Original Research Paper

Prediction of Covid-19 Outbreak in India by Employing Epidemiological Models

[1]**Sudheer Reddy K**, [2]**Anji Reddy V**, [3]**Mamatha K** and [4]**Santhosh Kumar C.**

[1]*Researcher, Hyderabad, India*
[2]*Department of Computer Science and Engineering, Lendi Institute of Engineering and Technology, Vizianagaram, India*
[3]*Department of Business Management, Anurag Group of Institutions, Hyderabad, India*
[4]*Department of Computer Science and Engineering, Anurag Engineering College, Kodad, India*

Abstract: The novel Corona virus (Covid-19) pandemic has frightened an unprecedented challenge to continents. The Covid-19 pandemic is a highly contagious viral disease that has been progressing across the globe. The global thought leaders, doctors, scientists and other intellects have enforced the lockdown to reduce the spread of the pandemic. The objective is to study the impact of lockdown and to compare its effects from an epidemiological standpoint. Authors employed Exponential and SIR (Susceptible, Infected and Removed) epidemiological models for study and analysis. The study reveals that India will enter equilibrium by the end of June with a prediction of 135,000 cases approximately.

Keywords: Novel Corona Virus, COVID-19, Outbreak, Epidemiological Model, SIR Model

Introduction

The International Committee on Taxonomy of Viruses (ICTV) titled the virus as Severe Acute Respiratory Syndrome Corona Virus 2 (SARS-CoV-2) (Healthline, 2020; He *et al*., 2020). As of 30th April 2020, over 3.5 million positive cases of Covid-19 and about over 200,000 deaths recorded globally. About 60% of the population across the globe is locked down to prevent the spread of theCovid-19. The characteristic features of COVID-19 are presented in (Singhal, 2020; Zunyou and McGoogan, 2020).

India, the second-largest populated country in the world, is at a humongous risk due to its larger density of population and insufficient medical infrastructure to cater to the massive loads of demands. In India, the first COVID-19 positive case reported on 30th January 2020, soon after, it spreads across the nation (India Press Release, 2020). However, the percentage of cases and deaths is found to be lower as compared with other nations. As of 30th April 2020, the total cases reported 33610 in which 24162 active, 8373 recoveries and 1075 deaths are recorded (Twitter Inc, 2020), graphical data is presented in Fig. 1.

On 25th March 2020, the Country had taken an early, strategic and historic decision of National lockdown, when there were only 536 cases and 11 mortalities. Even after, positive cases have been increasing exponentially and however with marginal death rate of about 2%. Hence, the lockdown period is further extended until 3rd May 2020, with fewer relaxations to certain regions where lesser or no cases reported. The nationwide Covid-19 data is collected, compiled and analyzed. It is observed that the basic reproduction number is within the expected range of 1.5-4.7. Exponential and classic SIR models are employed to make predictions on the available data (Ranjan, 2020; Tomar and Gupta, 2020). Due to certain limitations of the experimental model, the authors worked extensively on the SIR model.

The authors of this paper introduced the popular epidemiological models for predicting the outbreak. The study is to compare lockdown effects from an epidemiological standpoint. The SIR model depicts that the nation will enter equilibrium by the end of June with an estimate of 135 k infectious cases approximately. These estimates would be null if in case India enters into community transmission. It is pertinent to mention that the proposed model can be applied to any country to estimate the affects. The exponential, logistic and SIR models are presented for the study. However, SIR model was employed for the better analysis and prediction of the outbreak.

Materials and Methods

The day-wise data on the number of Covid-19 positive cases, recovered and deaths are collated from various public domains (including the official Covid-19 site, Wikipedia) during the timeline, 1st March till 30th April 2020. In addition to this data, the authors are also collected data and figures from the official portal of the Govt. of India. COVID-19 India figures as on 30th April 2020 are presented in the Fig. 1 (MHFW, 2020a; 2020b; Covid19.in, 2020; Twitter Inc, 2020).

Epidemiological Models

In this section, mathematical formulations for various epidemiological models are presented as they have predictive ability to estimate the growth of the epidemic.

Exponential, Logistic and SIR are the three categories of epidemiological models are discussed here.

Exponential Model

This model is an exponential fit for the short term predictions. Exponential models are usually predicting initial growth of epidemic diseases, but does not account for eventual decay. At the beginning of the epidemic, the infections growth rate with time is exponential (Hethcote, 2000).

Let us assume that $I(t)$ number of infections over time t, then the growth rate n can be evaluated by using the following equation (1):
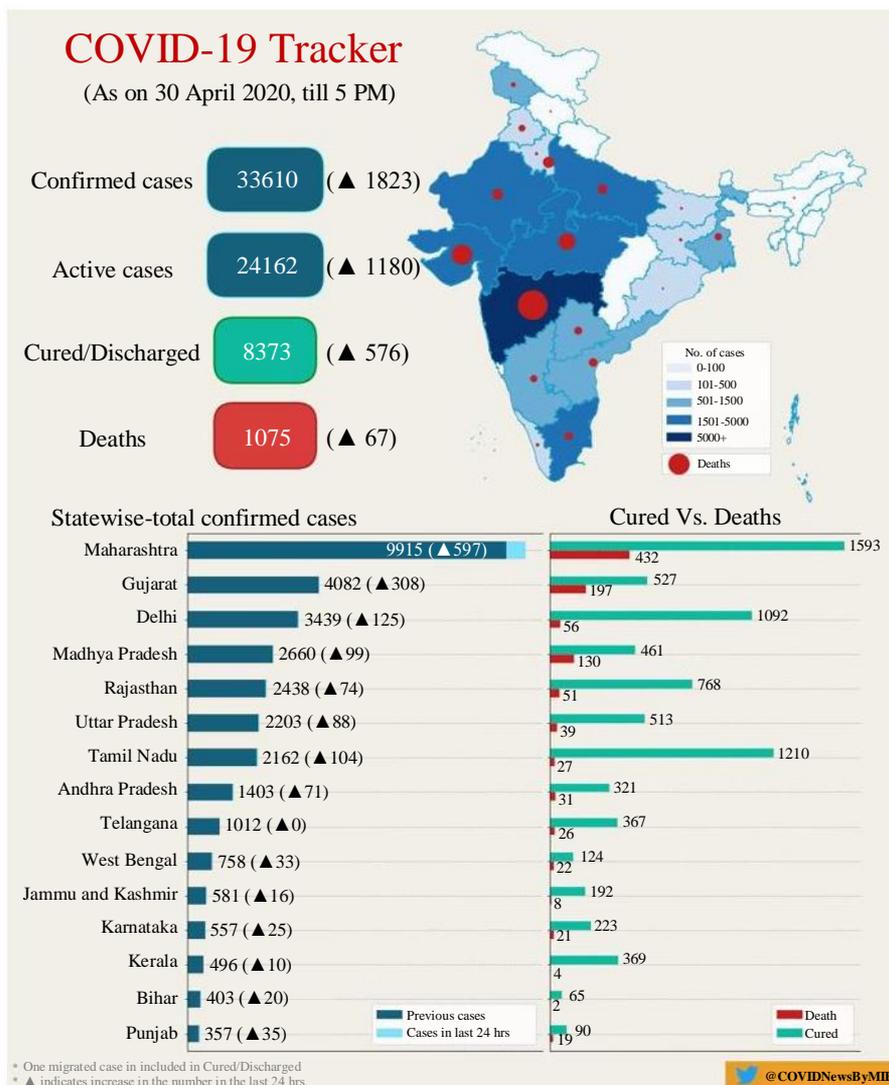
$$\frac{dI}{dt} = nI \tag{1}$$

Fig. 1: COVID-19, India figures as on 30th April 2020 (Source: COVIDNewsByMIB tweet)

After integration, $I(t)$ can be calculated as:

$$I(t) = I_0 \exp(nt) \tag{2}$$

Where:
I = Cumulative number of diagnosis,
Constant I_0 = For fitting the curve.

Logistic Model

While exponential model doesn't support decay, the Logistic model predicts the eventual decay. However, this model is unsuitable during initial stages. In short, the logistic regression is meant for the long term predictions (Hethcote, 2000). The following mathematical formulas can be used in logistic model:

$$I(t) \frac{m}{1 + C \exp(nt)} \tag{3}$$

Where:

$$C = \left(\frac{m}{I_0}\right) - 1$$

Assuming that $m > I_0$ and hence $C > 1$. When time t is small, one can choose the exponential model. However, the Logistic model is not fit for solving non-linear problems since its decision surface is linear.

SIR Model

The SIR model is employed to predict susceptible, infectious, recovered and/or deceased individuals (Ranjan, 2020; Twitter Inc, 2020):

$$\frac{dk}{dt} = -\beta k I \tag{4}$$

$$\frac{dI}{dt} = \beta k I - \gamma I \tag{5}$$

$$\frac{dS}{dt} = \gamma I \tag{6}$$

where, k is the fraction of susceptible individual, I is fraction of Infectious individual and S indicates the recovered number. The rate of transmission is β and the average recovery rate is γ.

From the Equation 6, after linearizing (by employing the disease free equilibrium), we can get:

$$\frac{dI}{dt} = (\beta - \gamma) I \tag{7}$$

$I(t)$ grows exponentially when $\beta - \gamma > 0$.

Assume that a pandemic free equilibrium for a completely susceptible population is zero.

In the current approach, the initial guess of β and γ are obtained by setting $\frac{dI}{dt} = 0$ and then the SIR equations are solved (Hethcote, 2000).

Results

In this section, the results of the exponential and SIR epidemiological models are presented. Figure 1 presented the growth rate in India until 30th April as a ready reference.

The Spatiotemporal Epidemiological Modeler (STEM) tool is employed to understand and predict the potential spread of COVID-19. STEM is an open source software tool by the Eclipse Foundation. The equations in the SIR model were solved using STEM's implementation of an adaptive integration algorithm, allowing the authors to specify a high level of precision.

The exponential model is renowned and best suits for short term predictions. In Table 1, the estimated coefficients for the exponential model are given. The co-efficient for this exponential fit is R2 = 1.0 and the p-value (~0:000) is almost negligible. These figures indicate that all the regression parameters are statistically significant [10]. This model predicts that the nation could have around 7000 cases by April 14. However in actual figure is close to 11 k that is 30% over and above the prediction. As the results are not encouraging for long term predictions through exponential fit, the authors consider SIR model for better prediction.

Earlier studies suggest that, a slightly longer duration is desired to obtain a rational estimate prediction for the SIR model (Hethcote, 2000; Elavarasan and Pugazhendhi, 2020). Hence, data from 25th March to 30th April, a total of 37 days were considered for SIR evaluation.

The SIR model is the best suited model for long term predictions. Table 2 has tabulated with the statistical parameters for SIR, with a coefficient of R2 = 1.3 and the p-value = 0 (approximately), β = 0.98 and γ = 0.72 indicating high statistical significance. The results are graphically presented in fig. 2.

Table 1: Nonlinear regression model: $y = a_1 \exp(a_2 x)$

Magnification	Estimate	95% of CI	Std. Error	tSat	p-Value
a1	47.8100	26.7800	6.21	11.070	5.8142e-06
a2	0.2236	0.3128	0,0126	19.631	9.1783e-08

Table 2: Statistical parameters for Epidemiological models

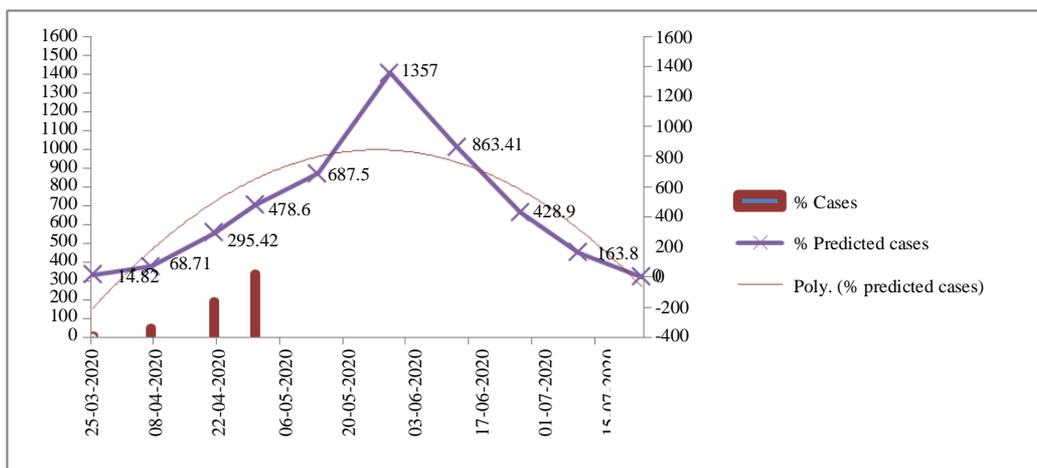| Model	Exponential model	SIR model
No. of observations	21	37
Degrees of Freedom	16	21
R-Squared	1.037	1.235
Adjusted R-Squared	1.0	1.3
p-Value	11.10e-10	8.34e-19
Root Mean Squared Error	21.62	46.4821
F-statistics vs. zero model	1076	971

Fig. 2: Covid-19 Predictions in India using SIR model

Figure 2 shows that India would reach a peak epidemic size of 135,000 cases around the end of June 2020. The acceleration phase has begins on 22nd April and continue to grow till the date. The slowdown phase begins after 3rd June and expected to reach zero by end of July 2020. The asymptotic slow growth persists between 3rd June and 30th July 2020. The predicted estimates of SIR model have reached to a plateau at 135k cases and this is final epidemic size.

Further, the Figure 2 reveals that the virus is expected touch the equilibrium by end of June as per the data available till 30 April 2020. However, there are high chances that the number may jump to a larger departure in the trend. Although, none of the said models will be valid if in case the nation enters into community transmission.

Another observation from the Fig. 1, 2 and Tables 1 and 2, the limited number of cases is due to the early imposing of the national lockdown.

Apart from this, the strict measures during the lockdown and social distancing would certainly help in restrict the spread of the pandemic.

Conclusion

In this study, a data-driven Exponential, Logistic and classic SIR forecasting methods have been explored and SIR model has been employed to predict the spread of COVID-19 in India. The SIR model assumes cases to be symptomatic, which may result in the values, underestimate the actuals. The model depicts that the country will touch the equilibrium by the end of June with a prediction of 135,000 positive cases. The proposed model can also be applied to various countries as well.

Declaration and Funding Information

Author's Contributions

Sudheer Reddy K: Conceptualization, Methodology, Software, Writing and Reviewing.
Anji Reddy V: Data curation, Validation, Editing.
Mamatha K: Investigation, Visualization.
Santhosh Kumar C: Investigation, Data curation.

Ethics

Authors acknowledge that there are no ethical issues that may arise after the publication of this manuscript.

References

Covid19.in, 2020. Official page: https://covid19.in/

Elavarasan, R. M. and Pugazhendhi, R. (2020). Restructured society and environment: A review on potential technological strategies to control the COVID-19 pandemic. Science of The Total Environment, 138858.

He, X., Lau, E. H., Wu, P., Deng, X., Wang, J., Hao, X., ... and Mo, X. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. Nature medicine, 26(5), 672-675.

Healthline, (2020). Healthline Media a Red Ventures Company.

Hethcote, H. W. (2000). The mathematics of infectious diseases. SIAM review, 42(4), 599-653.

India Press Release, (2020). Press Release. https://pib.gov.in/pressreleaseiframepage.aspx?prid=1601095

MHFW, (2020a). Ministry of Health and Family Welfare, India.

MHFW, (2020b). Ministry of Health and Family Welfare, India, 2020 Covid-19.

Ranjan, R. (2020). Predictions for COVID-19 outbreak in India using Epidemiological models. medRxiv.

Singhal, T. (2020). A review of coronavirus disease-2019 (COVID-19). The Indian Journal of Pediatrics, 1-6.

Tomar, A. and Gupta, N. (2020). Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. Science of The Total Environment, 138762.

Twitter Inc, (2020). Twitter handle - #COVIDNewsByMIB.

Zunyou, W. and McGoogan, J. M. (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 for Disease Control and Prevention. Jama, 323(13), 1239-1242.