Prediction of the Peak and Total Infected of the Coronavirus Disease in India

Parth Shah and Shubhadeep Nag

(Dated: 25 March 2020)

I. INTRODUCTION

January 30, 2020 marked the first identified case of the respiratory disease cased by a novel coronavirus in India. The World Health Organization (WHO) declared the disease a pandemic and named it the coronavirus disease 2019 (COVID-19) on February 11, 2020. The total number of confirmed cases in India and the daily increase in cases is depicted below.

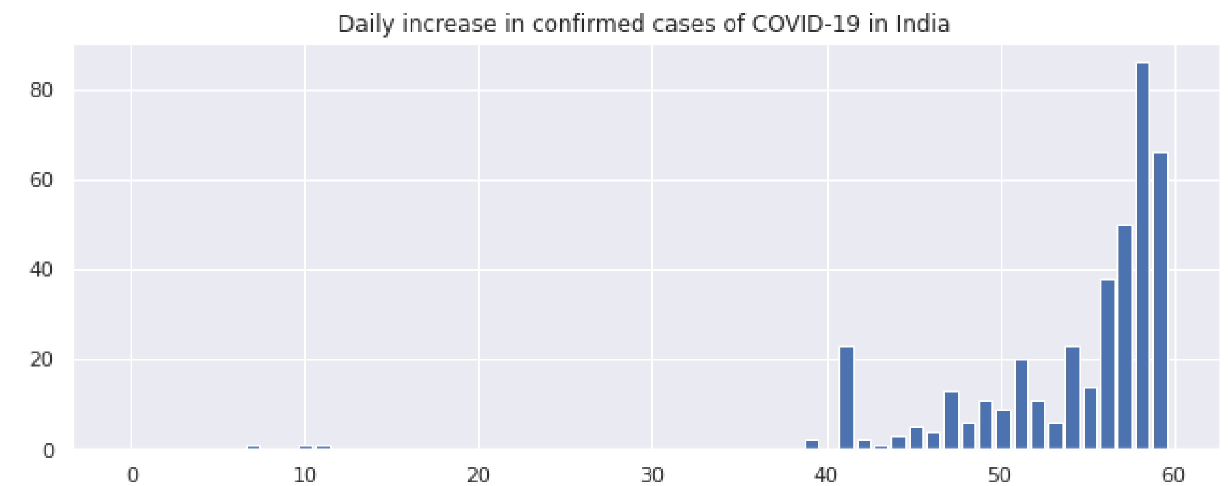

FIG. 1: Daily increase in confirmed cases of COVID-19 in India. Day 0 is January 22, 2020.

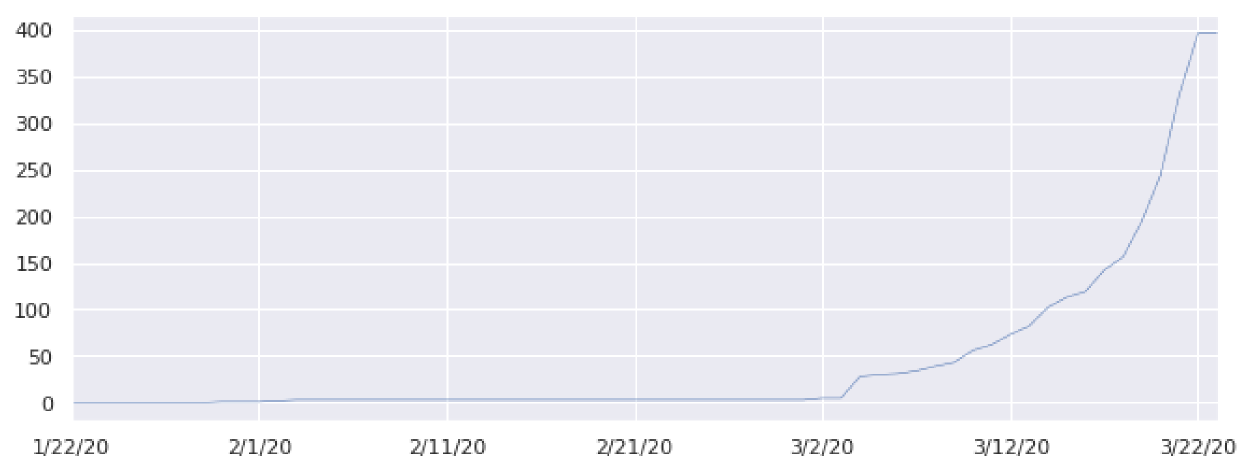

FIG. 2: Cumulative confirmed cases of COVID-19 in India.

The growth of the number of confirmed cases can be approximated by an exponential function. The growth factor r at day t can be calculated as

$$r = \frac{\text{Cases on day t}}{\text{Cases on day t-1}} \tag{1}$$

March 24, 2020, India is reporting an average growth rate of 1.230 calculated over the past 5 days. Stricter rules and measures of ensuring social distancing have begun. India is leading the fight against this disease and setting an example for the world by ensuring social distancing is protocol early on when the number of infected individuals are comparatively low.

II. METHODS

1. Model

A SEIR compartmental model is applied for the predictions. The assumption that once an individual contracts the virus and recovers, this individual is immune to this virus is adopted.

$$\begin{aligned}
S^{'}(t) &= -\beta S(t)I(t) \\
E^{'}(t) &= \beta S(t)I(t) - \epsilon E(t) \\
I^{'}(t) &= \epsilon E(t) - \gamma I(t) \\
R^{'}(t) &= \gamma I(t)
\end{aligned} \tag{2}$$

where $S(t), E(t), I(t), R(t)$ denote the susceptible, exposed, infected and removed populations at time t, respectively. β, ϵ, γ denote the infection rate, the onset rate and the removal rate, respectively. $1/\epsilon$, $1/\gamma$ are the average incubation period and the average infectious period respectively. The unit time is 1 day. We fix $1/\epsilon$ to 5 and $1/\gamma$ to 10. We also fix $S+E+I+R$ to 1 so that calculations are a proportion of the total. The assumption that 1 infected individual amongst a total population of $N = 1339.2 \times 10^6$ individuals is identified at $t = 0$ is adopted. Therefore

$$Y(t) = pI(t) \times N \tag{3}$$

are the confirmed number of infected individuals who are identified at time t. To obtain the initial conditions of the model, we assume that there are no exposed or removed populations at $t = 0$. Therefore

$$\begin{aligned}
&E(0) = 0 \\
&I(0) = \frac{1}{p \times N} \\
&R(0) = 0 \\
&S(0) = 1 - E(0) - I(0) - R(0) = 1 - \frac{1}{p \times N}
\end{aligned} \tag{4}$$

are the initial conditions. We fix $0.01 < p < 0.1$ based on multiple local reports of the density of affected people in a particular region. A reproduction number, R_0 is calculated. This is the expected value of secondary cases produced by one infected individual. Based on existing methodology

$$R_0 = \frac{\beta S(0)}{\gamma} = \frac{\beta}{\gamma}(1 - \frac{1}{p \times N}) \tag{5}$$

a. *Estimation of the infection rate,* β

Let $y(t), t = 0, 1, 2...60$ be the daily confirmed cases of COVID-19 in India from January 22, 2020 ($t = 0$) to March 23, 202 ($t = 60$). Using the least square approach with Poisson noise to estimate the infection rate, the following steps are adopted. With Poisson noise, equation 3 is modified to

$$\hat{Y}(t) = Y(t) + \sqrt{Y(t)}\epsilon(t) \tag{6}$$

$\epsilon(t), t = 0, 1, 2...60$ are random variables from a standard normal distribution.

1. For $\beta > 0$, calculate $Y(t), t = 0, 1, 2...60$ using equation 3.

2. Calculate using equation 6.

3. Calculate $J(\beta) = \sum_{t=0}^{60}[y(t) - \hat{Y}(t)]^2$

4. Run step 1 to step 3 for $0.2 \leq \beta \leq 0.4$ and find β^* such that $J(\beta^*) = min_{0.2 \leq \beta \leq 0.4} J(\beta)$

5. Repeat step 1 to step 4 10000 times and obtain the distribution of β^*. Approximate the same by a normal distribution and obtain the 95% confidence interval.

We obtain a value of β equal to 0.258 and the 95% confidence interval as $0.250 - 0.266$. Also, R_0 is equal to 2.58 and the 95% confidence interval as $2.50 - 2.66$.

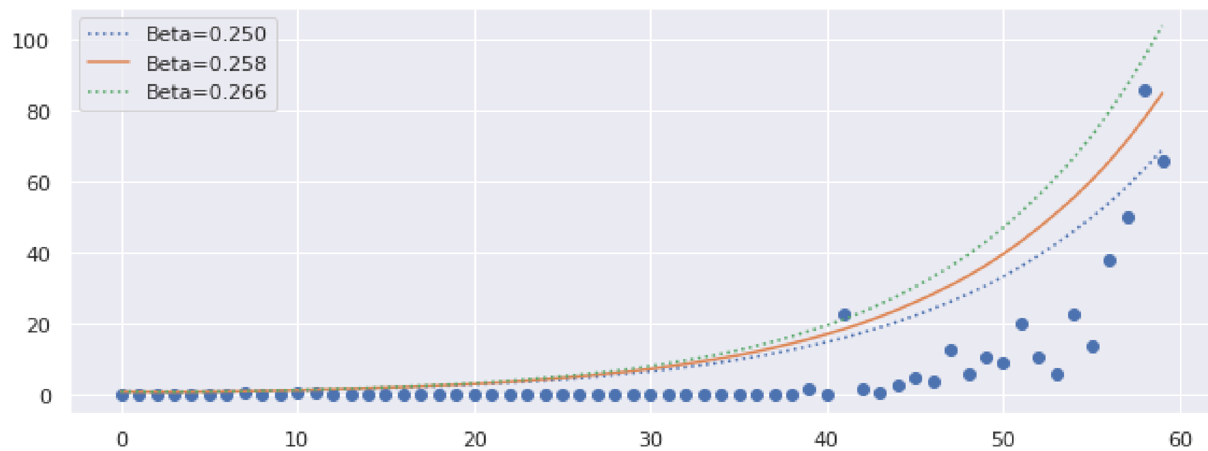

FIG. 3: Comparison of daily confirmed cases and $\hat{y}$ in India from $t = 0$ to $t = 60$

TABLE I: Parameters

Parameter	Description	Value
β	Infection rate	0.258
R_0	Reproduction number	2.58
ϵ	Onset rate	0.2
γ	Removal rate	0.1
N	Total population of India	1339.2×10^6
p	Identification rate	$0.01 - 0.1$

III. RESULTS

1. Peak Prediction

The epidemic peak, t^* is defined as the maximum value Y in a period of 1.5 years or $Y(t^*) = \max_{0 \leq t \leq 550}$. As the epidemic peak and size are sensitive to the identification rate p, we report the following.

For $p = 0.1$, the estimated peak is $t^* = 229$ with a 95% confidence interval of 221-238.

That is, starting with January 22, 2020 ($t = 0$), the estimated peak is September 8, 2020 ($t = 229$) and interval ranging from August 3, 2020 to September 16, 2020.

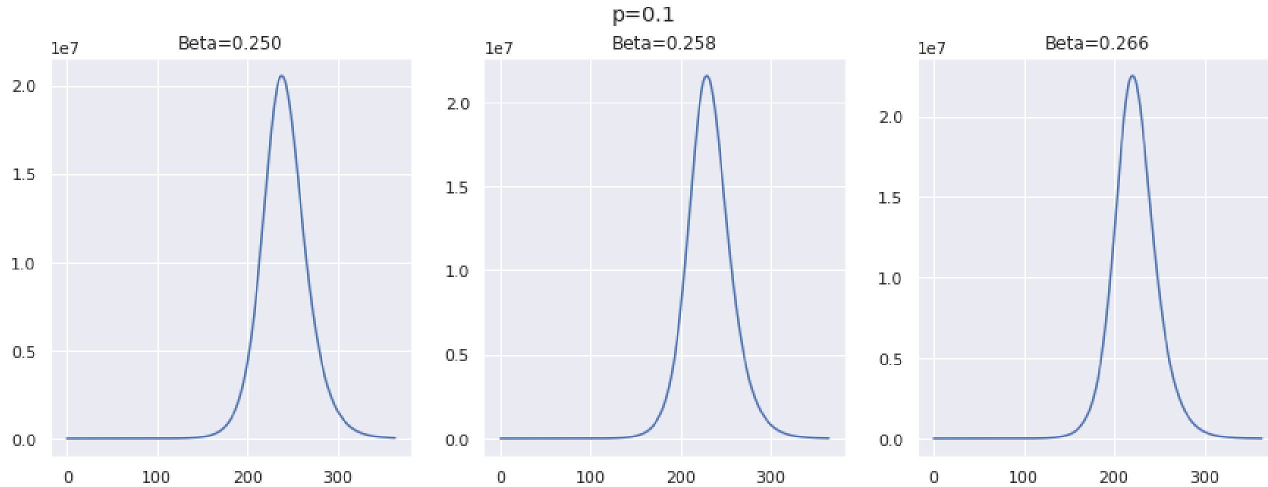

FIG. 4: Infected individuals for time t, $0 \leq t \leq 540$ for $p = 0.1$

For $p = 0.01$, the estimated peak is $t^* = 202$ with a 95% confidence interval of 194-210. That is, starting with January 22, 2020 ($t = 0$), the estimated peak is August 12, 2020 ($t = 202$) and interval ranging from August 4, 2020 to August 19, 2020.

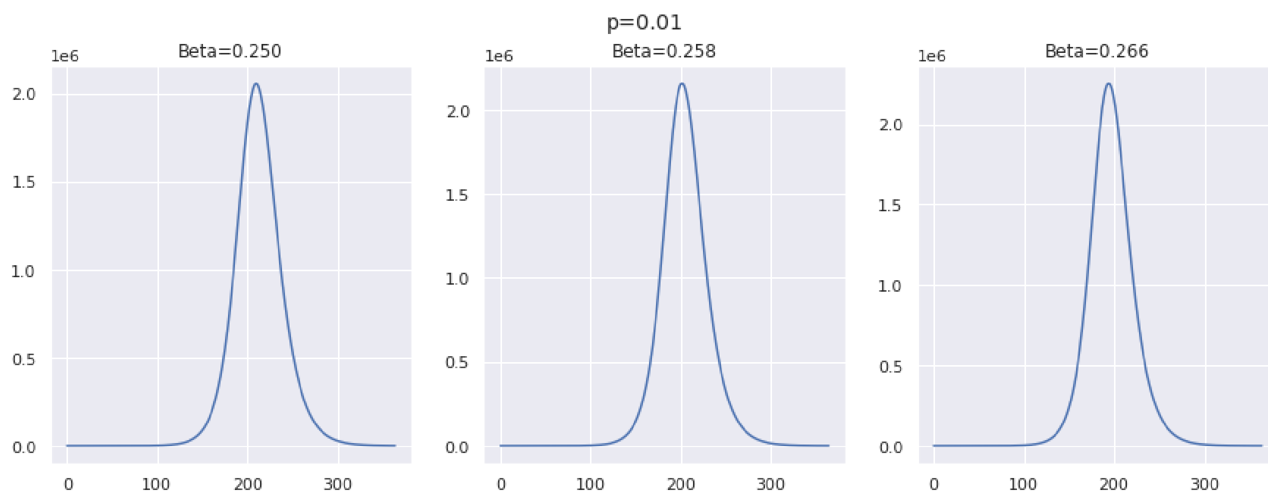

FIG. 5: Infected individuals for time t, $0 \leq t \leq 540$ for $p = 0.01$

2. Effect of Intervention

India has announced a 21 day complete shutdown starting March 25, 2020. For our calculations, we assume that this shutdown will reduce the infection rate to 75%, 50%

or 25% for a time period of 30 days or one month and 180 days or six months. Possibly due to the low number of tests conducted across India, the current identification rate is low. We set p to 0.01 as this closely matches available data for the next set of calculations.

For a 75% reduction in the infection rate and one month of intervention, the estimated peak is pushed from 229 to 241. A 12 day delay. Six months of intervention, the estimated peak is pushed from 229 to 295. A 66 day delay.

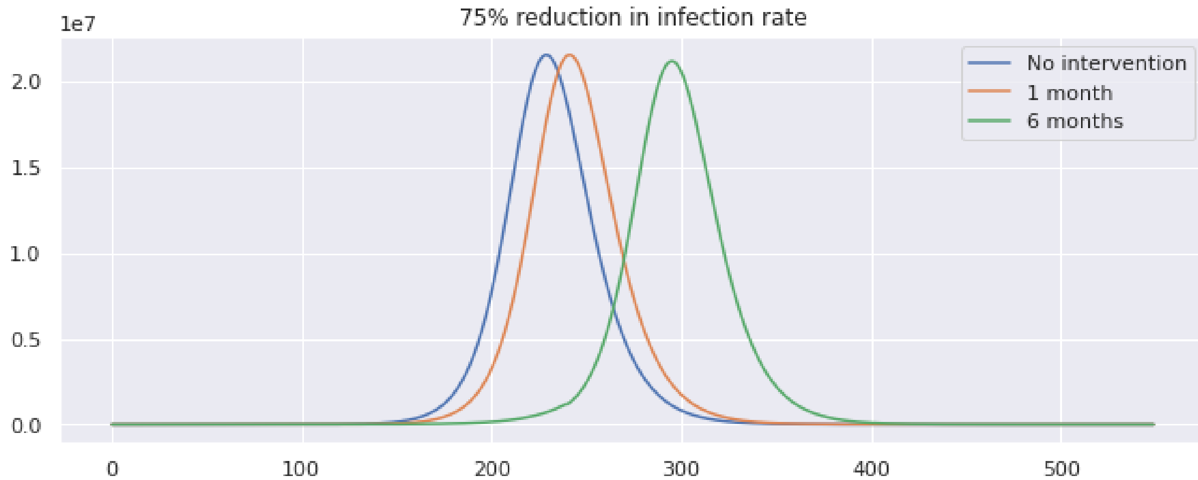

FIG. 6: Variation in $Y(t)$ for time t, $0 \leq t \leq 540$ with no intervention, 1 month of intervention and 6 months of intervention with assumption of $\beta = 0.75 \times 0.258$

For a 50% reduction in the infection rate and one month of intervention, the estimated peak is pushed from 229 to 253. A 24 day delay. Six months of intervention, the estimated peak is pushed from 229 to 370. A 141 day delay.

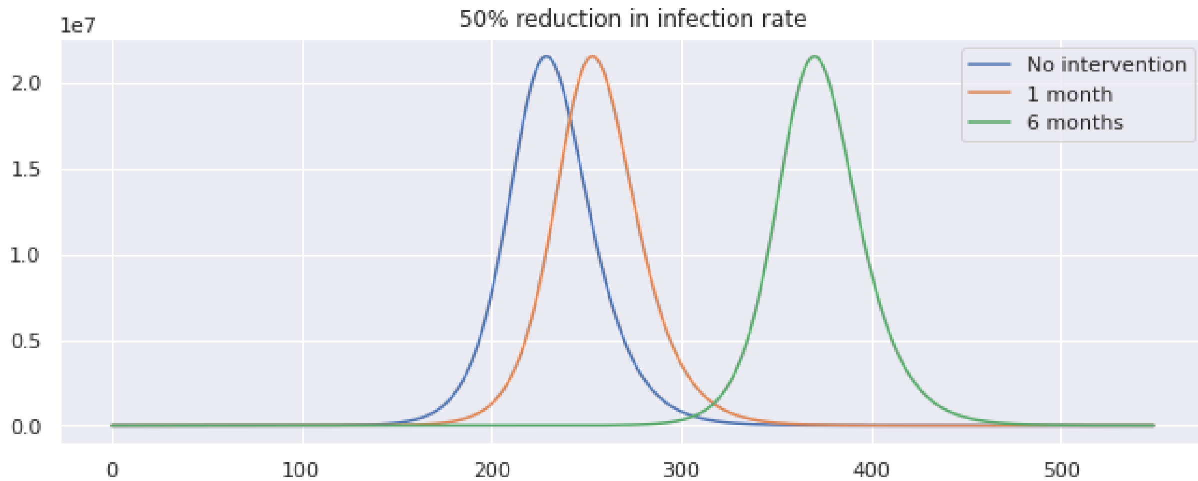

FIG. 7: Variation in $Y(t)$ for time t, $0 \leq t \leq 540$ with no intervention, 1 month of intervention and 6 months of intervention with assumption of $\beta = 0.50 \times 0.258$

For a 25% reduction in the infection rate and one month of intervention, the estimated peak is pushed from 229 to 268. A 39 day delay. Six months of intervention, the estimated peak is pushed from 229 to 461. A 232 day delay.

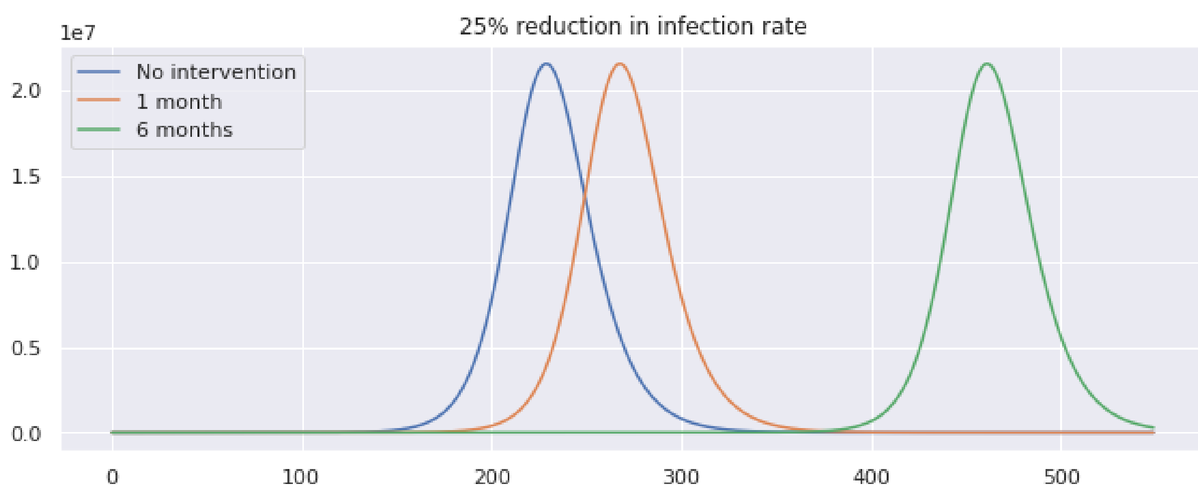

FIG. 8: Variation in $Y(t)$ for time t, $0 \leq t \leq 540$ with no intervention, 1 month of intervention and 6 months of intervention with assumption of $\beta = 0.25 \times 0.258$

TABLE II: Summary of change in estimated peak. Units are days. The 1 month and 6 months delays are the delays in the estimated peaks after respective intervention periods. The dates indicated are the revised dates of the estimated peaks.

Infection Rate Drop	1 Month Delay	Date	6 Months Delay	Date
75%	12	September 20	66	November 13
50%	24	October 2	141	January 27
25%	39	October 17	461	April 28

3. Total Infected

Upon integrating Y over time t, we obtain an approximate of the total number of infected individuals over 540 days. With no intervention, the approximate total number of infected individuals will be 1.2087×10^9 which is 90% of the total population of India. With one month of intervention and 75% reduction in the infection rate, the approximate total number of infected individuals will be the same. But with 75% reduction in the infection rate, the approximate total number of infected individuals will be 1.2079×10^9 which means approximately 739,260 individuals will not be infected. The change in these numbers for a higher reduction in infection rate does not significantly change the approximate total number of infected individuals (a reduction of 155 infected individuals). It only pushes the approximate epidemic peak to a later date.

IV. DISCUSSION

By applying the SEIR compartmentalization model, it is clear that the epidemic peak can easily reach summer. This prediction is sensitive to changes in the behavior of not only the virus but of people and their practice of social distancing. The WHO has issued similar statements reiterating the importance of practicing social distancing to ensure that national health care systems are not strained. Moreover, since India lacks a robust and healthcare system, it may be ill equipped to deal with a large number of cases immediately. Delaying the peak will ensure adequate medical equipment and personnel for all her citizens.

REFERENCES