

Forecasting the spread of COVID-19 pandemic in Bangladesh using ARIMA model

Lakshmi Rani Kundu^a, Most. Zannatul Ferdous^{a*}, Ummay Soumayia Islam^a, Marjia Sultana^b

^aDepartment of Public Health and Informatics, Jahangirnagar University, Savar, Dhaka-1342, Bangladesh

^bDepartment of Computer Science and Engineering, Begum Rokeya University, Rangpur-5400, Bangladesh

*Corresponding author

Email: m.zannatul.ferdous@juniv.edu

Author contributions

Lakshmi Rani Kundu: Conceptualization, Data searching, Statistical analysis, Writing-original draft, Editing, Validation; Most. Zannatul Ferdous: Conceptualization, Data Searching, Writing-original draft, Editing, and Validation, Ummay Soumayia Islam: Writing-original draft, Editing, and Validation; Marjia Sultana: Conceptualization, Editing and validation.

Abstract

Background: COVID-19 is one of the most serious global public health threats creating an alarming situation. Therefore, there is an urgent need for investigating and predicting COVID-19 incidence to control its spread more effectively. This study aim to forecast the expected number of daily total confirmed cases, total confirmed new cases, total deaths and total new deaths of COVID-19 in Bangladesh for next 30 days.

Methods: The number of daily total confirmed cases, total confirmed new cases, total deaths and total new deaths of COVID-19 from 8 March 2020 to 16 October, 2020 was collected to fit an Autoregressive Integrated Moving Average (ARIMA) model to forecast the spread of COVID-19 in Bangladesh from 17th October 2020 to 15th November 2020. All statistical analyses were conducted using R-3.6.3 software with a significant level of $p < 0.05$.

Results: The ARIMA (0,2,1) and ARIMA (0,1,1) model was adopted for forecasting the number of daily total confirmed cases, total deaths and total confirmed new cases, new deaths of COVID-19, respectively. The results showed that an upward trend for the total confirmed cases and total deaths, while total confirmed new cases and total new death, will become stable in the next 30 days if prevention measures are strictly followed to limit the spread of COVID-19.

Conclusions: The forecasting results of COVID-19 will not be dreadful for upcoming month in Bangladesh. However, the government and health authorities should take new approaches and keep strong monitoring of the existing strategies to control the further spread of this pandemic.

Keywords: COVID-19, Confirmed cases, Deaths, Forecast, ARIMA, Bangladesh

Introduction

The COVID-19 pandemic is an ongoing public health threat which is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) ¹. It is a viral infection and highly infectious disease considering to be transmitted from wild animals (bats) to human and first identified in Wuhan city of China ^{2,3}. The disease was then a local epidemic of China, but soon it became expanded all over the world by international travelers ⁴. Previously SARS-CoV emerged from China in 2003, respectively MERS-CoV emerged from Middle East in 2012, developed severe symptoms ^{5,6}. Now this virus is the 7th of coronaviruses that represents a serious public health threat for which World Health Organization (WHO) declared COVID-19 as a pandemic in March 2020 ⁷. Since then, the pandemic has spread all over the globe as days go by. Approximately, 190 countries have been affected, where major outbreaks occurred in USA, Italy, Spain, France, China ⁴. As of 14th October, 2020, COVID-19 cases have been exceeded 38 million including 1,083,234 deaths worldwide ⁸. According to a study conducted in China, COVID-19 can be asymptomatic including mild or moderate symptoms ⁹. Hence, by looking into these facts, we should have to imagine the heaviness of this pandemic globally and its impacts on public health.

In Bangladesh, Institute of Epidemiology, Disease Control and Research (IEDCR), declared the first three confirmed cases of COVID-19 on 8th March, 2020 ¹⁰. The present scenario of coronavirus cases in Bangladesh is 3,91586 confirmed cases, with 5699 deaths, and 3,07141 recoveries on 20 October 2020 ¹¹. As Bangladesh is one of the most densely populated countries around the world, it has a great risk of exposure due to COVID-19. Like other countries, government of Bangladesh has also already adopted several measures such as informing COVID-19 hotspot areas, maintaining social distance and increasing mass awareness by social media or televisions, setting lockdown of school, college and office ¹² to minimize the situation. However, these available control measures are significantly influenced by the knowledge, attitudes, and practices (KAP) towards COVID-19 ¹³. Furthermore, it is really a challenging task for the people of overcrowded Bangladesh where the chance of the COVID-19 spreading is much more than non crowded place and in a Bangladesh survey from March 29 to April 29 found that 98.7% reported wearing a face mask in crowded places ¹³. As the incubation period of COVID-19 is up to 14 days, the virus can be transmitted to other people during this time period ^{14,15}. Again, there is ambiguity about the proper decline and fall of the contagious disease ¹⁶.

Therefore, in this critical time, smart planning with sufficient preparation for mitigating the incidence and prevalence of disease including designing the future prospect is very important. Because evidence on the management approaches of current COVID-19 pandemic is still limited though the numbers of affected countries are increasing as the days go by ¹⁷. By modeling a future forecast which estimates the regular number of confirmed cases might help to implement new rules. Further, a statistical forecast model might also be beneficial for predicting future epidemic threat as well as better management of societal, economic, cultural and public health matters ^{18,19}.

The aim of this study is to predict the spread and the final size of COVID-19 epidemic in Bangladesh by using Auto Regressive Integrated Moving Average (ARIMA) model. The ARIMA model is generally known as Box-Jenkins methodology used to forecast and analysis in a time series modeling approach ²⁰. Recently this model has been used in the mostly affected 15 countries of the world to forecast the flow of COVID-19 which revealed similar number according to the current situation of those countries ²¹. Another study conducted in Italy and Spain showed accurate regular number of cases in these countries ²². Though modeling cannot always forecast the accurate number of cases it may help to summarize the future prospect of the pandemic by showing the acceptable number of occurrence happening for the next 30 days.

Methods and materials

Data Source and data description

The data extracted from the official website of the World Health Organization COVID-19 situation reports ²³ and website of the Humanitarian Data Exchange (<https://data.humdata.org/dataset/coronavirus-COVID-19-cases-and-death>).

This study considered the daily total confirmed cases, daily total confirmed new cases, total deaths and total new deaths of COVID-19 from 8th March (day 1) to 16th October 2020 (day 223) in Bangladesh. Then, this data were used to fit a best ARIMA model for forecasting next 30 days total confirmed cases or new cases as well as total deaths or new deaths from 17th October (day 224) to 15th November (day 253) of COVID-19 in Bangladesh.

Statistical Analysis

All analyses were done with R software (version 3.6.2), the stationarity check was conducted using 'tseries' package and ARIMA model was fitted using 'forecast' package. A *p*-value of less than 0.05 was considered statistically significant.

ARIMA Model description

In this study, a linear parametric Autoregressive Integrated Moving Average (ARIMA) model was applied for prediction purpose. If a time series Y_t is followed by $Y_t = \theta + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + u_t$, where u_t is a white noise with mean zero and variance σ^2 , then it is called an autoregressive process of order p and is denoted by $AR(p)$. If Y_t is defined by $Y_t = \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_p u_{t-p}$, then it is called a moving average process of order q and is denoted by $MA(q)$.

The combination of AR and MA models are known as ARMA model. An $ARMA(p,q)$ model is given by $Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_p u_{t-p}$ ²⁴. A time series Y_t is said to follow an Autoregressive Integrated Moving Average (ARIMA) model if the d^{th} difference $W_t = \nabla^d Y_t$ is a stationary ARMA process. If W_t follows an $ARMA(p,q)$ model, then we say that Y_t is an $ARIMA(p,d,q)$ process. For practical purposes, taking $d=1$ or at most 2²⁵.

Thus, an $ARIMA(p, 1, q)$ process with $W_t = Y_t - Y_{t-1}$ can be written as

$$W_t = \alpha_1 W_{t-1} + \alpha_2 W_{t-2} + \dots + \alpha_p W_{t-p} + u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_p u_{t-p}.$$

Box-Jenkins Method

ARIMA model also known as Box-Jenkins method which is widely used for the analysis of time series and forecasting²⁶. In recent years, this model has been widely used in the prediction of epidemic trend of infectious diseases²⁷. The fitting of the ARIMA model or Box-Jenkins methodology consists of the following steps:

Test of Stationarity

An Augmented Dickey-Fuller (ADF) test was conducted for testing whether the original series is stationary or not²⁸. Difference or logarithmic transformation was adapted to transformed nonstationary time series into a stationary time series. Achieving stationary is a precondition for establishing an ARIMA model.

Model Identification

An appropriate value of p , d and q of ARIMA model was identified. The value of d was identified according to the number of differentials. From the autocorrelation function (ACF) and partial autocorrelation function (PACF) plot against the lag length, the AR and MA parameters was selected. Nevertheless, there have been some model selection criteria such as Akaike information criterion (AIC), Bayesian information criteria (BIC). The optimal model was chosen based on smallest value of AIC and BIC²⁹.

Estimation of the fitted Model

After the identification of the appropriate values of p and q , the next stage was to estimate the parameter of the autoregressive and moving average terms included in the model. This was done using the maximum likelihood estimation method.

Diagnosing Checking

Having identified the fitted ARIMA model and parameter estimate, the Ljung-Box Q test was applied to check whether the residual series is a white noise. If so, then the fitted model was accepted. The Ljung-Box Q statistics is defined as

$$Q = T(T + 2) \sum_{k=1}^s (T - k)^{-1} r_k^2$$

where, T is the number of observations, s is length of coefficients to test autocorrelation, r_k is autocorrelation coefficient for lag k . This statistic Q approximately follows the chi-square distribution with $(k-q)$ degrees of freedom, where q is the number of parameter should be estimated in the model³⁰.

Forecasting

Finally, the future value was predicted by using the fitted model. The steps are presented in the following diagram:

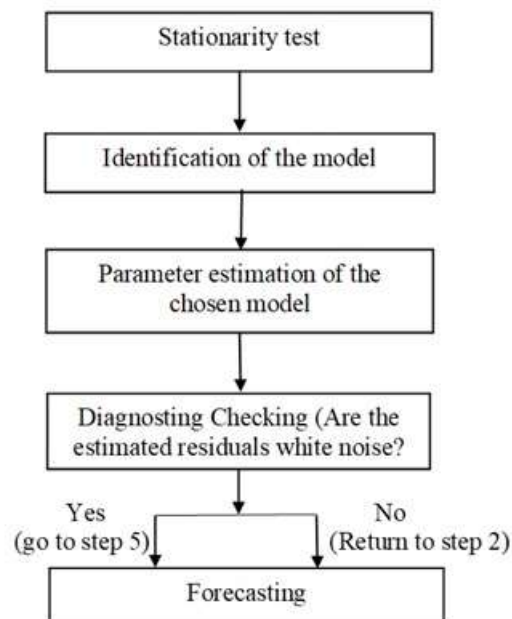


Figure 1: Scheme for the use of Box-Jenkins methodology³¹

Ethics

As all data were obtained from secondary data collection source, no formal ethical assessment was required.

Results

The time series plot of daily COVID-19 total confirmed new cases (A1), total confirmed cases (A2), total new deaths (B1) and total deaths (B2) in Bangladesh from March 8 to October 16, 2020 are presented in figure 2. During the study period, a total of 384,559 confirmed cases and 5,608 deaths were detected and there were maximum cases in 3rd July 2020 of 4019 cases and maximum deaths in 1st July 2020 of 64 deaths. The graphical inspection showed that the original series are in increasing trend and sometimes are in decreasing trend and the variance is not stable which leads the variables were nonstationary and need to be transformed into a stationary process.

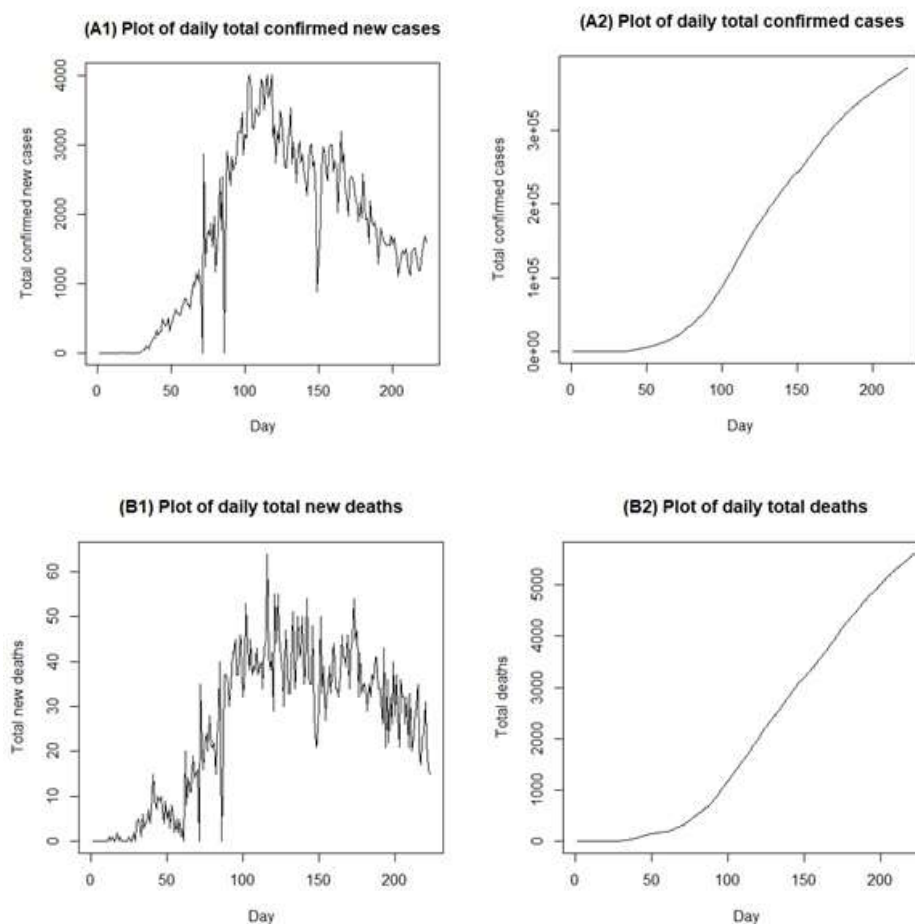


Figure 2. Time series plot displays for confirmed new cases (A1), total confirmed cases (A2), new deaths (B1) and total deaths (B2) of COVID-19 in Bangladesh.

Before fitting the ARIMA model, it is necessary to confirmed that the series must be stationary. Augmented Dickey Fuller (ADF) test was applied to check the stationarity of the series. The ADF unit root test results are presented in Table 1.

The findings indicated that all the series are nonstationary at their level ($p\text{-value} > 0.05$), but after taking 1st difference, total confirmed new cases and deaths achieved stable variance. On the other hand, total confirmed cases and deaths achieved stable variance after 2nd differences. This ensured that the series are stationary at 5% level of significance and ready for the Box-Jenkins ARIMA modeling approach.

Table 1. Results of Augmented Dickey Fuller unit root test

Variables	Original data		At 1 st difference		At 2 nd difference	
	ADF	<i>p</i> -value	ADF	<i>p</i> -value	ADF	<i>p</i> -value
Total confirmed new cases	-0.4148	0.985	-8.135	0.01*	-	-
Total new Deaths	-0.280	0.9831	-7.709	0.01*	-	-
Total confirmed cases	-3.609	0.063	-0.429	0.984	-8.122	0.01*
Total Deaths	-2.775	0.25	-0.290	0.99	-7.698	0.01*

The null hypothesis is that the series is non-stationary, or contain a unit root. Decision Rule: Reject the null hypothesis if the $p\text{-value} < \alpha = 0.05$; * indicates the rejection of the null hypothesis.

After achieving the stationary series, a list of potential models was formulated based on the significant spikes observed from the autocorrelation function (ACF) and partial autocorrelation function (PACF). Among the candidate models ARIMA(0,2,1) model produces the lowest AIC, AICc, BIC for the total confirmed cases and total deaths and ARIMA(0,1,1) model exhibits the lowest AIC, AICc, BIC for confirmed new cases and new deaths which are demonstrated in Table 2.

Table 2: Results of each selected ARIMA model and Ljung-Box test

ARIMA model	Coefficients mal	Standard error	AIC	AICc	BIC	p-value
Total confirmed cases ARIMA(0,2,1)	-0.6418	0.0527	3254.58	3254.63	3261.38	0.46
Total confirmed new cases ARIMA(0,1,1)	-0.6419	0.0526	3268.28	3268.34	3275.09	0.46
Total deaths ARIMA(0,2,1)	-0.7489	0.0391	1484.75	1484.81	1491.55	0.55
Total new deaths ARIMA(0,1,1)	-0.749	0.039	1490.45	1490.5	1497.25	0.55

The Ljung-Box Q test suggested that the residuals series of the ARIMA (0,2,1) and ARIMA (0,1,1) model are purely white noise (p -value>0.05) at 95% confidence level. Therefore, these selected models are probably adequate for the data. Then these models were applied to forecast the daily confirmed cases and deaths of COVID-19 in Bangladesh.

Table 3 and figure 3 showed the predicted values from 17th October to 15th November 2020 for all variables using the fitted ARIMA(0,2,1) and ARIMA(0,1,1) model with 95% confidence interval (CI). The forecasted value (in blue) based on fitted ARIMA model for daily confirmed cases or new cases and deaths or new deaths of COVID-19 for the next 30 days and the current number of confirmed cases and deaths from March 8, 2020 to October 16, 2020 (in black) are shown in figure 3. The results showed that an upward trend for daily total confirmed cases and total deaths in Bangladesh by 15th November, 2020 has a point forecast of 430828 (95% CI 402683-458974) and 6228 (95% CI 5849-6608), while total confirmed new cases and total new death, possible become stable. However, Bangladesh is hopeful to control this pandemic at the middle of November.

Table 3: Forecasting of daily total confirmed cases, total confirmed new cases, total deaths, and total new deaths in Bangladesh for the next 30 days according to ARIMA models with 95% CI

Date	Total confirmed cases ARIMA(0,2,1)	Total confirmed new cases ARIMA(0,1,1)	Total deaths ARIMA(0,2,1)	Total new deaths ARIMA(0,1,1)
17/10/2020	386101 (385359,386843)	1542 (802,2282)	5628 (5615,5642)	20 (7,34)
18/10/2020	387643 (386392,388895)	1542 (755,2328)	5649 (5627,5671)	20 (6,34)
19/10/2020	389186 (387400,390971)	1542 (712,2372)	5670 (5640,5699)	20 (6,35)
20/10/2020	390728 (388371,393085)	1542 (671,2413)	5690 (5652,5728)	20 (5,35)
21/10/2020	392270 (389301,395239)	1542 (631,2452)	5711 (5664,5758)	20 (5,35)
22/10/2020	393812 (390193,397432)	1542 (593,2490)	5732 (5676,5787)	20 (5,36)
23/10/2020	395355 (391047,399663)	1542 (557,2527)	5752 (5687,5818)	20 (4,36)
24/10/2020	396897 (391864,401930)	1542 (522,2562)	5773 (5698,5848)	20 (4,36)
25/10/2020	398439 (392647,404232)	1542 (488,2596)	5794 (5708,5879)	20 (4,37)
26/10/2020	399982 (393396,406568)	1542 (455,2628)	5814 (5718,5911)	20 (4,37)
27/10/2020	401524 (394113,408936)	1542 (423,2660)	5835 (5728,5942)	20 (3,37)
28/10/2020	403066 (394798,411335)	1542 (352,2691)	5856 (5737,5975)	20 (3,37)
29/10/2020	4046093(95454,413764)	1542 (362,2721)	5877 (5746,6007)	20 (3,38)
30/10/2020	406151 (396080,416222)	1542 (333,2751)	5897 (5755,6040)	20 (2,38)
31/10/2020	407693 (396678,418709)	1542 (304,2780)	5918 (5763,6073)	20 (2,38)
1/11/2020	409236 (397249,421223)	1542 (276,2808)	5939 (5771,6107)	20 (2,39)
2/11/2020	410778 (397792,423764)	1542 (249,2835)	5959 (5778,6141)	20 (1,39)
3/11/2020	412320 (398310,426331)	1542 (222,2862)	5980 (5785,6175)	20 (1,39)
4/11/2020	413863 (398802,428923)	1542 (195,2888)	6001 (5792,6209)	20 (1,40)
5/11/2020	415405 (399269,431540)	1542 (169,2914)	6021 (5799,6244)	20 (0,40)
6/11/2020	416947 (399712,434182)	1542 (144,2940)	6042 (5805,6279)	20 (0,40)
7/11/2020	418490 (400132,436847)	1542 (119,2965)	6063 (5811,6315)	20 (0,40)
8/11/2020	420032 (400528,439536)	1542 (95,2989)	6084 (5817, 6350)	20 (0,41)
9/11/2020	421574 (400901,442248)	1542 (71,3013)	6104 (5822,6386)	20 (-1,41)
10/11/2020	423117 (401252,444982)	1542 (47,3037)	6125 (5827,6423)	20 (-1,41)
11/11/2020	424659 (401580,447737)	1542 (24,3060)	6146 (5832,6459)	20 (-1,42)
12/11/2020	426201 (401887,450515)	1542 (1,3083)	6166 (5837, 6496)	20 (-1,42)
13/11/2020	427743 (402173,453314)	1542 (-21,3106)	6187 (5841,6533)	20 (-2,42)
14/11/2020	429286 (402438,456133)	1542 (-43,3128)	6208 (5845, 6570)	20 (-2,43)
15/11/2020	430828 (402683,458974)	1542 (-65,3150)	6228 (5849,6608)	20 (-2,43)

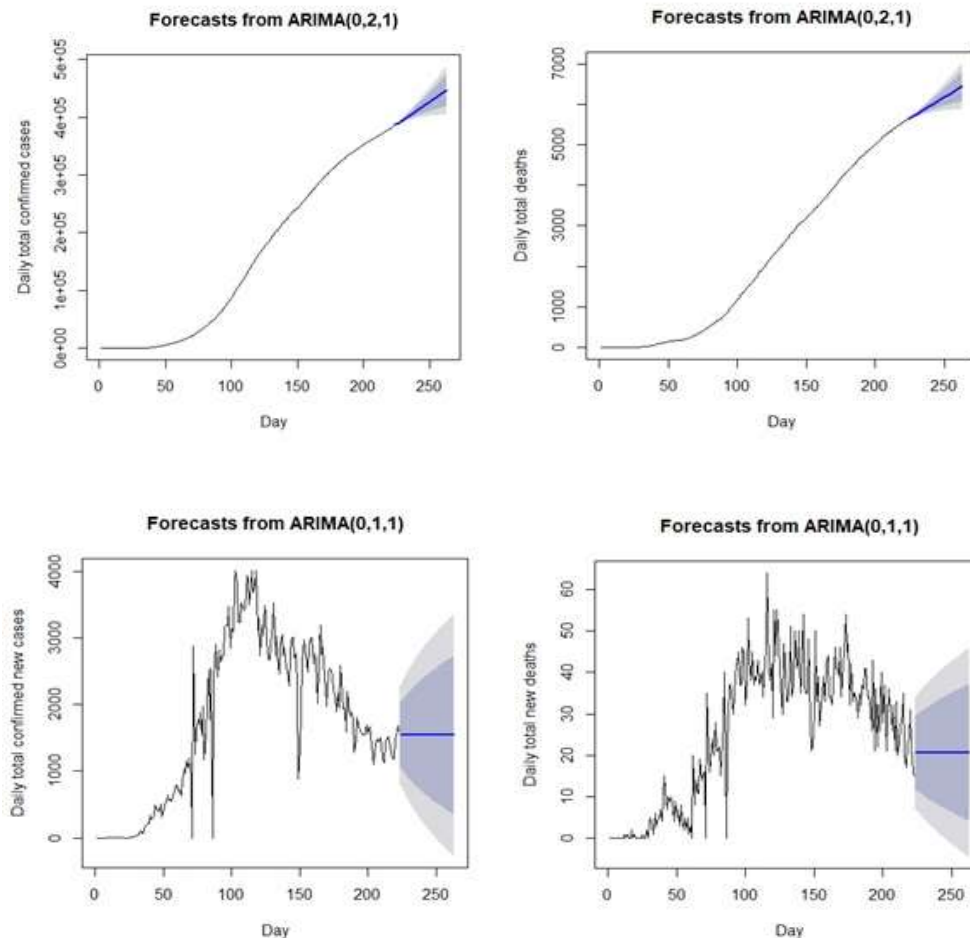


Figure 3: Predictive and confidence intervals of daily total confirmed cases, total confirmed new cases, total deaths, total new deaths of COVID-19 using fitted model (Black line: actual data, Blue line: 30 days forecast, Gray zone: 80% of CI, White zone: 95% of CI).

Discussion

It is alarming that the number of novel coronavirus (SARS-Cov-2) cases continues to escalate across the globe. The estimation of infectious disease like COVID-19 management by developing hypothesis for interpreting the observed situation can be done via time series analysis³². Time series health researchers widely use ARIMA model because of the importance of ‘time’ for disease management studies^{33,34}. Previous study revealed that ARIMA is one of the most suitable models as it has higher fitting and forecasting accuracy^{35,36}. In recent years, it is also an useful model for predicting the incidence of infectious disease²⁰. The main purpose of this work is to monitor and forecast the expected number of the new COVID-2019 patients in Bangladesh

by applying a commonly used time series model, known as an ARIMA model, based on the data of the total confirmed daily cases and deaths and the new confirmed cases and new deaths officially announced by the Institute of Epidemiology, Disease Control and Research.

The prediction indicates an upward trend for daily total confirmed cases along with total deaths. At the same time, total confirmed new cases, and total new deaths most probably become stable. According to the study the total confirmed cases of coronavirus in 18th October, 2020 was estimated 388,569 and total death was estimated 5660 which was nearly similar to the actual scenario³⁷. The present reveals that, Bangladesh is hopeful for controlling the pandemic at the mid of November 2020 if the spreading pattern of the disease remains the same. A study conducted in Iran reported that, an upward trend for total confirmed case and total death while the other variables such as total confirmed new cases, total new deaths possibly became steady which is similar to our study³⁸. Another study conducted for forecasting different countries COVID-19 trend demonstrated stable condition for China, stationary trend for South Korea while Thailand showed controlled condition¹⁸. Based on a study, for predicting the end of COVID-19 by using this model expected that top countries COVID-19 infection will slow down by October, 2020³⁹. In contrast, in Saudi Arabia and Nigeria this model forecasts highly increased of daily case with cumulative daily cases within one month^{40,41}. Similarly another study conducted India revealed explicit rising of infection in coming days especially in west and south Indian regions are more at risk^{42,43}.

Although, ARIMA model usually gives better forecast, few drawbacks exists. Firstly, it does not have automatic updates. Secondly, if more data added in the study then model gives different forecast results. Thirdly, the prediction accuracy has a direct relation with the number of observation.

Conclusion

In this study, the trend of COVID-19 outbreak in Bangladesh was observed. The results found that the best prediction model is ARIMA (0,1,1) for forecasting the trend of the number of daily new confirmed cases and deaths in Bangladesh. In this model the number of daily confirmed cases and deaths showed upward trend. But surprisingly model presents the number of new confirmed cases and new deaths will become stable in next 30 days. It is hopeful for Bangladesh, a developing country, though having inadequate medical facilities, becoming successful to

control this pandemic within last few days. However, to prevent the COVID-19 pandemic permanently until proper vaccine or medicine is developed, public health authorities, government, and non-government should take hard decisions to control the further increase of this pandemic. Besides all the authorities, the general public should maintain social distance and undertake all necessary preventive measures to stay free from the disease and control its spread.

References

1. WHO. Naming the coronavirus disease (COVID-19) and the virus that causes it. *WHO* (2020).
2. De Wit, E., Van Doremalen, N., Falzarano, D. & Munster, V. J. SARS and MERS: Recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **14**, 523–534 (2016).
3. Paules, C. I., Marston, H. D. & Fauci, A. S. Coronavirus Infections-More Than Just the Common Cold. *JAMA - Journal of the American Medical Association* vol. 323 707–708 (2020).
4. Chintalapudi, N., Battineni, G. & Amenta, F. COVID-19 virus outbreak forecasting of registered and recovered cases after sixty day lockdown in Italy: A data driven model approach. *J. Microbiol. Immunol. Infect.* **53**, 396–403 (2020).
5. Hui, D. S. C. & Zumla, A. Severe Acute Respiratory Syndrome: Historical, Epidemiologic, and Clinical Features. *Infectious Disease Clinics of North America* vol. 33 869–889 (2019).
6. Azhar, E. I., Hui, D. S. C., Memish, Z. A., Drosten, C. & Zumla, A. The Middle East Respiratory Syndrome (MERS). *Infectious Disease Clinics of North America* vol. 33 891–905 (2019).
7. Cascella, M., Rajnik, M., Cuomo, A., Dulebohn, S. C. & Di Napoli, R. *Features, Evaluation and Treatment Coronavirus (COVID-19)*. StatPearls (StatPearls Publishing, 2020).
8. WHO. WHO Coronavirus Disease (COVID-19) Dashboard | WHO Coronavirus Disease (COVID-19) Dashboard. *WHO* (2020).
9. Hua, W. *et al.* Consideration on the strategies during epidemic stage changing from emergency response to continuous prevention and control. *Chinese J. Endem.* **41**, 297–300 (2020).
10. Paul, R. Bangladesh confirms its first three cases of coronavirus | Reuters. *Reuters* (2020).
11. Worldometer. Bangladesh Coronavirus: 382,959 Cases and 5,593 Deaths - Worldometer. *worldometer* (2020).
12. Haque, A. The COVID-19 pandemic and the public health challenges in Bangladesh: a commentary. *J. Heal. Res.* **34**, 563–567 (2020).
13. Ferdous, M. Z., Islam, M. S., Sikder, T. & Syed, A. COVID-19 outbreak in Bangladesh : An online- based cross-sectional study. 1–17 (2020) doi:10.1371/journal.pone.0239254.
14. Al-Qaness, M. A. A., Ewees, A. A., Fan, H. & Aziz, M. A. El. Optimization method for forecasting confirmed cases of COVID-19 in China. *Appl. Sci.* **9**, 674 (2020).

15. Moftakhar, L. & Seif, M. The exponentially increasing rate of patients infected with COVID-19 in Iran. *Arch. Iran. Med.* **23**, 235–238 (2020).
16. Yousaf, M., Zahir, S., Riaz, M., Hussain, S. M. & Shah, K. Statistical Analysis of Forecasting COVID-19 for Upcoming Month in Pakistan. *Chaos, Solitons and Fractals* (2020) doi:10.1016/j.chaos.2020.109926.
17. Ferdous, M. Z., Kundu, L. R., Sultana, M. & Jafrin, S. J. Original Article Regional differences in COVID-19 attack and case fatality rates in the first quarter of 2020 : a comparative study. *IMC J. Med. Sci.* **14**, 3–10 (2020).
18. Dehesh, T., Mardani-Fard, H. A. & Dehesh, P. Forecasting of COVID-19 Confirmed Cases in Different Countries with ARIMA Models. *medRxiv* 1–12 (2020) doi:10.1101/2020.03.13.20035345.
19. Petropoulos, F. & Makridakis, S. Forecasting the novel coronavirus COVID-19. *PLoS One* **15**, e0231236 (2020).
20. Inoue, M., Hasegawa, S. & Suyama, A. P1-177 Development and evaluation of a forecasting model for infectious diseases in Japan using time-series analysis. *J. Epidemiol. Community Heal.* **65**, A115–A115 (2011).
21. Kumar, P. *et al.* Forecasting the dynamics of COVID-19 Pandemic in Top 15 countries in April 2020: ARIMA Model with Machine Learning Approach. *medRxiv* (2020) doi:10.1101/2020.03.30.20046227.
22. Monllor, P., Su, Z., Gabrieli, L., Montoro, A. & Taltavull de La Paz, M. de la P. COVID-19 Infection Process in Italy and Spain: Are the Data Talking? *SSRN Electron. J.* (2020) doi:10.2139/ssrn.3566150.
23. WHO. Coronavirus Disease (COVID-19) Situation Reports. *accessed on 31 July 2020* (2020).
24. Gujarati, D. N. & Porter, D. C. *Basic Econometrics*. (McGraw-Hill, 2008).
25. Kane, M. J., Price, N., Scotch, M. & Rabinowitz, P. Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinformatics* **15**, (2014).
26. Box, G. E. P. & Jenkins, G. M. *Time Series Analysis: Forecasting and Control*. (Holden Day, San Francisco (575 pages), 1976).
27. Duan, Y. *et al.* Impact of meteorological changes on the incidence of scarlet fever in Hefei City, China. *Int. J. Biometeorol.* **60**, 1543–1550 (2016).
28. Dickey, D. A. & Fuller, W. A. Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *J. Am. Stat. Assoc.* **74**, 427 (1979).
29. Akaike, H. A New Look at the Statistical Model Identification. *IEEE Trans. Automat. Contr.* **19**, 716–723 (1974).
30. Ljung, G. M. & Box, G. E. P. On a measure of lack of fit in time series models. *Biometrika* **65**, 297–303 (1978).
31. Makridakis, S. G., Wheelwright, S. C. & Hyndman, R. J. *Forecasting methods and applications*.

- Wiley (Wiley, 1997).
32. Sato, R. C. esa. Disease management with ARIMA model in time series. *Einstein (Sao Paulo)*. **11**, 128–131 (2013).
 33. Haines, L. M., Munoz, W. P. & Van Gelderen, C. J. ARIMA modelling of birth data. *J. Appl. Stat.* **16**, 55–67 (1989).
 34. Choi, K. & Thacker, S. B. An evaluation of influenza mortality surveillance, 1962-1979: I. Time series forecasts of expected pneumonia and influenza deaths. *Am. J. Epidemiol.* **113**, 215–226 (1981).
 35. Chen, P., Yuan, H. & Shu, X. *Forecasting crime using the ARIMA model. Proceedings - 5th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2008* vol. 5 (2008).
 36. Hue, H. T. T., Pradit, S., Lim, A., Goncalo, C. & Nitiratsuwan, T. Shrimp and fish catch landing trends in songkhla lagoon, Thailand during 2003-2016. *Appl. Ecol. Environ. Res.* **16**, 3061–3078 (2018).
 37. Worldometer. Bangladesh Coronavirus: 388,569 Cases and 5,660 Deaths - Worldometer. *worldometer* (2020).
 38. Tran, T. T., Pham, L. T. & Ngo, Q. X. Forecasting epidemic spread of SARS-CoV-2 using ARIMA model (Case study: Iran). *Glob. J. Environ. Sci. Manag.* **6**, 1–10 (2020).
 39. Ewis, A. *et al.* ARIMA Models for Predicting the End of COVID-19 Pandemic and the Risk of a Second Rebound. (2020) doi:10.21203/rs.3.rs-34702/v1.
 40. Alzahrani, S. I., Aljamaan, I. A. & Al-Fakih, E. A. Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. *J. Infect. Public Health* **13**, 914–919 (2020).
 41. Ibrahim, R. R. & OLADIPO, O. H. Forecasting the spread of COVID-19 in Nigeria using Box-Jenkins Modeling Procedure. *medRxiv* 2020.05.05.20091686 (2020) doi:10.1101/2020.05.05.20091686.
 42. Roy, S., Bhunia, G. S. & Shit, P. K. Spatial prediction of COVID-19 epidemic using ARIMA techniques in India. *Model. Earth Syst. Environ.* **1**, 3 (2020).
 43. Verma, P. *et al.* Forecasting the covid-19 outbreak: an application of arima and fuzzy time series models. *Res. Sq.* 1–15 (2020).