

Evolving trajectories of COVID-19 curves in India: Prediction using autoregressive integrated moving average modeling.

Sudhir Bhandari

SMS Medical College and Hospitals, Jaipur, Rajasthan, India

Amit Tak (✉ dramittak@gmail.com)

SMS Medical College and Hospitals, Jaipur, Rajasthan, India <https://orcid.org/0000-0003-2509-2311>

Jitendra Gupta

SMS Medical College and Hospitals, Jaipur, Rajasthan, India

Bhoopendra Patel

Government Medical College, Barmer, Rajasthan, India

Jyotsna Shukla

SMS Medical College and Hospitals, Jaipur, Rajasthan, India

Ajit Singh Shaktawat

SMS Medical College and Hospitals, Jaipur, Rajasthan, India

Sanjay Singhal

SMS Medical College and Hospitals, Jaipur, Rajasthan, India

Abhishek Saini

SMS Medical College and Hospitals, Jaipur, Rajasthan, India

Shivankan Kakkar

SMS Medical College and Hospitals, Jaipur, Rajasthan, India

Amitabh Dube

SMS Medical College and Hospitals, Jaipur, Rajasthan, India

Sunita Dia

Medstar Washington Hospital Center, Washington DC-20010, USA.

Mahendra Dia

North Carolina State University, Raleigh, NC 27695-7609, USA.

Todd C Wehner

North Carolina State University, Raleigh, NC 27695-7609, USA.

Research Article

Keywords: autoregressive integrated moving average, COVID-19, epidemic curve, forecast, mathematical modeling, prediction

Posted Date: July 7th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-40385/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Scripta Medica on July 7th, 2020. See the published version at <https://doi.org/10.5937/scriptamed52-29893>.

Abstract

The forecasting of Coronavirus Disease-19 (COVID-19) dynamics is a centerpiece in evidence based disease management. Numerous approaches that use mathematical modeling have been used to predict the outcome of the pandemic, including data driven models, empirical and hybrid models. This study was aimed at prediction COVID-19 evolution in India using a model based on autoregressive integrated moving average (ARIMA). Retrieving real time data from the Johns Hopkins dashboard from 11 Mar 2020 to 25 Jun 2020 (N = 107 time points) to fit the model. The ARIMA (1,3,2) and ARIMA (3,3,1) model fit best for cumulative cases and deaths respectively with minimum Akaike Information Criteria. The prediction of cumulative cases and deaths for next 10 days from 26 Jun 2020 to 05 Jul 2020 showed a trend toward continuous increment. The predicted root mean square error (PredRMSE) and base root mean square error (BaseRMSE) of ARIMA(1,3,2) model was 21137 and 166330 respectively. Similarly, PredRMSE and BaseRMSE of ARIMA(3,3,1) model was 668.7 and 5431 respectively. We propose that data on COVID-19 be collected continuously, and that forecasting continue in real time. The COVID-19 forecast assist government in resource optimization and evidence based decision making for a subsequent state of affairs.

Introduction

Coronavirus Disease-19 (COVID-19) caused by novel *Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)* has caused a pandemic with global devastation to human life and health. SARS-CoV-2 is closely related to two bat-derived SARS like coronaviruses, bat-SL-CoVZC45 and bat-SL-CoVZXC21. It is transmitted by human-to-human transmission via droplets or direct contact, and mean incubation period is 6.4 days [1]. According to World Health Organization, 10,185,374 confirmed cases and 503,862 deaths have been recorded by 1 Jul 2020 [2]. Further, India recorded 568,092 confirmed cases and 17400 deaths by the same time [3]. The novelty and rapid spread of *SARS CoV-2* has challenged medical science across the disciplines of epidemiology, clinical signs and symptoms, pathophysiology, disease progression and evolution and management and its preventive protocol. Meanwhile, mathematical models have been used to predict the course of disease and mortality [4]. Such models have potential in disease management as well as preventive protocols that are cost effective that can help optimum allocation of resources to manage the disease.

An econometric model has been proposed in the present study to predict and extrapolate transmission of COVID-19. The model makes use of *autoregressive integrated moving average modeling* on the epidemiological dataset of *COVID-19*. Our objective is to estimate forecast of COVID-19 cumulative cases and deaths for India. We propose that data on COVID-19 be collected continuously, and that forecasting continue in real time in order to assist governments in evidence based decision making.

Methods

A retrospective cohort study was run to forecast coronavirus–19 disease evolution in Indian subcontinent. The time series of COVID–19 cumulative cases and deaths from 11 Mar 2020 to 25 Jun 2020 (N = 107 time points) was used in prediction using autoregressive integrated moving average modeling. A useful ARIMA model depends on the number of sample time points, and a good series would have more than 50 sample points [5].

Data acquisition:

The Indian data of COVID–19 cumulative cases and deaths from 11 March 2020 to 25 Jun 2020 were sourced from official website of Johns Hopkins University Center for Systems Science and Engineering (<https://systems.jhu.edu/>) and the repository (<https://github.com/CSSEGISandData/COVID-19>) [6]. Excel 2010 was used to build the database.

To validate the model, dataset was distributed into a training set and a validation set. The training dataset from 11 March 2020 to 04 Jun 2020 (N = 86) was used to fit autoregressive integrated moving average (ARIMA) model and estimation of parameters. The validation dataset from 05 Jun 2020 to 25 Jun 2020 N = 21 was used to validate of the model (Table 1).

Model description

The autoregressive integrated moving average model was proposed to estimate the forecast of COVID–19 evolution [7][8][9]. We followed Box-Jenkins methodology (Figure 1). The methodology encompasses three phases—identification, estimation and testing and, and application. The Identification phase involves data preparation and model selection. The data were analyzed for trends and seasonal components. The Augmented Dickey-Fuller (ADF) unit root test was performed on time series, to examine stationarity. Logarithmic transformation and differencing operations were performed to stabilize the time series. Selection of ARIMA model, was required to establish the order of the autoregressive (AR) process, 'p', the differencing operator, 'd', and the order of moving average (MA) process, 'q'. The succeeding system of mathematical equations delineates the ARIMA (p,d,q) model:

$$W_t = \alpha_1 W_{t-1} + \alpha_2 W_{t-2} + \dots + \alpha_{t-p} W_{t-p} + a_t + \beta_1 a_{t-1} + \beta_2 a_{t-2} + \dots + \beta_q a_{t-q}$$

$$W_t = \nabla^d X_t$$

where, $\{X_t\}$ is the original time series, W_t is the time series acquired after differencing operation on $\{X_t\}$, ∇^d is the difference operator, α, β, q are parameters and a_t is the white noise. Since,

$$\nabla^d \equiv (1-B)^d$$

where B is the backward shift operator⁴, a series $\{X_t\}$ is integrated of order d if

$$W_t = (1-B)^d X_t$$

The autocorrelation function (ACF) and partial autocorrelation function (PACF) were used to guide the autoregressive and moving average order respectively. The Akaike Information Criterion (AIC) was used to select the best model. The AIC contemplate the maximum log likelihood estimation (Log L) and number of parameters, as the criterion for best model selection. The minimization of the following equation is required:

$$AIC = -2\log\text{maximum likelihood} + 2k$$

where, $k = p+q+1$, $2k$ serves as a penalty function. The model having minimum AIC value was considered the best. Estimation of parameters was accomplished using the method of maximum likelihood estimation. The assumptions of the model were scrutinized with residual diagnostics. The forecast of cumulative cases and deaths were performed from the best selected ARIMA model. The validation of the model was performed from validation dataset by enumerating Predicted Root Mean Square Error (PredRMSE) and comparing it with Base Root Mean Square Error (Base RMSE). The forecast of cumulative cases and deaths was performed from 26 Jun to 05 Jul 2020.

Statistical analysis

The estimates from the ARIMA model follows a normal distribution, and variation of the forecast was considered within 95% confidence intervals. The MS Excel 2010 was used to maintain the database and MATLAB 2016a (version 9.0.0.341360) was used for analysis [10].

Results

In order to achieve stationarity, the logarithmic transformation of datasets (cumulative cases and deaths) were performed as they were showing upward trends. Further difference operations were performed on cumulative cases ($d = 3$) and death ($d = 2$) datasets. The Augmented Dickey-Fulller (ADF) test showed stationarity (P value < 0.05). The autocorrelation function and partial autocorrelation functions were calculated and correlograms were used to guide the autoregressive order and the moving average order of the ARIMA models (Supplemental Figure 1 and 2). The ARIMA (1,3,2) and ARIMA (3,3,1) model for cumulative cases (AIC = 2.32) and death cases (AIC = 108.75) was chosen respectively with lowest AIC values. Goodness of fit and model assumptions were tested with residual analysis. The graph of standardized residuals showed random distribution of residuals. The middle values of the QQ plot were on a straight line. The autocorrelation function and partial autocorrelation functions within random limits, showed normal distribution and independence (Supplemental Figure 3 and 4). The parameters of both the ARIMA models were estimated (Supplemental Table S1 and S2).

Model validation was conducted by detecting the differences between the observed values and predicted values from the validation dataset. The PredRMSE and BaseRMSE of ARIMA(1,3,2) model for time series

of cumulative cases was 21137 and 166330 respectively. Similarly, PredRMSE and BaseRMSE of ARIMA(3,3,1) model for time series of death cases was 668.7 and 5431 respectively. Lower values of PredRMSE than Base RMSE indicated a good fit (Table 2). The estimation of forecast for cumulative cases and death cases were performed at 95% confidence interval (Figure 2 and 3; Table 3).

Discussion

The dynamics of any infectious disease involves interactions of three elements - host, agent and environment [11]. India lies in both the Northern and Eastern hemispheres, with latitudes 8°4' to 37°6' north and longitudes 68°7' to 97°25' east results in high environmental variability, and subsequent effects on human behavior and society [12]. Mathematical models are a centerpiece to forecasting the dynamics of infectious disease pandemics. The fundamental models used in prediction include data driven ([13][14][15][16][17]) empirical [18][19][20], hybrid [21][22] models. The parameters of empirical models (incubation period, attack rate and recovery rate) possess probability distributions such as Erlang and the Poisson distribution [11]. The data driven models include ARIMA, single-input and single-output (SISO) model [23] and AI based models (machine learning and deep learning techniques) [14]. The significance of ARIMA model lies in modeling non-stationary time series. The ARIMA model assumes the trends will continue in the future indefinitely as against the empirical models which assumes convergence. Few studies used non-parametric models like fourier decomposition methods to predict turn around dates of the epidemic and the results were found to agree with popular SIR models [24].

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) expressed its presence in India with the first case diagnosed on 30 Jan 2020. On 01 Jul 2020 the cumulative cases and deaths in India reached 568,092 and 17400 respectively [3]. India ranked 57th among the 100 countries in Global Health Security Index 2019, a scale to gauge preparedness for the outbreak of serious infectious diseases [25]. The forecast of COVID-19 cases helps government agencies in early preparedness to combat subsequent state of affairs. The patients of COVID-19 can be classified into confirmed cases, recovered cases, admitted and death cases. All categories have differential importance from the management point of view. The requirement of hospital beds, medical equipment, hospital staff is a function of number of admitted cases. Similarly, procurement of plasma for antibody therapy is a function of number of recovered cases.

The present study forecast the cumulative cases and deaths for India from 26 Jun 2020 to 05 Jul 2020. The forecast shows continuously increasing trends for both cumulative cases and deaths and shows no decreasing trends till October 2020.

The ARIMA model was proposed by various authors for forecasting the COVID-19 evolution. The data from 10 Jan 2020 to 20 Feb 2020 was used to fit ARIMA (1,0,4) and ARIMA (1,0,3) model for cumulative diagnoses and newer diagnoses to forecast next two days [26]. The ARIMA and wavelet hybrid model was proposed to forecast ten time points of cumulative cases from 05 Apr 2020 to 14 Apr 2020 for India. The forecast showed oscillations, may be due to effects of lockdown [22]. In another data driven model,

using bi-directional LSTM (long short term memory) model, 15 days prediction of actual cases in India from 30 Apr 2020 to 14 May 2020 showed error of less than 3% [21]. Susceptible-Exposed-Infectious-Recovered (SEIR) model was used to predict cumulative cases of COVID-19 in India during lockdown and post lockdown. The model predicts peak of cumulative cases around 43,000 in mid-May. However, 7–21% increase in peak value of cases was predicted for post lockdown period, reflecting relaxation in control strategies [20]. The evolution of COVID-19 in topmost affected states of India was done using SISO model [23]. The most severely affected states were Maharashtra, Gujarat, Tamil Nadu, Delhi and Rajasthan [23][27][28][29]. One time series analysis used genetic programming to predict the COVID-19 evolution in India. On 13 May 2020, the cumulative cases and death cases were 80,000 and 2500 respectively. The prediction for next ten days was done with 142,000 and 4,200 cumulative cases and deaths respectively on 23 May 2020 [30]. Sujath used linear regression, multilayer perceptron and vector autoregressive method and 80 time points till 10 Apr 2020, to predict confirmed cases, deaths and recovered cases from 11 Apr 2020 to 18 Jun 2020. Although prediction varies across the methods and did not seem very accurate [31]. Yadav used six regression analysis based machine learning models for prediction and found six degree polynomial model predict very close to observed data [32].

As the epidemic continues, the effects of different interventional strategies inherited in the time series, and thus data driven models seem to be more accurate than empirical models. The real-time data modeling has been proposed for monitoring trends of the Indian subcontinent. Evidence based interventions should be implemented to control the pandemic.

Conclusion

The present study produces encouraging results with the potential to serve as a good adjuvant to existing models for continuous predictive monitoring of the COVID-19 pandemic. The forecast of COVID-19 may assist public health authorities and governmental agencies for early preparedness and evidence based decision making.

Limitations Of The Study

Firstly, ARIMA model used in forecasting does not capture the non-linear and chaotic dynamics of the pandemic. Secondly, parameter selection procedure requires repetition with each time series update.

Declarations

Contribution of Authors: All authors contributed equally.

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

Data Availability— Data is available on reasonable request from corresponding author.

Ethics Declarations: The present study used de-identified data from public domain, thus exempted from ethical review by Institutional Ethics Committee.

References

- [1] Lai C-C, Shih T-P, Ko W-C, Tang H-J, Hsueh P-R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *Int. J. Antimicrob. Agents* [Internet]. 2020 Mar;55(3):105924. Available from: <http://dx.doi.org/10.1016/j.ijantimicag.2020.105924>
- [2] World Health Organization, WHO Coronavirus Disease (COVID-19) Dashboard, cited on: 01 Jul 2020; Available from: <https://covid19.who.int/>
- [3] Ministry of Health and Family Welfare, Government of India, COVID-19 India; cited on: 01 Jul 2020; Available from: <https://www.mohfw.gov.in/>
- [4] Bhandari S, Shaktawat AS, Tak A, Patel B, Shukla J, Singhal S, et al. Logistic regression analysis to predict mortality risk in COVID-19 patients from routine hematologic parameters. *Ibnosina J Med Biomed Sci* [serial online] 2020 [cited 2020 Jul 1];12:123-9. Available from: <http://www.ijmbs.org/text.asp?2020/12/2/123/288204>
- [5] Box GEP, Tiao GC. Intervention Analysis with Applications to Economic and Environmental Problems. *J. Am. Stat. Assoc* [Internet]. 1975 Mar;70(349):70-9. Available from: <http://dx.doi.org/10.1080/01621459.1975.10480264>
- [6] Johns Hopkins University Center for Systems Science and Engineering, 2019. (Accessed: 25th Jun 2020). Available from: <https://github.com/CSSEGISandData/COVID-19>
- [7] Cryer JD, Chan KS. Models for Non-stationary time series in Time series analysis: with applications in R (2nd edition) 98-99 (Springer Science & Business Media, 2008).
- [8] Indrayan, Abhaya, Rajeev Kumar Malhotra Relationships: Quantitative Outcome in *Medical biostatistics* (4th edition) 456 (CRC Press, 2018).
- [9] Metcalfe AV, Cowpertwait PS. Non-stationary models in Introductory time series with R 137-140 (Springer-Verlag New York, 2009).
- [10] MATLAB Team, Statistics and Machine Learning Toolbox 10.2. version 9.0.0.341360 (R 2016a). Natick, Massachusetts: The Mathworks Inc
- [11] Shinde GR, Kalamkar AB, Mahalle PN, Dey N, Chaki J, Hassanien AE. Forecasting Models for Coronavirus Disease (COVID-19): A Survey of the State-of-the-Art. *SN Computer Science* [Internet]. 2020 Jun 11;1(4). Available from: <http://dx.doi.org/10.1007/s42979-020-00209-9>

- [12]National Portal of India (cited on: 27 April 2020). Available from: <https://www.india.gov.in/india-glance/profile>
- [13]Yang C, Wang J. A mathematical model for the novel coronavirus epidemic in Wuhan, China *Math Biosci Eng* [Internet]. 2020;17(3):2708–24. Available from: <http://dx.doi.org/10.3934/mbe.2020148>
- [14]Chatterjee A, Gerdes MW, Martinez SG. Statistical Explorations and Univariate Timeseries Analysis on COVID–19 Datasets to Understand the Trend of Disease Spreading and Death. *Sensors* [Internet]. 2020 May 29;20(11):3089. Available from: <http://dx.doi.org/10.3390/s20113089>
- [15]Tiwari S, Kumar S, Guleria K. Outbreak Trends of Coronavirus Disease–2019 in India: A Prediction. *Disaster Med Public Health Prep* [Internet]. 2020 Apr 22;1–6. Available from: <http://dx.doi.org/10.1017/dmp.2020.115>
- [16]Tomar A, Gupta N. Prediction for the spread of COVID–19 in India and effectiveness of preventive measures. *Sci. Total Environ* [Internet]. 2020 Aug;728:138762. Available from: <http://dx.doi.org/10.1016/j.scitotenv.2020.138762>
- [17]Tuli S, Tuli S, Tuli R, Gill SS. Predicting the growth and trend of COVID–19 pandemic using machine learning and cloud computing. *Internet of Things* [Internet]. 2020 Sep;11:100222. Available from: <http://dx.doi.org/10.1016/j.iot.2020.100222>
- [18]Giordano, Giulia, Franco Blanchini, Raffaele Bruno, Patrizio Colaneri, Alessandro Di Filippo, Angela Di Matteo, and Marta Colaneri. Modelling the COVID–19 epidemic and implementation of population-wide interventions in Italy *Nat. Med.* (2020): 1–6.
- [19]Mandal M, Jana S, Nandi SK, Khatua A, Adak S, Kar TK. A model based study on the dynamics of COVID–19: Prediction and control. *Chaos Soliton Fract* [Internet]. 2020 Jul;136:109889. Available from: <http://dx.doi.org/10.1016/j.chaos.2020.109889>
- [20]Pai C, Bhaskar A, Rawoot V. Investigating the dynamics of COVID–19 pandemic in India under lockdown. *Chaos Soliton Fract* [Internet]. 2020 Sep;138:109988. Available from: <http://dx.doi.org/10.1016/j.chaos.2020.109988>
- [21]Arora P, Kumar H, Panigrahi BK. Prediction and analysis of COVID–19 positive cases using deep learning models: A descriptive case study of India. *Chaos Soliton Fract* [Internet]. 2020 Oct;139:110017. Available from: <http://dx.doi.org/10.1016/j.chaos.2020.110017>
- [22]Chakraborty T, Ghosh I. Real-time forecasts and risk assessment of novel coronavirus (COVID–19) cases: A data-driven analysis. *Chaos Soliton Fract* [Internet]. 2020 Jun;135:109850. Available from: <http://dx.doi.org/10.1016/j.chaos.2020.109850>
- [23]Rafiq D, Suhail SA, Bazaz MA. Evaluation and prediction of COVID–19 in India: A case study of worst hit states. *Chaos Soliton Fract* [Internet]. 2020 Oct;139:110014. Available from:

<http://dx.doi.org/10.1016/j.chaos.2020.110014>

[24]Singhal A, Singh P, Lall B, Joshi SD. Modeling and prediction of COVID–19 pandemic using Gaussian mixture model. *Chaos Soliton Fract* [Internet]. 2020 Sep;138:110023. Available from:

<http://dx.doi.org/10.1016/j.chaos.2020.110023>

[25]Tabish SA. The COVID–19 pandemic: Emerging perspectives and future trends. *J Public Health Res* [Internet]. 2020 Jun 4;9(1). Available from: <http://dx.doi.org/10.4081/jphr.2020.1786>

[26]Benvenuto D, Giovanetti M, Vassallo L, Angeletti S, Ciccozzi M. Application of the ARIMA model on the COVID–2019 epidemic dataset. *Data Brief* [Internet]. 2020 Apr;29:105340. Available from:

<http://dx.doi.org/10.1016/j.dib.2020.105340>

[27]Bhandari S, Shaktawat AS, Tak A, Patel B, Gupta K, Gupta J, et al. A multistate ecological study comparing evolution of cumulative cases (trends) in top eight COVID–19 hit Indian states with regression modeling. *Int J Acad Med* [serial online] 2020 [cited 2020 Jul 1];6:91–5. Available from: <http://www.ijam-web.org/text.asp?2020/6/2/91/287965>

[28]Kakkar S, Bhandari S, Shaktawat A, Sharma R, Dube A, Banerjee S, et al. A preliminary clinico-epidemiological portrayal of COVID–19 pandemic at a premier medical institution of North India. *Ann Thorac Med* [Internet]. 2020;15(3):146. Available from: http://dx.doi.org/10.4103/atm.ATM_182_20

[29]Bhandari S, Sharma R, Singh Shaktawat A, Banerjee S, Patel B, Tak A, et al. COVID–19 related mortality profile at a tertiary care centre: a descriptive study. *Scr Med* 2020;51(2):69–73.

DOI:10.5937/scriptamed51–27126

[30]Salgotra R, Gandomi M, Gandomi AH. Time Series Analysis and Forecast of the COVID–19 Pandemic in India using Genetic Programming *Chaos Soliton Fract* [Internet]. 2020 Sep;138:109945. Available from:

<http://dx.doi.org/10.1016/j.chaos.2020.109945>

[31]Sujath R, Chatterjee JM, Hassanien AE. A machine learning forecasting model for COVID–19 pandemic in India. *Stoch Env Res Risk A* [Internet]. 2020 May 30;34(7):959–72. Available from:

<http://dx.doi.org/10.1007/s00477–020–01827–8>

[32]Yadav RS. Data analysis of COVID–2019 epidemic using machine learning methods: a case study of India. *International Journal of Information Technology* [Internet]. 2020 May 26; Available from:

<http://dx.doi.org/10.1007/s41870–020–00484-y>

Tables

Table 1. Distribution of complete datasets of cumulative cases and death cases into training and validation subsets.

| Sub-datasets | Cumulative cases | Death cases |
|----------------|---------------------------------|---------------------------------|
| Training set | 11 Mar - 04 Jun 2020 (N=86) | 11 Mar - 04 Jun 2020 (N=86) |
| Validation set | 05 Jun - 25 Jun 2020 (N= 21) | 05 Jun - 25 Jun 2020 (N= 21) |

Table 2. Goodness of fitted models for cumulative cases and deaths cases using prediction root mean square error and base root mean square error.

| Dataset | Model | PredRMSE | BaseRMSE |
|------------------|---------------|----------|----------|
| Cumulative cases | ARIMA (1,3,2) | 21137 | 166330 |
| Death cases | ARIMA (3,3,1) | 668.70 | 5431 |

*ARIMA: autoregressive integrated moving average; PredRMSE: predicted root mean square error; BaseRMSE: base root mean square error.

Table 3. Forecast of cumulative cases and death cases using ARIMA(1,3,2) and ARIMA (3,3,1) models respectively. The dataset used for prediction is from 11 Mar 2020 to 25 Jun 2020 (N= 107 time points).

| Date | Cumulative cases | | | Cumulative deaths | | |
|-----------|------------------|--------|--------|-------------------|--------|-------|
| | Forecast data | 95% CI | | Forecast | 95% CI | |
| | | LL | UL | | LL | UL |
| 26 Jun 20 | 508210 | 378270 | 638150 | 15732 | 12399 | 19065 |
| 27 Jun 20 | 526760 | 396820 | 656700 | 16243 | 12910 | 19575 |
| 28 Jun 20 | 546010 | 416070 | 675950 | 16711 | 13379 | 20044 |
| 29 Jun 20 | 566050 | 436100 | 695990 | 17198 | 13865 | 20530 |
| 30 Jun 20 | 586900 | 456960 | 716840 | 17740 | 14408 | 21073 |
| 1 Jul 20 | 608630 | 478690 | 738570 | 18314 | 14982 | 21647 |
| 2 Jul 20 | 631290 | 501350 | 761230 | 18873 | 15540 | 22205 |
| 3 Jul 20 | 654940 | 525000 | 784890 | 19480 | 16147 | 22813 |
| 4 Jul 20 | 679660 | 549710 | 809600 | 20141 | 16808 | 23474 |
| 5 Jul 20 | 705500 | 575560 | 835440 | 20818 | 17485 | 24150 |

*CI: confidence interval; LL: lower limit of CI; UL: upper limit of CI.

Figures

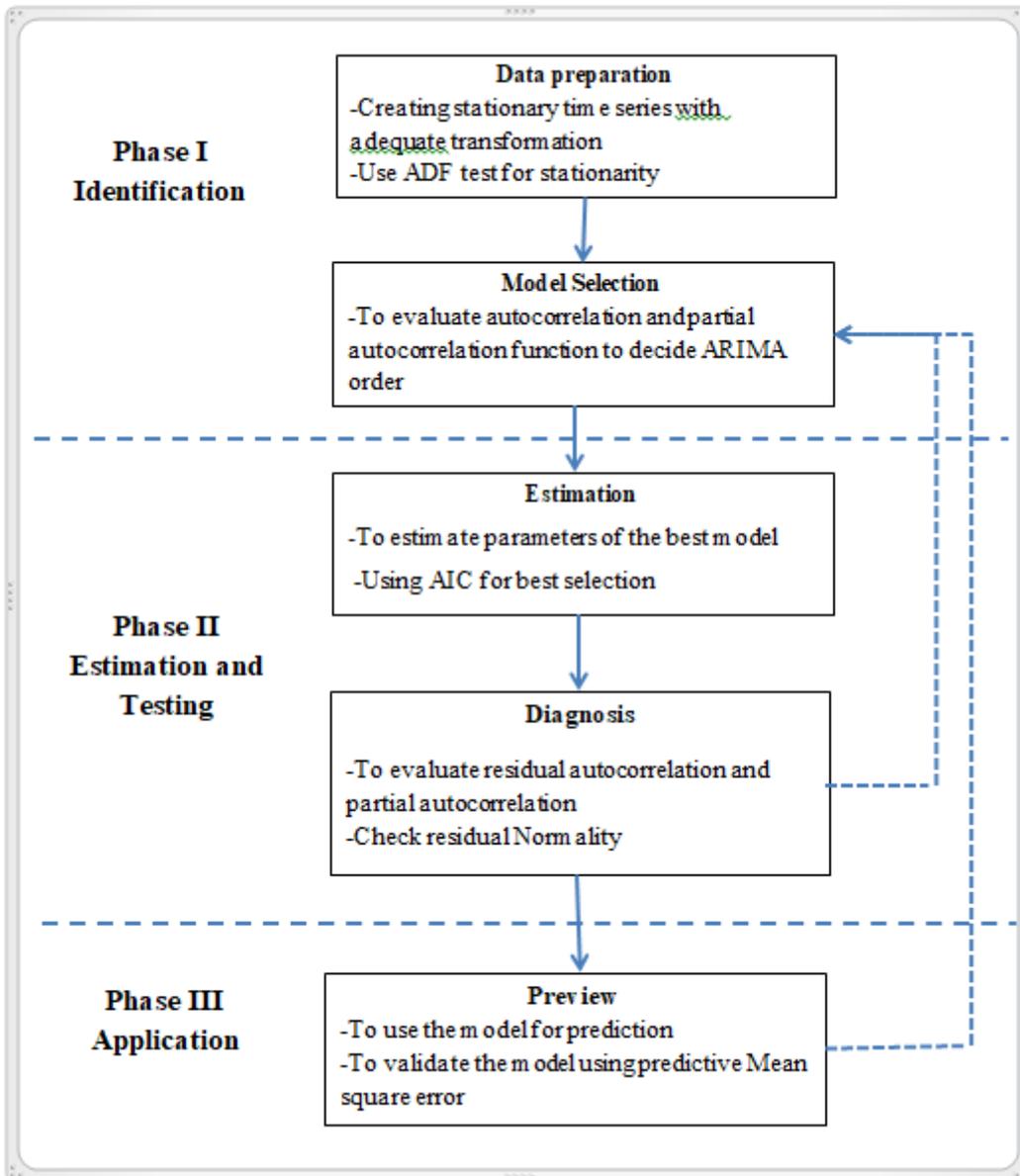


Figure 1

Flowchart outline Box Jenkins Methodology for implantation of autoregressive integrated moving average modeling

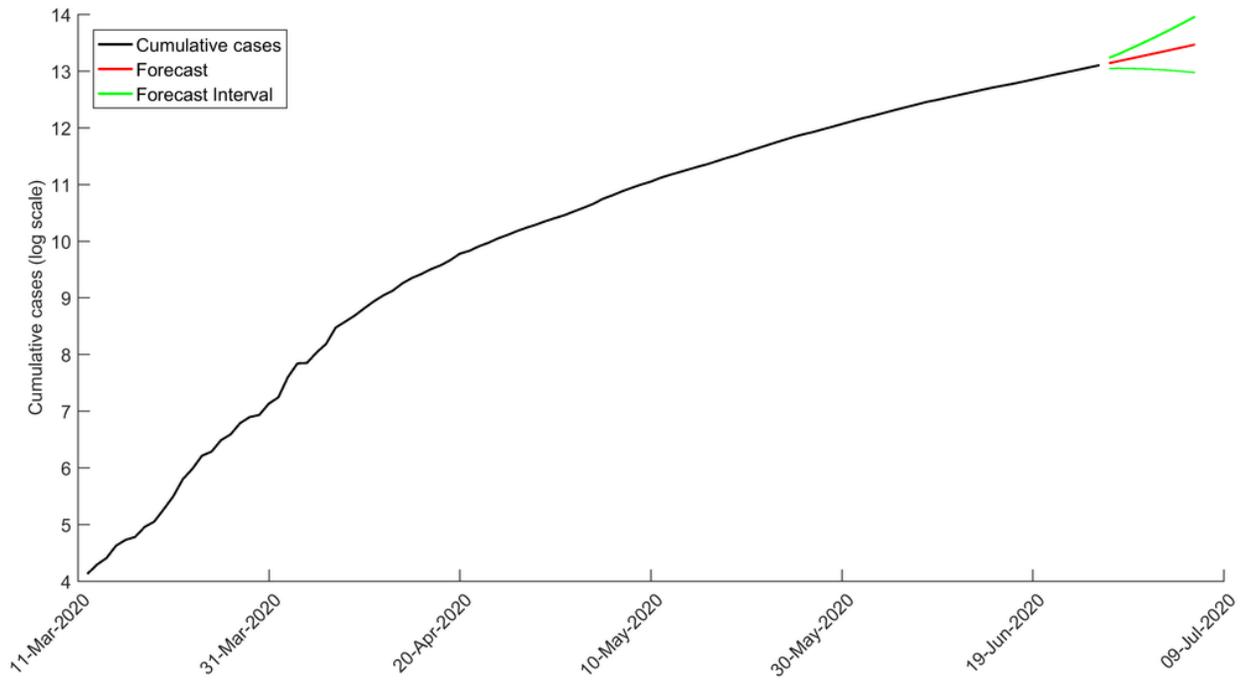


Figure 2

Line plot shows observed (black line plot) and predicted (magenta line plot) cumulative cases of COVID-19 with 95% confidence interval (green line plot). The forecast is based on ARIMA (1,3,2) model fitted with time series of cumulative cases from 11 March to 25 Jun 2020.

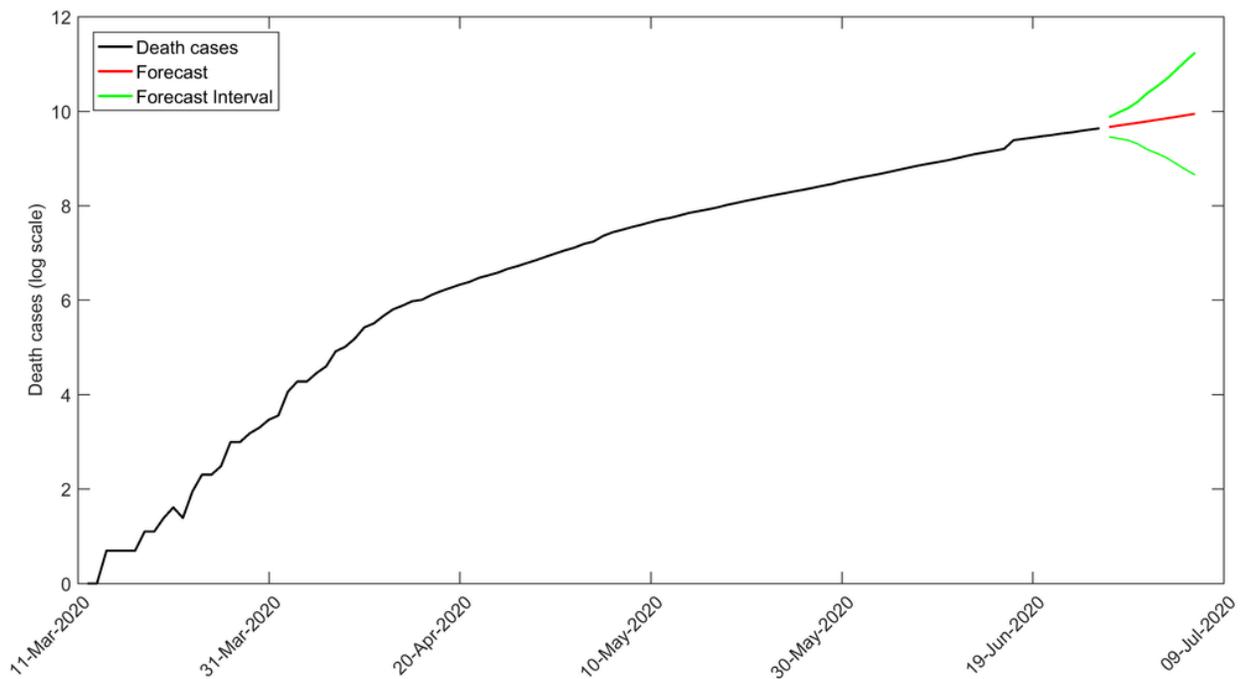


Figure 3

Line plot shows observed (black line plot) and predicted (magenta line plot) death cases of COVID-19 with 95% confidence interval (green line plot). The forecast is based on ARIMA (3,3,1) model fitted with time series of death cases from 11 March to 25 Jun 2020.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformationFile.docx](#)