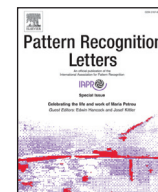




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

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



# Intelligent computing on time-series data analysis and prediction of COVID-19 pandemics

Sujata Dash<sup>a</sup>, Chinmay Chakraborty<sup>b,\*</sup>, Sourav K. Giri<sup>a</sup>, Subhendu Kumar Pani<sup>c</sup>

<sup>a</sup> Maharaja Sriram Chandra BhanjaDeo University (Erstwhile North Orissa University) Takatpur, Baripada, India

<sup>b</sup> Birla Institute of Technology, Mesra, Jharkhand, India

<sup>c</sup> Krupajal Computer Academy, Bhubaneswar, Odisha, India



## ARTICLE INFO

### Article history:

Received 6 June 2021

Revised 11 July 2021

Accepted 28 July 2021

Available online 14 August 2021

Edited by

### Keywords:

Facebook Prophet Model

Changepoints

Infectious Disease

Pandemics

Logistic growth function

High Incidence

Additive model

## ABSTRACT

Covid-19 disease caused by novel coronavirus (SARS-CoV-2) is a highly contagious epidemic that originated in Wuhan, Hubei Province of China in late December 2019. World Health Organization (WHO) declared Covid-19 as a pandemic on 12th March 2020. Researchers and policy makers are designing strategies to control the pandemic in order to minimize its impact on human health and economy round the clock. The SARS-CoV-2 virus transmits mostly through respiratory droplets and through contaminated surfaces in human body. Securing an appropriate level of safety during the pandemic situation is a highly problematic issue which resulted from the transportation sector which has been hit hard by COVID-19. This paper focuses on developing an intelligent computing model for forecasting the outbreak of COVID-19. The Facebook Prophet model predicts 90 days future values including the peak date of the confirmed cases of COVID-19 for six worst hit countries of the world including India and six high incidence states of India. The model also identifies five significant changepoints in the growth curve of confirmed cases of India which indicate the impact of the interventions imposed by Government of India on the growth rate of the infection. The goodness-of-fit of the model measures 85% MAPE for all six countries and all six states of India. The above computational analysis may be able to throw some light on planning and management of healthcare system and infrastructure.

© 2021 Published by Elsevier B.V.

## 1. Introduction

COVID-19 infectious disease declared by World health organization (WHO) as a global pandemic in March 2020 has infected millions of people worldwide since it emerged from China in December 2019. It has high mutating capability, spread very fast, very similar to respiratory tract infections and causes serious illness for the people suffering from chronic diseases like cardiovascular disease or diabetes or having weak immune system or being older in age [1,27]. The infected person takes long time to show the symptoms or no sign of the infection becomes a challenge to contain the disease. As proper medicines and vaccines are not available, some preventive and awareness measures like social distancing to break the chain of spread, testing at large scale to identify positive cases, isolating infected person to contain the spreading of the virus, travel constraints as guidelines and suggesting people to maintain a personal hygiene are made compulsory rules to be adopted by everyone in the society [2,28]. The spread of the

disease can be categorized into three different states [3] such as local outbreak, community transmission, and large-scale transmission. The local outbreak can be controlled by tracing and finding the source of infection whereas the source of chain of infection cannot be found in case of community transmission. In large-scale transmission, the virus spreads to different regions of the country due to uncontrolled movement of people. To curb these spreading national governments have imposed lockdown and other interventions [4].

COVID-19 situation report [5] as on March 31st, 2021, published by World Health Organization (WHO), more than 128 million people worldwide have been infected along with USA reporting the highest number of cases (30,462,210), followed by Brazil (12,748,747), India (12,221,665), France (4,705,068), Russia (4,494,234) and United Kingdom (4,359,982) on the sixth position. The death toll stands at 2,815,939 worldwide with the maximum number of death cases reported from USA (5,52,352) followed by Brazil, Mexico, India, UK, Italy while the recovered cases is 73,111,302. The time series data is collected from [6] and 'Johns Hopkins University Corona virus Data Stream that combines World Health Organization (WHO) and Centre for Disease Control and Prevention (CDC) case data'.

\* Corresponding author.

E-mail address: [cchakraborty@bitmesra.ac.in](mailto:cchakraborty@bitmesra.ac.in) (C. Chakraborty).

The above statistics motivated to develop models for six most affected nations including India and for six severely affected states of India to study the spreading behavior of the disease and the impact of interventions imposed on the states by government of India to manage various services and resources for the public. Many extensive researches [8,9] for forecasting the escalation of corona virus have been carried out employing machine learning algorithms [7] namely, neural networks for deep learning, polynomial fitting and exponential smoothing to contain the spread of the virus, and highlight impacts in the coming days. However, these studies were carried out at the early stage of the outbreak so has some limitations like the amount of information were not adequate. Currently, many predictive models are proposed [8,11] for providing analysis to support health services and governments to plan and contain the spread of the infections [10] with abundant information.

The paper is organized as follows: Section 1 introduces the research work; Section 2 outlines objectives and contributions of the study. Section 3 reviews the literatures, and Section 4 presents the proposed models and methodologies to be used in this research. Section 5 analyses the experimental findings of the problem. Section 6 provides discussion of the study followed by conclusion and limitations in Section 7. Section 8 represents an exhaustive bibliography.

## 2. Objectives & Contributions of the Study

The objectives and contributions of the present study are outlined below:

Objectives:

- 1) Studying the growth pattern and forecast the outbreak of six most affected states of India from 18th March 2020 to 31st March 2021 and top five most affected countries from 30th January 2020 to 31st March 2021 employing Prophet Model.
- 2) Studying the impact of imposed interventions on the incidence pattern of the disease only for India.

Contributions:

- 1) Prophet model provides understanding of the number of people affected daily by this disease.
- 2) Prophet model detects the significant changes in the growth pattern of the disease due to the interventions by using 'changepoints' in the epidemic curve of India.
- 3) Prophet Model forecasts 90 days ahead future trend of the disease and finds the peak time for all the six countries and six states of India.
- 4) The model has achieved around 85% in terms of accuracy for all the six countries and the six states of India.

This analysis will help to identify the risk factors for decision-makers and health authorities for taking proper precautions and strategies to prioritize the challenges.

## 3. Literature Review

Time series data predicts the future values using statistical and mathematical techniques which are based on some certain hypothesis considered for the underlying system [4,12]. Various problems from different fields of science viz, information systems [13], electrical engineering [14], medical diagnosis applications [15,16] are addressed successfully by time series forecasting models. The time series forecasting models are of two types – short-term and long-term forecasting model. The short-term model performs an exhaustive analysis and computation of the underlying assumptions and generates a dependable prediction even for few hours ahead future trend. The long-term model predicts a long future trend by

using the trend of the time series and the parameters of the problem.

Machine learning algorithms [17,18] such as neural networks for deep learning [20], polynomial fitting and exponential smoothing are used by many researchers for predicting the spread of COVID-19. In addition, artificial intelligence (AI) has proved its mettle by solving complex healthcare problems like cancer [7], neurological disorders [19] etc. AI driven models [20,19] can be applied for predicting the growth of infection of the virus and also can help to contain the disease. But, due to the scarcity of the statistical data neural network has faced overfitting problem. The reason is because the trend of the spreading of the disease changes in different phases of the lockdown over time, these model shows variations in their pattern. Therefore, many researchers have worked extensively applying mathematical and statistical models [4,23-26] to understand the spatio-temporal dynamics of COVID-19 outbreaks.

The reason is because the trend of the spreading of the disease changes in different phases of the lockdown over time, these model shows variations in their pattern. Therefore, many researchers have worked extensively applying mathematical and statistical models [12,13] to understand the spatio-temporal dynamics of COVID-19 outbreaks. These models gave a new impetus to understand the public policy for proper selection and allocation of resources and public health interventions [4] during the pandemics. Reliable predictions for healthcare resources and manpower can be made if the measures such as peak time, peak height, and enormity of the pandemic can be forecasted in time. Therefore, the trends of the outbreak and the epidemiological stage of the country can be forecasted effectively using the regression models such as Facebook Prophet [8], SEIR model [21], ARIMA model [12,15] prediction rules [17] etc.

Prophet, an additive model introduced by Facebook is very popular for forecasting time series data [22]. It detects separately the non-linear trends in the time series like yearly, weekly, daily seasonality and holiday effects and then combines them together to obtain the forecast value. The method employs a decomposable time series model with three components namely, trend, seasonality, and holidays where trend represents growth, seasonality introduce periodic effects within the trend and holiday effects capture sudden events that are predictable in time. In addition, it is robust to missing data, outliers and also captures the changepoints in the trend. It functions like a generalized additive model with time as a regressor parameter and fits the linear and non-linear functions of time as components. Essentially, it achieves suitable estimates of the mixed data without putting much effort [8,9]. The above-specified reasons are the motivation of selecting Prophet as the forecasting tool for this study. Moreover, Prophet has its own data frame for handling seasonality in the time series data that helps the data analyst to use time granularity for the prediction purpose [22].

## 4. Proposed Model Description & Methodology

The three main variables of this study are number of confirmed cases, number of recovery, and number of deaths. This study mainly focuses on developing the statistical regression models for forecasting the incidence, peak date of the outbreak and the changepoints in the infection growth curve.

### 4.1. Proposed Prophet Forecasting Model

This model is well known for describing and forecasting business problems. In this study, a decomposable time series model [22] with two components known as trend and seasonality is used to study the prevalence and incidence of COVID-19 pandemic in six

countries namely USA, Brazil, India, Russia, South Africa and Peru from 30th January 2020 to 31st March 2021 and for six states of India from 18th March 2020 to 31st March 2021. This procedure is similar to a generalized additive model with time as a regressor and is represented in its simplest form as follows:

$$y(t) = g(t) + s(t) + \epsilon(t) \quad (1)$$

In eq. (1),  $y(t)$  represents the time series of the confirmed cases,  $g(t)$  the nonlinear saturating trend function that models the non-periodic changes of the series,  $s(t)$  the seasonality component fits only yearly periodic changes in the trend and the last term  $\epsilon(t)$  corresponds to any unusual changes that cannot be accommodated by the model.

The nonlinear saturating trend function is modeled using logistic growth function and mathematically represented as:

$$g(t) = \frac{c(t)}{1 + e^{-k(t-m)}} \quad (2)$$

where  $C(t)$  is a time-varying capacity that represents the maximum number of corona patients added per day,  $k$  denotes a varying growth rate and  $m$  as offset parameter. In addition to this, the effects of the interventions on the trend of the growth are explicitly examined by introducing five changepoints in the model. Suppose  $S$  changepoints are defined at time then a vector of rate of change adjustments is defined by:

$$\delta \in \mathbb{R}^S$$

Where  $\delta_j$  represents the rate of change at time  $s_j$  then the rate of change at any time  $t$  corresponds to the base rate  $k$  plus the rate of change of adjustments happened till that time can be represented as follows:

$$\delta_t = k + \sum_{j:t>s_j} \delta_j$$

By defining a vector that can be represented as:

$$a_j(t) = \begin{cases} 1, & \text{if } t \geq s_j \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Hence, the rate of change at time  $t$  can be set to. Then by adjusting  $k$  and the offset parameter  $m$ , the correct adjustment at the changepoints  $j$  can be written as:

$$\gamma_j = \left( s_j - m - \sum_{l<j} \gamma_l \right) \left( 1 - \frac{k + \sum_{l<j} \delta_l}{k + \sum_{l \leq j} \delta_l} \right) \quad (4)$$

Now a piecewise logistic model can be obtained:

$$g(t) = \frac{C(t)}{1 + e^{-(k+a(t)^T \delta)(t-(m+a(t)^T \gamma))}} \quad (5)$$

In this model,  $C(t)$  refers to the expected capacity of the system at any given point of time. As this problem deals with daily cases, the value of the capacity  $C(t)$  is fixed considering the current growth pattern of the countries and states to study the future growth pattern, peak and end of the disease.

The periodic effect of yearly seasonal variations is modeled using Fourier series. Therefore, an approximate smooth seasonal effect is tied with a standard Fourier series which can be represented as:

$$s(t) = \sum_{n=1}^N a_n \cos(2\pi nt/p) + b_n \sin(2\pi nt/p) \quad (6)$$

where for yearly seasonality  $P = 365.25$  and  $N =$  varies from 1 to 2. However, the dataset is split into 80: 20 ratios with 80% data being used to train the model and 20% used for testing and predicting the future values of the pandemic after minimizing the variance and bias error. The workflow diagram of the model is shown in Fig. 1.

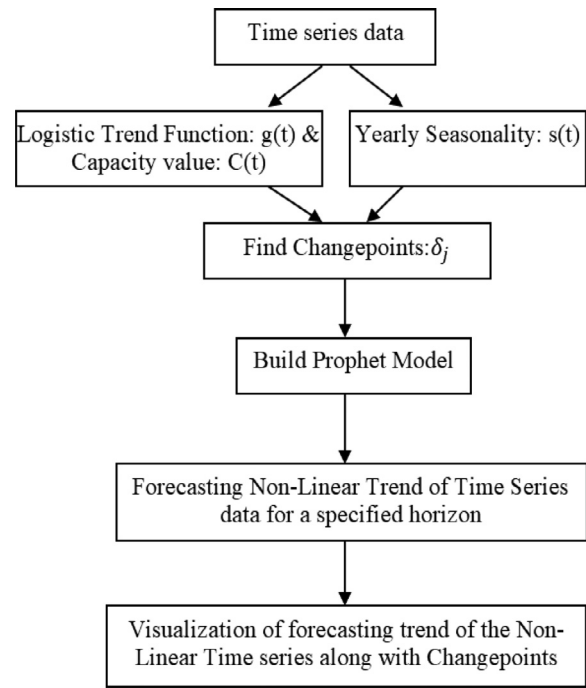


Fig. 1. Work Flow Diagram of Prophet Model

## 5. Experimental Setup and Analysis

All of the above methods are implemented using Python 3.8 version in Jupyter Notebook and executed in Windows 10, Intel(R), core-i7-7500U CPU @2.70GHz and 12.0 GB RAM. The packages which are used for prediction and visual representation of the findings are as follows: NumPy, pandas, SciPy, Matplotlib, Facebook Prophet, sklearn, matplotlib, seaborn and statsmodels.

### 5.1. Dataset Description

All the datasets used in this analysis for predicting state wise disease status collected from the sources namely, <https://www.covid19india.org/> [6] and for the six different countries collected from 'Johns Hopkins University Corona virus Data Stream that combines World Health Organization (WHO) and Centre for Disease Control and Prevention (CDC) case data'. The time series datasets were imported using .CSV format.

### 5.2. Analysis of the Experimental Results

The changepoints in prophet model are the points where there are sudden and abrupt changes occur in the trend. In this study these points refer to the time slots where number of infections increases or decreases. Through these changepoints the effect of lockdown, unlock and other interventions on the trend of the confirmed cases of India is analyzed. However, the five potential changepoints shown in Fig. 2 are placed uniformly in the first 80% of the time series data to avoid over-fitting fluctuations at the end of the series. Prophet detects automatically these changepoints and permits the trend of the series to adapt aptly. The analysis of the five changepoints dates are given in Table 2.

Prophet model works best for the time series data that have many seasons of past data. The present study has taken 435 and 383 days of data which is not adequate for making robust model to forecast 90 days future values of the countries and states respectively. With this information, the model is used to predict expected date of the peak of the disease after 31<sup>st</sup> March 2021 for

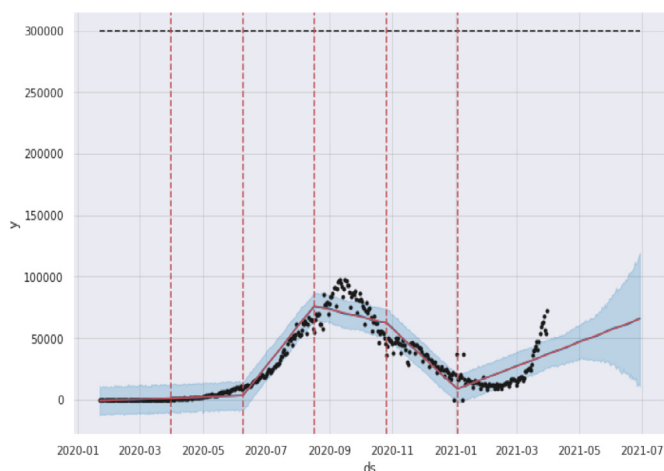


Fig. 2. : Significant Change Points of Confirmed cases of India IIndiaCOVID-19 in INDIA

most affected top six countries of the world and six severely affected states of India. The forecast plots depicting the peak time of pandemic for the six countries with 95% confidence interval and is shown in Fig. 3(a, b, c, d, e, f).

USA may reach the peak in the first part of January 2021 with more than 3 lakhs per day infection and ends around Second week May 2021 shown in Fig. 3(a). Brazil shows an increasing trend till the end of June 2021 with per day cases around 1.2 lakh (Fig. 3(b)). India has attained its peak on mid of September 2020 with per day cases around 1 lakh and then started declining till 25<sup>th</sup> January with number of cases around 9098 per day. The growth remained in the same state till the beginning of March 2021 and then the prediction curve shows an increase in the number of cases at around 50,000 by the end of March 2021. The forecast window shown in Fig. 3 (c) depicts a sharp growth in infection rate till the end of June 2021 with per day infection cases more than 2.5 lakhs.

The prediction curve of France shows an increasing trend of the disease till the end of June 2021 with per day infection cases more than 20000 which are shown in Fig. 3(d). Likewise the prediction curve (Fig. 3(e)) of Russia has already reached a peak in the end of January 2021 with per day cases around 30000 and then declines up to end of March 2021. The curve shows a sharp increase in the number of cases till June 2021 with per day cases around 50000. Fig. 3(f) shows UK has already reached the peak in the beginning of January 2021 with per day cases around 68000 and then declined till end of April 2021. After the end of April 2021 there is a sudden rise in the number infection cases and then the curve flattens.

### 5.3. Performance Metrics of Prophet Models

Kerala shows an increase in the infection cases and attained the peak in the beginning of October 2020 with per day cases around 12000 then the curve oscillates touched the ground by the end of March 2021 shown in Fig. 4(b).

Then suddenly the predicted curve shows increasing in the number of infection cases and it reaches a second peak in the end of April. After that the curve declines and touches the ground by the end of May 2021. Karnataka shows a steep rise in the number of cases and reached the first peak in the beginning of October 2020 with per day cases around 12000 shown in Fig. 4(c). Then the curve declines and touched the ground in the beginning of February 2021 and then the curve moves upward showing an increase in the number of cases. The curve reached the second peak in the

Table 1  
Performance Metrics of Prophet Models

| Country | Model   | MSE          | RMSE     | MAE      | MAPE % |
|---------|---------|--------------|----------|----------|--------|
| US      | Prophet | 3.807926e+08 | 19513.90 | 16718.97 | 20.27  |
| Brazil  | Prophet | 2.319667e+08 | 15230.45 | 14968.04 | 23.49  |
| India   | Prophet | 1.879391e+08 | 13709.08 | 9573.82  | 17.74  |
| France  | Prophet | 2.464339e+08 | 15698.21 | 11385.83 | 24.87  |
| Russia  | Prophet | 730823.77    | 854.88   | 823.22   | 5.89   |
| UK      | Prophet | 730823.77    | 854.88   | 823.22   | 5.89   |

mid of April 2021 with number of cases around 10000 and then starts declining.

The prediction curve of Tamil Nadu shown in Fig. 4(d) shows an increasing trend of the cases up to beginning of September 2020 with number of cases around 7000 and then declines steadily up to end of February 2021. Again the number of cases shows a sharp increase from the end of March 2021 with per day cases around 25000. Fig. 4(e) shows the trend of Andhra Pradesh that reaches the first peak in the beginning of September 2020 with per day cases around 10000 and touched the ground at the beginning of January 2021. Again, the prediction curve shows an increase in the number of cases in the beginning of March 2021 and the growth of the curve follows a zig-zag rise and fall. Fig. 4(f) shows Uttar Pradesh has attained its first peak in the beginning of September 2020 with per day cases around 7000 and has reached the ground in the beginning of February 2021. The curve shows an increase in the number of cases from March till end of April 2021 and then declined up to mid of May 2021. The curve then shows a sharp increase with per day cases around 6000.

Finally, discussing about the estimated value of the performance metrics of Prophet Model which is recorded in Table 1 establish that Prophet Model strongly forecast the sudden changes happened due to the imposed interventions in the trends of the disease and peak time. The forecasting accuracy of Prophet is 95% for Russia and UK. However, the accuracy is 80% for USA, 77% for Brazil, 76% for France and 83% for India. This confirms the efficiency and efficacy of Prophet Model for being used as an epidemiological model to study the incidence of the disease.

## 6. Discussion

Suitable forecasting models capture the information from the time series data thoroughly and provide better understanding of the spread of the disease across the population which helps to decide pertinent measures to control the transmission of the infection and increase the capacity of healthcare system. At this moment, epidemiological solutions are highly essential rather than the pharmaceutical solutions. However, it is crucial to assess the efficacy of the applied interventions for taking timely actions to alleviate the pandemic. These timely actions need precise information about the ongoing disease, accurate growth predictions and reassessment of the implemented interventions. The present study endeavored to forecast the current scenario using regression models considering the data from 30<sup>th</sup> January 2020 to 31<sup>st</sup> March 2021 for five countries and 18<sup>th</sup> March 2020 to 31<sup>st</sup> March 2021 for India and for six states of India which projected the daily cases very close to the observed cases.

The Prophet Prediction curves indicate the infection rate will still increase even after June 2021 for the states namely, Maharashtra, Tamil Nadu and Uttar Pradesh. The first lockdown was one of the major interventions imposed by the Government relatively bit early along with other public health precautions to alleviate the transmission of the pandemic. It raises an apparent question on the effectiveness of lockdown over the incidence cases. The effectiveness of the interventions [4] is measured by many studies

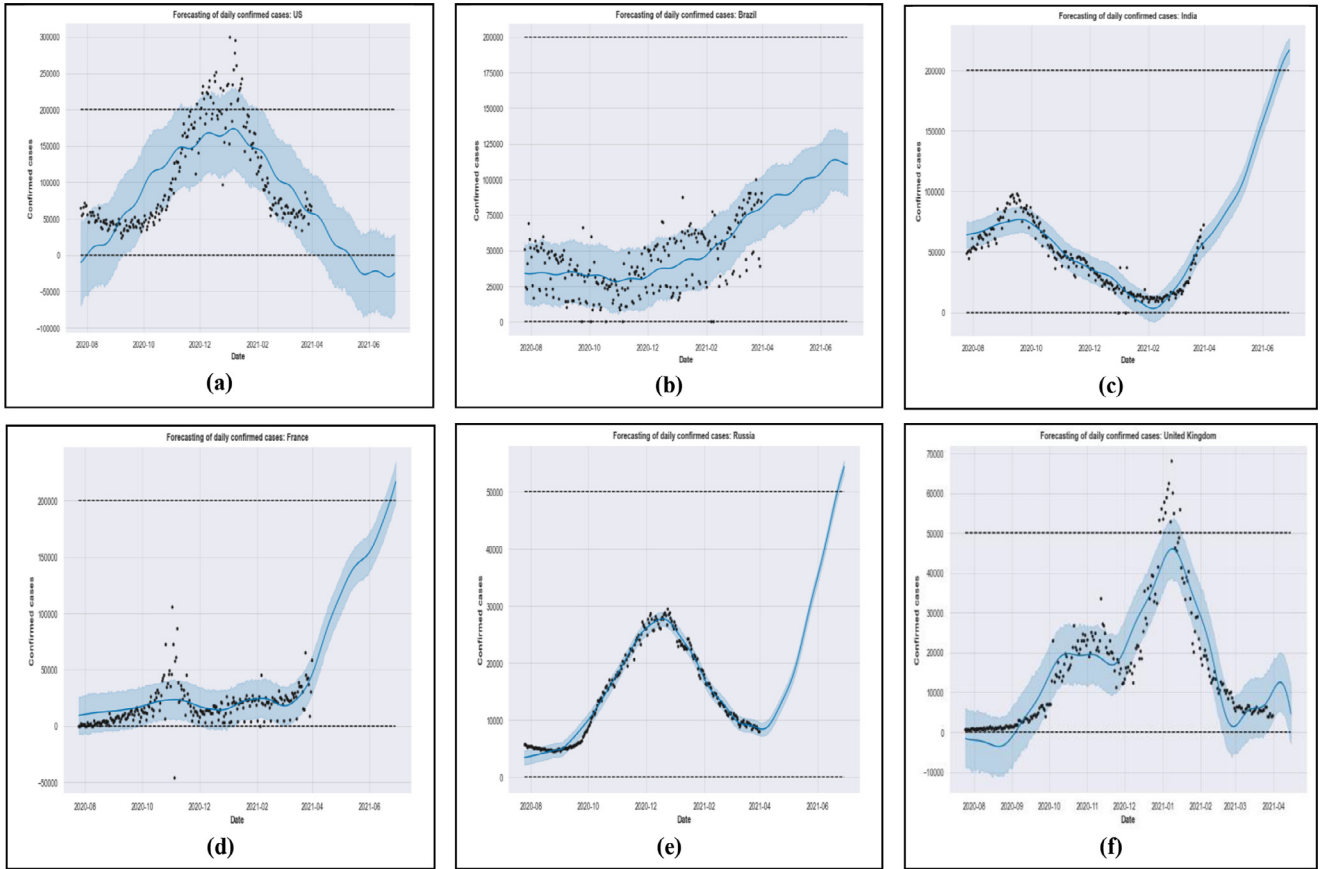


Fig. 3. Forecasting Trends of Six most affected Countries using Prophet Model

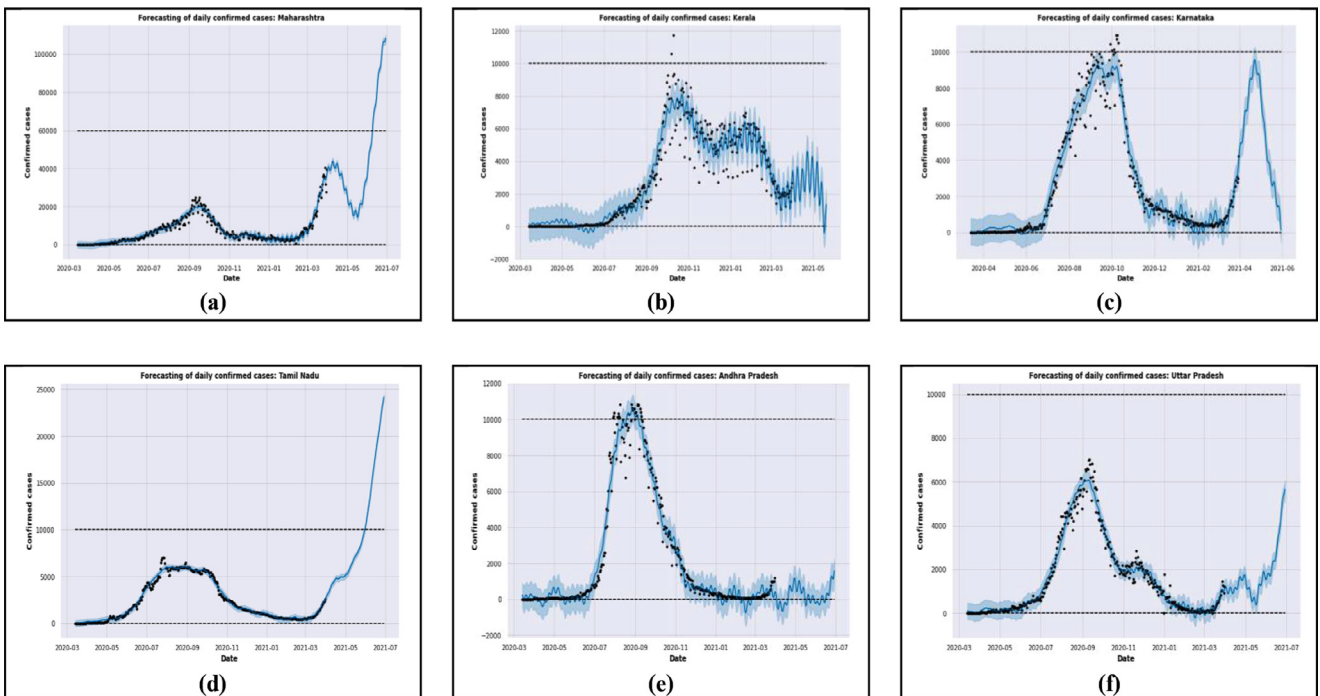


Fig. 4. Forecasting of top six states of India

**Table 2**  
Five significant change points of Confirmed series of INDIA

| Change Point Date   | Lockdown Period   | Possible reason  |
|---|---|--|
| 29-03-2020<br>1 <sup>st</sup> Change points shown in Fig. 2 | Lockdown 1.0 continued from 23 <sup>rd</sup> March to 14 <sup>th</sup> April 2020 | This point falls in the Lockdown 1.0. From covid19india.org, it is found that there is a sudden drop in the number of corona cases on 29 <sup>th</sup> March viz. only 3.6% patients were identified between 28 <sup>th</sup> (987 cases) to 29 <sup>th</sup> (1024) March. This can be attributed to the increase in COVID testing. Towards the end of the first lockdown, the rate of growth of COVID infections in India had significantly slowed down from a rate of doubling every three days before the lockdown to one of doubling every eight days on 18 <sup>th</sup> April 2020.   |
| 07-06-2020<br>2 <sup>nd</sup> Change points shown in Fig. 2 | Unlock 1.0 (1 <sup>st</sup> - 30 <sup>th</sup> June 2020)                         | This change point falls in the Unlock 1.0 period. The change points shows there is a sharp rise of the confirmed cases and this can be attributed to the Unlock 1.0. As the Ministry of Home Affairs lifted the lockdown restrictions and imposed lockdown in the containment zones, while allowed activities in other zones in a phased manner. Unlock 1.0 reopened shopping malls, religious places, hotels, and restaurants from 8 <sup>th</sup> June 2020. Large gatherings were still banned but there were no restrictions on interstate travel. Night curfews were in effect from 9 p.m. to 5 a.m. in all areas and state governments were allowed to impose suitable restrictions on all activities. |
| 17-08-2020<br>3 <sup>rd</sup> Change points shown in Fig. 2 | Unlock 3.0 (1 <sup>st</sup> - 31 <sup>st</sup> August 2020)                       | This point shows a steady decline of the confirmed cases. This is because; Unlock 3.0 started maintaining strict parameter control such as intensive contact tracing, house-to-house surveillance and other clinical interventions. The guidelines for the usage of ArogyaSetu, social distancing and masks were strengthened. Maharashtra and Tamil Nadu imposed a lockdown for the whole month, while West Bengal imposed lockdowns twice in a week. Medical infrastructures were strengthened that impacted in the increase in recovery rate.   |
| 29-10-2020<br>4 <sup>th</sup> Change points shown in Fig. 2 | Unlock 5.0 (1 <sup>st</sup> - 31 <sup>st</sup> October 2020)                      | This Change points shows a sharp decrease in the number of confirmed cases. The information available in covid19india.org validates the changes illustrated by 4 <sup>th</sup> Change points of Fig. 2. It is observed that rate of recovery was very close to rate of confirmed cases. It was made possible as Government strictly maintained the covid control parameters and made compulsory face coverings/masks in public places, workplaces and during transport.  |
| 01-01-2021<br>5 <sup>th</sup> Change points shown in Fig. 2 | 5 <sup>th</sup> Change points shows flattening of the disease trend               | The 5 <sup>th</sup> Change points shows flattening of the disease curve for a short while and then a gradual increase in the number of cases observed after February 2020. This situation can be attributed to the withdrawal of lockdowns and opening of Universities, colleges and public places.  |

with varying level of outcomes. In this paper, five change points are computed by Prophet that indicates the sudden changes happened in the growth curve of the disease. The fifth Change points clearly gives an indication of second wave in India after March 2021. Both the prediction models of France and Russia indicate for a second wave with much severity. However, as discussed above necessitates the revision of the forecast model in a regular interval as and when the disease data gets available. Cross-country performance was hard to explain and interpret. However, important factors that should be noted include discrepancies among the different countries in terms of climatic and geographical characteristics; in terms of population-related characteristics such as density; in terms of COVID-19 measures and testing procedures; and in terms of timing, duration and severity of any social distancing measures if any that could be implemented will enhance the prediction in a more effective way. Cross-country performance was hard to explain and interpret. However, important factors that should be noted include discrepancies among the different countries in terms of climatic and geographical characteristics; in terms of population-related characteristics such as density; in terms of COVID-19 measures and testing procedures; and in terms of timing, duration and severity of any social distancing measures if any that could be implemented will enhance the prediction in a more effective way.

## 7. Conclusion and Limitations

Facebook Prophet Model forecasts the sudden changes in the growth rate using Change points. The significant effect of the lockdown and unlock over the growth of the disease assessed by Prophet Change points is precisely discussed in Table 2. The Prophet Prediction curves are very close to the observed values for India, France, and Russia except US, Brazil and UK. Likewise the forecast made by Prophet for the states are not very close to the actual observations. However the statistical performance metrics proves the

efficiency of Prophet Models. The findings of the models may be used for making plans for possible interventions to strengthen the healthcare system for better management of the infected people in India and other countries.

The limitations of the proposed models not only depend on the underlying assumptions but many factors viz, density of population, healthcare system, interventions imposed by administration, economic and socio-demographic situation. If the data on testing and screening strategies, policies adopted by different countries, information about the access of pre-exposure drug profile and robustness of the healthcare system would be available for analyzing the existing information, one can incorporate these statistics to develop a robust predictive model.

## Declaration of Competing Interest

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

## References

- [1] C. Sohrabi, Z. Alsafi, N. O'Neill, M. Khan, A. Kerwan, A. Al-Jabir, C. Iosifidis, R. Agha, World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19), *International journal of surgery (London, England)* 76 (2020) 71–76, doi:10.1016/j.ijso.2020.02.034.
- [2] Ensheng Dong, Hongru Du, Lauren Gardner, An interactive web based dashboard to track covid-19 in real time, *The Lancet infectious diseases* 20 (5) (2020) 533–534.
- [3] Lin Jia, Kewen Li, Yu Jiang, Xin Guo, et al., 2020. Prediction and analysis of coronavirus disease 2019, arXiv preprint arXiv:2003.05447, 2020.
- [4] S. Mandal, T. Bhatnagar, N. Arinaminpathy, A. Agarwal, A. Chowdhury, M. Murhekar, R.R. Gangakhedkar, S. Sarkar, Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: A mathematical model-based approach, *Indian J Med Res* 151 (2020) 190–199.
- [5] World Health Organization, 2020, Coronavirus disease 2019 (COVID-19): situation report, 123.
- [6] <https://www.covid19india.org/>

- [7] Q.H. Ye, L.X. Qin, M. Forgues, P. He, J.W. Kim, A.C. Peng, R. Simon, Y. Li, A.I. Robles, Y. Chen, Z.C. Ma, Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning, *Nature medicine* 9 (4) (2003) 416–423.
- [8] G. Battineni, N. Chintalapudi, F. Amenta, Forecasting of COVID-19 epidemic size in four high hitting nations (USA, Brazil, India and Russia) by Fb-Prophet machine learning model, *Applied Computing and Informatics* (2020) Vol. ahead-of-print No. ahead-of-print, doi:10.1108/ACI-09-2020-0059.
- [9] T. Chafiq, M. Ouadoud, K. Elboukhari, Covid-19 forecasting in Morocco using FBprophet Facebook's Framework in Python, *International Journal of Advanced Trends in Computer Science and Engineering* 9 (5) (2020) 8654–8660.
- [10] IHME COVID, Christopher J.L. Murray, et al., 2020. Forecasting covid-19 impact on hospital bed-days, icu-days, ventilator-days and deaths by us state in the next 4 months. MedRxiv.
- [11] A.B. Banu, P. Thirumalaikolundusubramanian, Time Series Analysis For Predicting Covid-19 Infection Using Facebook Prophet Model, *Turkish Journal of Computer and Mathematics Education* 12 (7) (2021) 1680–1685.
- [12] Q. Yang, J. Wang, H. Ma, X. Wang, Research on COVID-19 based on ARIMA model $\Delta$ -Taking Hubei, China as an example to see the epidemic in Italy, *Journal of infection and public health* 13 (10) (2020) 1415–1418, doi:10.1016/j.jiph.2020.06.019.
- [13] B. Spencer, O. Alfandi, F. Al-Obeidat, A refinement of LASSO regression applied to temperature forecasting, *Procedia Computer Science* 130 (2018) 728–735, doi:10.1016/j.procs.2018.04.127.
- [14] N. Ludwig, S. Feuerriegel, D. Neumann, Putting Big Data analytics to work: Feature selection for forecasting electricity prices using the LASSO and random forests, *Journal of Decision Systems* 24 (1) (2015) 19–36, doi:10.1080/12460125.2015.994290.
- [15] F.M. Khan, R. Gupta, ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India, *Journal of Safety Science and Resilience* 1 (1) (2020) 12–18.
- [16] G. Purcaro, C.A. Rees, W.F. Wieland-Alter, M.J. Schneider, X. Wang, P.H. Stefanuto, P.F. Wright, R.I. Enelow, J.E. Hill, Volatile fingerprinting of human respiratory viruses from cell culture, *J Breath Res* 12 (2) (2018) 026015 MarPMID: 29199638; PMID: PMC5912890, doi:10.1088/1752-7163/aa9eef.
- [17] S. Dash, Hill-climber Based Fuzzy-Rough Feature Extraction with an Application to Cancer Classification, *Journal of Network and Innovative Computing* 2 (2014) 178–186 ISSN 2160-21742014.
- [18] S. Dash, C. Chakraborty, KG Sourav, KP Subhendu, F Jaroslav, BIFM: Big-data Driven Intelligent Forecasting Model for COVID-19, *IEEE Access* (2021) 1–13, doi:10.1109/ACCESS.2021.3094658.
- [19] S. Dash, A. Abraham, A. Kr Luhach, JM. Pietraszko, JJPC Rodrigues, Hybrid chaotic firefly decision-making model for Parkinson's disease diagnosis, *International Journal of Distributed Sensor Networks* 15 (12) (2019) 2019.
- [20] S. Biswas, S. Dash, LSTM-CNN Deep learning based hybrid system for real time covid-19 data analysis and prediction using twitter data, in: S. Pani, S. Dash, Wellington dos Santos, S.A. Chan Bukhari, F. Flammini (Eds.), *Assessing covid-19 and other pandemics and epidemics using computational modelling and data analysis* 2021, Springer, 2021 Accepted.
- [21] G. Pandey, P. Chaudhary, R. Gupta, and S. Pal, 2020. SEIR and Regression Model based COVID-19 outbreak predictions in India, Tech. rep. arXiv:2004.00958. 2020.
- [22] S.J. Taylor, B. Letham, Forecasting at Scale, *The American Statistician* 72 (1) (2018) 37–45 2017, doi:10.1080/00031305.2017.1380080.
- [23] KB Hemanta, C Chinmay, S Yogesh, KP Suvendu, COVID-19 Diagnosis System by Deep Learning Approaches, *Expert Systems*, 1–12, 10.1111/exsy.12776
- [24] A Bakhtawar, RJ Abdul, C. Chinmay, N Jamel, R Saira, R Muhammad, Blockchain and ANFIS empowered IoMT Application for Privacy Preserved Contact Tracing in COVID-19 Pandemic, *Personal and Ubiquitous Computing* (2021), doi:10.1007/s00779-021-01596-3.
- [25] R. Vinayakumar, N Harini, C. Chinmay, D.P. Tuan, Deep learning based Meta-classifier Approach for COVID-19 Classification using CT scan and Chest X-ray Images, *Multimedia Systems* (2021) 1–15, doi:10.1007/s00530-021-00826-1.
- [26] R. Sukanya, C. Chinmay, Panic Buying Situation during COVID-19 Global Pandemic, *Journal of Information Technology Management* 13 (2) (2021) 231–244, doi:10.22059/jitm.2021.80625.
- [27] C. Wu, M. Khishe, M. Mohammadi, H. Sarkhel, T. Karim, T.A. Rashid, Evolving deep convolutional neural network by hybrid sine-cosine and extreme learning machine for real-time COVID19 diagnosis from X-ray images, *Soft Comput* (2021) 2021, doi:10.1007/s00500-021-05839-6.
- [28] C. Wu, M. Khishe, M. Mohammadi, H.T. Sarkhel, A.T. Karim, Rashid Evolving deep convolutional neural network by hybrid sine-cosine and extreme learning machine for real-time COVID19 diagnosis from X-ray images, *Soft Comput* (2021), doi:10.1007/s00500-021-05839-6.