases in the world, Sha et al. [17] also proposed a model. They developed a discrete-time stochastic epidemic model with binomial distributions to study the transmission of the disease. Their simulation result shows that the total confirmed cases will reach the peak around the end of February 2020 though the reported confirmed cases violated it as the number of confirmed cases continued to increase after the end of February 2020 and spread to different countries including Ivory Coast. Abdon [18] used Italy’s COVID-19 data as a guide to develop a new fractal-fractional differential operators taking into account the possibility of transmission of COVID-19 from dead bodies to humans and the effect of lockdown. The model showed that if lock-down recommendations are observed, the threat of COVID-19 can be reduced to zero within a short period of time. The log-cubic model was proposed to forecast the cumulative confirmed COVID-19 cases in China, Italy, South Korea and Turkey [19].

Methodology

Test models

We modelled the reported cumulative number of confirmed, discharged and death cases using different regression models so as to select the best model to forecast for future cases. The models considered are Linear Regression Model, Logarithm Regression Model, Inverse Regression Model, Compound Regression Model, S-Curve Regression Model and Growth Regression Model. The simple, quadratic, cubic and quartic form of each of the models was also considered to give twenty-four model estimators.

**Linear regression Model:** The simple form of the model is given as:

\[ y_t = \beta_0 + \beta_1 x_t + \epsilon_t \]  

The Quadratic Regression Model is given as:

\[ y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \epsilon_t \]  

The Cubic Regression Model is given as:

\[ y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \epsilon_t \]  

The Quartic Regression Model is given as:

\[ y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \beta_4 x_t^4 + \epsilon_t \]  

![Fig. 1. Ivory Coast confirmed, recovered and death cases for the first 173 days.](image-url)
Logarithm regression Model: The simple form of the model is given as:

\[ y_t = \beta_0 + \beta_1 \ln x_t + \varepsilon_t \]  

The Logarithm Quadratic Regression Model is given as:

\[ y_t = \beta_0 + \beta_1 \ln x_t + \beta_2 \ln x_t^2 + \varepsilon_t \]  

The Logarithm Cubic Regression Model is given as:

\[ y_t = \beta_0 + \beta_1 \ln x_t + \beta_2 \ln x_t^2 + \beta_3 \ln x_t^3 + \varepsilon_t \]  

The Logarithm Quartic Regression Model is given as:

\[ y_t = \beta_0 + \beta_1 \ln x_t + \beta_2 \ln x_t^2 + \beta_3 \ln x_t^3 + \beta_4 \ln x_t^4 + \varepsilon_t \]  

Inverse regression Model: The simple form of the model is given as:

\[ y_t = \exp(\beta_0 + \beta_1 \frac{1}{x_t}) \varepsilon_t \]  

Taking the \( \ln \) of both sides gives

\[ \ln y_t = \beta_0 + \beta_1 \frac{1}{x_t} + \ln \varepsilon_t \]  

Following the same technique, the Quadratic, Cubic and Quartic forms are respectively given as:

\[ \ln y_t = \ln \beta_0 + x_t \ln \beta_1 + x_t^2 \ln \beta_2 + \ln \varepsilon_t \]  

\[ \ln y_t = \ln \beta_0 + x_t \ln \beta_1 + x_t^2 \ln \beta_2 + \ln \varepsilon_t \]  

\[ \ln y_t = \ln \beta_0 + x_t \ln \beta_1 + x_t^2 \ln \beta_2 + x_t^3 \ln \beta_3 + \ln \varepsilon_t \]  

Growth regression Model: The simple form of the model is given as:

\[ y_t = \exp(\beta_0 + \beta_1 x_t) \varepsilon_t \]  

Taking the \( \ln \) of both sides gives

\[ \ln y_t = \beta_0 + \beta_1 x_t + \ln \varepsilon_t \]  

Following the same technique, the Quadratic, Cubic and Quartic forms are respectively given as:

\[ \ln y_t = \ln \beta_0 + x_t \ln \beta_1 + \beta_2 x_t + \ln \varepsilon_t \]  

\[ \ln y_t = \ln \beta_0 + x_t \ln \beta_1 + \beta_2 x_t + \beta_3 x_t^2 + \ln \varepsilon_t \]  

\[ \ln y_t = \ln \beta_0 + x_t \ln \beta_1 + \beta_2 x_t + \beta_3 x_t^2 + \beta_4 x_t^3 + \ln \varepsilon_t \]
\[ y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \text{Int} + \in_t \]  
\[ y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \text{Int} + \in_t \]  
\[
\text{where}
\]
\[ y_t \] is the confirmed, discharged, and death COVID-19 cases
\[ x_t \] is the time
\[ \beta_0, \beta_1, \beta_2, \beta_3 \] and \[ \beta_4 \] are the regression coefficients of the models
\[ \varepsilon_t \] is error term, where \[ \varepsilon_t \sim N(0, \sigma^2) \].

The regression coefficients were estimated using the ordinary least squares (OLS) method which minimizes the residual sum of squares. The efficiency of each of the models depends on the estimate of the regression coefficients. The OLS method depends on the normality and autocorrelation assumptions. The errors of the model are to follow a normal distribution and must be independent of each other. In a case when these assumptions are violated, OLS will not provide efficient estimates of the regression coefficients and most of the regression coefficients will be insignificant. This has made many researchers to develop other models such as Cochrane Orcutt [11], Least Absolute Deviation [12], etc. to solve the problem of OLS. The Runs test and Durbin Watson test developed by Durbin [13] were employed to test for the autocorrelation assumption while the Shapiro Wilk test developed by Shapiro and Wilk [14] was used to test for the normality assumption.

**Autoregressive integrated moving average (ARIMA) model**

The ARIMA model developed by Box and Jenkins [15] was also considered (the 25th model) and compared with the other linear models already stated to select the best model to forecast for future COVID-19 cases. The ARIMA model is applied when the data show evidence of non-stationarity, where an initial differencing step (corresponding to the “integrated” part of the model) can be applied one or more times to eliminate the non-stationarity. The AR part of ARIMA indicates that the evolving variable of interest is regressed on its own lagged (i.e., prior) values. The MA part indicates that the regression error is actually a linear combination of error terms whose values occurred contemporaneously and at various times in the past. The 1 (for “integrated”) indicates that the data values have been replaced with the difference between their values and the previous values (and this differencing process may have been performed more than once). The purpose of each of these features is to make the model fit the data as well as possible. The ARIMA models are generally denoted ARIMA (p, d, q) where parameters p, d, and q are non-negative integers, p is the order (number of time lags) of the autoregressive model, d is the degree of differencing (the number of times the data have had past values subtracted), and q is the order of the moving-average model. Given a time series data \[ y_t \], the Auto Regressive (AR) model is one where \[ y_t \] depends only on its own lags. That is, \[ y_t \] is a function of the “lags of \[ y_t \]”. The AR model is given below:

\[ y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + \varepsilon_t \]  
\[ y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + \varepsilon_t \]  
\[ y_{t-1} = \beta_1 y_{t-2} + \beta_2 y_{t-3} + \ldots + \beta_p y_{t-p} + \varepsilon_{t-1} \]  
\[ y_{t-1} = \beta_1 y_{t-2} + \beta_2 y_{t-3} + \ldots + \beta_p y_{t-p} + \varepsilon_{t-1} \]  

The ARIMA model is the combination of the AR and MA model (Eqs. (28) and (29)) respectively that have been differenced to make it stationary. The ARIMA model is given as:

\[ y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + \varepsilon_t \]  
\[ y_{t-1} = \beta_1 y_{t-2} + \beta_2 y_{t-3} + \ldots + \beta_p y_{t-p} + \varepsilon_{t-1} \]  

Predicted \[ y_t = \text{Constant} + \text{linear combination lags of } y \text{ (up to } p \text{ lags)} + \text{linear combination of lagged forecast errors (up to } q \text{ lags)}

Steps in choosing an ARIMA model

Finding the order of differencing (d) in the ARIMA (p, d, q) model

The purpose of differencing is to make the time series stationary. The right order of differencing is the minimum differencing required to get a time series to be stationary. We test for stationarity using the graphical method and the Augmented Dickey Fuller test.

Finding the order of the AR term (p) in the ARIMA (p, d, q) model

The next step is to identify if there is an AR term and the order it follows. We find out the order “p” of the AR model by inspecting the Partial Autocorrelation (PACF) plot.

The PACF is the correlation between the series and its lag, after excluding the contributions from the intermediate lags. It conveys the pure correlation between a lag and the series making it possible to know if the lag is needed in the AR term or not. The order of the AR term is equal to the number of lags that crosses the significance limit in the PACF plot.

Finding the order of the MA term (q) in the ARIMA (p, d, q) model

The next step is to identify the number of MA terms using the ACF plot just as we looked at the PACF plot.

Try the chosen model and use the Akaike Information Criterion (AIC) and Root Mean Square Error (RMSE) to search for a better model. Check the residuals for the chosen model by plotting the ACF of the residuals, and doing a test of the residuals if they are white noise. Once the residuals look like white noise, calculate the forecasts.

**Performance measures**

The R-squared and RMSE were used to select the best among the various models considered in forecasting future confirmed, recovered and death COVID-19 cases.

**Results and discussion**

**Confirmed COVID-19 cases**

The quartic form was preferred for all the regression models based on their R-squared and RMSE except for the logarithm and compound models that preferred the simple form as shown in Table 1.

The linear model was the best among the different regression models as it has the lowest RMSE and the highest R-squared. The Runs and Shapiro-Wilk tests showed that the errors of all the models failed to satisfy the autocorrelation and normality assumptions. The coefficients of the preferred model are given in Table 2. The problem of nonconformity to the OLS assumptions led us to use the Cochrane Orcutt, Least Absolute Deviation and ARIMA to address the problem to get better estimates as shown in Table 2.
The Cochrane Orcutt and Least Absolute Deviation models did not solve the assumption problem and still performed less than the preferred OLS model. The ARIMA model with order (8,2,10) was selected as the best model among all the models considered. It explains 99.98% of the variation in the confirmed COVID-19 cases. The model was used to forecast for July and August 2020 (testing set) and compared with the true values. The results were similar to the true values of the confirmed COVID-19 cases as reported in appendix A with a 90% R-squared value (about 90% of the variation in the forecasted confirmed COVID-19 cases is explained by the model). The ARIMA (8,2,10) model was further used to forecast for October 2020 to February 2021. It is expected that the number of confirmed COVID-19 cases at the end of October 2020 to February 2021 will be 41,521, 49,378, 57,493, 65,606 and 72,936 respectively as shown in Fig. 3. The expected number of confirmed COVID-19 cases is alarming. There is a need for serious tactical and robust approach to control the spread of COVID-19 in the country.

### Table 1
The preferred form for the different regression models for confirmed COVID-19 cases.

| Models   | Form   | R-squared | RMSE   | Runs test | SW test |
|----------|--------|-----------|--------|-----------|---------|
|          |        |           |        | Value     | P-value | Value     | P-value |
| Linear   | Quartic| 0.996     | 168,099| −9.111    | 0.000   | 0.884     | 0.000   |
| Logarithm| Simple | 0.644     | 1859,555| −10.209   | 0.000   | 0.748     | 0.000   |
| Inverse  | Quartic| 0.470     | 1848,696| −9.605    | 0.000   | 0.822     | 0.000   |
| Compound | Simple | 0.799     | 3450,223| −10.248   | 0.000   | 0.891     | 0.000   |
| S-Curve  | Quartic| 0.380     | 3450,528| −8.797    | 0.000   | 0.841     | 0.000   |
| Growth   | Simple | 0.799     | 3450,223| −10.248   | 0.000   | 0.891     | 0.000   |

*Note: (i) + implies that other forms of the model were not considered because of multicollinearity. (ii) Bold implies the most preferred model.*

### Table 2
Result of the different models for confirmed COVID-19 cases.

| Estimator | OLS | CORC | LAD | ARIMA(8,2,10) |
|-----------|-----|------|-----|--------------|
| β₀        | Estimate | 117.60 | 24412000.00 | 321.32 | ar1 | −1.06 | ma2 | 0.57 |
|           | Std. Error | 85.81 | 6589100.00 | 101.56 | Std. Error | 0.33 | Std. Error | 0.37 |
|           | T-value | 1.37 | 3.71 | 3.16 | ar2 | −1.15 | ma3 | 0.66 |
|           | P-value | 0.17 | 0.00 | 0.00 | Std. Error | 0.52 | Std. Error | 0.29 |
| β₁        | Estimate | −39.12 | 17088.00 | −70.74 | ar3 | −1.19 | ma4 | 0.47 |
|           | Std. Error | 10.42 | 46122.00 | 12.34 | Std. Error | 0.52 | Std. Error | 0.32 |
|           | T-value | −3.75 | 3.71 | −5.73 | ar4 | −1.08 | ma5 | 0.50 |
|           | P-value | 0.00 | 0.00 | 0.00 | Std. Error | 0.54 | Std. Error | 0.38 |
| β₂        | Estimate | 2.90 | 604.19 | 4.13 | ar5 | −1.03 | ma6 | 0.64 |
|           | Std. Error | 37.27 | 162.91 | 9.36 | Std. Error | 0.55 | Std. Error | 0.32 |
|           | T-value | 7.77 | 3.71 | 9.36 | ar6 | −1.07 | ma7 | 0.04 |
|           | P-value | 0.00 | 0.00 | 0.00 | Std. Error | 0.51 | Std. Error | 0.36 |
| β₃        | Estimate | −0.05 | 1.26 | −0.07 | ar7 | −0.39 | ma8 | −0.11 |
|           | Std. Error | 0.00 | 0.34 | 0.01 | Std. Error | 0.50 | Std. Error | 0.26 |
|           | T-value | −9.68 | 3.67 | −11.15 | ar8 | −0.02 | ma9 | 0.10 |
|           | P-value | 0.00 | 0.00 | 0.00 | Std. Error | 0.35 | Std. Error | 0.18 |
| β₄        | Estimate | 0.00 | 0.00 | 0.00 | Std. Error | 0.57 | ma10 | −0.41 |
|           | Std. Error | 0.00 | 0.00 | 0.00 | Std. Error | 0.32 | Std. Error | 0.17 |
|           | T-value | 13.27 | 3.81 | 14.31 |
|           | P-value | 0.00 | 0.00 | 0.00 |
| DW Stat value | 0.13 | 1.11 | 0.13 |
| P-value | 0.00 | 0.00 |
| SW Stat value | 0.88 | 0.95 | 0.79 |
| P-value | 0.00 | 0.00 |
| Training set R² | 0.9956 | 0.6460 | 0.9950 | 0.9998 |
| RMSE | 168.10 | 181.30 | 38.67 |
| Testing set R² | 0.75 | 32562.60 | 9.90 | 3569.83 |
the variation in the discharged COVID-19 cases. The model was used to forecast for July and August 2020 (testing set) and compared with the true values. The results were similar to the true values of the discharged COVID-19 cases as reported in appendix B with a 99% R-squared value (about 90% of the variation in the forecasted discharged COVID-19 cases is explained by the model). The ARIMA (2,2,3) model was further used to forecast for October 2020 to February 2021. It is expected that the number of discharged COVID-19 cases at the end of October 2020 to February 2021 will be 32,130, 39,025, 46,150, 53,275 and 59,711 respectively as shown in Fig. 4. The expected number of discharged COVID-19 cases is encouraging. If the existing measures for taking care of COVID-19 patients are maintained and improved upon, the spread of COVID-19 in the country will be controlled.

Death Covid-19 cases

After using the OLS method on the six regression models, the quartic and simple form were preferred for all the regression models in the death COVID-19 cases based on their highest R-squared and least RMSE. Three of the models – Linear, Inverse and S-Curve – preferred the quartic form while the other three models preferred the simple form as shown in Table 5.

The Linear regression model was the most preferred model though the Growth and Compound models performed well. The Runs test and the Shapiro-Wilk tests showed that all the models failed to satisfy both the normality and autocorrelation assumptions except for the linear model that satisfied the normality assumption (Shapiro-Wilk P-value > 0.05). The most preferred model had autocorrelated errors like the other models and this made us to consider the Cochrane Orcutt, Least Absolute Deviation and ARIMA models to address the problem. The estimate of the regression coefficients for all the models are shown in Table 6.

Conclusion

The present study identified a high-performance/best model (ARIMA) and was used to forecast the number of confirmed, discharged and death COVID-19 cases in Ivory Coast for October 2020 to February 2021. The expected numbers of confirmed and deaths COVID-19 cases are alarming. There is need to strictly adhere to the preventive measures already pronounced and as may be updated from time to time. Good planning, tactical and robust approaches with honest interventions are required to control the spread of COVID-19 in the country. The application of useful suggestions from this study (after early warning and risk assessment) will ensure that COVID-19 does not overwhelm the Ivory Coast’s health, economic, social and political systems.

Declaration

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Ugochinyere Ihuoma Nwosu: Conceptualization, Validation, Investigation, Data curation, Writing - original draft. Chukwudi Paul Obite: Methodology, Software, Visualization, Formal analysis, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Table 3

The preferred form for the different regression models for discharged COVID-19 cases.

| Models       | Form    | R square | RMSE  | Run test Value | P-value | SW test Value | P-value |
|--------------|---------|----------|-------|----------------|---------|---------------|---------|
| Linear       | Quartic | 0.998    | 55.65 | −7.804         | 0.000   | 0.978         | 0.078   |
| Logarithm    | Simple+ | 0.516    | 785.54| −9.970         | 0.000   | 0.815         | 0.000   |
| Inverse      | Quartic | 0.542    | 764.53| −9.362         | 0.000   | 0.866         | 0.000   |
| Compound     | Simple  | 0.864    | 1562.11| −10.004        | 0.000   | 0.947         | 0.000   |
| S-Curve      | Quartic | 0.483    | 1562.39| −8.247         | 0.000   | −0.889        | 0.000   |
| Growth       | Simple  | 0.864    | 1562.11| −8.247         | 0.000   | 0.889         | 0.000   |

Note: (i) + implies that other forms of the model were not considered because of multicollinearity. (ii) Bold implies the most preferred model.


Table 4

Result of the different models for discharged COVID-19 cases.

| Estimator | OLS | CORC | LAD | ARIMA(2,2,3) |
|-----------|-----|------|-----|-------------|
| $\beta_0$ Estimate | 83.99 | 402.34 | 117.83 | ar1 -0.09 |
| Std. Error | 29.17 | 243.36 | 44.55 | Std. error 0.21 |
| T-value | 2.88 | 1.65 | 2.64 | ar2 0.36 |
| P-value | 0.00 | 0.10 | 0.01 | Std. error 0.22 |
| $\beta_1$ Estimate | -24.47 | -52.50 | -28.08 | ma1 -0.68 |
| Std. Error | 3.71 | 22.27 | 5.66 | Std. error 0.17 |
| T-value | -6.60 | -2.36 | -4.96 | ma2 -0.60 |
| P-value | 0.00 | 0.02 | 0.00 | Std. error 0.19 |
| $\beta_2$ Estimate | 1.43 | 2.25 | 1.56 | ma3 0.66 |
| Std. Error | 0.14 | 0.67 | 0.21 | Std. error 0.14 |
| T-value | 10.31 | 3.37 | 7.38 | |
| P-value | 0.00 | 0.00 | 0.00 | |
| $\beta_3$ Estimate | -0.02 | -0.03 | -0.02 | |
| Std. Error | 0.00 | 0.01 | 0.00 | |
| T-value | -10.50 | -3.75 | -7.48 | |
| P-value | 0.00 | 0.00 | 0.00 | |
| $\beta_4$ Estimate | 1.16E-04 | 1.56E-04 | 1.00E-04 | |
| Std. Error | 0.00 | 0.00 | 0.00 | |
| T-value | 13.07 | 4.73 | 9.10 | |
| P-value | 0.00 | 0.00 | 0.00 | |
| DW Stat value | 0.29 | 1.74 | 0.28 | |
| P-value | 0.00 | 1.74 | 0.28 | |
| SW Stat value | 0.98 | 0.92 | 0.98 | |
| P-value | 0.08 | 0.00 | 0.04 | |
| Training set R2 | 0.9976 | 0.9790 | 0.9975 | 0.9992 |
| RMSE | 55.65 | 85.66 | 57.22 | 31.50 |
| Testing set R2 | 0.91 | 0.99 | 0.99 | 0.99 |
| RMSE | 6244.88 | | 751.72 | |

Fig. 4. Discharged Covid-19 cases for March to June 2020 and its forecast for July 2020 to February 2021.

Appendix A. True value and forecast of confirmed COVID-19 cases for July and August 2020

| Date  | TV       | Point Forecast | Lo 80   | Hi 80   | Lo 95   | Hi 95   |
|-------|----------|----------------|---------|---------|---------|---------|
| 1-Jul | 9703     | 9833.267       | 9782.042| 9884.492| 9754.925| 9911.609|
| 2-Jul | 9993     | 10090.254      | 9997.642| 10182.87| 9948.616| 10231.89|
| 3-Jul | 10,245   | 10430.878      | 10294.22| 10567.53| 10221.88| 10639.87|
| 4-Jul | 10,463   | 10688.604      | 10501.13| 10876.08| 10401.89| 10975.32|
| 5-Jul | 10,773   | 10935.916      | 10693.92| 11177.91| 10565.81| 11306.02|
| 6-Jul | 10,967   | 11085.54       | 10784.46| 11386.62| 10625.07| 11546.01|
| 7-Jul | 11,195   | 11348.89       | 10982.67| 11715.11| 10788.81| 11908.97|

(continued on next page)
Table 5
The preferred form for the different regression models for death COVID-19 cases.

| Models    | Form     | R square | RMSE  |
|-----------|----------|----------|-------|
| Linear    | Quartic  | 0.996    | 1.17  |
| Logarithm | Simple+  | 0.690    | 10.29 |
| Inverse   | Quartic  | 0.733    | 9.55  |
| Compound  | Simple   | 0.961    | 29.02 |
| S-Curve   | Quartic  | 0.656    | 29.10 |
| Growth    | Simple   | 0.961    | 29.02 |

| Run test | SW test |
|----------|---------|
| Value    | P-value |
| −4.525   | 0.000   |
| −9.298   | 0.000   |
| −8.645   | 0.000   |
| −9.332   | 0.000   |
| −7.252   | 0.000   |
| −9.332   | 0.000   |

Note: (i) + implies that other forms of the model were not considered because of multicollinearity. (ii) Bold implies the most preferred model.

Table 6
Result of the different models for death COVID-19 cases.

| Estimator | Estimation Methods | OLS | CORC | LAD | ARIMA(4,2,1) |
|-----------|--------------------|-----|------|-----|--------------|
| β0        | Estimate           | 0.68| −0.49| 1.0326| ar1          |
|           | Std. Error         | 0.66| 2.46 | 0.9736| std. error   |
|           | T-value            | 1.03| −0.20| 1.0605| ar2          |
|           | P-value            | 0.31| 0.84 | 0.2889| std. error   |
| β1        | Estimate           | −0.02| 0.14| −0.0987| ar3          |
|           | Std. Error         | 0.10| 0.30 | 0.1404| std. error   |
|           | T-value            | −0.23| 0.48| −0.7027| ar4          |
|           | P-value            | 0.62| 0.64 | 0.4822| std. error   |
| β2        | Estimate           | 0.03| 0.02 | 0.0309| ma1          |
|           | Std. Error         | 0.00| 0.01 | 0.0060| std. error   |
|           | T-value            | 6.88| 1.80 | 5.1794|                |
|           | P-value            | 0.00| 0.08 | 0.0000|                |
| β3        | Estimate           | −5.09E-04| −3.96E-04| 0.0006|
|           | Std. Error         | 0.00| 0.00 | 0.0001|                |
|           | T-value            | −7.97| −2.33| −5.8713|                |
|           | P-value            | 0.00| 0.02 | 0.0000|                |
| β4        | Estimate           | 3.18E-06| 2.58E-06| 8.9E-07|
|           | Std. Error         | 0.00| 0.00 | 0.0000|                |
|           | T-value            | 9.54| 3.07 | 6.8953|                |
|           | P-value            | 0.00| 0.00 | 0.0000|                |

| DW Stat value | P-value | SW Stat value | P-value |
|---------------|---------|---------------|---------|
| 0.29          | 0.00    | 0.98          | 0.08    |
| 1.95          | 0.64    | 0.98          | 0.98    |
| 0.01          | 0.10    |               |         |

| Training set | R2      | RMSE          | Lo 80   | Hi 80   | Lo 95   | Hi 95   |
|--------------|---------|---------------|---------|---------|---------|---------|
| 0.9960       | 0.9958  | 0.9960        | 0.9973  |         |         |         |
| 1.166        | 1.20    | 1.175         | 0.98    |         |         |         |
| 0.82         | 0.96    |               |         |         |         |         |

| Testing set  | R2      | RMSE          | Lo 80   | Hi 80   | Lo 95   | Hi 95   |
|--------------|---------|---------------|---------|---------|---------|---------|
| 214.58       | 18.09   |               |         |         |         |         |

Fig. 5. Death Covid-19 cases for March to June 2020 and its forecast for July 2020 to February 2021.
Appendix B. True value and forecast of discharged COVID-19 cases for July and August 2020

| Date       | TV     | Point Forecast | Lo 80  | Hi 80  | Lo 95  | Hi 95  |
|------------|--------|----------------|--------|--------|--------|--------|
| 1-Jul      | 9703   | 4443.335       | 4402.78| 4484.291| 4381.206| 4505.865|
| 2-Jul      | 9993   | 4603.329       | 4538.665| 4607.994| 4504.434| 4702.225|
| 3-Jul      | 10,245 | 4818.146       | 4734.887| 4901.405| 4690.812| 4945.48|
| 4-Jul      | 10,463 | 5024.268       | 4914.802| 5133.734| 4856.854| 5191.682|
| 5-Jul      | 10,773 | 5250.829       | 5111.825| 5389.833| 5038.241| 5463.417|
| 6-Jul      | 10,967 | 5472.479       | 5327.378| 5534.24| 5287.93| 5638.96|
| 7-Jul      | 11,195 | 5701.87        | 5486.552| 5917.188| 5375.569| 5817.01|
| 8-Jul      | 11,505 | 5928.821       | 5668.375| 6189.268| 5530.503| 6327.14|
| 9-Jul      | 11,751 | 6158.756       | 5849.46| 6468.052| 5685.729| 6631.783|
| 10-Jul     | 12,053 | 6387.555       | 6025.571| 6749.54| 5833.947| 6941.163|
| 11-Jul     | 12,444 | 6617.521       | 6199.587| 7035.455| 5978.346| 7256.696|
| 12-Jul     | 12,767 | 6846.987       | 7324.064| 6173.34| 7576.617| 6862.66|
| 13-Jul     | 12,873 | 7076.898       | 6537.764| 7616.031| 6252.364| 7901.432|
| 14-Jul     | 13,038 | 7306.594       | 6702.609| 7910.579| 6382.878| 8230.31|
| 15-Jul     | 13,404 | 7536.475       | 6865.015| 8207.936| 6509.565| 8563.386|
| 16-Jul     | 13,555 | 7766.261       | 7024.791| 8507.73| 6632.281| 8900.241|
| 17-Jul     | 13,657 | 7996.121       | 7182.227| 8810.014| 6751.378| 9240.863|
| 18-Jul     | 13,913 | 8225.94       | 7337.286| 9114.593| 6866.862| 9585.018|
| 19-Jul     | 14,120 | 8455.789       | 7490.124| 9421.454| 6978.932| 9932.646|
| 20-Jul     | 14,313 | 8685.621       | 7640.761| 9730.482| 7087.645| 10283.6|
| 21-Jul     | 14,532 | 8915.466       | 7789.292| 10041.64| 7193.132| 10638.7|
| 22-Jul     | 14,734 | 9145.303       | 7935.754| 10354.85| 7295.458| 10995.15|
| 23-Jul     | 15,002 | 9375.145       | 8080.215| 10670.08| 7394.72| 11355.57|
| 24-Jul     | 15,254 | 9604.984       | 8222.713| 10987.26| 7490.982| 11718.99|
| 25-Jul     | 15,495 | 9834.825       | 8363.299| 11306.35| 7584.321| 12085.33|

(continued on next page)
### Appendix C. True value and forecast of death COVID-19 cases for July and August 2020.

| Date     | TV    | Forecast | Lo 80 | Hi 80 | Lo 95 | Hi 95 |
|----------|-------|----------|-------|-------|-------|-------|
| 1-Jul    | 68    | 69,45784 | 68,19171 | 70,72398 | 67,52145 | 71,39423 |
| 2-Jul    | 68    | 70,9912  | 69,20611 | 72,77629 | 68,26114 | 73,72126 |
| 3-Jul    | 70    | 72,31262 | 70,04306 | 74,58218 | 68,84163 | 75,78361 |
| 4-Jul    | 72    | 73,69977 | 71,03287 | 76,36686 | 69,6211 | 77,77845 |
| 5-Jul    | 74    | 75,07591 | 71,99777 | 78,15406 | 70,36827 | 79,78354 |
| 6-Jul    | 75    | 76,47427 | 72,98103 | 79,68573 | 71,13818 | 81,81672 |
| 7-Jul    | 76    | 77,86486 | 73,94336 | 81,78635 | 71,86745 | 83,82267 |
| 8-Jul    | 78    | 79,25696 | 74,90051 | 83,61341 | 72,59434 | 85,91958 |
| 9-Jul    | 79    | 80,64671 | 75,84639 | 85,44703 | 73,30525 | 87,98817 |
| 10-Jul   | 81    | 82,03737 | 76,78457 | 87,29016 | 74,00391 | 90,07082 |
| 11-Jul   | 82    | 83,42783 | 77,71298 | 89,14267 | 74,68773 | 91,16793 |
| 12-Jul   | 84    | 84,81854 | 78,63209 | 90,01948 | 75,35719 | 94,27988 |
| 13-Jul   | 84    | 86,20914 | 79,54141 | 92,87687 | 76,01733 | 96,40655 |
| 14-Jul   | 87    | 87,59977 | 80,44114 | 94,7584 | 76,65159 | 98,54795 |
| 15-Jul   | 87    | 88,99037 | 81,33124 | 96,64953 | 77,27674 | 100,7004 |
| 16-Jul   | 87    | 90,38099 | 82,21181 | 98,55016 | 77,88731 | 102,85774 |
| 17-Jul   | 87    | 91,7716  | 83,08291 | 100,4603 | 78,48343 | 105,0599 |
| 18-Jul   | 91    | 93,16221 | 83,94464 | 102,3798 | 79,06515 | 107,2593 |
| 19-Jul   | 92    | 94,55283 | 84,79709 | 104,3086 | 79,63272 | 109,4729 |
| 20-Jul   | 92    | 95,94344 | 85,64036 | 106,2465 | 80,18624 | 111,7006 |
| 21-Jul   | 93    | 97,33405 | 86,47456 | 108,1935 | 80,72589 | 113,9422 |
| 22-Jul   | 93    | 98,72467 | 87,29979 | 110,1495 | 81,25183 | 116,1975 |
| 23-Jul   | 93    | 100,11522 | 88,11617 | 112,1144 | 81,76422 | 118,4663 |
| 24-Jul   | 94    | 101,50589 | 89,82379 | 114,088 | 82,36222 | 120,7486 |
| 25-Jul   | 94    | 102,89651 | 89,72275 | 116,0703 | 82,78499 | 123,044 |
| 26-Jul   | 96    | 104,28712 | 90,51318 | 118,0611 | 83,22169 | 125,3526 |
| 27-Jul   | 96    | 105,67773 | 91,29515 | 120,0603 | 83,86148 | 127,674 |
| 28-Jul   | 98    | 107,08635 | 92,06879 | 122,0679 | 84,12850 | 129,8404 |
| 29-Jul   | 99    | 108,45896 | 92,83417 | 124,0838 | 84,5692 | 132,355 |
| 30-Jul   | 100   | 109,84579 | 93,59114 | 126,1078 | 84,98464 | 134,7143 |
| 31-Jul   | 102   | 111,24019 | 94,34056 | 128,1398 | 85,39444 | 137,0859 |
| 1-Aug    | 102   | 112,63086 | 95,08176 | 130,1798 | 85,79186 | 139,4697 |
| 2-Aug    | 102   | 114,02141 | 95,81507 | 132,2278 | 86,17722 | 141,8656 |
| 3-Aug    | 102   | 115,41203 | 96,54058 | 134,2835 | 86,55065 | 144,2734 |

(continued on next page)
Appendix D. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rinp.2020.103763.

References

[1] Ayinde K, Lukman AF, Rauf IR, Alabi OO, Okon CE, Ayinde OE. Modeling Nigerian Covid-19 cases: a comparative analysis of models and estimators. Chaos Solitons Fractals 2020;138:109911.

[2] Mayo Clinic (2020). Coronavirus disease 2019 (COVID-19). Accessed September 19, 2020.

[3] Reynolds, M. (2020).

[4] WHO (2020). World Health Organization Novel Coronavirus (2019-nCoV). Available at https://www.who.int/20200425-sitrep-96-19.pdf. Accessed August 31, 2020.

[5] Wu Y-C, Chen C-S, Chan Y-J. The outbreak of COVID-19: an overview. J Chin Med Council, Tuesday, March 31, 2020. https://www.atlanticcouncil.org/blogs/africa-visions/coronavirus/symptoms-causes/syc-20479963.

[6] Bedford J, Enria D, Giesecke J, Heymann DL, Ihekweazu C, Kobinger G. COVID-19: a statistical modeling of the course of COVID-19 (SARS-CoV-2) outbreak: a comparative analysis. Asia Pac J Public Health 2020;32(4):157–60. https://doi.org/10.1177/1010539520928180.

[7] Sha H, Sanyi T, Libin R. A discrete stochastic model of the COVID-19 outbreak: forecast and control (2020). Mathemat Biosci Eng 2020;17(4):2792–804. https://doi.org/10.3934/mbe.2020153.

[8] Al Jazeera News (2020). Which countries have not reported any coronavirus cases?

[9] Bone, M. (2020). Watching the spread of coronavirus in Cote d ’Ivoire. Atlantic Council, Tuesday, March 31, 2020. https://www.atlanticcouncil.org/blogs/africa-source/watching-the-spread-of-coronavirus-in-cote-divoire/, https://www.aljazeera.com/news/2020/04/countries-reported-coronavirus-cases-200412093314762.html. Accessed September 19, 2020. https://www.mayoclinic.org/diseases-conditions/coronavirus/symptoms-causes/syc-20479963.

[10] Anjorin AbdulAzeezA. The coronavirus disease 2019 (COVID-19) pandemic: a review and an update on cases in Africa. Asian Pac J Trop Med 2020;13(5):199. https://doi.org/10.1016/S0140-6736(20)31804.https://doi.org/10.1177/1010539520928180.

[11] Cochrane D, Orcutt GH. Application of least square to relationship containing autocorrelated error terms. J Am Stat Assoc 1949;44:32–61.

[12] Taylor LD. Estimation by minimizing the sum of absolute errors. Frontiers of Econometrics. New York: Academic Press; 1974.

[13] Durbin J. Estimation of parameters in time-series regression models. J Roy Stat Soc: Ser B (Methodol) 1960;22(1):139–61.

[14] Shapiro SS, Wilk MB. An analysis of variance test for normality. Biometrika 1965;52(3):591–611.

[15] Box GEP, Jenkins GM. Time Series Analysis: Forecasting and control (2020). Mathemat Biosci Eng 2020;17(4):2792–804. https://doi.org/10.3934/mbe.2020153.

[16] Atangana A. Modelling the spread of COVID-19 with new fractal-fractional operators: can the lockdown save mankind before vaccination? Chaos Solitons Fractals 2020;136:109860. https://doi.org/10.1016/j.chaos.2020.109860.

[17] Ankarali H, Ankarali S, Caskurlu H, Cag Y, Arslan F, Erdem H, Vahaboglu H. Estimation by minimizing the sum of absolute errors. Frontiers of Econometrics. New York: Academic Press; 1974.