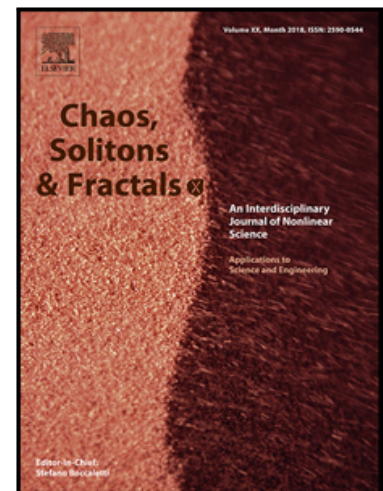


Journal Pre-proof

Prediction of Epidemic Trends in COVID-19 with Logistic Model and Machine Learning Technics

Peipei Wang, Xinqi Zheng, Jiayang Li, Bangren Zhu

PII: S0960-0779(20)30455-0
DOI: <https://doi.org/10.1016/j.chaos.2020.110058>
Reference: CHAOS 110058



To appear in: *Chaos, Solitons and Fractals*

Received date: 19 June 2020

Accepted date: 23 June 2020

Please cite this article as: Peipei Wang, Xinqi Zheng, Jiayang Li, Bangren Zhu, Prediction of Epidemic Trends in COVID-19 with Logistic Model and Machine Learning Technics, *Chaos, Solitons and Fractals* (2020), doi: <https://doi.org/10.1016/j.chaos.2020.110058>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier Ltd.

Highlights:

- A hybrid prediction model for COVID-19 based on Logistic and Prophet is proposed.
- The epidemic trend of COVID-19 in global, Brazil, Russia, India, Peru and Indonesia are predicted by our proposed model.
- Three significant points are summarized from our modeling results.
- The number of accumulated infections in global by late October is estimated to be 14.12 million.

Prediction of Epidemic Trends in COVID-19 with Logistic Model and Machine Learning Technics

Peipei Wang^a, Xinqi Zheng^{a,b,*}, Jiayang Li^a, Bangren Zhu^a

^a*School of Information Engineering, China University of Geosciences, Beijing 100083, China*

^b*Technology Innovation Center of Territory Spatial Big-data, MNR of China, Beijing 100036, China*

Abstract

COVID-19 has now had a huge impact in the world, and more than 8 million people in more than 100 countries are infected. To contain its spread, a number of countries published control measures. However, its not known when the epidemic will end in global and various countries. Predicting the trend of COVID-19 is an extremely important challenge. We integrate the most updated COVID-19 epidemiological data before June 16, 2020 into the Logistic model to fit the cap of epidemic trend, and then feed the cap value into Fbprophet model, a machine learning based time series prediction model to derive the epidemic curve and predict the trend of the epidemic. Three significant points are summarized from our modeling results for global, Brazil, Russia, India, Peru and Indonesia. Under mathematical estimation, the global outbreak will peak in late October, with an estimated 14.12 million people infected cumulatively.

Keywords: Coronavirus, COVID-19, Epidemic, Logistic, FbProphet, Modeling, Forecasting

2020 MSC: 00-01, 99-00

*Corresponding author

Email address: zhengxq@cugb.edu.cn (Xinqi Zheng)

1. Introduction

An outbreak of atypical pneumonia [coronavirus disease 2019 (COVID-19)] caused by Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) since late December 2019 has made huge impact on people life and work. The virus may spread from bats to humans through another intermediate host and cause severe respiratory syndrome [1], characterized by strong human-to-human transmission through the air [2, 3]. The world health organization (WHO) declared an international emergency on 31 January, 2020. Since initial identification, despite of strict control, now it becomes a pandemic in global, which is a big threat and challenge to world health and economy [4], the disease has spread to over 100 countries across the world (Figure 1). As of 3:33 PM on 16 June, 2020, a total of 8,044,683 COVID-19 cases have been reported worldwide, with 437,131 deaths and 3,883,243 survivors, with an overall case fatality rate of 5.43% [5]. John Hopkins University is offering the current data [6]. The infectivity of COVID-19 is greater than that of influenza, with an estimated R_0 value (the basic reproduction number, representing viral infectivity) of 2.28 [7].

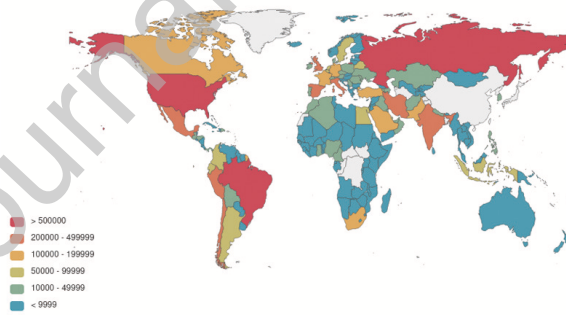


Figure 1: All infected countries with COVID-19 pandemic by June 16, 2020, the warmer the color is, the more infections are in the country.

According to the statistics of WHO, the United States of America (US), Brazil, Russia, India, and the United Kingdom (UK) are the top 5 countries with the highest number of infections in the world by 16 June, 2020 (Figure

20 2). More to the point, the number of infections worldwide is still increasing. A
 number of measures were taken to prevent the spread of the disease, including
 social distance, the closure of businesses and schools, and bans on travel and
 outdoor activities, beginning in the cited areas and later applied throughout
 the country. The degree of infection varies from country to country, and the
 25 control strategy and degree vary according to national conditions, too. How the
 global epidemic will peak or diminish is the most concerned problem worldwide.
 Therefore, it is of striking significance to predict the pandemic trends of infection
 worldwide.

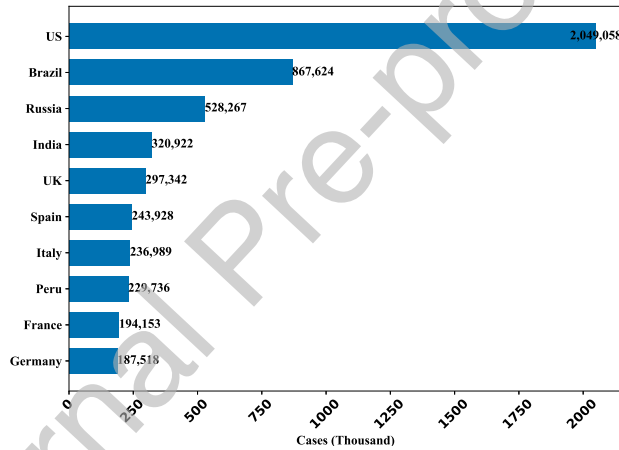


Figure 2: The world's top 10 countries with the number of confirmed COVID-19 cases.

Many scholars have developed a number of predicting methods for the trend
 30 forecasting of COVID-19, in some severe countries and global [8, 9], debating
 about mathematical model, infectious disease model, and artificial intelligence
 model. The models based on mathematical statistics, machine learning and
 deep learning have been applied to the prediction of time series of epidemic
 development [10, 11]. Logistic is often used in regression fitting of time series
 35 data due to its simple principle and efficient calculation. For example, in the
 Coronavirus case, Logistic growth is characterized by a slow increase in growth

at the beginning, fast growth phase approaching the peak of the incidence curve, and a slow growth phase approaching the end of the outbreak, i.e., the maximum of infections. Wu et al. [12] have calibrated the logistic growth model, the generalized logistic growth model, the generalized growth model and the generalized Richards model to the reported number of infected cases in the COVID-19 epidemics, and their different models imply that Logistic model could provide upper and lower bounds of our scenario predictions. When the sample size of machine learning model is small, the model is easy to overfit, and the model performs well in the historical data set. Therefore, the prediction effect is not good enough in real scene. For instance, the short-term forecasts of machine learning are good, but long-term ones are inferior.

In order to obtain the long-term prediction curve, Yang et al. [8] train the Long Short-Term Memory (LSTM) model based on the 2003 severe acute respiratory syndrome (SARS) data because the times series data of SARS is complete. However, SARS and COVID-19 are different in many ways, such as incubation rate, incubation period and the development degree of the world at the time of the outbreak, so the simulation may be not reliable. FbProphet is an open source library of Facebook [13]. It uses the method of time series decomposition and machine learning fitting, and can predict time series with high accuracy with simple and intuitive parameters. However, it only has two types of prediction, linear and Logistic, which are more accurate for short-term prediction. If we want to use Prophet with Logistic growth for long-term prediction, a cap value must be specified. Therefore, we integrate the most updated COVID-19 epidemiological data before June 16 into the Logistic model to fit the cap of epidemic trend, and then feed the cap value into Fbprophet model, a time series prediction model to derive the epidemic curve and predict the long term of epidemic.

This paper aim at constructing a reliable model based on time series data to predict the long-time period trend of COVID-19 in global, also in some heavily infected countries. Forecasts about when the pandemic will reach the turning point and global maximum number of infections may be of great usefulness for

public decision-makers and humans, as it could relieve peoples emotions and give policy-makers the time to respond to the changes that the outbreak has brought.

The rest of this paper is structured as follows. In Section 2, data sources and our proposed method are introduced detailed. Experiments and results analysis are given in Section 3. Section 4 presents the discussion of our research. Finally, in Section 5, the main contribution of this paper is summarized.

2. Methodology

2.1. Data Sources

The most recent daily COVID-19 outbreak epidemical data in global are collected from John Hopkins University, which released a dashboard at the country level. Data are available at an ArcGIS platform [6]. Therefore, we collected the time series at County level for the periodical analysis from January 22, 2020 to June 16, 2020. The 2003 SARS epidemic data between April and June 2003 across the whole of China retrieved from an archived news-site (SOHU) [14] was used for Logistic training.

2.2. A Logistic growth forecasting model

The Logistic model originated from the modeling of population growth in ecology [15]. As an improvement on the Malthus population model [16], in 1838, Pierre Franois Verhulst published the logistic equation:

$$\frac{dQ}{dt} = rQ(1 - \frac{Q}{K}) \quad (1)$$

where Q , r and K indicate the population size, intrinsic growth rate and maximum population size that the environment could carry, respectively. dQ/dt represents the growth of the population. r and K are constants number and the value of Q derives with time to produce an S-shaped curve in this Logistic equation.

The Logistic growth model was first used by LotKa [17]. Nowadays, Logical functions have been applied in a series time series prediction problem beyond

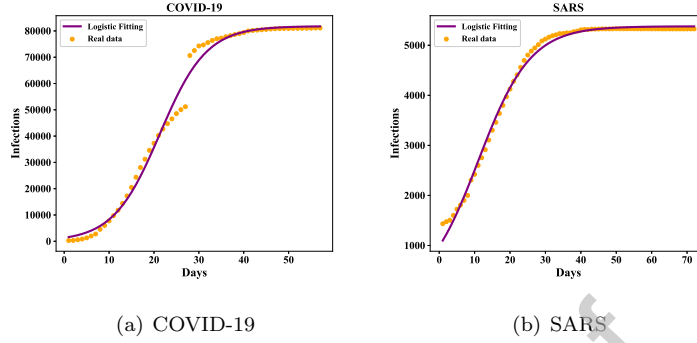


Figure 3: The Logistic growth modeling of COVID-19 and SARS 2003 in China.

95 their ecological roots, like epidemiology modeling. Logistic Growth is characterized by increasing growth in the beginning period, but a decreasing growth at a later stage, as you get closer to a maximum. Zhou and Yan [18] used the Richards model (a logistics type of model) to fit the cumulative number of SARS cases reported daily in Singapore, Hong Kong and Beijing. As is shown in Figure 3, the cumulative number of SARS and COVID-19 in China are always be
 100 S-shaped curve which could be well described by Logistic fitting. At the beginning of the outbreak, people did not take strict measures, and the initial number of people infected is small, so the number of infections slowly increased. When the infection base reached a certain proportion, the epidemic situation showed
 105 an explosive exponential growth trend, and then with the government's regulation and public's cooperation, the epidemic situation gradually slowed down the spread speed, finally reached the maximal number of cumulative infected people. The key point is the time at which the cumulative situation curve turns, i.e., when rapid increases in the number of cases are replaced by slow increases.
 110 Because this inflection point specifies the point at which the daily number of cases increases to a maximum, this moment marks a critical turning point at which transmission of the disease begins to decline. As long as the data include this key point and the time interval shortly thereafter, the curve fitting and prediction of the future number of cases will be fairly accurate.

115 2.3. *FbProphet Model*

Machine learning techniques for forecasting algorithms are a branch of computer science that is trained from historical data, such as artificial neural networks, deep learning, decision trees and Bayesian networks [19, 20, 21]. The idea of the algorithm is to select a suitable training model according to the characteristics of historical data and use it to predict the future observation results. We applied this technique to COVID-19 forecasting in the real world. Prophet is an open source framework of Facebook for time series forecasting based on additive model which is opened up to the public in 2017 [13, 22]. The non-linear trends of Prophet are fitted with yearly, weekly, and daily seasonality, plus holiday effects. The perfect Prophet function can not only predict the future, but also fill in missing values and detect anomalies. In Prophet, the prediction model consists of superpositions $y(t) = g(t) + s(t) + h(t) + \epsilon_t$, where $g(t)$ is a trend function used to analyze the non-periodic changes of time series. $s(t)$ is a periodic term, reflecting the periodic change, such as the periodicity of a week or a year. $h(t)$ is the influence of an occasional day or days, such as a holiday. ϵ_t is an error term, on behalf of the failed to consider the effect of the error of the model. Here in our research, we only consider about the non-periodic changes of time series.

We create an instance of the Prophet class and then call its fit and predict methods. The input to Prophet is always a time series with two features: date ds and value y . Here in our study, ds is the date of day, and y is the accumulated cases in a particular country.

2.4. *Formulation of the proposed method*

In this paper, We integrate the most updated COVID-19 epidemiological data before Jun 16, 2020 into the Logistic model to fit the cap of epidemic trend, and then feed the cap value into Fbprophet model, a machine learning based time series prediction model to derive the epidemic curve and predict the trend of epidemic. The main steps of our method are described in Figure 4.

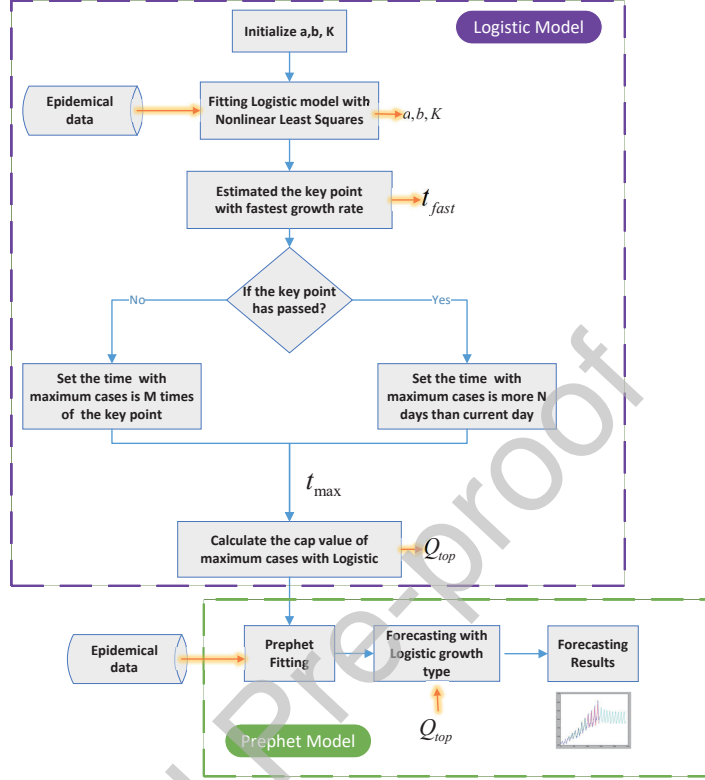


Figure 4: Framework of forecasting the trend of COVID-19 with Logistic model and Prophet. The key point with fastest growth rate t_{fast} is fitted from Logistic model based on epidemical data of COVID-19. By estimating the time t_{max} with maximum cases, the cap value of epidemic size Q_{top} is calculated by using Logistic modeling method, then the cap value is used to the Prophet model, and model the full epidemic trend of COVID-19.

To modeling the Logistic growth of COVID-19, Q , r and K in Equation (1) indicate the size of accumulates infected cases, intrinsic growth rate and maximum cases size that the world or country could carry, respectively. dQ/dt represents the growth of the scale. r and K are constants number and the value of Q derives with time to produce an S-shaped curve in this Logistic equation. Logistic Growth is characterized as follows [23]:

$$Q(t) = \frac{K}{1 + a * e^{bt}} \quad (2)$$

where $Q(t)$ is the number of cases at time t , a is constant, b could be considered as incubation rate and K is the cap value, the maximal number of cases for $Q(t)$. Therefore, the number of cases at the very beginning is $K/(1+a)$, and the key point is $\ln a/b$ at which the cumulative situation curve turns, when rapid increases in the number of cases are replaced by slow increases. We initialize a, b and K randomly and update it by using Nonlinear Least Squares [24]. The upper bound of b is set according to the pandemic situation for each country. Hereafter, we could model the growth of COVID-19 with Logistic formula and learn the parameters a, b and K . Using calculated parameters from the Logistic growth model, Prophet is applied to forecast the trend of COVID-19.

Based on the key point calculated from Logistic model fitting with Nonlinear Least Squares method, we can forecast the maximal number of cases with a time interval P . A number of measures were taken to prevent the spread of the disease and the situation of infection varies from country to country. For example, the growth of COVID-19 in China has already stabled, while Brazil is still increasing at a scale. Therefore, the first step we need to do is identifying whether the virus is still growing sharply or have passed the exponential growth period and starts to reach the maximal cases size in a particular country. The key point t_{fast} and the accumulated cases $Q(t_{fast})$ at that time can be written as follow:

$$t_{fast} = \frac{\ln(a)}{b} \quad (3)$$

$$Q(t_{fast}) = \frac{K}{2} \quad (4)$$

If the current day in time series is less than t_{fast} , it means the key point is still ahead and the growth is increasing maybe exponential. Otherwise, it means that virus spread has been controlled and growth is going towards the end. We set the time t_{max} with maximum cases is three times of t_{fast} for the first kind of growth while it is more 20 days than current day for the second kind of growth. Then the estimated top number of infections Q_{top} can be calculated as follows:

$$Q_{top} = \frac{K}{1 + a * e^{(-bt_{max})}} \quad (5)$$

The aforementioned top number of infections Q_{top} will be feed into Prophet model with actual time series data. We perform around five to six months ahead forecasting by using Prophet, with 95% prediction intervals and logistic growth type. No tweaking of seasonality-related parameters and additional regressors are performed.

3. Results

In this section, experiments for forecasting the trend of COVID-19 in global and particular countries based on the epidemiological data (from January 22, 2020 to June 16, 2020) are conducted to analyze the effectiveness of our proposed model. There are 4 parameters in our experiments. The lower bounds of a, b and K in Logistic model are all set to be 0 and the upper bounds of them are set to be 1000,000, 5, and 7,000,000,000, respectively. The calibrated model is then employed to generate 160-day-ahead to 200-day-ahead (time interval P) dynamic forecast over the post-sample period: from June 16, 2020, to December 23, 2020 or January 2, 2021. All of our experiments are executed on the Colaboratory platform provided by Google.

3.1. Forecasting the epidemic trend in Global

Three time series are constructed from our collected data, namely confirmed cases, recovered cases and death cases series. In our experiment, we assume that each of these three sequences has a peak, in other words, the epidemic will end eventually. Obviously, the number of active confirmed cases is equal to the number of accumulated confirmed cases minus the number of recovered and deaths. We first apply Logistic model to fit the curve and calculate the time with fastest growing rate, then use Prophet to make a prediction. Figure 5 (a) shows the predicting results of Prophet model for the total case number in Global. We then plotted the number of accumulated, recovered, death and

active confirmed cases derived from the hybrid Logistic and Prophet model and
 205 the actual reported total case numbers for global. There was a remarkable fit
 between the actual number of confirmed cases and the predicted curve between
 January 22 and the June 16.

It is worth noting that there are three significant points in our forecasting
 results, as is described in Table 1. The first is the maximum number of existing
 210 infections (the time when the blue line reach peak in Figure 5), i.e., epidemic
 peak point. The second point is the fastest growing point of the epidemic,
 after which, the epidemic gradually slows down and finally becomes stabled.
 The third point is the time when the number of cumulative cured (green line)
 exceeds the number of active confirmed cases (blue line), which marks an early
 215 victory in the control of the epidemic. Of course, the maximum number of total
 infections can also be obtained from the experimental results. Table 2 is given
 to summarize the major turning points in global, Brazil, Russia, India, Peru
 and Indonesia, respectively. With current control, the total epidemic size in
 global is predicted to peak on June 21 with 3,603,111 active infections. The
 220 total epidemic size is predicted to be 14,117,911 by the end of October (Figure
 5 (a)). The fastest growth point has passed on May 16. The actual reported
 cumulative number of infections is also in good agreement with our predicted
 curve.

Table 1: Implication of main points.

Main point name	Implication
Epidemic peak point	This peak means that the active infections has reach the top value and since then, the number of active cases will decrease.
The fastest growth point	After this point, the epidemic gradually slows down and finally becomes stabled.
Turn point	This point occurred when the number of cumulative cured exceeds the number of active confirmed cases, marks an early victory in the control of the epidemic.

3.2. Forecasting the epidemic trend in particular countries

225 We chose Brazil, Russia, India, Peru and Indonesia as the forecast countries,
 because these countries are among the highest in the world in the total num-
 ber of confirmed COVID-19 infections and are the most concerned countries

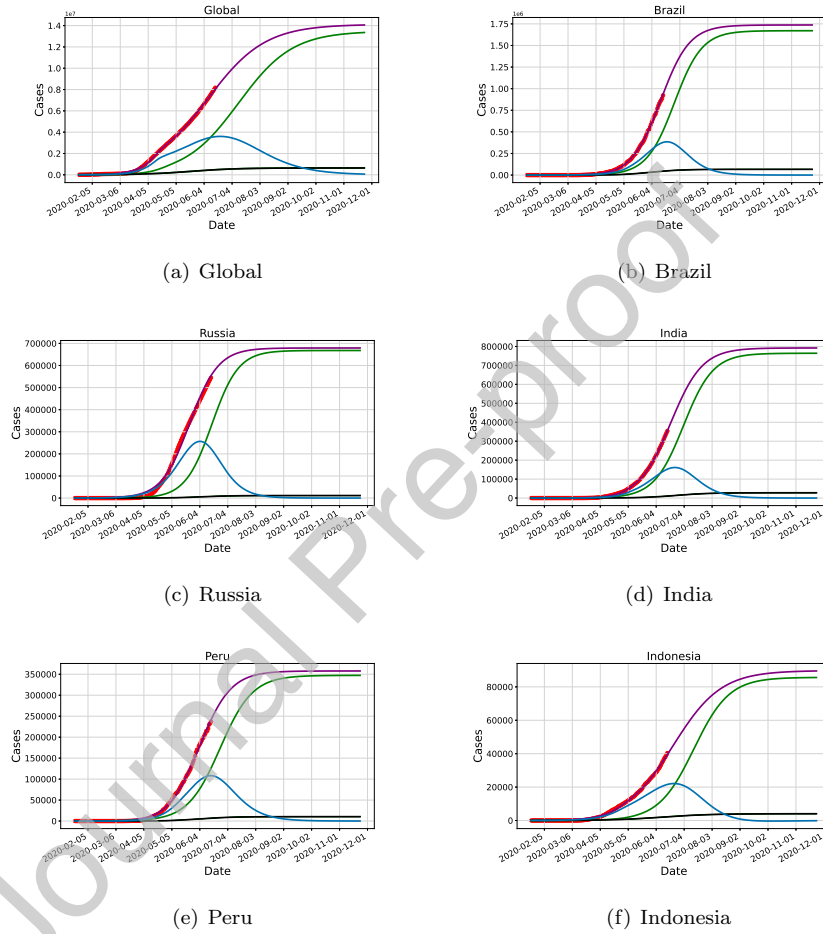


Figure 5: Number of accumulated confirmed cases (purple), recovered (green), death (black), and active confirmed cases (blue) by the hybrid Logistic and Prophet model for Global (a), Brazil(b), Russia (c), India (d), Peru (e), and Indonesia (f). Actual data of accumulated confirmed infections were fitted onto the curve (red circles).

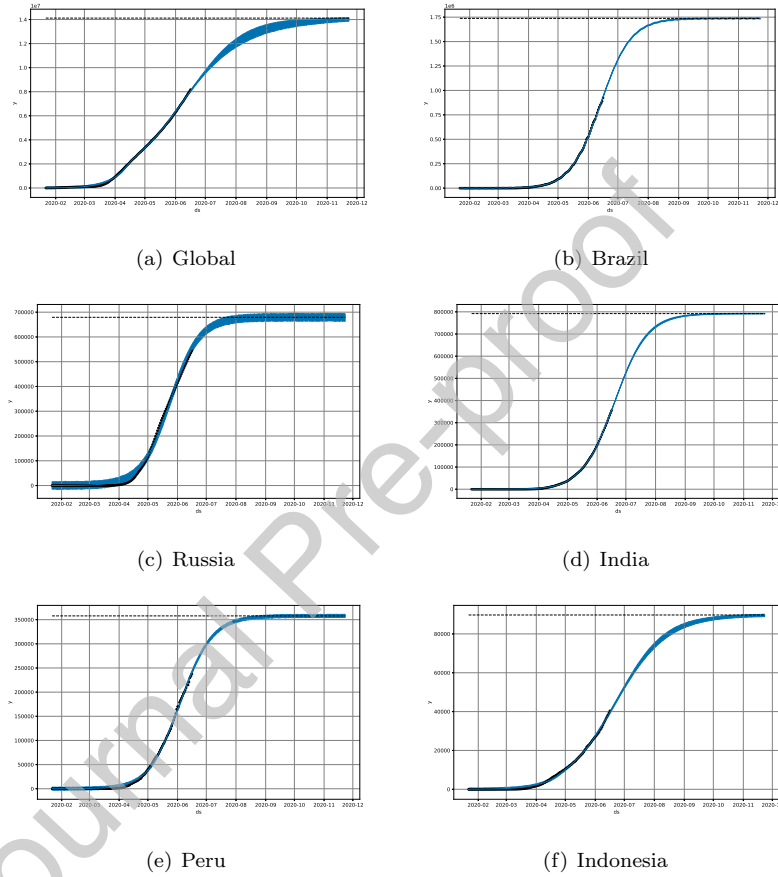


Figure 6: The predicting results of Prophet model for the total confirmed case number in Global (a), Brazil (b), Russia (c), India (d), Peru (e), and Indonesia (f). The line with a 95% confidence interval is confirmed cases and the true reported values are marked as circles. The horizontal axis represents the date, and the vertical axis represents the cumulative number of infections.

Table 2: Summary of the main turning points from our study.

Area	Epidemic peak		The fastest growth point		Turn point		Epidemic Siz
	Cumulative active infections		Cumulative infections		Cumulative Recovered		
	Time	Number	Time	Number	Time	Number	
Global	2020/6/21	3,603,111	2020/5/16	4,706,052	2020/6/11	3,524,598	14,117,911
Brazil	2020/6/20	384,998	2020/6/10	742,424	2020/6/7	324,872	1,737,272
Russia	2020/6/4	256,535	2020/5/18	283,029	2020/6/10	252,753	679,269
India	2020/6/24	161,185	2020/6/14	330,043	2020/6/9	135,419	792,103
Peru	2020/6/16	108,217	2020/5/29	149,088	2020/6/12	107,906	357,812
Indonesia	2020/6/23	22,127	2020/6/4	29,914	2020/6/25	22,668	89,741

worldwide. We could summarize our basic predictions as follows:

As is shown in Table 2, the fastest growth point of those five countries has already passed. With current intervention, the total epidemic size in Brazil is predicted to peak on late June with 384,998 active infections. The total epidemic size is predicted to be 1,737,272 and the fastest growth point came on June 10. The turning point at which the cumulative recovered cases would greater than active ones came by the middle of June.

For Russia, the fastest growth point and epidemic peak has passed on May 18 and June 4, with an 283,029 and 256,535 active infections, respectively. The turning point where cumulated recovered cases exceeded active infections came on June 10. In terms of India, the epidemic peak is still ahead, which is predicted on June 24. At that time, the number of active infections predicted by our model is 161,185. The fastest growth point came on June 14, with a total epidemic size of 330,043. And the turning point in India is predicted in early June.

Our hybrid Logistic and Prophet model predicted that the epidemic of Peru peaked on middle June, resulting in 108,217 active cases, and the scale of epidemic size will reach 357,812. We also predicted a peak of 22,127 active cases in Indonesia on June 23, and the epidemic size is expected to be 89,741. The cumulative recovered cases will be greater than active cases by late June according to the prediction of our model.

We plotted the actual reported cumulative active infections (red circles in Figure 5 (b), (c), (d), (e), (f)) up to June 16 2020 for each country onto our predicted curve and the actual cumulative number of confirmed and predicted values were compared separately (Figure 6). The results found that there was

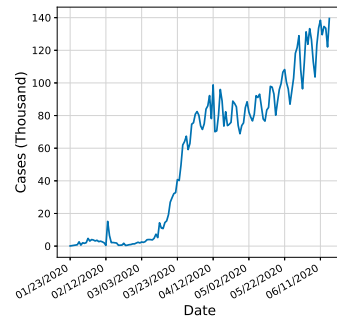
overall a good fit between our projected and reported data.

4. Discussion

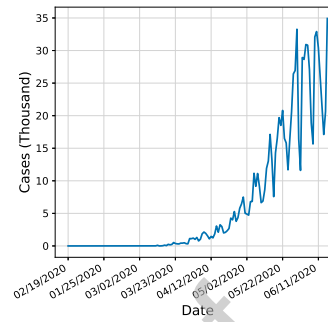
According to the experimental results, the time at which the important
 255 inflection point arrives vary from country to country. The epidemic is projected
 to begin in early-August with the number of new daily deaths in Brazil edging
 toward 4600. This result is in accordance with the result of Institute for health
 measurement and evaluation, university of Washington [25]. Brazil has viewed a
 massive spread of the disease, in addition to the large number of deaths caused
 260 by COVID-19, the epidemic of Brazil will burden hospitals far beyond their
 current management capacity, especially in ICU care [26]. Once the outbreak
 exceeds the combination of national health resources, it will take a long time to
 recover. In our predicted results, by the late July or mid of August 2020, the
 healing rate will increase and Brazil will have about 1.7 million confirmed cases
 265 at the end.

Our study highlighted another key point, the strict control measures adopted
 since March by many countries are effective, leading to the decline in growth
 rate in global around middle May. But as outbreaks continue in some particular
 countries, the worldwide epidemic may rebound. Because when the epidemic
 270 goes beyond a certain scale, it becomes very difficult to control or get back to the
 previous level. Russia, for example, reported its first confirmed case on March 2
 and launched prevention and control measures in early March. According to the
 Russian Health Ministry, as of 19 April 2020, the first decrease in the number
 of daily confirmed cases in Russia, to 4,268. This may be due to the fact that
 275 severe control and isolation has been implemented in Russia, which shows that
 isolation is effective in reducing human exposure and controlling the epidemic.
 But in early May, there was a massive outbreak in this country, with 11,656
 new cases in a single day, leading to a modest slowdown despite strict control
 measures.

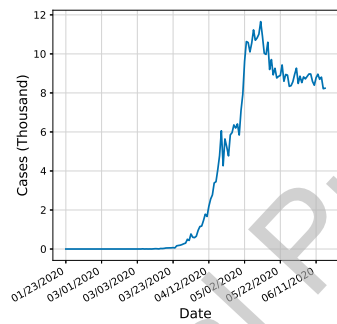
280 As of June 16, 2020, the case fatality rates in Russia, Brazil and Peru are



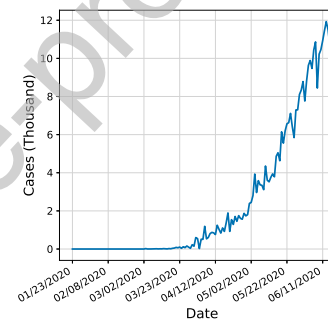
(a) Global



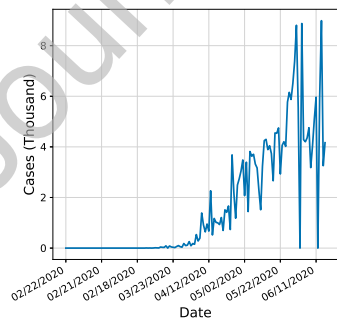
(b) Brazil



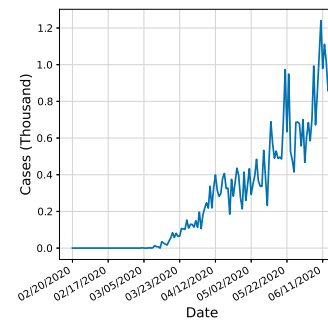
(c) Russia



(d) India



(e) Peru



(f) Indonesia

Figure 7: Number of Daily confirmed cases in Global (a), Brazil (b), Russia (c), India (d), Peru (e), and Indonesia (f) by June 16, 2020.

1.35%, 4.90% and 2.98%, respectively and there are still some increasing cases every day (Figure 7). Because of the rapid outbreak in South American and European countries, the medical system nearly collapsed in a short period of time [27]. After strict control, the growth rate of the epidemic gradually slowed
 285 down, and the days when the cumulative cure was greater than the existing diagnosis came in mid-June, which means that South American and European countries still have a long way to fight in strictly controlling the outbreak.

While more data is needed to make more detailed predictions, these models could help predict future confirmed cases if the spread of the virus does not
 290 change in a way beyond expectation. As we all know, this virus is new and has the ability to spread seriously. This characteristic may affect all our predictions, but to our best knowledge at the time we spent writing this paper, the proposed model is effective.

5. Conclusion

295 In this article, a forecasting method with Logistic and Prophet model is proposed to analysis the COVID-19. The cap value is fitted by Logistic model through determining the fastest growing point, which is then feed into Prophet model for forecasting. We conducted experiments for global pandemic and also in some particular countries to forecast the epidemic peak, growing fasted point
 300 and the turning point of recovered. The Results in Section 3 plot the predictive trend of global and five particular countries, and demonstrate the effectiveness of our model to predicting the turning point and epidemic size of COVID-19.

However, as is shown in Figure 6, all of our predictions are based on the assumption of there will be a maximum of outbreak, and the epidemic curve is
 305 modeled based on a full Logistic curve. In real world, there maybe some small peak during the pandemic due to different intervention of the government and different public cooperation. Besides, when we forecasting the epidemic in some countries, the effects of input cases and spatial influence between countries are not taken into account.

310 To address the aforementioned limitations, the following aspects are worthy
 for our future research. First, piecewise prediction with a variant of Logistic
 model deserves further study. Second, Prophet is a perfect model based on
 additive model, and the non-linear trends of Prophet are fitted with yearly,
 weekly, and daily seasonality, plus holiday effects. In our research, we only
 315 applied times series data as trend term into the model, therefore, converting
 the control measures, like travel bans as holiday effects into Prophet model is a
 significant research.

As the research in this paper shows that a hybrid Logistic and Prophet model
 has a valuable advantage in terms of forecasting the epidemic trend, using our
 320 model can significantly improve estimates of the number of infections in global
 and particular countries and help public policymakers better plan health policy
 interventions.

Acknowledgments

This research was supported by the Funds for Basic Scientific Research of
 325 Central Universities under Grant 2652020002 and 2652020004.

Declaration of Competing Interest

There is no conflict of interest in this work.

References

- 330 [1] X. Li, J. Zai, Q. Zhao, Q. Nie, Y. Li, B. T. Foley, A. Chaillon, Evolutionary
 history, potential intermediate animal host, and cross-species analyses of
 sars-cov-2, Journal of Medical Virology doi:10.1002/jmv.25731.
- [2] W. Guan, Z. Ni, Y. Hu, W. Liang, C. Ou, J. He, L. Liu, H. Shan,
 C. Lei, D. S. C. Hui, et al., Clinical characteristics of coronavirus dis-
 ease 2019 in china, The New England Journal of Medicine doi:10.1056/
 335 NEJMoa2002032.

- [3] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, L. Ksm, L. Ehy, J. Y. Wong, et al., Early transmission dynamics in wuhan, china, of novel coronavirusinfected pneumonia, The New England Journal of Medicinedoi: 10.1056/NEJMoA2001316.
- 340 [4] K. H. Jacobsen, Will covid-19 generate global preparedness?, The Lancetdoi:10.1016/S0140-6736(20)30559-6.
- [5] W. H. Organization, Coronavirus disease (covid-19) situation report 147, [EB/OL], [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200615-covid-19-sitrep-147.pdf?sfvrsn=](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200615-covid-19-sitrep-147.pdf?sfvrsn=2497a605_2)
345 2497a605_2 Accessed June 16, 2020.
- [6] E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track covid-19 in real time, Lancet Infectious Diseasesdoi:10.1016/S1473-3099(20)30120-1.
- [7] S. Zhang, M. Diao, W. Yu, L. Pei, Z. Lin, D. Chen, Estimation of
350 the reproductive number of novel coronavirus (covid-19) and the probable outbreak size on the diamond princess cruise ship: A data-driven analysis, International Journal of Infectious Diseases 93 (2020) 201–204. doi:10.1016/j.ijid.2020.02.033.
- [8] Z. Yang, Z. Zeng, K. Wang, S. Wong, W. Liang, M. Zanin, P. Liu, X. Cao,
355 Z. Gao, Z. Mai, et al., Modified seir and ai prediction of the epidemics trend of covid-19 in china under public health interventions, Journal of Thoracic Disease 12 (2) (2020) 165–174. doi:10.21037/jtd.2020.02.64.
- [9] D. Giuliani, M. M. Dickson, G. Espa, F. Santi, Modelling and predicting the spatio-temporal spread of coronavirus disease 2019 (covid-19) in italydoi:
360 dx.doi.org/10.2139/ssrn.3559569.
- [10] P. F. Fanelli D, Analysis and forecast of covid-19 spreading in china, italy and france, Chaos, Solitons & Fractalsdoi:<https://doi.org/10.1016/j.chaos.2020.109761>.

- [11] R. M. e. a. Yousaf M, Zahir S, Statistical analysis of forecasting covid-19
 365 for upcoming month in pakistan, *Chaos, Solitons & Fractals* doi:<https://doi.org/10.1016/j.chaos.2020.109926>.
- [12] K. Wu, D. Darcet, Q. Wang, D. Sornette, Generalized logistic growth modeling of the covid-19 outbreak in 29 provinces in china and in the rest of the world, *arXiv: Populations and Evolution*.
- [13] S. J. Taylor, B. Letham, Forecasting at scale, *The American Statistician*
 370 72 (1) (2018) 37–45. doi:10.1080/00031305.2017.1380080.
- [14] Combatting sars (in chinese) 2003, [EB/OL], <http://news.sohu.com/57/26/subject206252657.shtml>.
- [15] Q. Yu, J. Liu, Y. Zhang, J. Li, Simulation of rice biomass accumulation
 375 by an extended logistic model including influence of meteorological factors, *International Journal of Biometeorology* 46 (4) (2002) 185–191. doi:10.1007/s00484-002-0141-3.
- [16] T. R. Malthus, *An essay on the principle of population*, University of Michigan Press,, 1986.
- [17] A. J. Lotka, *Elements of Physical Biology*, Dover Publications, 1956. doi:
 380 10.2105/AJPH.15.9.812-b.
- [18] Y.-H. Hsieh, J.-Y. Lee, H.-L. Chang, Sars epidemiology modeling, *Emerging Infectious Diseases* 10 (6) (2004) 1165–1167. doi:10.3201/eid1006.031023.
- [19] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, J. Liu, Lstm network: a deep
 385 learning approach for short-term traffic forecast, *Iet Intelligent Transport Systems* 11 (2) (2017) 68–75. doi:10.1049/iet-its.2016.0208.
- [20] G. W. Chang, H. J. Lu, Y. R. Chang, Y. D. Lee, An improved neural network-based approach for short-term wind speed and power forecast,

- 390 Renewable Energy 105 (2017) 301–311. doi:10.1016/j.renene.2016.12.071.
- [21] P. Wang, X. Zheng, J. Ku, C. Wang, Multiple-instance learning approach via bayesian extreme learning machine, *IEEE Access* PP (99) (2020) 1–1. doi:10.1109/ACCESS.2020.2984271.
- 395 [22] Prophet: Automatic forecasting procedure, [EB/OL], <https://facebook.github.io/prophet/docs/orhttps://github.com/facebook/prophet>.
- [23] Y. Yan, E. Williams, Logistic model-based forecast of sales and generation of obsolete computers in the u.s., *Technological Forecasting & Social Change* 76 (8) (2009) 1105–1114. doi:10.1016/j.techfore.2009.03.004.
- 400 [24] J. E. Dennis, R. E. Walsh, An adaptive nonlinear least-squares algorithm, *ACM Transactions on Mathematical Software* 7 (3) (1981) 348–368. doi:10.1145/355958.355965.
- [25] T. I. for Health Metrics, Evaluation, Covid-19 projections, [EB/OL], <https://covid19.healthdata.org/united-states-of-america>.
- 405 [26] T. F. F. Ribeiro V S T, Joo Paulo Telles, Arboviral diseases and covid19 in brazil: Concerns regarding climatic, sanitation and endemic scenario, Arboviral diseases and COVID19 in Brazil: Concerns regarding climatic, sanitation and endemic scenari doi:<https://doi.org/10.1002/jmv.26079>.
- 410 [27] t. E. C. The WHO Regional Office for Europe, the European Observatory on Health Systems, Policies, Covid-19 health system response monitor (hsrm), [EB/OL], <https://www.covid19healthsystem.org/mainpage.aspx>.