Peertechz



Bioinformatics and Biostatistics

DOI: https://dx.doi.org/10.17352/ojbb

Research Article

Forecast number of new cases of Corona Virus Disease (COVID-19) in Ethiopia, using the case-based autoregressive integrated moving average model

Alemu Bekele Eticha*

Department of Statistics, College of Natural and Computational Science, Mizan-Tepi University, Tepi, Ethiopia

Received: 21 October, 2020 Accepted: 30 December, 2020 Published: 31 December, 2020

*Corresponding author: Alemu Bekele Eticha, Department of Statistics, College of Natural and Computational Science, Mizan-Tepi University, Tepi, Ethiopia, E-mail: alemubekele1@gmail.com

Keywords: COVID-19 trend; Time series forecaster of new case of Covid-19; ARIMA (2,2,2); New case confirmed of Corona virus

https://www.peertechz.com



Abstract

After the initial outbreak in Ethiopia, the dispersion of SARS-CoV-2 is elevated number of cases. Literally, reported results for confirmed cases peaked in August 2020 and declined after that time, as evidenced by the contestd responses that have invested in pandemic control in the country.

ARIMA models are a most widely used approaches to time series forecasting and provide harmonizing approaches to the problem of forecasting. ARIMA models aim to describe autocorrelations in the data. Thus, in this study, the Autoregressive Integrated Moving Average (ARIMA) method is used to predict the number of new coronavirus cases. In short, Auto regression uses the dependent relationship between observation and lagged observations; Integrated using the difference in raw observations; and Moving Average relies on the dependency between observation and residual error.

ARIMA (2, 2, 2) predicts the number of confirmed cases of COVID-19, based on the period between March 2020 and December 2020 at 95% confidence intervals. The result revealed that the maximum expected new case per day was 807 and the minimum forecast was 410 cases per day in the next two months. In addition, the total number of confirmed COVID-19 expected cases could reach about 160585 by end-February 2021.

In general, if the government of Ethiopia stops controlling the COVID-19 mechanisms the pandemic may relapse severely and affect the country more. The study therefore proposed that the constructive stepladder incorporates control mechanisms. Thus, depending on the results of the report, all the organizations involved will establish policies.

Introduction

It is understood that human coronavirus infections are responsible for mild respiratory diseases. The three global coronavirus outbreaks that contributed to significant mortality and morbidity were SARS CoV-1, MERS-CoV and SARS-CoV2. The first outbreak of the twenty first century caused by coronavirus was SARS CoV-1.

As the most common, SARS COV-1 infection had a wide range of respiratory and gastrointestinal symptoms and the

last confirmed case was reported in 2004. The second outbreak in 2012 was caused by Middle East respiratory coronavirus syndrome (MERS-CoV), and case fatality was substantially higher than that of SARS-COV-1. MERS-CoV has a wide variety of mild, moderate to serious physical symptoms and some patients have acute respiratory distress syndrome. On December 2019 the third and most recent outbreak of serious acute respiratory syndrome coronavirus-2 (SARS-CoV-2) began, leading to a global pandemic. Patients with SARS-CoV2 infection may be asymptomatic or have the most typical signs

017

of fever, cough, and shortness of breath. SARS-Cov-2, literally called COVID-19 [1], was the subject of this study.

The latest outbreak (COVID-19) became a global pandemic as of 4 May 2020 and is still ongoing [2]. The pandemic was identified in Wuhan, China in December 2019 [3]. Later, the rapid spread of this pandemic from corner to corner was a sudden shock to the entire planet. As of 31 December 2020, more than 83.8 million cases of COVID-19 have been registered in about 2010 countries and regions, resulting in more than 1.8 million deaths, while more than 59.3 million people have recovered [4].

In Africa, the first case of coronavirus was reported in Egypt on 14 February 2020. As of 31 December 2020, the five top African countries reporting most cases of Covid–19 were South Africa (1,057,161), Morocco (439,193), Tunisia (139,140), Egypt (138,062) and Ethiopia (124,264). By the end of 2020, the number of confirmed cases reported from Ethiopia (124,264) and the number of deaths reach 1923. Thus, Ethiopia is one of the countries that have been seriously affected by coronavirus disease (COVID–19) since the first case was identified on 13 March 2020. Then, to date, the country's case report passes one hundred and twenty–four thousand. At the end of 2020, over 124264 confirmed COVID–19 cases were recorded from Ethiopia resulted in over 1923 deaths, with more than 112096 recoveries [4,5].

While well known, interventions such as hand washing, maintaining social distances and wearing face masks suggested by public health workers to monitor the spread of coronavirus, it's transmission never stopped still in Ethiopia [6]. In the first three months, the case confirmed, and the rate of death in the country grew over the initial three months [7,8]. Since COVID-19 continues to be a global public health concern and requires the utmost effort to track the prevalence of the virus, it is argued that the fundamental prerequisite for successful monitoring of the prevalence rate is the use of pandemic predictive methods [9]. Thus, the study predicted a new case of COVID-19 in Ethiopia for the next 60 days with the goal of notifying all concerned entities to monitor the prevalence, and hazard of this pandemic.

Methods

Data set

Across the world, researchers and policy makers are looking at confirmed cases and deaths to understand and compare the spread of the COVID-19. In this study, data of COVID-19 were extracted and organized from the daily report of the public health institute until 31 December 2020. The data presented and analyzed throughout Minitab Version 13. The time series forecasting method was used to show the country's future 60-day. The study design was a registry-based trend series analysis from the Federal Ministry of Health on a daily basis from March to the end of 2020.

Analytical methods

ARIMA modeling is one of the best modeling techniques

in the time series [2,9]. ARIMA methods with the help of a number of parameters and a model expressed as ARIMA (p, d, q). Here, p stands for the order of self-regression, d stands for the degree of difference in trends, and q stands for the order of the moving average [10]. For the confirmed COVID-19 cases in Ethiopia, I have used ARIMA (2,2,2) technique. The model for predicting possible reported cases of COVID-19 was shown as;

$$ARIMA(p,d,f) = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + Z_t$$
$$Z_t = X_t - X_{t-1}$$

Where X (t) is the average number of confirmed cases of COVID-19 at ith day; parameters are α_i , α_2 , β_i and β_2 while Zt is the residual term for ith day. In conjunction with already verified events, the trend analysis can be estimated and the time series analysis conducted for this purpose. The prediction of forecast refers to the application of a formula to estimate future figures on the basis of previous results. The ARIMA (2, 2, 2) was used in this study to identify patterns in reported cases of COVID-19 in Ethiopia based on data 2020. The statistical significance amount is set at 0.05. A diagram shows confirming cases regarding time to verify the model's efficiency [11], for actual confirmed cases.

Results

Descriptive statistics

The description of the corona virus spread in Ethiopia by months was displayed on Table 1 and Figure 1.The Ethiopian Federal Ministry of Health reported a first COVID-19 case on 13 March 2020. By the end of the month, there were 25 confirmed cases, two recoveries, and no deaths, leaving 23 positive cases of pandemics in April 2020. By the end of April, there had been 106 new cases and bringing the total number of confirmed cases to 131. There were three deaths. The number of people recovered increased to 59, leaving 69 active cases of pandemic at end of that month.

In May, 1041 new cases were identified as a pandemic, taking the total number of cases to 1172. Only eleven people died. The number of recovered patients rose to 209, leaving 952 active cases of coronavirus. There were 4,674 new cases in June, taking the total number of confirmed cases to 5846. The death toll has risen to 103. The number of people recovered rose to 2430, leaving 3313 active cases at the end of June. In July, there were 11,684 new cases, bringing the total number of confirmed cases to 17,530. The death toll has risen to 274. The number of people recovered increased to 6950. At the end of the month, there were 10,306 active cases. There were also 34,601 new cases in August, increasing the total number of confirmed cases to 52,131. The number of deaths nearly tripled to 809. At the end of the August there were 32,328 active cases of pandemic.

There were 23,237 new cases in September, taking the total number of confirmed cases to 75,368. The death toll increased to 1198. The number of people recovered increased to 30,952, leaving 42,441 active pandemic cases at the end of the month.

6

Table 1: The descriptive result of corona virus cases, deaths, and recovery of ten months, 2020.

· ·			-	-		-					
Months	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Total
New Cases	25	106	1041	4674	11684	34601	23237	20801	13905	14190	124264
New deaths	0	3	8	92	171	535	389	271	237	217	1923
Recovery	2	57	150	2221	4520	11792	12210	21701	21162	38281	112096

Similarly, 20801 new cases and 271 additional deaths occurred in the end of October. The event of a pandemic increased to 96,169, and the death toll increased to 1469.

There were 13905, new cases in November, raising the total number of confirmed cases to 110,074. The death toll rose to 1706. The number of recovered patients increased to 73,815, at the end of the month. In the last month of 2020, there were 14190 new cases and 217 additional deaths. The pandemic case increased to 124,264, and the death toll rose to 1,923. The total number of new cases, deaths and recoveries was shown in Table 1.

The line of new case report decreased after it peaked in August 2020. The pattern of case is not obvious from the fuigure1. The death chart is somewhat constant while the recovery chart show increase with month.

Measures of model accuracy

In this study, the time series model encompasses to forecast COVID-19 cases in the coming 60 days. The results for the measure of model accuracy for ARIMA, Linear Trend, Quadratic Linear, and S-Curve Trend, Moving Average, and Exponential model had displayed in Table 2. Look at the mean absolute percent error (MAPE), mean absolute deviation (MAD), and the mean square of deviation (MSD) values suggest that ARIMA(2,2,2) is the most accurate of all for forecasting future values as it possesses the least point for all the measures of the models.

Then, parameters are estimated for the ARIMA (2, 2, 2) model and displayed in Table 3. Then it is observed that AR (2) and MA (2) parameters have a p-value of 0.000, 0.000, 0.000, and 0.0001 respectively, indicating that the parameters are significant in the model at a 5% level of significance except for intercept.

The prediction of new cases of COVID-19 in Ethiopia, shown in Table 4, with a 95% confidence interval. According to the expected result, the number of confirmed COVID-19 new cases will increase slightly over the next 60 days. This increase is evidenced by the unrestricted response that has been invested in pandemic control in the country. However, the estimated predicted values were high, requiring more effort to minimize the spread of this pandemic across the country. The key problem of the outbreak is the explanation why a few people have not shown any signs of the virus spreading the virus to others without understanding the test.

The outcome of the forecasted maximum and minimum number of new Covid-19 cases will be 410 and 807 in one day. In addition, the cumulative confirmed cases reached 160,585 at the end of February 2021. Basically, it is displayed in Table



Figure 1: The line chart of coronavirus cases, deaths and recoveries in Ethiopia, 2020.

Table 2: The measure of model accuracy of time series of COVID-19 cases, Ethiopia, 2020.

Models	MAPE	MAD	MSD
Single exponential smooth	37.0	88.2	22722.1
Double exponential smooth	47.4	100.3	31970.1
Moving average (MA)	40.2	141.9	52113.4
S-Curve Trend Model	98.0	175.0	96537.7
Quadratic Trend Model	1169.5	211.9	86022.4
Linear Trend Model	973.0	244	117723
ARMA(2,2,2)	35.7	75.7	22714.8

Туре	Coefficients	SE Coeff.	t-statistic	P-value
AR(1)	-1.0083	0.0785	-12.84	0.000
AR(2)	-0.4116	0.0569	-7.23	0.000
MA(1)	0.5270	0.0535	9.86	0.000
MA(2)	0.4625	0.0769	6.01	0.000
Constant	-0.1216	0.3221	-0.38	0.706

4. Thus, more prevention measures and more resources will be introduced by the government; unless the coronavirus relapses and affects the country more.

Trend analysis presents the related statistics for reported COVID-19 data in Figure 2. The time series graph of the confirmed corona virus cases from 14 March 2020 to the end of December 2020 is shown in Figure 2. From the story, it is obvious that the time series is not stationary. The growing pattern of time series plot and model is a decline with time after peak in August 2020. The predicted value points out that the pandemic would quickly relapse.

Figures 3–5 show the residual plots for confirmed COVID–19 cases in Ethiopia from 13 March 2020 to 31 December 2020. Minor and insignificant residual deviations were shown from the straight line on the gauss probability plot (Figure 4). This also implies that the errors are somewhat near normal due to

 Table 4: Forecast point estimated values for COVID-19 cases with their lower and upper intervals for first two months of 2021 at 95% Confidence Intervals.

Data		New cases	
Date	Forecast	Lower limits	Upper limits
1-Jan-21	409.52	102.98	716.06
2-Jan-21	443.47	106.84	780.11
3-Jan-21	424.95	44.68	805.22
4-Jan-21	441.6	7.81	875.4
5-Jan-21	446.43	-18.86	911.72
6-Jan-21	448 7	-55 93	953.33
7- Jan-21	457.83	-80.97	996.63
8- Jan-21	461.8	-108 55	1032.16
0-lan-21	467.01	-134 44	1032.10
10- Jan-21	47/ 12	-157 78	1106.02
10 Jan 21	474.12	-181 36	11/0.38
12- Jan-21	485.67	-203 33	1174.67
12 Jan 21	403.07	-224.62	1207.66
13 Jan 21	491.32	-245.2	1207.00
14-Jan-21	502.40	-245.3	1240.10
15-Jdll-21	505.49	-203.10	12/2.14
10-Jan-21	509.47	-284.53	1303.40
17-Jan-21	515.53	-303.33	1334.4
18-Jan-21	521.62	-321.62	1364.87
19-Jan-21	527.73	-339.49	1394.95
20-Jan-21	533.89	-356.91	1424.69
21-Jan-21	540.07	-3/3.96	1454.11
22-Jan-21	546.29	-390.65	1483.23
23-Jan-21	552.54	-407	1512.08
24-Jan-21	558.83	-423.03	1540.69
25-Jan-21	565.14	-438.78	1569.07
26-Jan-21	571.49	-454.25	1597.24
27-Jan-21	577.88	-469.47	1625.22
28-Jan-21	584.29	-484.44	1653.02
29-Jan-21	590.74	-499.18	1680.66
30-Jan-21	597.22	-513.7	1708.14
31-Jan-21	603.74	-528.02	1735.49
1-Feb-21	610.29	-542.14	1762.71
2-Feb-21	616.87	-556.07	1789.81
3-Feb-21	623.48	-569.83	1816.79
4-Feb-21	630.13	-583.42	1843.67
5-Feb-21	636.81	-596.85	1870.46
6-Feb-21	643.52	-610.13	1897.17
7-Feb-21	650.26	-623.26	1923.79
8-Feb-21	657.04	-636.25	1950.33
9-Feb-21	663.85	-649.1	1976.81
10-Feb-21	670.7	-661.83	2003.22
11-Feb-21	677.57	-674.43	2029.58
12-Feb-21	684.49	-686.91	2055.88
13-Feb-21	691.43	-699.27	2082.13
14-Feb-21	698.4	-711.53	2108.34
15-Feb-21	705.41	-723.68	2134.51
16-Feb-21	712.46	-735.73	2160.64
17-Feb-21	719.53	-747.68	2186.74
18-Feb-21	726.64	-759.53	2212.8
19-Feb-21	733.78	-771.29	2238.85
20-Feb-21	740.95	-782.96	2264.87
21-Feb-21	748 16	-794 54	2290.86
22-Feb-21	755.4	-806.04	2316.84
22-Feb-21	762.67	-817 46	2342.81
24-Feb-21	769.98	-828.8	2368 76
2-1 CD 21	777 22	-840 07	2300.70
20-FED-21	701 60	-040.07	2074./
20-Feb-21	702.00	-001.20	2420.04
27-FEU-21	700 50	-002.30	2440.37
20-Feb-21	20.56 L	-0/3.43	2472.49
29-F6D-21	00/	-004.41	2498.41

a few outliers, such as the August report. Thus, the normality assumption may be slightly followed, but the residual histogram followed the normality assumption (Figure 5). The graph between the residuals and fitted values shows a small dispersion at very early times, and about 19 percent of the new case values are zero (See Figure 3). This implies that the assumption of constant variance is also satisfied, except the smaller (zero) data values are dropped.

Figures 6,7 demonstrate the diagnostic plots for the expected cases of COVID-19 in Ethiopia. ACF is a (complete) auto-correlation function that gives us the auto-correlation values of any sequence with its lagged values. It defines in simple terms how well the present value of the sequence is connected to its previous values (Figure 6). PACF is a partial auto-correlation function, finds a residual correlation (remaining after removing the effects that have already been explained by the earlier lag(s)). Then next lag value, therefore 'partial' and not 'complete' as we remove the found variations before we find the next correlation (Figure 7).



Figure 2: Times series plot for confirmed COVID-19 infections in Ethiopia from 14 March 2020 to December 2020 (Black line represents actual confirmed cases and colored lines represent case forecasts).



Figure 3: Residual verses fitted values of the Corona Virus in Ethiopia, 2020.



Figure 4: Normal probability plot of the residual of the Corona Virus in Ethiopia, 2020.

020



Figure 5: Histogram of the residual graph of the Corona Virus in Ethiopia, 2020.



Figure 6: PACF of Residual for case of COVID-19, in Ethiopia, 2020.



Discussion

The main objective of this analysis was to predict confirmed cases of COVID-19 in Ethiopia based on past recorded cases. It is important to build a reliable and effective predictive model that can enable governments and other stakeholders to monitor the further spread of COVID-19. ARIMA models are predictive technique that offers a good forecast and has been widely used for the rapid trend of infectious diseases [2,9,10,12].

The time series model and plot indicated a small rise in the number of future cases in Ethiopia. The number of reported cases in the country is expected to exceed 160585 in the two months 2021. At the end of October, Inline projected cumulative infections across Ethiopia to hit 56,610 on average [9], while the real value until 19 October was just 89,137 cases and 1352 deaths [5]. This means that the number of people who are corona-virus-positive in Ethiopia can increase more than predicted [9]. However, in this analysis, the total number of new cases reached 124,264 at the end of 2020. As a result, this analysis was estimated to have a more reliable value than other researchers, since the disparity is 36, 321 events. Similar to the predicted values in this study, monthly variation of pandemics 8

in Ethiopia in the first three months has been shown to increase [9].

The scholars disprove the effect of geographical difference and temperature in reducing the distribution of COVID-19 (16). However, Covid-19 has caused serious global social and economic distress. Thus, COVID-19, regulated only by preprotective strategies. However, no aspect is evaluated in this study due to the unavailability of exposure data.

Pandemic patterns in the US and India have risen from 12 July 2020 to 11 September 2020 [10,13]. Countries are now the first and second highly confirmed cases of the pandemic. In the case of India, the case of COVID-19 was predicted using the ARIMA method and suggested the implementation of lockdown [9]. The security measures proposed were not, however, well implemented, and now the possibility of a pandemic in that country is uncontrollable. Ethiopia was 5th in the number of COVID-19 confirmed cases in Africa, according to the WHO report [5]. It is presumed that if and only if the standard WHO disease prevention and control measures are not taken [9], the worst-case scenario could occur. This study therefore showed slight elevate in cases Ethiopia.

Like other scholars, this research was carried out to forecast the potential estimation of cases of COVID-19 using the ARIMA model [9-12]. In line with this analysis, the scholars have found that ARIMA technique is suitable for predicting the prevalence of Covid-19 [13-16]. The ARIMA model predicted a pandemic outbreak in India and indicated important corona viruses in the country prior to a rise in confirmed cases [10]. Since the same model used in this analysis, the ARIMA model is useful in predicting potential cases of COVID-19. Ethiopia must also introduce controlling mechanisms and establish even more pandemic prevention policies.

Conclusion

Since the study shows a small increase in the number of cases in Ethiopia, more attention has been paid to the control of Covid-19. Unless the Government of Ethiopia implements a mechanism to control the pandemic, it may relapse and affect the country more. The study therefore suggested that proactive and control mechanisms should be implemented on a continuous basis. Since ARIMA model is an effort to predict the future forecast of the distribution of COVID-19, based on current data, so that the institutions have to formulate policies from now on the result of the study.

References

- Worldometer (2020) Ethiopia Corona Virus –Worldometer. WHO. Link: https://bit.ly/3pRr6cU
- Unicef. Ethiopia (2020) Socio-economic impacts of COVID-19. UNICEF Ethiopia. 1. Link: https://uni.cf/3rTH08D
- Takele R (2020) Stochastic modelling for predicting COVID-19 prevalence in East Africa Countries. Infectious Disease Modelling 5: 598-607. Link: https://bit.ly/3pQdlex
- Ayenew B, Yitayew M, Pandey D (2020) Challenges and opportunities to tackle COVID-19 spread in Ethiopia. Peer Scientist 2: e1000014. Link: https://bit.ly/2LqUg3B

021

- Abate L, Bekele A, Bedada B (2020) Status of distribution of coronavirus disease (COVID-19) in Ethiopia within first three months. AJRSP 2: 156-166. Link: https://bit.ly/2X8V8wu
- Na J, Tibebu H, De Silva V, Kondoz A, Caine M (2020) Probabilistic approximation of effective reproduction number of COVID-19 using daily death statistics. Chaos, Solitons and Fractals 140: 110181. Link: https://bit.ly/35beJ3o
- Takele R (2020) Stochastic modelling for predicting COVID-19 prevalence in East Africa Countries. Infect Dis Model 5: 598-607. Link: https://bit.ly/39170q2
- Tandon H, Ranjan P, Chakraborty T, Suhag V (2020) Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future. Link: https://bit.ly/3oijNdV
- Das RC (2020) Forecasting incidences of COVID-19 using Box-Jenkins method for the period July 12-Septembert 11, 2020: A study on highly affected countries. Chaos, Solitons and Fractals 140: 110248. Link: https://bit.ly/2MBS4XR
- Khana FM, Gupta R (2020) ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India. Journal of Safety Science and Resilience 1: 12-18. Link: https://bit.ly/3ot9YtC

- 11. Zeynep C (2020) Estimation of COVID-19 prevalence in Italy, Spain, and France. Sci Total Environ 729: 138817. Link: https://bit.ly/38dyucA
- Papastefanopoulos V, Linardatos P, Kotsiantis S (2020) COVID-19: A Comparison of Time Series Methods to Forecast Percentage of Active Cases per Population. Applied sciences 10: 3880. Link: https://bit.ly/35bxKTy
- Lu H, Stratton CW, Tang YW (2020) Outbreak of Pneumonia of Unknown Etiology in Wuhan China: The Mystery and the Miracle. J Med Virol 92: 401– 402. Link: https://bit.ly/358SWcY
- CSSE (2020) Corona virus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Link: https://bit.ly/35cYvH7
- Mann R, Perisetti A, Gajendran M, Gandhi Z, Umapathy C, et al. (2020) Clinical Characteristics, Diagnosis, and Treatment of Major Coronavirus Outbreaks. Front Med 7. Link: https://bit.ly/2MAjkG9
- Fattorini D, Regoli F (2020) Role of the chronic air pollution levels in the Covid-19 outbreak risk in Italy. Environ Pollu 264: 114732. Link: https://bit.ly/2X8Sve6

Discover a bigger Impact and Visibility of your article publication with Peertechz Publications

Highlights

- Signatory publisher of ORCID
- Signatory Publisher of DORA (San Francisco Declaration on Research Assessment)
- Articles archived in worlds' renowned service providers such as Portico, CNKI, AGRIS, TDNet, Base (Bielefeld University Library), CrossRef, Scilit, J-Gate etc.
- Survey of the second se
- OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting)
- Dedicated Editorial Board for every journal
- Accurate and rapid peer-review process
- Increased citations of published articles through promotions
- Reduced timeline for article publication
- Submit your articles and experience a new surge in publication services

(https://www.peertechz.com/submission).

Peertechz journals wishes everlasting success in your every endeavours.

Copyright: © 2020 Eticha AB. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

022