Journal of Agriculture, Biology and Applied Statistics
Vol. 1, No. 1, 2022, pp. 1-24

URL : www.arfjournals.com
https://doi.org/10.46791/jabas.2022.v01i01.01

State of the Art in COVID-19 in the SAARC Countries and China using BATS,TBATS, Holt's Linear and ARIMA Model

Pradeep Mishra[1*], Mostafa Abotaleb[2], Kadir Karakaya[3], Amr Mostafa[4], Harun Yonar[5], Hazhar Talaat Abubaer Blbas[6], Umme Habibah Rahman[7] and S S Das[8]

[1]*College of Agriculture,Powarkheda,J.N.K.V.V. (M.P.) India*
[2]*Department of Electrical Engineering and Computer Science,South Ural State University, Chelyabinsk, Russia*
[3]*Selçuk University, Faculty of Science, Department of Statistics, Konya, Turkey*
[4]*University of New England, Faculty of Science, School of Science and Technology, Australia*
[5]*Selçuk University, Faculty of Veterinary Medicine,Deparment of Biostatistics, Konya, Turkey*
[6]*Statistics Department, College of Administration and Economics, Salahaddin University-Erbil, Erbil, Iraq*
[7]*Departments of Statistics, Assam University, Silchar, India*
[8]*Ranchi Agriculture College, Ranchi ,Birsa Agricultural University,India*
[*]*Corresponding E-mail:pradeepjnkvv@gmail.com*

To cite this article

Pradeep Mishra, Mostafa Abotaleb, Kadir Karakaya, Amr Mostafa, Harun Yonar, Hazhar Talaat Abubaer Blbas, Umme Habibah Rahman and S S Das. (2022). State of the Art in COVID-19 in the SAARC Counties and China using BATS,TBATS, Holt's Linear and ARIMA Model. *Journal of Agriculture, Biology and Applied Statistics*. Vol. 1, No. 1, pp. 1-24.

Abstract: Corona Virus is the biggest global health disease and it is an epidemic according to the World Health Organization (WHO). The purpose of this paper is to identify the best fitted model among BATS, TBATS, Holt's linear trend and ARIMA based on the minimum value of AIC and MAE and forecasting data about corona virus from SAARC and China countries. Cumulative daily data about covid19 has been collected from 3 January, 2020 to 10 March, 2021in the world health organization (WHO) in both infection and death cases. BATS, TBATS, Holt's linear trend and ARIMA models with some selection criteria are used to forecast the next Covid-19 situation in SAARC and China using R Language Program.

According to Akaike Information Criterion (AIC), Holt's linear trend model is the best model and better than both BATS and TBATS models in both infection and death cases. Depending on MAE criteria, ARIMA is the best fitted model for Nepal, Pakistan, Sri Lanka, and Bangladesh countries for both death and infection cases while Holt is the best fitted model for both India and Myanmar countries in both cases. On the other hand, BATS is the best fitted model for both death and infection cases in China.

Keywords: ARIMA,BATS, BATS, Holts Linear, MAE, AIC, COVID-19, Forecasting.

Article History Received : 05 January 2022; Revised : 25 January 2022; Accepted : 19 February 2022; Published : 30 June 2022

Introduction

The coronavirus disease 2019 (COVID-19) outbreak has crossed the globe since December 2019, causing widespread hurt and rising hospitalizations (Wu and McGoogan 2020; Zhu *et al.* 2020, Blbas *et al.* 2020).Covid19 is one of the biggest challenges facing the world at the present time. Since the spread of Covid19 in China, it quickly spread to the rest of the world, leaving 117332,262 cases of infection and 2605356 deaths until March 10, 2021, according to WHO reports, these numbers refer to How dangerous is Covid19, so in this article, we are modeling cases of infection, Covid 19, as well as deaths, in SAARC and china area that includes (India - China - Pakistan - Nepal - Bangladesh - Maymar - Sri Lanka).It is evident from the World Health Organization data that the most of these countries have cases of infection. India is also ranked first in terms of the number of deaths, while Sri Lanka ranks last in terms of the number of infections and deaths.There are many studies that have predicted infections and death (Abotaleb, 2020). This study predicted cases of infection, death, and recovery in three countries, such as China, Italy, and the United States of America. It was concluded that the Holt's Linear trend model is the best modeland it is better than the ARIMA model for forecasting infection, deaths, and recovery cases in that three countries. Also, another study to predict cases of infection and death in India were used ARIMA model and this study concluded that there is an increase in the number of deaths and infections (Mishra P, 2020). There are many studies that did not touch on BATS and TBATS(Trigonometric seasonality, Box-Cox transformation, ARMA errors, Trend and Seasonal components) models to predict cases of infection and death cases with Coronavirus.We have used cumulative daily data set about covid19 infection and death

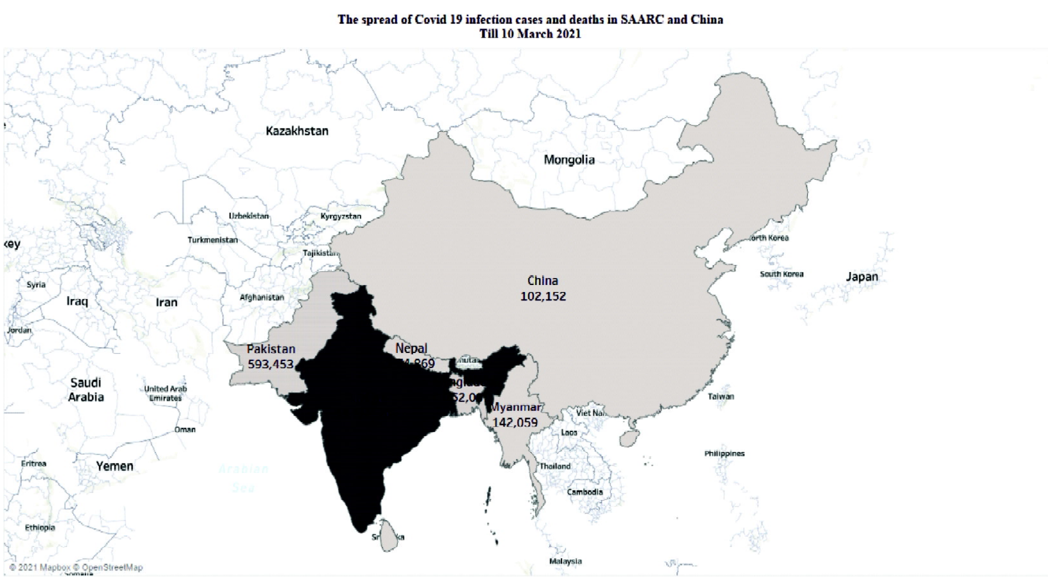

Figure A: Spread of COVID19 infection and death cases in the SAARC counties and China

cases in the world health organization (WHO) from 3 January, 2020 to 10 March, 2021, the data from 3 January, 2020 to 3 March, 2021 is used for training data while the last seven days from 4 March, 2020 to 10 March, 2021 is used for testing models.

Mishra *et al*. (2020) explored the COVID-19 database of India between 17th March to 1st July 2020 and estimated two nonlinear time series models such as ANN and FTS by comparing them with ARIMA model. Khan *et al* (2020) have modeled and forecasted new cases, deaths and recover cases of COVID-19 by using Vector Autoregressive model in Pakistan. Ramirez *et al* (2020) have applied the time series model to access the dynamics and forcasted monthly reports of abuse, neglect or exploitation of involving a vulnerable adult between January first 2014 to June 30th 2018 in South Carolina, USA. Rahman (2020) has forecasted Covid-19 data for the month of August and September of 2020 by using the data from the beginning that is January 30, 2020 to July 31, 2020 and used the time series method like ARIMA. Tekindal et ai (2020) Analyzed COVID-19 outbreak for Turkey Germany, United Kingdom, France, Italy, Russia, Canada and Japan with Curve Estimation Models, Box-Jenkins (ARIMA), Brown Linear Exponential Smoothing Method, Autoregressive Distributed Lag (ARDL) and SEIR Models. Mishra *et al* (2020) forecasted new and total deaths and occurrence of Covid-19 ARIMA and SARIMA found suitable respectively and forecasted for 1 September, 2020. Mohammed *et al* (2020) have used the data on new deaths, total deaths, total cases, new cases, collected on a daily basis from 13th March 2020 to 30th September 2020, obtained from the Ghana Health Services and forecasted by using ARIMA and SERIMA model.Ali *et al* (2020) forecasted COVID-19 in Pakistan from February 2020 to June 2020 by applying two different ARIMA models. Hierro *et al* (2020) predicted mortality for Covid 19 in the US using the delayed elasticity method. Lynch and Gore (2021) forecasted by using case count data provided by The New York Times as of April 22, 2020.

Material and Methods

Present investigation data is collect from (https://www.who.int/health-topics/coronavirus) for total cases and infection of SAARC counites along with China. Statistical analysis has been done using R statisitcal software.

BATS and TBATS Models

TBATS is an improvement modification of BATS that allows multiple seasonal incorrect cycles. TBATS has the following equation(De Livera, 2011)from figure number B that represented BATS and TBATS model

The first Equation (1) is a Box-Cox transformation, error modeled by ARMA

$$Y_t^{(\eta)} = \begin{cases} \dfrac{y_t^{(\eta)} - 1}{\eta} \eta \neq 0 \\ log y_t \eta = 0 \end{cases} \tag{1}$$

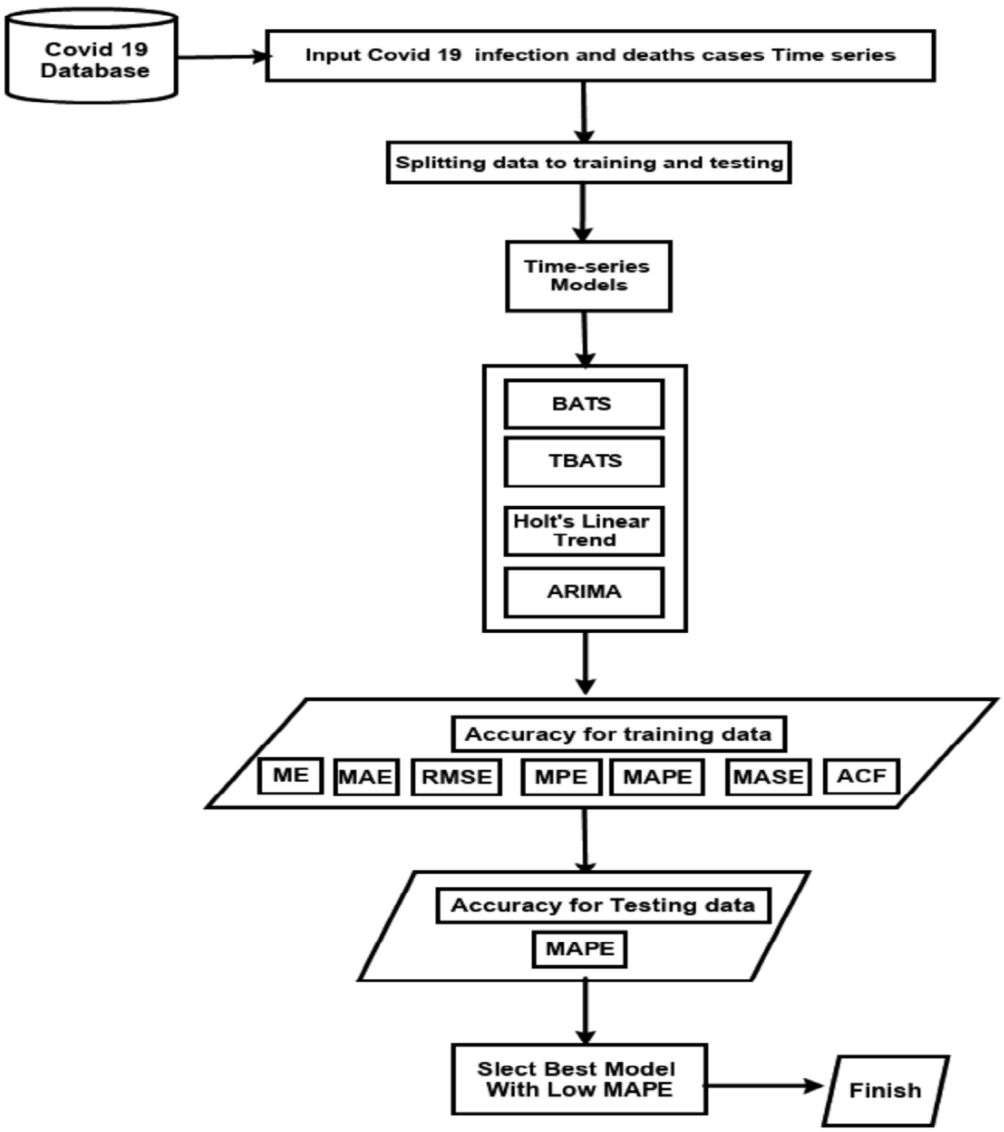

Figure B: Schema for select best model for forecasting

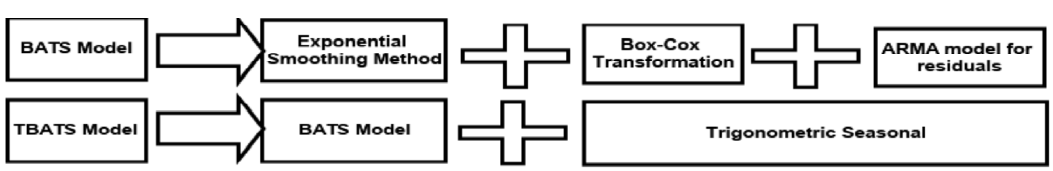

Figure B: BATS and TBATS model

The second Equation (2) represents the seasonal M pattern

$$Y_t^{(\eta)} = l_{t-1} + \xi Z_{t-1} + \sum_{i=1}^{T} s_{t-\rho_i}^{(i)} + d_t \tag{2}$$

global trends and local trends are Equations (3),(4), and (5)

$$l_t = l_{t-1} + \xi Z_{t-1} + \alpha d_t \tag{3}$$

$$b_t = \xi b_{t-1} + \beta d_t \tag{4}$$

$$s_t^{(i)} = s_{t-\rho_i}^{(i)} + \gamma_i d_t \tag{5}$$

Equation (6) error can be modeled by ARMA

$$d_t = \sum_{i=1}^{p} \varphi_i d_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_{t'} \tag{6}$$

From ρ_1,...to, ρ_T denote that seasonal period, level, and trend of components of time series can be denoted by l_t and Z_t at time t, the seasonal component can be denoted by $s_t^{(i)}$ at time t, d_t represents to ARMA(p, q) component and ε_t is white noise process.

The smoothing parameters are given by α, β, γ_i for $i = 1 \ldots T$ and ξ is the dampening parameter, which gives more control over trendextrapolation when the trend component is damped (Taylor, 2003). For seasonal data the following equations representing Trigonometric exponential smoothing models

$$s_t^{(i)} = \sum_{j=1}^{k_i} a_{j,t}^{(i)} \cos\left(\psi_j^{(i)} t\right) \tag{7}$$

$$a_{j,t}^{(i)} = a_{j,t-1}^{(i)} + k_1^{(i)} d_t \tag{8}$$

$$\beta_{j,t}^{(i)} = \beta_{j,t-1}^{(i)} + k_2^{(i)} d_t \tag{9}$$

The smoothing parametersare $k_1^{(i)}$ and $k_2^{(i)}$.

$\psi_j^{(i)} = 2\pi j / \rho_i$. This is an extended, modified single source of error version of single seasonal multiple sources of error representation suggested by (Hannan, 1970) (Harvey, 1990) and (Durbin, 2012)

$$a_{j,t}^{(i)} = s_{j,t}^{(i)} \cos\left(\psi_j^{(i)} t\right) - s_{jt}^{*(i)} \sin\left(\psi_j^{(i)} t\right) \tag{10}$$

$$\beta_{j,t}^{(i)} = s_{j,t}^{(i)} \sin\left(\psi_j^{(i)} t\right) - s_{jt}^{*(i)} \cos\left(\psi_j^{(i)} t\right) \tag{11}$$

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t'}^{(i)} \tag{12}$$

Where

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos \psi_j^{(i)} + s_{j,t-1}^{*(i)} \sin \psi_j^{(i)} + \left[k_1^{(i)} \cos(\psi_j^{(i)} t) + k_2^{(i)} \sin(\psi_j^{(i)} t)\right] d_t \tag{13}$$

$$s_{j,t}^{*(i)} = -s_{j,t-1} \sin \psi_j^{(i)} + s_{j,t-1}^{*(i)} \cos \psi_j^{(i)} + \left[k_2^{(i)} \cos(\psi_j t) - k_1^{(i)} \sin(\psi_j t)\right] d_t \tag{14}$$

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t'}^{(i)} \tag{15}$$

Equations (16) and (17) are seasonal patterns modeled by the Fourier model.

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos \psi_j^{(i)} + s_{j,t-1}^{*(i)} \sin \psi_j^{(i)} + \gamma_1^{(i)} d_t \tag{16}$$

$$s_{j,t}^{*(i)} = -s_{j,t-1} \sin \psi_j^{(i)} + s_{j,t-1}^{*(i)} \cos \psi_j^{(i)} + \gamma_2^{(i)} d_{t'} \tag{17}$$

The notation of TBATS $(p, q, \{\rho_1, k_1\}, \{\rho_2, k_2\}, \ldots, \{\rho_T, k_T\})$ is used for these trigonometric models.

Holt's Linear Trend Method

The exponentially weighted moving average is also the averages of smoothing random variability with the following advantages: (1) older data have a declining weightthat is very important that; (2) very simple to calculate; and (3) the most important for data set is that minimal data is needed. Holt, C.E. 1957 had given three equations for forecast, level, and trend(Holt, 1957) and Mishra *et al.* (2021)

Forecast Equation

$$\hat{\chi}_{t+\rho\backslash t} = \mathrm{M}_t + \rho\theta_t \tag{1}$$

Level Equation

$$\mathrm{M}_t = \upsilon\, \chi_t + (1-\omega)(\mathrm{M}_{t-1} + \theta_{t-1}) \tag{2}$$

Trend Equation

$$b_t = \gamma^*(\mathrm{M}_t - \mathrm{M}_{t-1}) + (1-\gamma^*)\theta_{t-1} \tag{3}$$

ARIMA Model

ARIMA model consist of three parts The first part is (AR) that is Autoregressive The second part is (I) integrated The third part is (MA) Moving Average so that model is named that Autoregressive integrated moving average (ARIMA). Sometimes data of time series not required integrated part to decline the seasonality and in that case ARIMA model represented as ARMA (p, q) model where p is the order of the autoregressive part and q is the order of the moving average and integrated part is equal zero ARIMA(p,0,q) that represented as ARMA (p, q).

The first part is Autoregressive model

Equation (1) The autoregressive model of order pis writtenasAR(p)

$$X_t = K + \sum_{i=1}^{P} \omega_i X_t + \varepsilon_t \tag{1}$$

Where ω_1, ω_2..... ω_p are the parameters of the model, K is a constant andSometimes the constant term is avoided is white noise.

The second part is Moving Average model

Equation (2) the moving average model of order (q) is written as MA (q)

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{2}$$

$$V\grave{}_t = K + \emptyset_1 V\grave{}_{t-1} + \emptyset_2 V\grave{}_{t-2} + \dots + \emptyset_p V\grave{}_{t-p} + e_t + \Theta_1 \varepsilon_{t-1} + \Theta_2 e_{t-2} + \cdots + \Theta_p e_{t-p} \tag{3}$$

After this using Ljung Box test performance of residual validated (Mishra,2021)and forecasting (Young, 1977), (Frain, 1992), (Kirchgässner, 2012), and (Chatfield, 2019).

Results

In Table 1, descriptive statistics such as mean, minimum, maximum, standard deviation, skewness and kurtosis for Covid-19 infection cases and Covid-19 death cases from 3 January 2020 to 3 March 2021 are given in detail. In this table the average of cumulative daily cases is highest in India with 4197111 while the lowest average of cumulative daily cases is in Myanmar with 37220. Also, while the average cumulative daily deaths are highest in India with 63878, the lowest cumulative average daily deaths are in Myanmar with 823.In all other countries except China has positive skewness for both Covid-19 infection cases and Covid-19 death cases. It is concluded from the coefficient of skewness that the right tail is longer and the distribution of the Covid-19 infection cases and Covid-19 death cases right-skewed for all other countries except China. It is also concluded from the coefficient of kurtosis in Table 1 that all countries have leptokurtic distribution namely which shows fatter tails.

From Table 2, BATS models for Covid-19 infection cases and deaths in SAARC and China from 3 January to 3 March are examined. The best-suited BATS models for Covid-19 infection cases are given respectively for countries India, China, Nepal, Pakistan, Sri Lanka, BangladeshandMyanmar: (1, {0,0}, 1, -), (1, {0,0}, 0.954, -), (1, {2,2}, 1, -), (1, {0,0}, 1, -), (1, {0,0}, 1, -),(1, {0,0}, 1, -), and (1, {0,0}, 1, -). Also, the best-suited BATS models for Covid-19 death cases are given respectively for countries India, China, Nepal, Pakistan, Sri Lanka, Bangladesh and Myanmar: (1, {0,0}, 1, -), (1, {0,0}, 0.959, -), (1, {0,0}, 1, -), (1, {0,0}, 1, -), (1, {0,0}, 1, -), (1, {0,0}, 1, -) and (1, {2,2}, 0.988, -).

In Table 3, TBATS models for Covid-19 infection cases and deaths in SAARC and China from 3 January to 3 March are studied. The best-suited TBATS models for Covid-19 infection cases are given respectively for countries India, China, Nepal, Pakistan, Sri Lanka, BangladeshandMyanmar:(1, {0,0}, 1, {<6,2>}), (1, {0,0}, 1, {<6,2>}), (1, {0,0}, 1, {<6,2>}), (1, {0,0}, 1, {<6,2>}), (1, {0,0}, 1, {<6,2>}), (1, {0,0}, 1, {<6,2>})and(1, {0,0}, 1, {<6,2>}).Also, the best-suited TBATS models for Covid-19 death cases are given respectively for countries India, China, Nepal, Pakistan, Sri Lanka, Bangladesh and Myanmar:(1, {0,0}, 1, {<6,2>}), (1, {0,0}, 1, {<6,2>}), (1, {0,0}, 1, {<6,2>}), (1, {0,0}, 1, {<6,2>}), (1, {0,0}, 1, {<6,2>}), (1, {0,0}, 1, {<6,2>})and(1, {0,0}, 1, {<6,2>}). Figure 1 and 2 giving the MAPE information of all testing models used in present study,

In Table 4, Holt's trend models are investigated for Covid-19 infection cases and deaths in SAARC and China. The Box-Cox transformations for Covid-19 infection cases are given respectively for countries India, China, Nepal, Pakistan, Sri Lanka, Bangladesh and Myanmar: 0.2949, 0.2224, 0.3729, 1.0000, 0.3267, 1.0000 and 0.3174. Also, the Box-Cox transformations for Covid-19 death cases are given respectively for countries India, China, Nepal, Pakistan, Sri Lanka, Bangladesh and Myanmar: 0.3777, 2.0000, 0.1929, 0.8821, 0.5738, 0.9975 and 0.8605.

In Table 5, the ARIMA models are examined in detailed for Covid-19 infection cases and deaths in SAARC and China. The best-fitted ARIMA models for Covid-19 infection cases are given respectively for countries India, China, Nepal, Pakistan, Sri Lanka, Bangladesh and Myanmar: ARIMA(1,2,3), ARIMA(0,1,2), ARIMA(2,2,3), ARIMA(1,2,3), ARIMA(2,2,3), ARIMA(1,2,2)andARIMA(3,2,2).Alsothebest-fitted ARIMA models for Covid-19 death cases are given respectively for countries India, China, Nepal, Pakistan, Sri Lanka, Bangladesh and Myanmar:ARIMA(2,2,3), ARIMA(0,2,1), ARIMA(0,2,2), ARIMA(0,2,5), ARIMA(2,2,2), ARIMA(1,2,2)and ARIMA(1,2,3).Forecasted figure of infection and total deaths cases presented in figures 3 to 16 for the SAARC counties and China,

In Table 6, the lowest values of the RMSE, MAE, and MAPE areused to copare among BATS, TBATS, Holt'slinear trend, and ARIMA models are given. InTable 8, MAPE for BATS, TBATS, Holt'slineartrend, and ARIMA models are calculated for Covid-19 infection cases and deaths in SAARC and China. According to Table 8, the best-fitted models are selected lowest values of MAPE and reported respectively for countries India, China, Nepal, Pakistan, Sri Lanka, Bangladesh and Myanmar as ARIMA(1,2,3), BATS, Holt Model, ARIMA(1,2,3), TBATS Model, BATS (or Holt Model) andARIMA(3,2,2). Also,the best-fitted models are selected lowest values of MAPEs and reported respectively for countries India, China, Nepal, Pakistan, Sri Lanka, Bangladesh and Myanmar:Holt Model, ARIMA(0,2,1), Holt Model, ARIMA(0,2,5), Holt Model, BATS (orHolt Model) andHolt Model.

Conclusions

In this study, Covid-19 is extensively analyzed and forecasted for SAARC and China using data from 3 January to 3 March 3. The BATS, TBATS, Holt's linear trend and ARIMA models were utilized for forecasting purposes. Some criteria such as ME, RMSE, MAE, MPE, MAPE and MASE were examined in order to select the best model. Very important results regarding the measures to be taken and the future were obtained in this study. In the future, we expect a decrease in both cases of infection and deaths in the short term, and this not means that quarantine and the measures imposed by the government to limit the spread of infection are abandoned, although the models used were able to identify the pattern of spread of Covid-19 in these countries with a high degree of accuracy, but there is also an expectation of an increase in cases of infection in the long run. Also has not been immunized with herd immunity, which is done by vaccinating from 60% to 70% of the country's population.

Refrences

Ali, M., Khan, D. M., Aamir, M., Khalil, U. and Khan, Z., Forecasting COVID-19 in Pakistan. Plos One, 2020, 1-13.

Blbas, H., Aziz, K. F., Nejad, S. H., & Barzinjy, A. A. (2020). Phenomenon of depression and anxiety related to precautions for prevention among population during the outbreak of COVID-19 in Kurdistan Region of Iraq: based on questionnaire survey.

Chatfield, C. A. (2019). *The analysis of time series: an introduction with R.* CRC press.

Frain, J. (1992). Lecture notes on univariate time series analysis and box jenkins forecasting. Economic. *Research and Publications*.

Hierro, L. A., Garzón, A. J., Montero, P. A. and Márquez, J. L. Predicting mortality for Covid 19 in the US using the delayed elasticity method. Scientifc Reports, 1-6.

Holt, C. E. (1957). Forecasting seasonals and trends by exponentially weighted averages (ONR Memorandum No. 52). *Carnegie Institute of Technology , Pittsburgh USA.*

Khana, F., Saeedb, A. and Al, S. Modelling and forecasting of new cases, deaths and recover cases of COVID-19 by using Vector Autoregressive model in Pakistan. Chaos, Solitons and Fractals Nonlinear Science, and Nonequilibrium and Complex Phenomena, 2020, 1-5.

Kirchgässner, G. J. (2012). *Introduction to modern time series analysis.* Springer Science & Business Media.

Lynch, C.J. and Gore, R., Application of one-, three-, and seven-day forecasts during early onset on the COVID-19 epidemic dataset using moving average, autoregressive, autoregressive moving average, autoregressive integrated moving average, and naïve forecasting methods. *Data in Brief*, 2021, 35, 1-12.

Makarovskikh Tatyana Anatolyevna, A. m. (2021). *Forecasting wind speed in England by using BATS, TBATS, and Holt's Linear trend model.* https://rpubs.com/abotalebmostafa/735338.

Mishra, P., Fatih, C., Rawat, D., Sahu, S., Pandey, S.S., Ray, M., Dubey, A. and Sanusi, O.M., Trajectory of COVID-19 Data in India: Investigation and Project Using Artificial Neural Network, Fuzzy Time Series and ARIMA Models. *Annual Research & Review in Biology, 2020,* 35(9), 46-54.

Mishra, P. , Khatib, A.M., Sardar, I., Mohammed, J., Ray,M., Kumar,M., Rawat,D., Pandey, S.A., Dubey, A., Feys, J. and Rono, K., Modelling and Forecasting of COVID-19 in India. *Journal of Infectious Diseases and Epidemiology*, 2020, 6, 1-11.

Mishra, P., Matuka, A., Abotaleb, M. S. A., Weerasinghe, W. P. M. C. N., Karakaya, K., and Das, S. S. 2021. Modeling and forecasting of milk production in the SAARC countries and China. *Modeling Earth Systems and Environment*, 1-13.

Mohammed, J., Khatib, A. M., Mishra, P., Adjei, P., Singh, P. K., Priya, S.R. K. and Das, S.S., Modeling and forecasting of Covid-19 from the context of Ghana. *African Review of Economics and Finance*, 2020, 1-18.

Rahmam, U.H. , Forecasting Covid-19 of India: An era. *Journal of Seybold Report*, 2020, 15, 2807-2818.

Ramirez, N. S., Odeku, J., Foxe, C., Flynn, C. and Tester, D. Applying Time Series modelling to assess the dynamics and forecast monthly reports of abuse, neglect and/or explotation involving a vulnerable adult. Achives of Public-Health, 1-8.

Tekindal, M. A., Yonar, H., Yonar, A., Tekindal, M., Cevrimli, M.B., Alkan, H., Inanc, Z.,S. and Mat, B. Analyzing COVID-19 outbreak for Turkey and Eight Country with Curve Estimation Models, Box-Jenkins (ARIMA), Brown Linear Exponential Smoothing Method, Autoregressive Distributed Lag (ARDL) and SEIR Models. Eurasian Journal of Veterinary Sciences, 2020, 142-155.

Wu Z, McGoogan JM (2020) Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China:summary of a report of 72 314 cases from the Chinese Center forDisease Control and Prevention. JAMA 323:1239–1242

Young, W. L. (1977). "The Box-Jenkins approach to time series analysis and forecasting: principles and applications. *RAIRO-Operations Research-Recherche Opérationnelle, 11*, 129-143.

Zeits chrift fur Gesund heitswissenschaften Journal of public health, 1–5. Advance online publication. https://doi.org/10.1007/s10389-020-01325-9

Zhu N *et al.* (2020) A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 382(8):727–733.

BATS, TBATS, Holt's Linear Trend, and ARIMA Models

- BATS, TBATS , Holt's Linear Trend, ARIMAModels were fitted by using R software.
- Data set From(3 January 2020) To (10 march 2021) cumulative daily data set of covid 19 infection cases and deaths and using the period from (3 January 2020) To (3 march 2021) for training data and using last 7 days from (4 march 2021) To (10 march 2021) for testing models

Table 1: Descriptive statistics COVID 19 infection cases and deaths in SAARC and China

Descriptive statistics COVID- 19 infection cases in SAARC and China						
Country	*Mean*	*Minimum*	*Maximum*	*Standard Deviation*	*Skewness*	*Kurtosis*
India	4197111	0	11262707	4425430	0.4240798	1.438539
China	81927	0	102152	24818.14	-2.529194	8.338625
Nepal	91496	0	274869	110122.1	0.7130723	1.725675
Pakistan	239231	0	593453	197614.8	0.1428546	1.717353
Sri Lanka	15312	0	86343	24462.55	1.650295	4.387297
Bangladesh	243065	0	552087	210690.6	0.09576721	1.398541
Myanmar	37220	0	142059	54069.89	1.031122	2.331919
Descriptive statistics COVID-19 deaths cases in SAARC and China						
India	63878	0	158063	63024.53	0.3240957	1.381315
China	4039	0	4848	1400.836	-1.988954	5.661521
Nepal	639.7	0	3012	852.0055	1.063637	2.784078
Pakistan	5029	0	13281	4218.113	0.2259007	1.847803
Sri Lanka	75.68	0	511	130.1267	1.890499	5.439092
Bangladesh	3513	0	8489	3134.428	0.2014343	1.469234
Myanmar	823	0	3200	1195.379	1.0359	2.380095

Table 2: BATS Model fitted for COVID-19 infection cases and deaths in SAARC and China using the period from (3 January 2020) To (3 march 2021) for training data

Country	*Model*	**Box-Cox transformation (Lambda)*	*Smoothing parameter*		*Damping Parametr For trend*	*ARMA Coefficients*		*prediction error*	
			Alpha	*Beta*		*AR coefficients*	*MA Cofficients*	*Sigma*	*AIC*
BATS Model fitted for Covid 19 infection cases in SAARC and China									
India	BATS(1, {0,0}, 1, -)	1	0.4266192	0.1583616	1	-	-	13180.27	10669.67
China	BATS(1, {0,0}, 0.954, -)	1	1.002489	0.3005035	0.954272	-	-	690.7027	8159.32
Nepal	BATS(1, {2,2}, 1, -)	1	1.733581	0.422903	1	-1.188685 -0.403793	0.382548 0.53659	352.8338	7601.021
Pakistan	BATS(1, {0,0}, 1, -)	1	0.5090203	0.2517022	1	-	-	826.526	8310.272
Sri Lanka	BATS(1, {0,0}, 1, -)	1	0.5589512	0.1789467	1	-	-	179.3466	7008.492
Bangladesh	BATS(1, {0,0}, 1, -)	1	0.8468071	0.252579	1	-	-	396.8151	7685.108
Myanmar	BATS(1, {0,0}, 1, -)	1	0.5875224	0.2053331	1	-	-	264.3873	7339.149
BATS Model fitted for Covid 19 deaths cases in SAARC and China									
India	BATS(1, {0,0}, 1, -)	1	0.5903846	0.1460504	1	-	-	192.6369	7069.399
China	BATS(1, {0,0}, 0.959, -)	1	0.964021	0.08668136	0.958916	-	-	66.08972	6159.934
Nepal	BATS(1, {0,0}, 1, -)	1	1.119837	0.01980915	1	-	-	30.07261	5487.071
Pakistan	BATS(1, {0,0}, 1, -)	1	0.7979561	0.1983245	1	-	-	22.57919	5242.9
Sri Lanka	BATS(1, {0,0}, 1, -)	1	0.8335172	0.102205	1	-	-	1.804009	3089.881
Bangladesh	BATS(1, {0,0}, 1, -)	1	0.7406643	0.210743	1	-	-	7.163127	4264.734
Myanmar	BATS(1, {2,2}, 0.988, -)	1	0.1691038	0.05230338	0.987871	0.579504 0.284813	0.032048 -0.109539	6.01962	4134.552

Table 3: TBATS Model fitted for COVID-19 infection cases and deaths in SAARC and China using the period from (3 January 2020) To (3 march 2021) for training data

Country	*Model*	**Box-Cox transformation (Lambda)*	*Smoothing parameter*				*Damping Parametr For trend*		*prediction error*
			Alpha	*Beta*	*Gamma-1 Values*	*Gamma-2 Values*		*Sigma*	*AIC*
TBATS Model fitted for Covid 19 infection cases in SAARC and China									
India	TBATS(1, {0,0}, 1, {<6,2>})	1	0.4319085	0.1605974	-0.003149591	0.005012398	1	12983.43	10668.85
China	TBATS(1, {0,0}, 1, {<6,2>})	1	1.015877	0.2982216	-0.001474908	0.0005871215	1	694.1428	8173.553
Nepal	TBATS(1, {0,0}, 1, {<6,2>})	1	0.8848285	0.397865	-0.003456192	0.002868234	1	371.3342	7640.562
Pakistan	TBATS(1, {0,0}, 1, {<6,2>})	1	0.5153473	0.2542032	-0.005445111	0.004315993	1	816.2209	8311.583
Sri Lanka	TBATS(1, {0,0}, 1, {<6,2>})	1	0.5590723	0.18113	0.001497917	0.004126179	1	177.8256	7013.236
Bangladesh	TBATS(1, {0,0}, 1, {<6,2>})	1	0.855153	0.2525394	-0.001465332	0.003255705	1	395.4919	7694.262
Myanmar	TBATS(1, {0,0}, 1, {<6,2>})	1	0.5866163	0.2060073	-0.0004500818	2.069486e-05	1	263.8409	7349.386
TBATS Model fitted for Covid 19 deaths cases in SAARC and China									
India	TBATS(1, {0,0}, 1, {<6,2>})	1	0.5912045	0.1470497	-0.002843247	0.004128876	1	190.2337	7070.703
China	TBATS(1, {0,0}, 1, {<6,2>})	1	0.9948524	0.06891627	-0.003860951	4.089986e-05	1	66.16334	6170.883
Nepal	TBATS(1, {0,0}, 1, {<6,2>})	1	1.121317	0.01984665	-0.001443719	0.0009320508	1	29.91364	5494.555
Pakistan	TBATS(1, {0,0}, 1, {<6,2>})	1	0.8048746	0.2006655	-0.004803871	0.002491345	1	22.27392	5243.302
Sri Lanka	TBATS(1, {0,0}, 1, {<6,2>})	1	0.8400533	0.1023039	-0.0002780891	0.006044675	1	1.795149	3097.686
Bangladesh	TBATS(1, {0,0}, 1, {<6,2>})	1	0.7452036	0.2117857	-0.003935554	0.003389093	1	7.120594	4271.66
Myanmar	TBATS(1, {0,0}, 1, {<6,2>})	1	0.6446687	0.1933936	-0.0002563382	-9.738689e-05	1	6.160466	4148.257

Table 4: Holt's Linear Trend Model fitted for COVID-19 infection cases and deaths in SAARC and China using the period from (3 January 2020) To (3 march 2021) for training data

Country	*Box-Cox transformation*	*Smoothingparameters*		*Initial states*			
	(Lambda)	*Alpha*	*Beta*	*L*	*B*	*Sigma*	*AIC*
Holt's Linear Trend Model fitted for Covid 19 infection cases in SAARC and China							
India	0.2949	0.7943	0.1578	-3.3946	0.0031	0.4689	1939.861
China	0.2224	0.9435	0.2444	-5.7708	1.4969	0.3634	1722.801
Nepal	0.3729	0.9369	0.2147	-2.6902	0.0078	0.4991	1993.138
Pakistan	1	0.5086	0.2517	-1.0616	-0.2	830.2095	8312.042
Sri Lanka	0.3267	0.9999	0.1378	-3.0772	0.0246	0.3834	1768.318
Bangladesh	1	0.8465	0.2526	-1.4292	-0.6518	398.5952	7686.903
Myanmar	0.3714	0.6505	0.2431	-2.6925	-1e-04	0.4232	1852.439
Holt's Linear TrendModel fitted for Covid 19 deaths cases in SAARC and China							
India	0.3777	0.8642	0.1275	-2.6471	2e-04	0.3753	1750.255
China	2	0.9999	0.0141	-1.0212	-0.3763	255041.2	13191.87
Nepal	0.1929	0.9999	0.0217	-5.3686	0.0073	0.2657	1455.914
Pakistan	0.8821	0.7772	0.1999	-0.6191	-0.2497	8.3911	4397.522
Sri Lanka	0.5738	0.7825	0.1645	-1.7427	3e-04	0.2311	1337.079
Bangladesh	0.9975	0.7409	0.2107	-1.294	-0.2004	7.0597	4250.321
Myanmar	0.8605	0.6599	0.2059	-1.6453	0.0059	2.3288	3305.412

Table 5: ARIMA fitted for COVID-19 infection cases and deaths in SAARC and China using the period from (3 January 2020) To (3 march 2021) for training data

Country	*Model*	*AR (1)*	*AR (2)*	*AR (3)*	*MA (1)*	*MA (2)*	*MA (3)*	*MA (4)*	*MA (5)*
ARIMA Model fitted for Covid 19 infection cases in SAARC and China									
India	ARIMA(1,2,3)	0.9522	-	-	-2.3512	1.8461	-0.4792	-	-
China	ARIMA(0,2,1)	-	-	-	-0.7019	-	-	-	-
Nepal	ARIMA(2,2,3)	-0.7573	-0.7334	-	0.0514	0.5202	-0.6138	-	-
Pakistan	ARIMA(1,2,3)	0.8210	-	-	-2.0320	1.4162	-0.3205	-	-
Sri Lanka	ARIMA(2,2,3)	-1.1584	-0.2481	-	-0.0766	-0.7743	0.2518		
Bangladesh	ARIMA(1,2,2)	0.9092	-	-	-1.8033	0.8453	-	-	-
Myanmar	ARIMA(3,2,2)	0.5099	0.1884	0.2113	-1.6910	0.7202	-	-	-
ARIMA Model fitted for Covid 19 deaths cases in SAARC and China									
India	ARIMA(2,2,3)	-0.5213	0.4419	-	-0.6850	-0.8120	0.6863	-	-
China	ARIMA(0,2,1)	-	-	-	-0.9331	-	-	-	-
Nepal	ARIMA(0,2,2)	-	-	-	-0.8581	-0.1194	-	-	-
Pakistan	ARIMA(0,2,5)	-	-	-	-1.0102	0.1981	0.0028	-0.0836	0.1170
Sri Lanka	ARIMA(2,2,2)	-1.1611	-0.2314	-	0.1060	-0.8528	-	-	-
Bangladesh	ARIMA(1,2,2)	0.4037	-	-	-1.4362	0.5922	-	-	-
Myanmar	ARIMA(1,2,3)	0.8371	-	-	-1.9963	1.2452	-0.2006	-	-

Table 6: BATS, TBATS, Holt's Linear Trend, and ARIMA Models fitted for COVID-19 infection cases and

Model	*ME*	*RMSE*	*MAE*	*MPE*	*MAPE*	*MASE*	*ACF1*
			Covid 19 Infection cases in India				
BATS	228.2447	13180.27	6055.134	-	-	0.2310183	0.07705743
TBATS	227.5411	12983.43	6475.681	-	-	0.2470632	0.07428195
Holt's linear Trend	-1784.803	14083.39	5355.172	-	-	0.204313	-0.2062556
ARIMA(1,2,3)	143.8378	12882.76	5688.494	1.372526	2.51792	0.2170301	0.0245581
			Covid 19 Deaths cases in India				
BATS	1.693479	192.6369	84.80609	-	-	0.2290658	0.05422684
TBATS	1.67981	190.2337	89.75114	-	-	0.2424226	0.06042829
Holt's linear Trend	-27.2825	200.4808	81.98246	-	-	0.221439	-0.1616733
ARIMA(2,2,3)	1.467933	188.2662	82.17214	1.436279	1.969348	0.2219514	-0.005500256
			Covid 19 Infection cases in China				
BATS	28.28899	690.7027	119.3941	-	-	0.4976313	-0.001735609
TBATS	-6.134403	694.1428	162.7885	-	-	0.6784984	-0.0004200286
Holt's linear Trend	-86.83598	776.5601	163.2435	-	-	0.6803946	0.2054061
ARIMA(0,2,1)	0.2233988	695.1634	123.5587	0.4655899	1.802854	0.5149896	0.008590153
			Covid 19 Deaths cases in China				
BATS	3.73803	66.08972	10.51737	-	-	0.9225763	-0.002973245
TBATS	-0.1555195	66.16334	16.25988	-	-	1.426305	0.0008538861
Holt's linear Trend	4.751824	67.64017	14.19265	-	-	1.244969	0.07591785
ARIMA(0,2,1)	-0.04610167	66.51379	13.50305	1.144144	1.852715	1.184478	-0.00705444
			Covid 19 Infection cases in Nepal				
BATS	0.6993527	352.8338	153.4554	-	-	0.2377635	-0.02548295
TBATS	0.4020773	371.3342	162.8087	-	-	0.2522556	-0.017856
Holt's linear Trend	-38.12568	392.728	165.6325	-	-	0.2566307	0.1446897
ARIMA(2,2,3)	0.5481111	347.116	152.5877	0.2473133	4.873593	0.2364192	0.007263339
			Covid 19 Deaths cases in Nepal				
BATS	2.040459	30.07261	4.419081	-	-	0.6763088	-0.006075391
TBATS	2.018866	29.91364	5.669233	-	-	0.8676356	-0.006111956
Holt's linear Trend	-1.68015	30.46538	5.672198	-	-	0.8680894	0.1153355
ARIMA(0,2,2)	1.895922	30.07482	4.355124	1.661352	2.399302	0.6665206	-0.007558469
			Covid 19 Infection cases in Pakistan				
BATS	12.1063	826.526	428.9502	-	-	0.3129529	0.03634994
TBATS	11.8149	816.2209	442.3373	-	-	0.3227199	0.04014466
Holt's linear Trend	12.08541	826.526	428.9765	-	-	0.3129721	0.03666592
ARIMA(1,2,3)	8.786337	818.3863	417.6971	0.8854473	2.169595	0.3047429	-0.008831837
			Covid 19 deaths cases in Pakistan				
BATS	0.43612	22.57919	11.20931	-	-	0.3682143	0.004630509
TBATS	0.4202794	22.27392	11.89566	-	-	0.3907601	0.005140045
Holt's linear Trend	0.1211489	22.58708	11.22829	-	-	0.3688378	0.02220533
ARIMA(0,2,5)	0.3840221	22.38443	11.06943	0.8231571	1.546422	0.3636195	0.002023223
			Covid 19 Infection cases in Sri Lanka				
BATS	5.187621	179.3466	81.99532	-	-	0.4155003	-0.0136032
TBATS	5.069633	177.8256	91.67795	-	-	0.4645658	-0.01215789
Holt's linear Trend	-9.841366	198.545	83.20154	-	-	0.4216127	-0.364559
ARIMA(2,2,3)	5.711783	175.977	81.40019	0.6727652	2.57963	0.4124846	-0.002542299

contd. table 6

			Covid 19 deaths cases in Sri Lanka				
BATS	0.11631	1.804009	0.809449	-	-	0.7122481	-0.01364533
TBATS	0.1160703	1.795149	0.8535513	-	-	0.7510544	-0.01427908
Holt's linear Trend	-0.00400315	1.827601	0.8297035	-	-	0.7300704	-0.001406146
ARIMA(2,2,2)	0.1125745	1.778707	0.8097604	0.7167823	2.062419	0.7125221	0.002775716
			Covid 19 Infection cases in Bangladesh				
BATS	4.527447	396.8151	214.2041	-	-	0.166333	0.02007944
TBATS	4.502196	395.4919	220.9963	-	-	0.1716073	0.01767431
Holt's linear Trend	4.511486	396.8147	214.1802	-	-	0.1663145	0.02030718
ARIMA(1,2,2)	3.016379	385.5906	205.0704	0.7562203	2.099683	0.1592406	-0.03324843
			Covid 19 deaths cases in Bangladesh				
BATS	0.08793135	7.163127	4.640486	-	-	0.2341454	0.02968924
TBATS	0.08725413	7.120594	4.706845	-	-	0.2374937	0.03070913
Holt's linear Trend	0.0854441	7.162863	4.640762	-	-	0.2341593	0.02937275
ARIMA(1,2,2)	0.06690964	7.098082	4.5671	0.6905658	1.797588	0.2304426	0.004784918
			Covid 19 Infection cases in Myanmar				
BATS	0.2785471	264.3873	105.2569	-	-	0.315107	0.02448542
TBATS	0.2845554	263.8409	112.6473	-	-	0.3372318	0.02442239
Holt's linear Trend	-24.40663	269.4462	99.67153	-	-	0.2983862	-0.06340266
ARIMA(3,2,2)	0.0001918449	262.0476	100.5691	1.072106	2.301411	0.3010731	-0.006985556
			Covid 19 deaths cases in Myanmar				
BATS	0.2518862	6.01962	2.575044	-	-	0.3421049	0.0004625826
TBATS	0.001782515	6.160466	2.796787	-	-	0.3715644	0.04029174
Holt's linear Trend	-0.1490585	6.159658	2.572228	-	-	0.3417308	0.004955476
ARIMA(1,2,3)	0.001345702	6.035088	2.577932	0.6592087	2.261941	0.3424887	-0.0006245787

deaths in SAARC and China using the period from (3 January 2020) To (3 march 2021) for training data

Table 7: MAPE for infection cases and deaths cases in SAARC and china using last 7 days from (4 march 2021) To (10 march 2021) for testing models

Country	*BATS*	*TBATS*	*Holt's Linear Trend Model*	*ARIMA*	*Best Model*
			MAPE for Covid 19 Infection cases		
India	0.067 %	0.068 %	0.1 %	0.029 %	ARIMA(1,2,3)
China	0.011 %	0.134 %	0.017 %	0.016 %	BATS
Nepal	0.014 %	0.033 %	0.007 %	0.015 %	Holt Model
Pakistan	0.159 %	0.184 %	0.159 %	0.141 %	ARIMA(1,2,3)
Sri Lanka	0.267 %	0.241 %	0.682 %	0.274 %	TBATS Model
Bangladesh	0.111 %	0.117 %	0.111 %	0.115 %	BATSHolt Model
Myanmar	0.023 %	0.032 %	0.025 %	0.014 %	ARIMA(3,2,2)
			MAPE for Covid 19 deaths cases		
India	0.016 %	0.009 %	0.008 %	0.029 %	Holt Model
China	0.019 %	0.15 %	0.043 %	0.013 %	ARIMA(0,2,1)
Nepal	5.53 %	5.588 %	4.314 %	5.415 %	Holt Model
Pakistan	0.547 %	0.589 %	0.55 %	0.53 %	ARIMA(0,2,5)
Sri Lanka	1.196 %	1.175 %	1.167 %	1.184 %	Holt Model
Bangladesh	0.056 %	0.057 %	0.056 %	0.06 %	BATSHolt Model
Myanmar	0.035 %	0.016 %	0.011 %	0.029 %	Holt Model

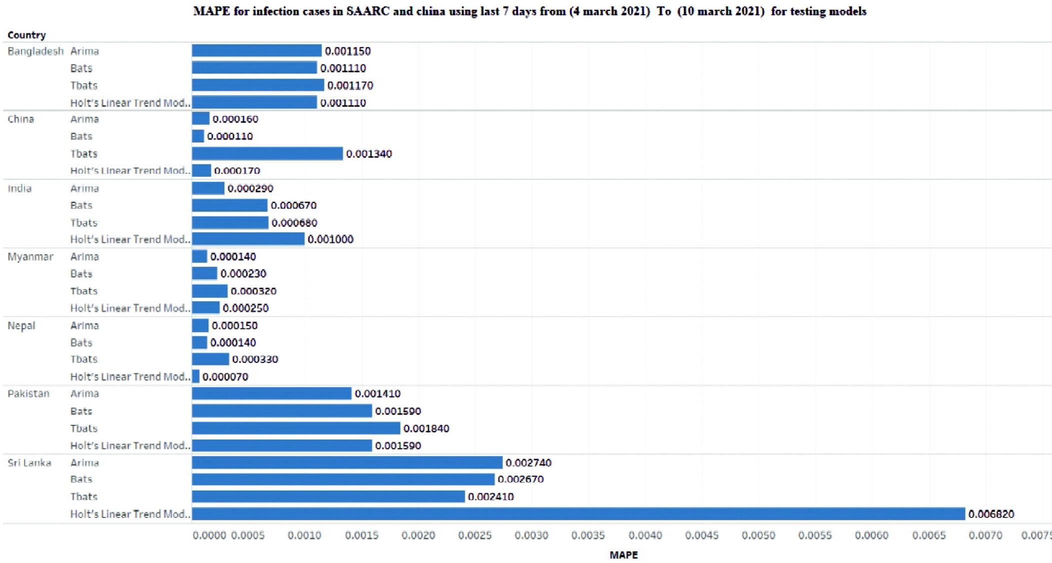

Figure 1: MAPE for infection cases in the SAARC and China using testing models

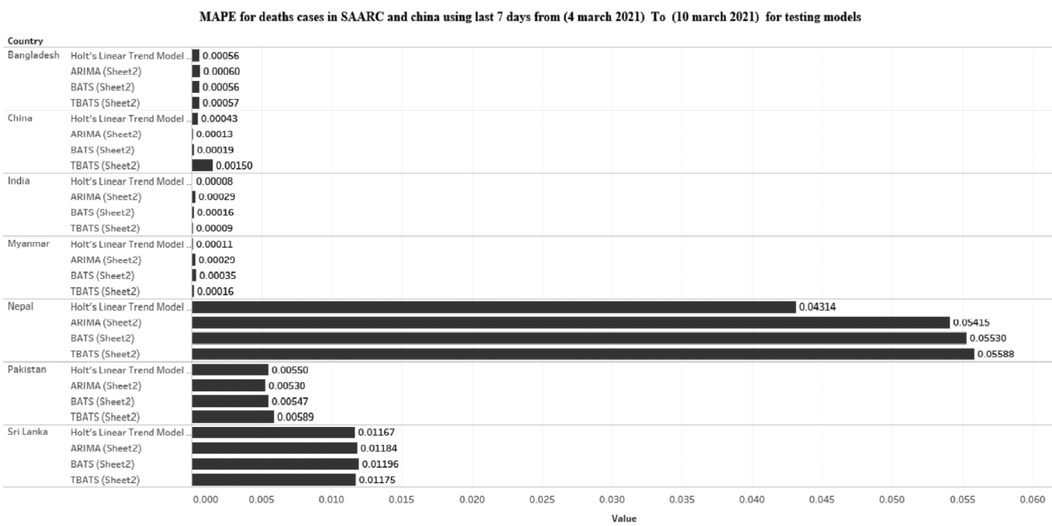

Figure 2: MAPE for total death cases in the SAARC and China using testing models

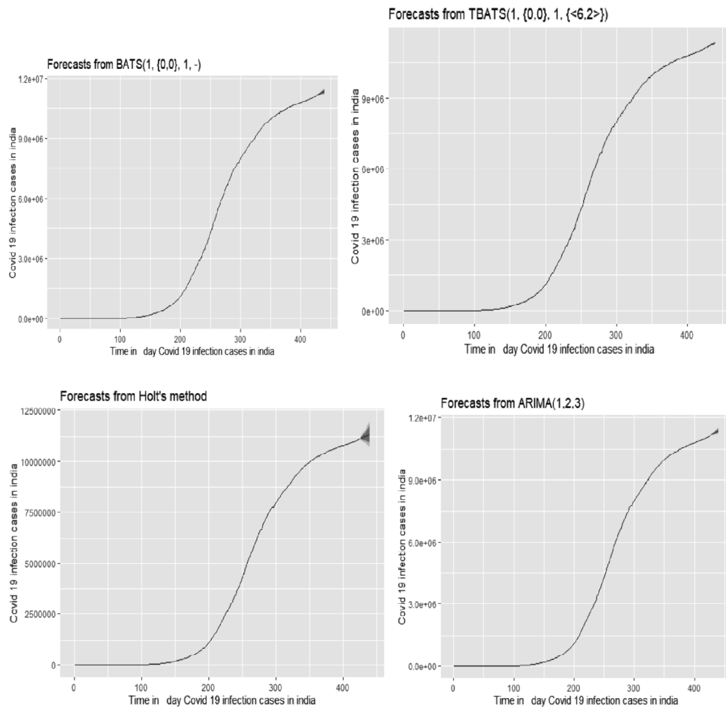

Figure 3: Forecasting of infection case using different models in India

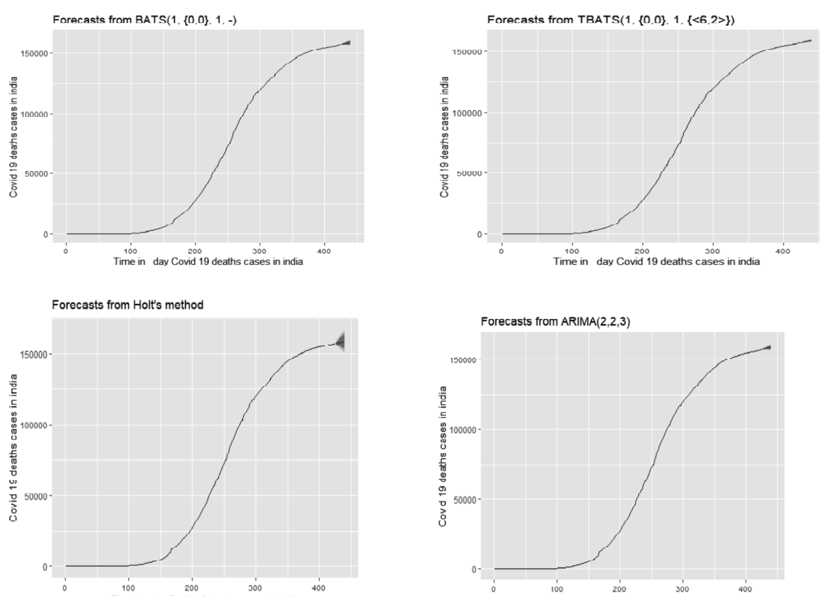

Figure 4: Forecasting of total death case using different models in India

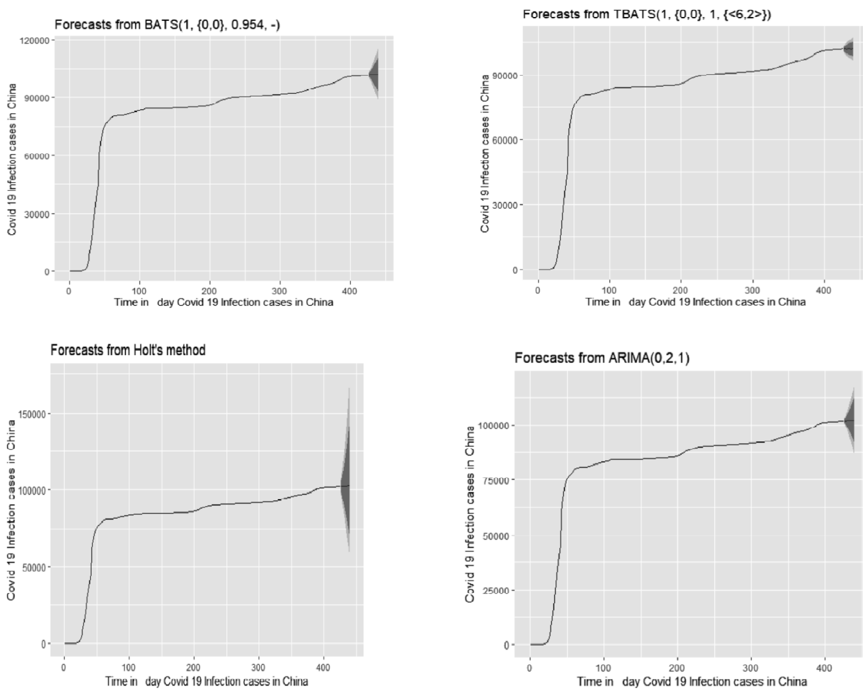

Figure 5: Forecasting of total infection case using different models in China

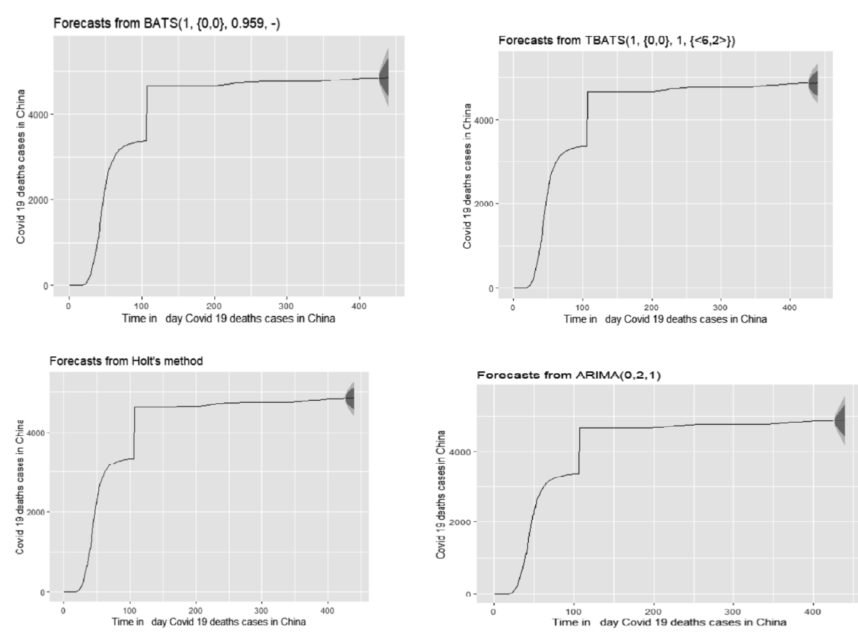

Figure 6: Forecasting of total death case using different models in China

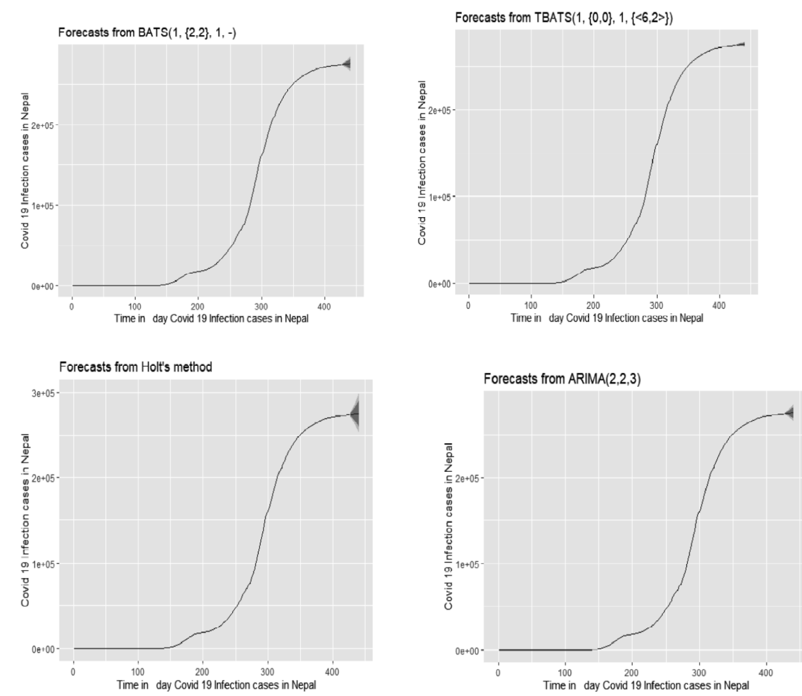

Figure 7: Forecasting of total infection case using different models in Nepal

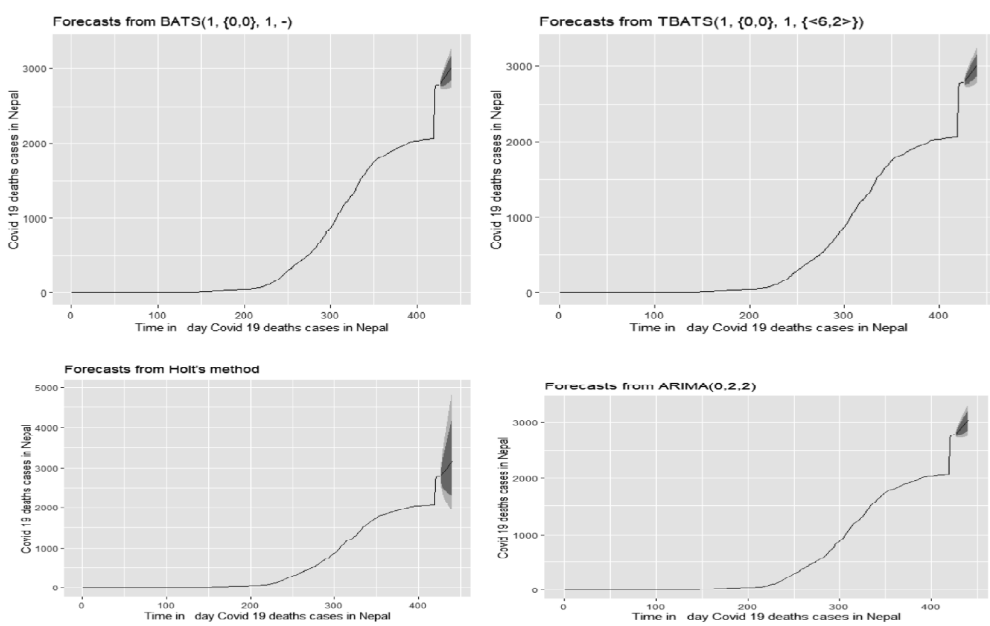

Figure 8: Forecasting of total death case using different models in Nepal

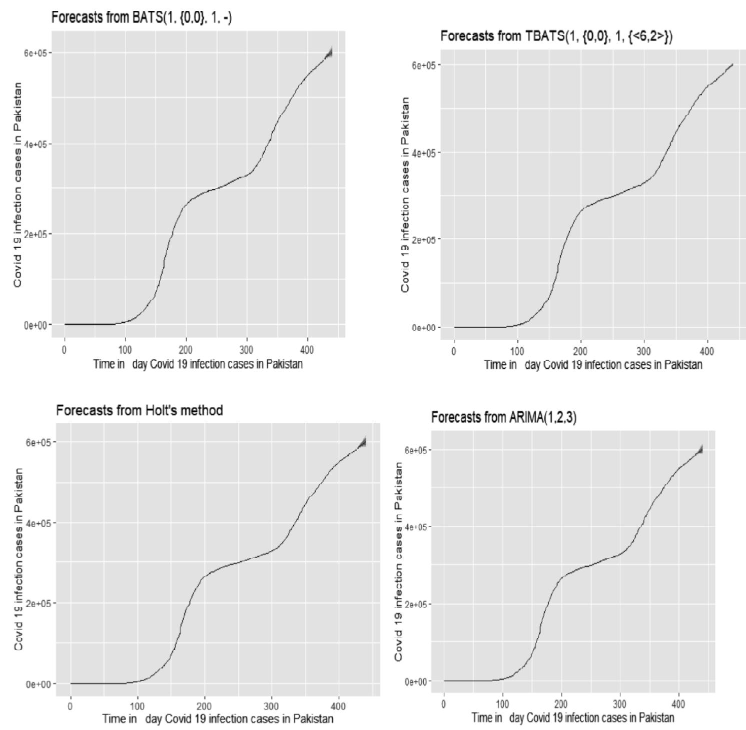

Figure 9: Forecasting of total infection case using different models in Pakistan

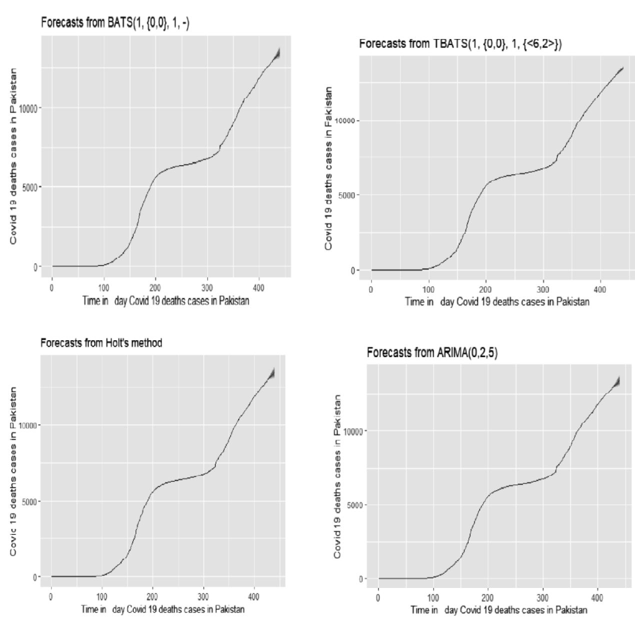

Figure 10: Forecasting of total death case using different models in Pakistan

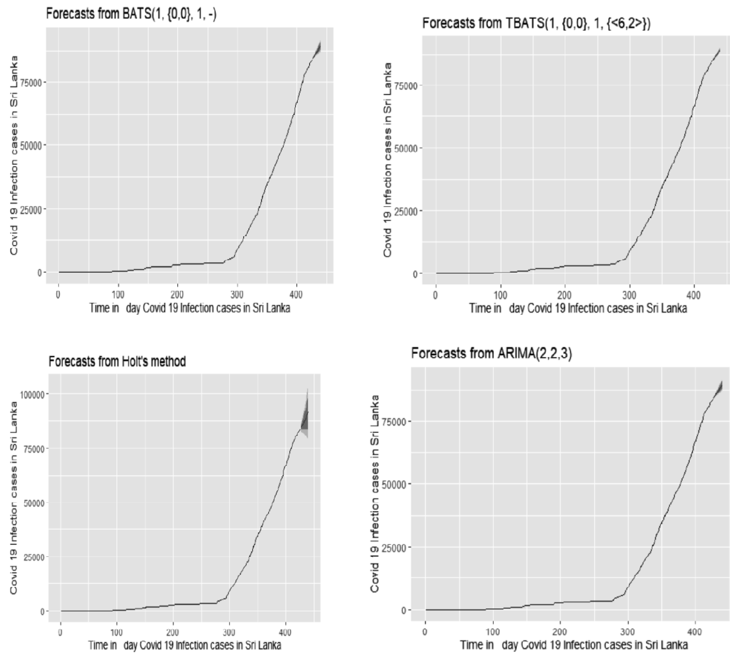

Figure 11: Forecasting of total infection case using different models in Sri Lanka

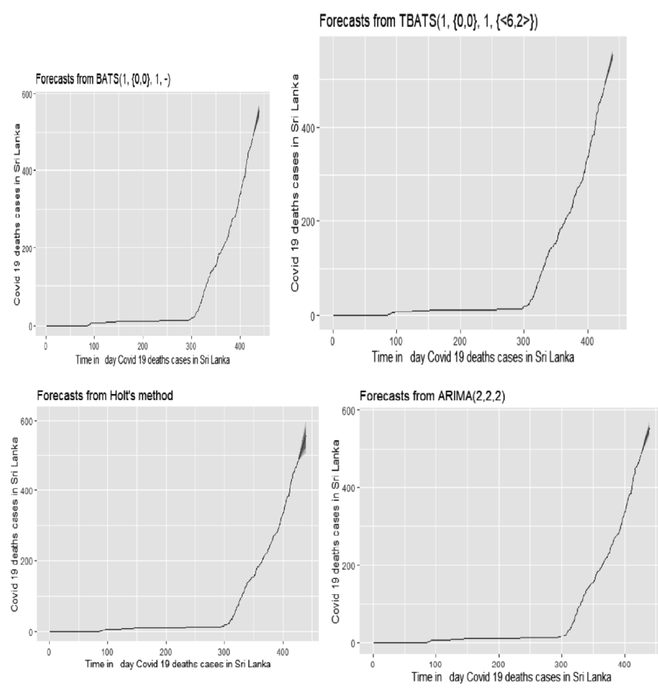

Figure 12: Forecasting of total death case using different models in Sri Lanka

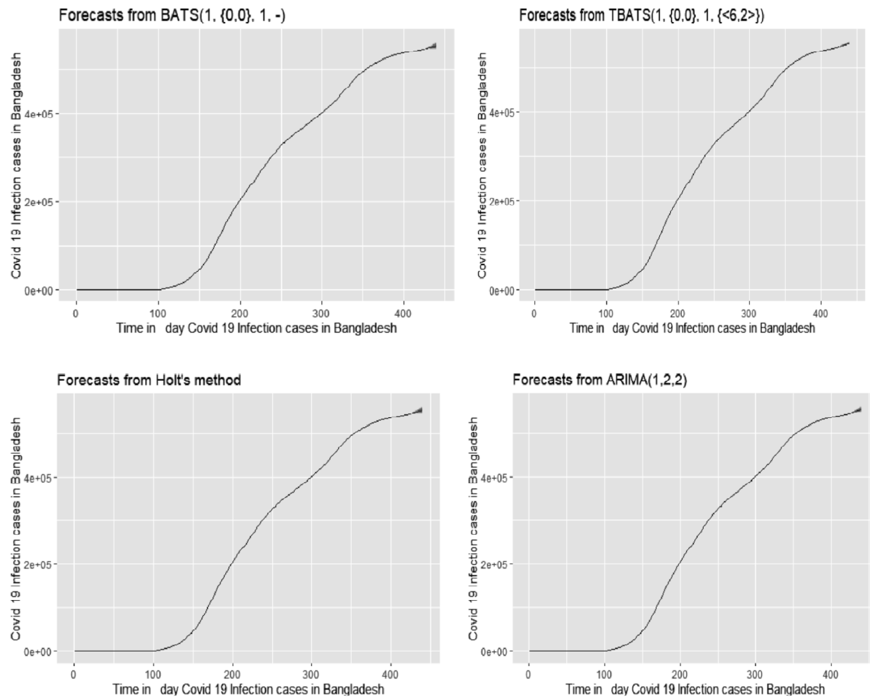

Figure 13: Forecasting of total infection case using different models in Bangladesh

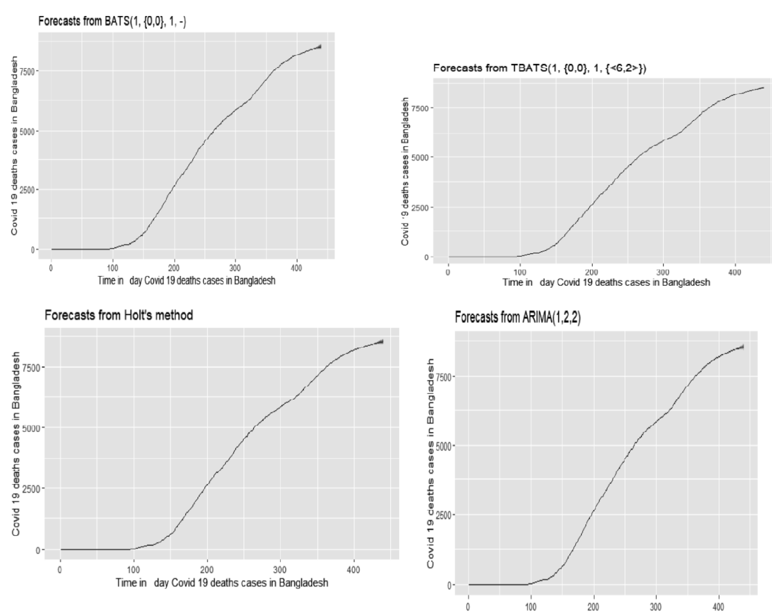

Figure 14: Forecasting of total death case using different models in Bangladesh

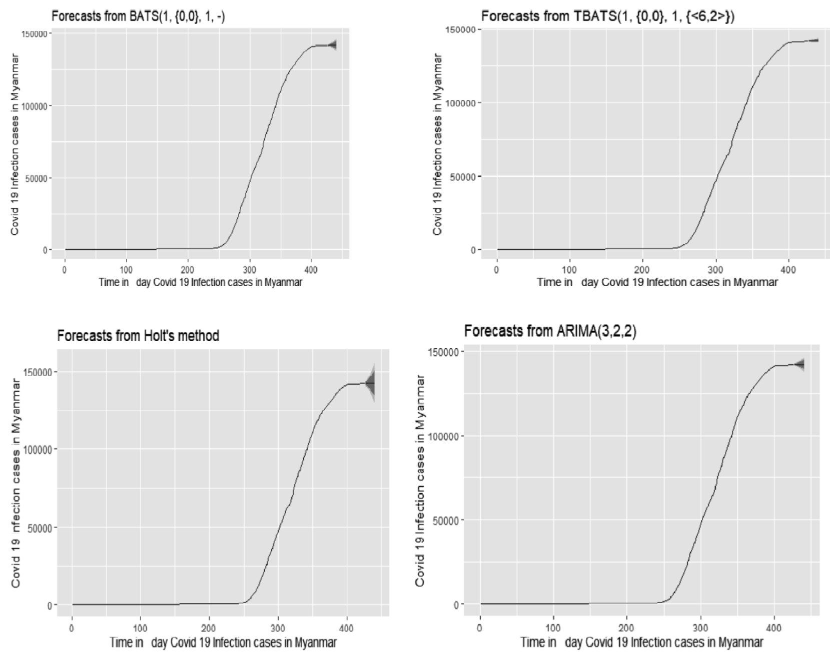

Figure 15: Forecasting of total infection case using different models in Myanmar

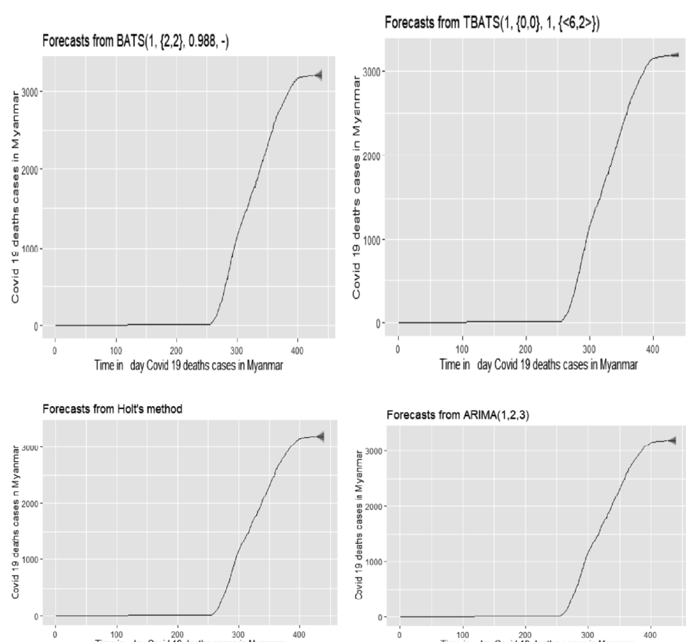

Figure 16: Forecasting of total death case using different models in Myanmar