Mathematical Modelling

MINI PROJECT – MODELING OF COVID19

Team Members:

Roshan Raj Ramesh (rrr13)

SanjayKumar Thiyagarajan (skt18)

Tarun Shekdar (ts386)

Table of Contents

Introduction

The Novel corona virus originated at Wuhan city in November 2019 [1], China took only three months to spread across the globe. The world has never been the same since then. On March 11, 2020, the WHO (World health organization) announced n-covid19 as a global pandemic. This led to various problems in the country's economy, work, education, and our daily life. People are restricted to get out of the house and a country wide lockdown was imposed by the government for most of the countries in the world to ensure the spread of n-covid19 is under control. People and their behaviour towards restrictions is not the same across the countries and the government has a role to ensure the restrictions are being followed properly. For example, China had only one wave in this pandemic episode. This might be a reason for the people of China and the government's strict measure towards the pandemic and lockdown.

Generally, SIR models are suggested and modelled towards the virus cases to predict the further spread of the n-covid19. This might help the government to impose the lockdown and to ensure the economy of the country stay balance during the pandemic period. Traditional SIR models do not include factors such as psychological state of the people which is one of the main factors to consider as discussed above.

Chapter 1

Selecting the countries

COVID 19 has been disrupting the normal life since December 2019. Analysis is made for following countries.

1. India
2. United States of America
3. The United Kingdom

Reason for the above choice is that it would yield interesting results comparing the effect of COVID 19 across the three countries as each country has their own vaccine. Also, the population of all the three countries varies a lot. So, it would be a metric to test how well the chosen model is modelling the cumulative cases for these countries.

Defining the start of each country

For the analysis purpose, the day after the cumulative cases hit hundred is taken for India, United States of America, and United Kingdom. Once the cumulative case hits the hundred marks, the spread of the covid19 becomes faster in all these countries.

Normalization

Normalization is done so that the new metric normalized cumulative case address the fact that each country has different population. This has been taken care by using normalized cumulative case for the analyses further in this study.

$$Normalized\ cumulative\ case = \frac{Cumulative\ case\ of\ the\ Country}{Population\ of\ the\ Country}$$

Plotting cumulative data for each country

Cumulative case plot

Following are the cumulative case plot for each country in the order of India, United Kingdom, United States of America.

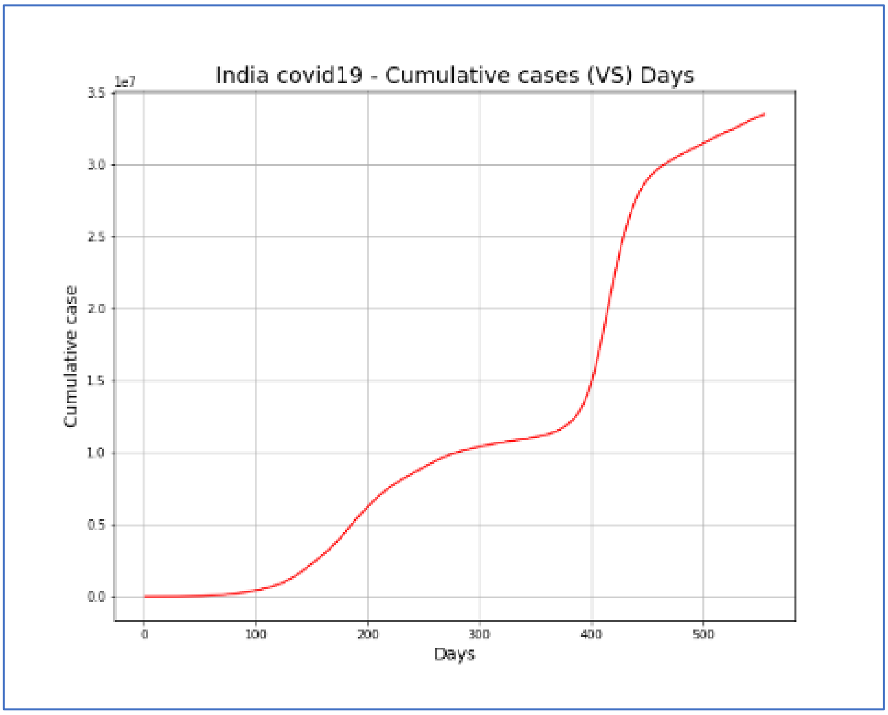

Figure 1: India cumulative case plot

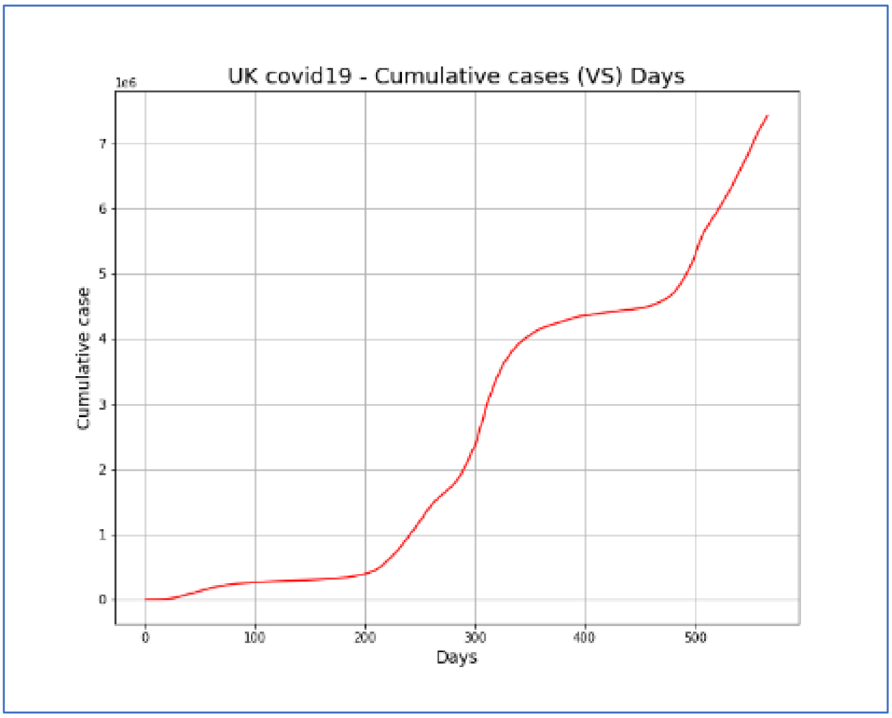

Figure 2: United Kingdom cumulative case plot

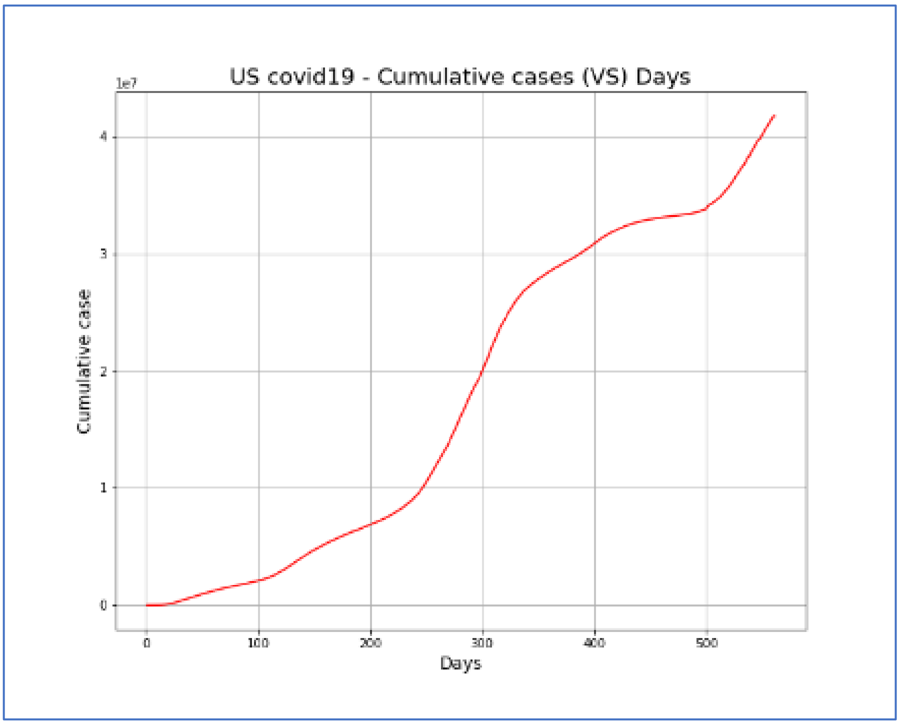

Figure 3: United States cumulative case plot

Normalized cumulative case plot

Following are the Normalized cumulative case plot for each country in the order of India, United Kingdom, United States of America. We can observe that the cumulative case lies between the zero to one and this metric can be compared with other country.

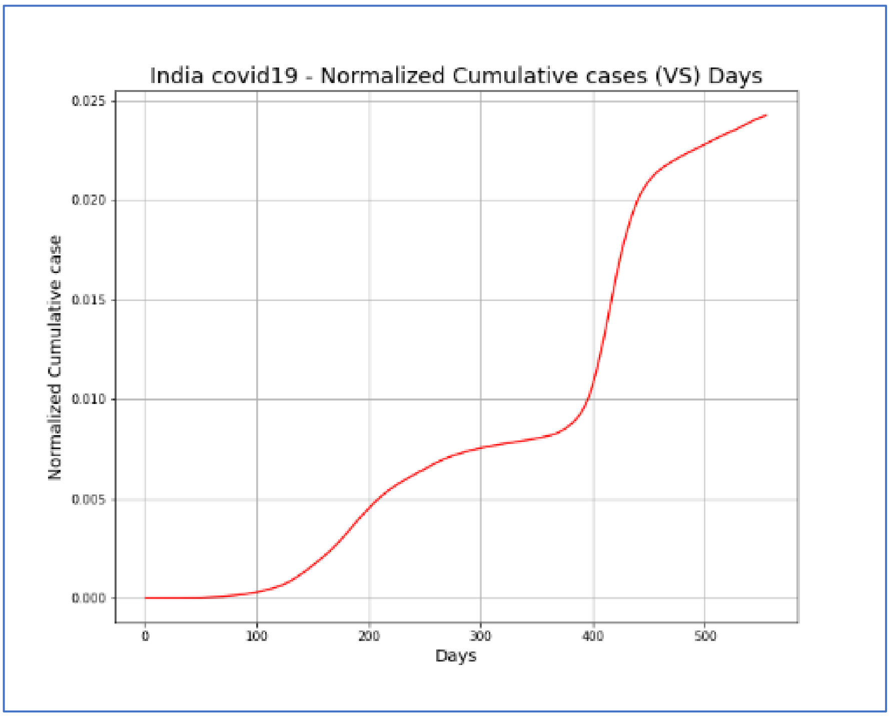

Figure 4: India Normalized cumulative case plot

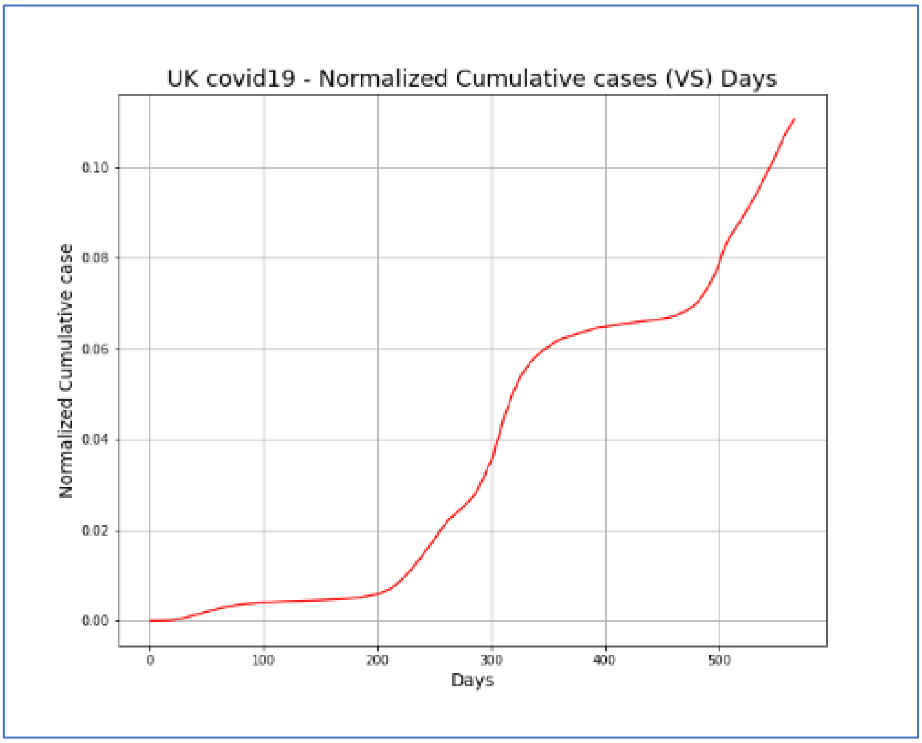

Figure 5: United Kingdom Normalized cumulative case plot

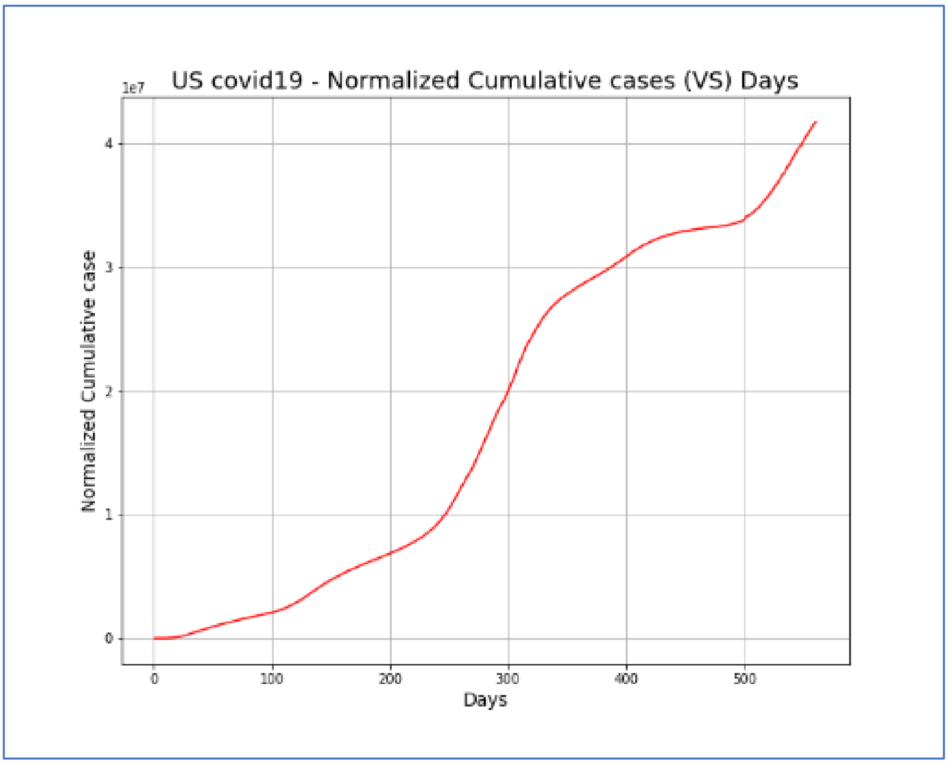

Figure 6: United States Normalized cumulative case plot

Wave analysis

Straight line fragments on logarithm graph (UK):

Let us consider logarithmic graph of United Kingdom. 5 fragments with good straight-line approximation of logarithmic data for United Kingdom. We can recognise three waves: the first from 1 to 202 (end of the first red line), the second from 202 to 480, and the third from 480 to the end of data.

For United Kingdom we have graph presented in Figure 7. We can see the first wave from 1 to 202, the second wave from 202 to 480 and the third wave from 480 to the end. However, slightly more careful analysis of shape shows that for one wave the first part (exponential grows) should have greater slope than saturation part. This means that point at 202 day is not end of wave. The end of the first wave is point 305. The second wave is on interval from 305 to the end of data. It is not clear which fragment should be used as exponential grows in the first wave:

1. From 1 to 30.
2. From 31 to 202.
3. From 1 to 202.

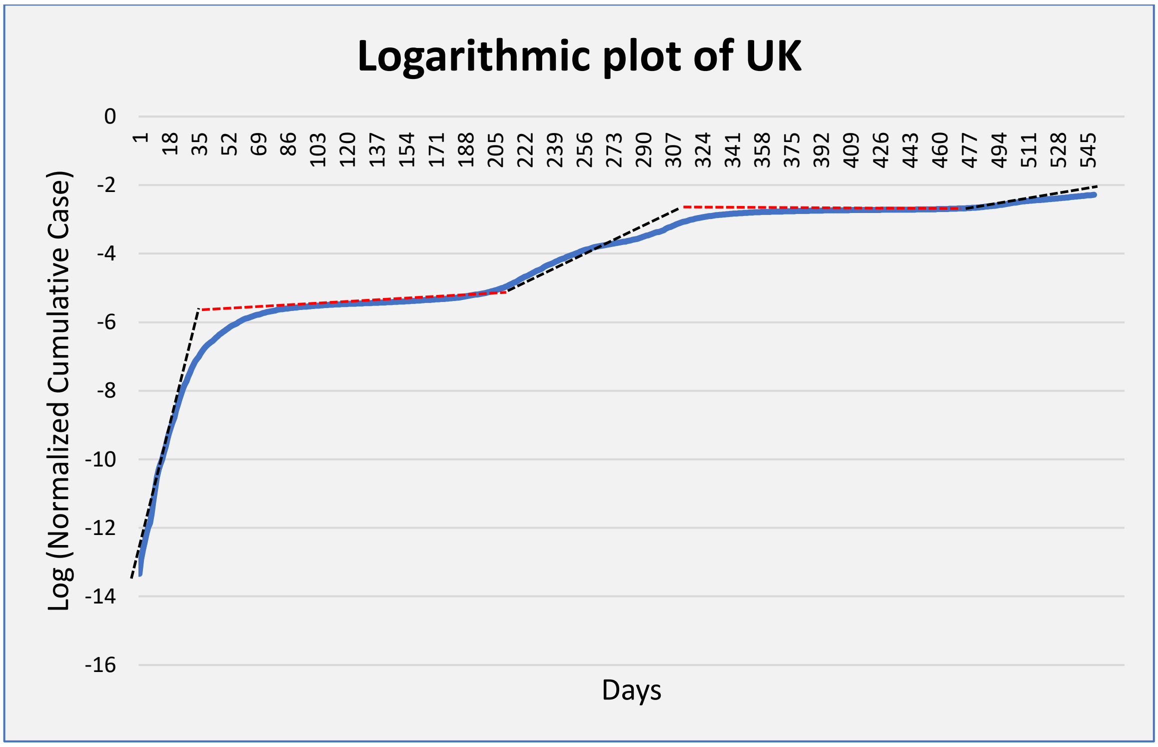

Figure 7: UK Logarithmic plot for wave analysis

Straight line fragments on logarithm graph (US):

Let us consider logarithmic graph of United States. 5 fragments with good straight-line approximation of logarithmic data for United States. We can recognise three waves: the first from 1 to 165 (end of the first red line), the second from 165 to 350, and the third from 350 to the end of data.

However, slightly more careful analysis of shape shows that for one wave the first part (exponential grows) should have greater slope than saturation part. It is not clear which fragment should be used as exponential grows in the first wave:

1. From 1 to 40.
2. From 40 to 165.
3. From 1 to 165.

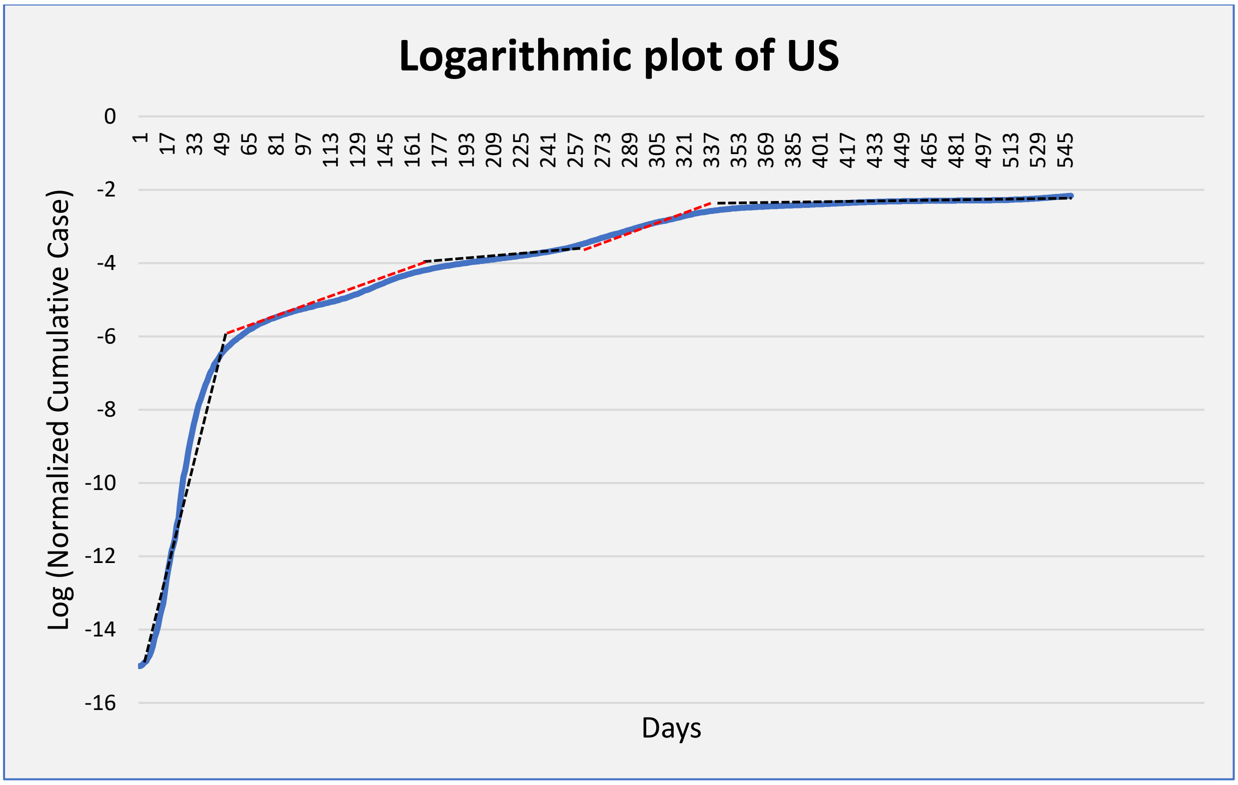

Figure 8: US Logarithmic plot for wave analysis

Inflection points for India: Inflection point is point where the second derivative change sign: convex function become concave or concave function become convex. For convex function on interval $[a, b]$ we always have

$$f\left(\frac{a+b}{2}\right) < \frac{f(a)+f(b)}{2}$$

and for concave function we have

$$f\left(\frac{a+b}{2}\right) > \frac{f(a)+f(b)}{2}.$$

Let us try to check changes of convexity for 5 points: point i is "convex" if

$$P_i < \frac{P_{i-2}+P_{i-1}+P_i+P_{i+1}+P_{i+2}}{5}$$

As we can conclude, there are many small fluctuations in infection rate and raw data are not appropriate for these purposes. Let us try to smooth data by mean filter:

$$\breve{P}_i = \frac{1}{5}\sum_{j=i-5}^{i+5} P_j.$$

To remove impulse noise, we can use median filter:

$$x_i = median(\{x_{i-5}, \dots, x_{i+5}\}).$$

After double used mean filter, we have graph presented in Figure 9. We can clearly find exponential fragment of the first wave from 1 to 135, the saturation fragment of the first wave is interval from 135 to 271. The second wave starts at day 271 and ends at day 450. India doesn't have a third wave yet.

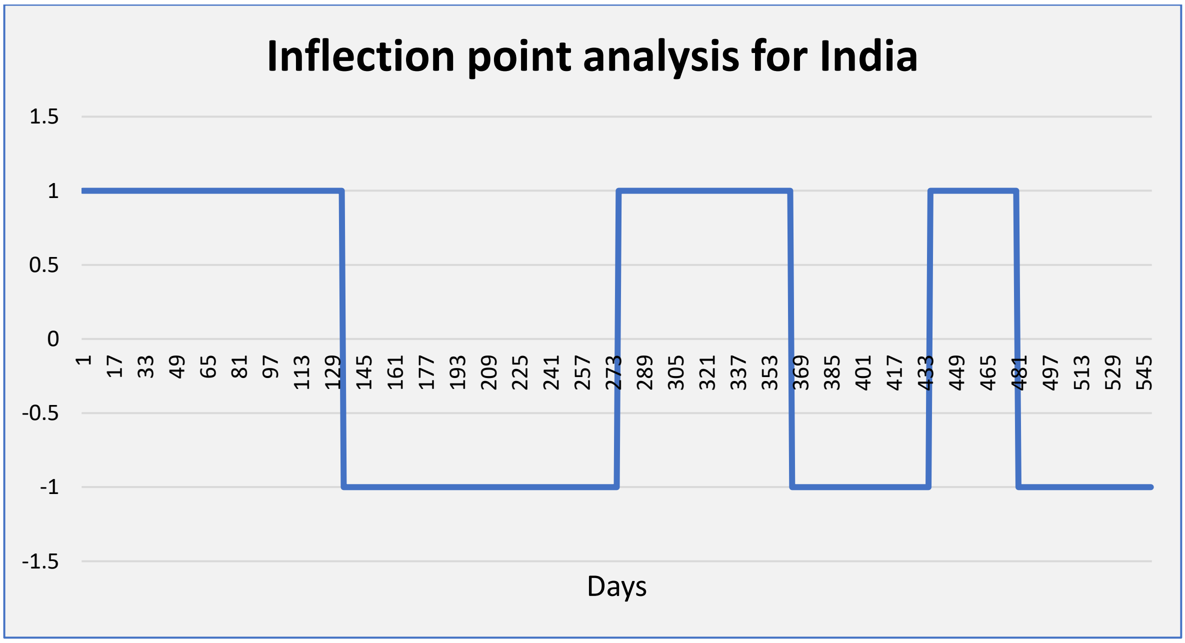

Figure 9: Inflection point analysis for India

The table below summarizes the wave analysis for United States of America, United Kingdom, India. As we can infer, United States of America and United Kingdom has three waves in total. Whereas India doesn't have a third wave and has only two waves. India might expect a third wave in 2022.

Table 1: Wave analysis for US, UK, and India

Wave	**United States**		**United Kingdom**		**India**	
	From	To	From	To	From	To
1	1	165	1	202	1	271
2	165	350	202	480	271	450
3	350	End	480	End		

Statistical analysis

The period of Exponential growth is selected from the wave analysis and taken for further analysis such as applying Exponential and Logistic growth. Coefficient and Intercept is obtained by fitting the exponential grows interval with line approximation (Linear regression). These numerical values are entered in the below table for each country separately.

$$\log P = a + rt \ \ (where\ a\ is\ the\ intercept\ and\ r\ is\ the\ \ coefficent$$

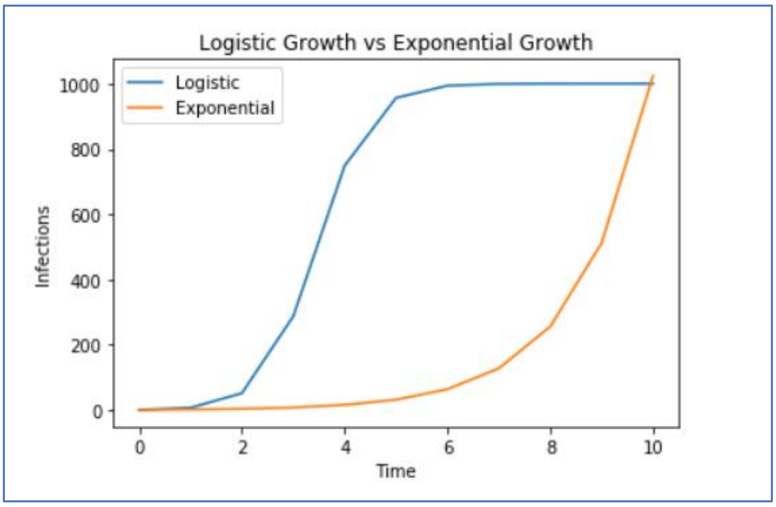

Figure 10: Logistic Growth vs Exponential growth [2]

The figure 10, shows the difference between logistic growth and exponential growth. Note that the exponential growth used by fitting linear regression to the log value of cumulative cases as shown in the above equation.

Table 2: Exponential growth and Logistic Growth for US

Wave	United States		Standard Error		Exponential Growth		Logistic Growth	
	From	To	Exponential Growth	Logistic Growth	Coefficient	Intercept	Coefficient	Intercept
1	1	165	0.0009975	0.0031569	0.0000847	-0.0022404	0.0098923	0.0021313
2	165	350	0.0061884	0.0834125	0.0003700	-0.0574645	0.0087123	0.0576564
3	350	End	0.0017116	0.0239876	0.0001322	0.0376622	0.0092131	0.0356345

Table 3: Exponential growth and Logistic Growth for UK

Wave	United Kingdom		Standard Error		Exponential Growth		Logistic Growth	
	From	To	Exponential Growth	Logistic Growth	Coefficient	Intercept	Coefficient	Intercept
1	1	202	0.0005467	0.0239471	0.0004188	0.0000283	0.0001243	0.0034524
2	202	480	0.0073725	0.0031763	0.0002576	-0.0417199	0.0097241	0.0432666
3	480	End	0.0005755	0.0632122	0.0004757	-0.1590529	0.0032144	0.0134500

Table 4: Exponential growth and Logistic Growth for India

Wave	India		Standard Error		Exponential Growth		Logistic Growth	
	From	To	Exponential Growth	Logistic Growth	Coefficient	Intercept	Coefficient	Intercept
1	1	271	0.0008533	0.0036851	0.0000299	-0.0017622	0.0067841	0.0546567
2	271	450	0.0021513	0.0613242	0.0000806	-0.0182431	0.0033411	0.0132456

The Plot of K(t) for all the countries are presented below. Here t is the days and K(t) is calculated using the equation below. The value of a and r is obtained from the exponential growth interval, and they are used to get the value of function K with respect to the days.

$$K(t) = P(t) * \frac{1 + \exp{(a + rt)}}{\exp{(a + rt)}}$$

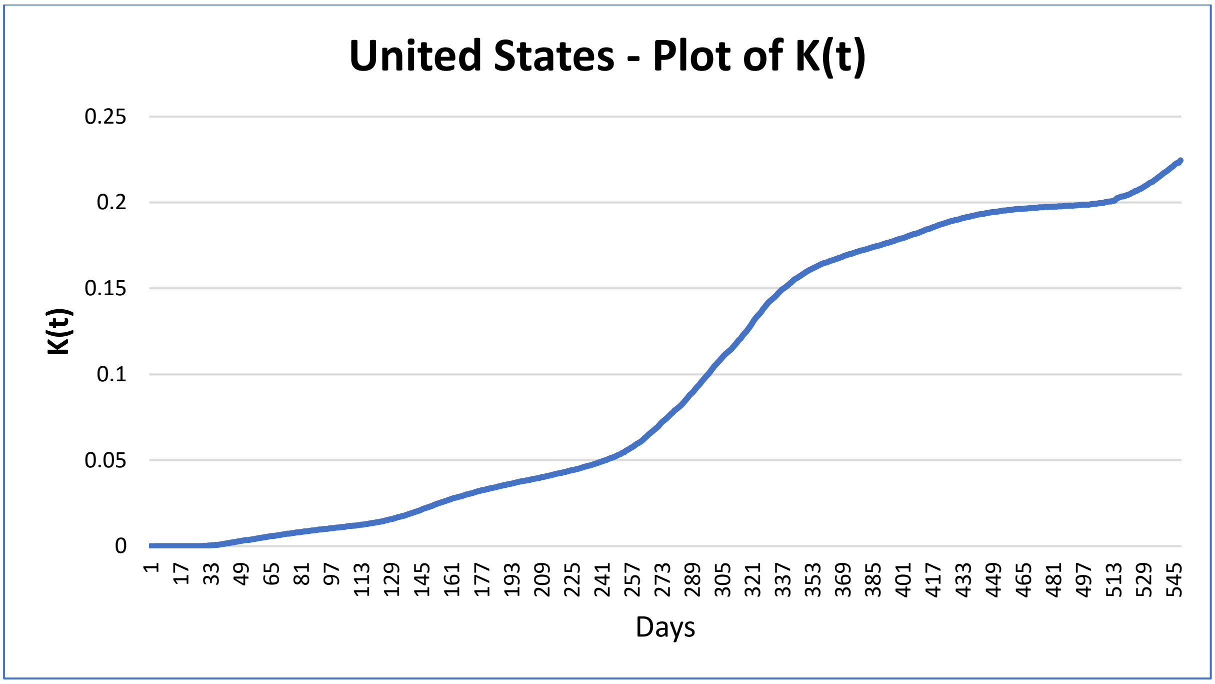

Figure 11: United States K(t) plot

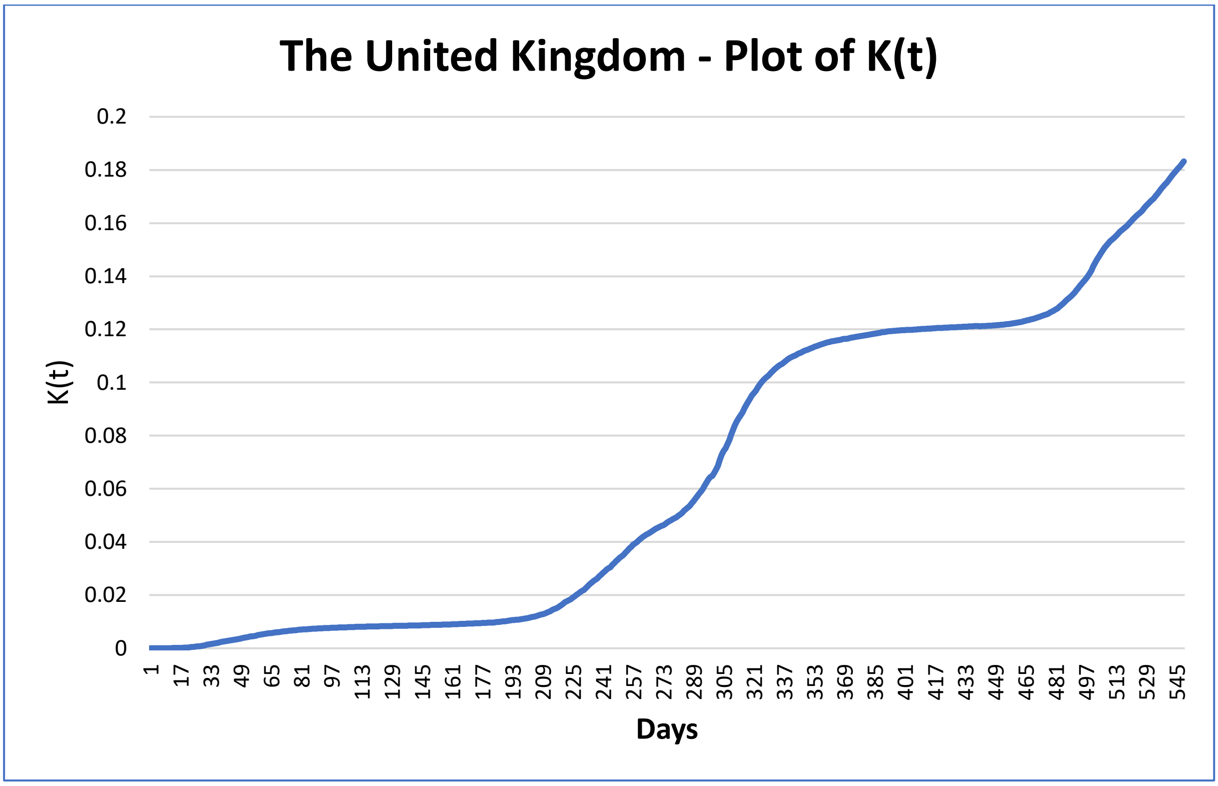

Figure 12: The United Kingdom K(t) plot

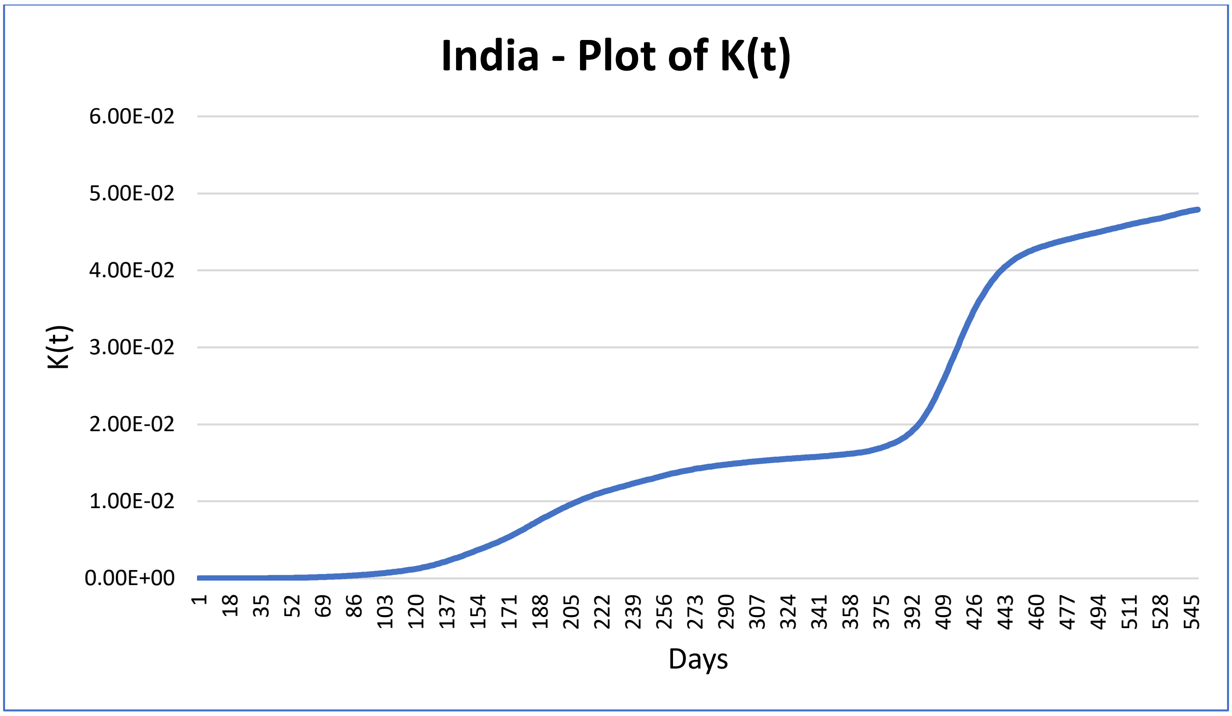

Figure 13: India K(t) plot

The K(t) function's behaviour is like that of normalized cumulative cases curve of the respective countries with days as x axis. Also, the wave behaviour of the countries is preserved in the K(t) plot, and it is reflected in the above visualization. The K(t) function is not constant as observed above in the visualization. This is as expected, because the values a and r obtained from the exponential growth is used to solve the K(t) as mentioned in the above K(t) equation. Thus, reciprocating it into the normalized cumulative cases of the respective countries.

Do all countries have the same number of waves? The United Kingdom and United States of America have three waves in total as discussed above. Whereas India have only two waves in total. These may be because of various factors including non-technical things like psychology of the people and the government measures and restriction to the response of covid19 spread in the respective country. India might experience a third wave later this year or in the beginning of 2022 as many other countries in the world are experiencing third wave of corona virus as the emergence of omicron variant is causing a buss around the world.

Are these waves synchronous? The United Kingdom and United States of America's waves are quite synchronous. Whereas India, has a delay of start in the waves. In general, we can conclude that United Kingdom and United States of America are having very common cumulative case trend and have same number of waves. Reason might be the government response to the preventive measures of covid19 and the people's mindset towards the disease.

Chapter 2

Log plot for countries

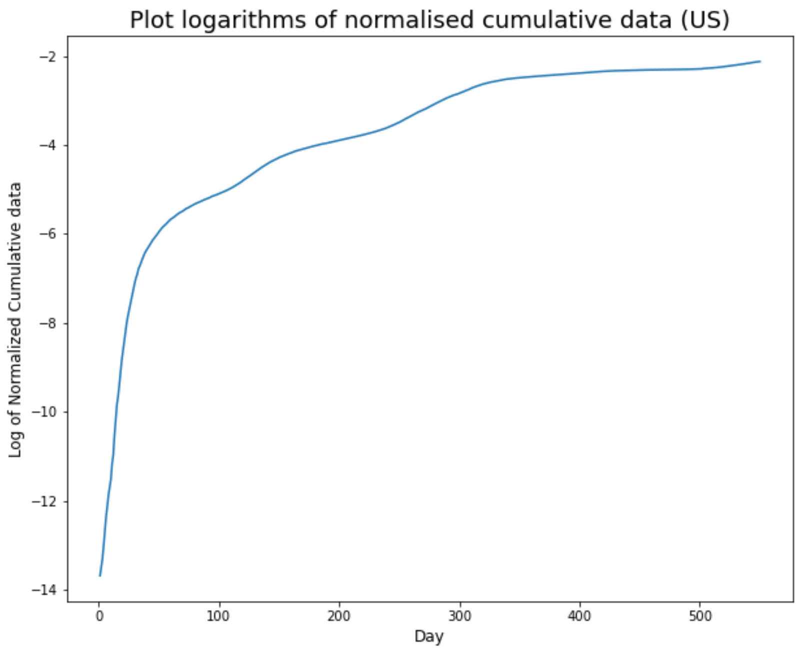

Figure 14: Logarithmic plot of United states

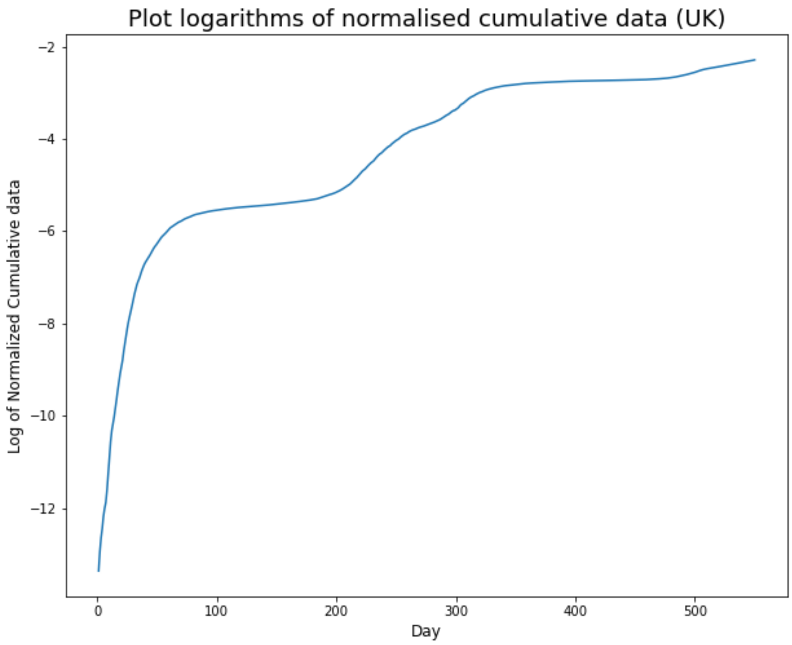

Figure 15: Logarithmic plot of United Kingdom

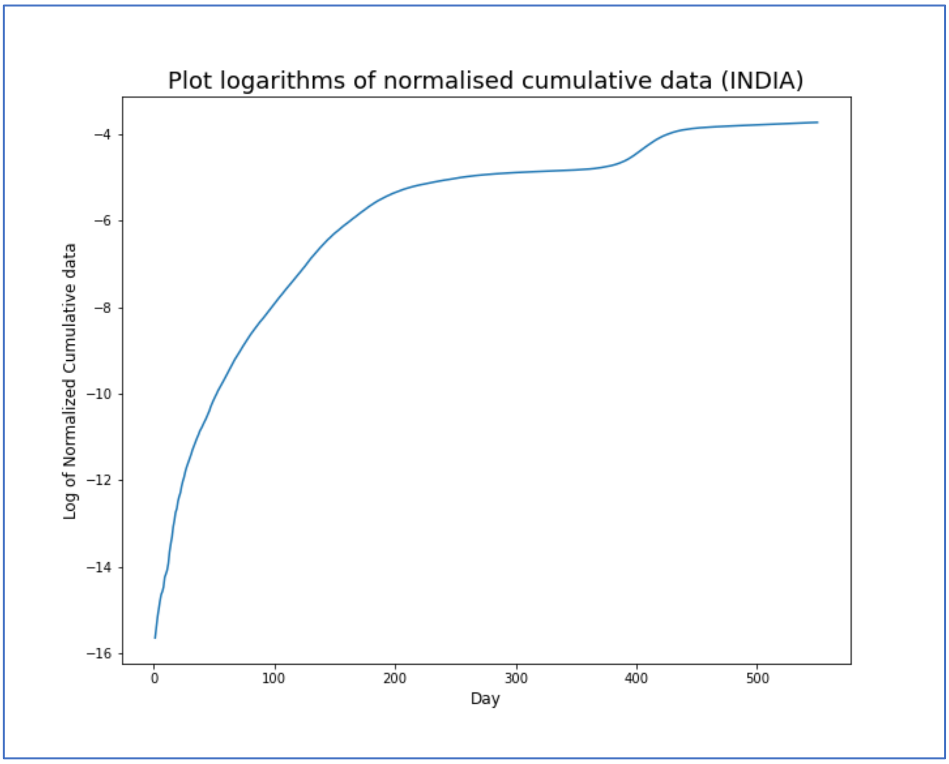

Figure 16: Logarithmic plot of India

Finding an optimum r and c for all the three countries is crucial for implementing SIR model and can be compared with the observed cumulative case of each country. This can be done by selecting an appropriate interval of days and applying exponential grows. For each interval MSE is calculated for the specific interval and for all days since the pandemic start. The interval which has very low MSE interval and MSE (all day) is taken, and their r and c values are chosen for SIR model calculation. In the below tabulation, the selected interval and their corresponding r, c, b, and a is highlighted.

Table 5: Finding optimum r and c for US

UNITED STATES OF AMERICA						
Interval	r	c	b	a	**MSE (Interval)**	**MSE (All Days)**
1 – 15	0.13923	-11.96262	0.1	0.23923	$6.21407*10^{-12}$	Too large value for MSE
1 – 216	0.12854	-9. 94235	0.1	0.22854	0.0001535143	2. 4762083
216 – 310	0.01492	-4.74212	0.1	0.11492	$6.316173*10^{-7}$	0.08613444

Table 6: Finding optimum r and c for UK

UNITED KINGDOM						
Interval	r	c	b	a	**MSE (Interval)**	**MSE (All Days)**
1 – 36	0.18390	-15.98621	0.1	0.28390	18.19800	0.000390421
1 – 57	0.12387	-11.94942	0.1	0.22387	0.17832	$4.40647*10^{-6}$
57 – 380	0.01143	-6.93300	0.1	0.11143	0.01145	$2.38674*10^{-8}$

Table 7: Finding optimum r and c for India

INDIA						
Interval	r	c	b	a	**MSE (Interval)**	**MSE (All Days)**
1 – 160	0.03192	-18.86234	0.1	0.13192	$2.1407*10^{-13}$	Too large value for MSE
1 – 315	0.02662	-11.43363	0.1	0.12662	0.00007453	20.8924469
165 – 315	0.00612	-6.63578	0.1	0.10612	$3.1228*10^{-7}$	0.00001165

Defining SIR model

States: S - susceptible, I - Infected/infector, R - Removed (recovered or dead)

System of transitions:

$$S \ \rightarrow I \ \rightarrow R$$

We consider constant population size: S + I + R = N = const.

Further we use dimensionless variables:

$$\{S, I, R\} = \left\{\frac{S}{N}, \frac{I}{N}, \frac{R}{N}\right\}$$

$$S + I + R = 1$$

Becoming infected depends on contact between S and I. Therefore, assume that intensity of transitions S→I is aI and the flux S→I is aSI, a = const. Assume that the intensity of recovering (or death) is constant (b) and, therefore, the flux I→ R is bI.

SIR differential equation is given below,

$$\frac{dS}{dt} = -aSI$$

$$\frac{dI}{dt} = aSI - bI$$

$$\frac{dR}{dt} = bI$$

Criteria

1. When S close to 1, exponent is r ≈ a-b, according to SIR model
2. Taking b = 0.1 (time is measured in days; b = 1/τ, where τ is, approximately, the time of virus spreading by an infected person; we take here τ ≈ 10 days)
3. Thus, we know b = 0.1 and a = r +b
4. Where r is the coefficient obtained from the model: log P ≈ c + rt

Initialization of S_0, I_0, R_0 for all three countries

$$R(t) \approx P(t-10)$$

$$While\ t = 0, \qquad R(0) \approx P(-10)$$

$$I(0) = Day(1) - R(0)$$

$$S(0) = 1 - I(0) - R(0)$$

Using the above equation, the initialization of S_0, I_0, R_0 for all three countries are calculated and documented in the below table,

$$R(0)US = \frac{74}{Population} = \frac{74}{330000000} = 2.24242 * 10^{-7}$$

$$R(0)UK = \frac{10}{Population} = \frac{10}{68000000} = 1.47058 * 10^{-7}$$

$$R(0)UK = \frac{29}{Population} = \frac{29}{1380000000} = 2.10144 * 10^{-8}$$

Table 8: SIR Initialization without normalization

Country	Population	S_0	I_0	R_0
United States	330000000	100	27	74
United Kingdom	68000000	100	98	10
India	1380000000	100	78	29

S_0 is taken as 100 for all the country as we are considering the start of pandemic as the hundredth mark of cumulative cases in respective countries. Integrating system of ODE and solving it we obtain S, I, R for each country. By using the function 'odeint' from the library 'scipy.integrate', the differential equation is solved and the results are stored in a array. The predicted number of normalized cumulative value is calculated by the following equation

$$P = 1 - S$$

Using the predicted normalized cumulative cases obtained from SIR model, the following comparison is between observed normalized cumulative cases and predicted normalized cumulative cases.

Table 9: SIR Model MSE with Observed cumulative case

SIR Model	
Country	**MSE**
United States	7. 7098304
United Kingdom	5. 8908234
India	0. 4762083

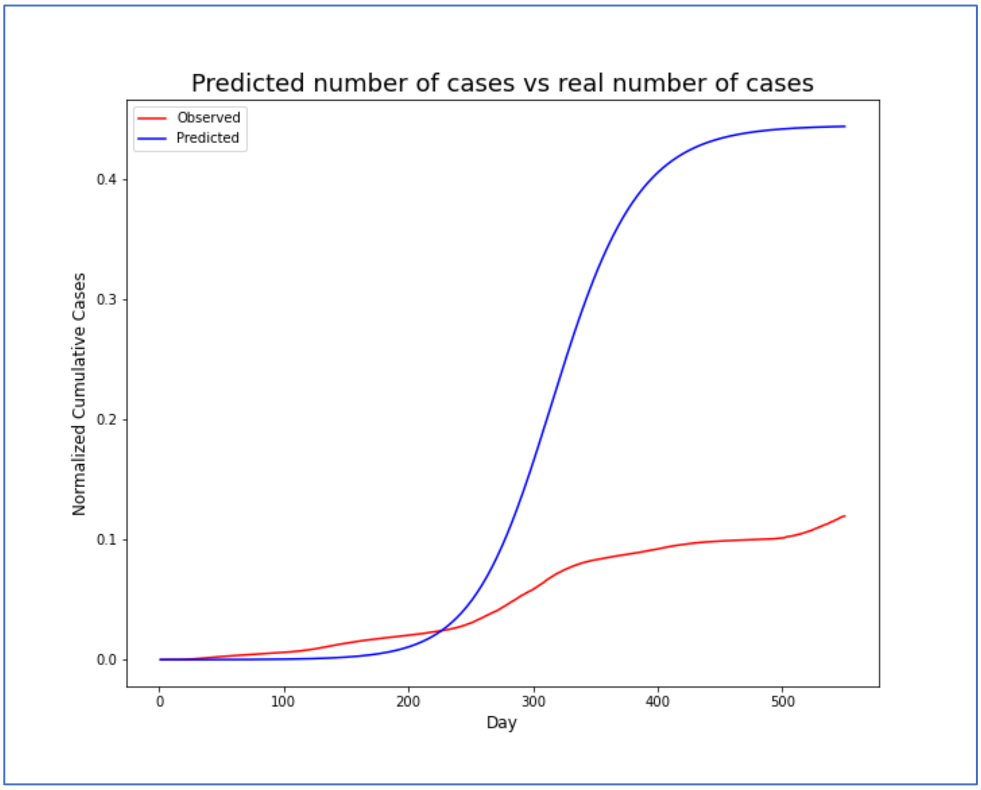

Figure 17: US SIR prediction (blue) and observed cumulative case (red)

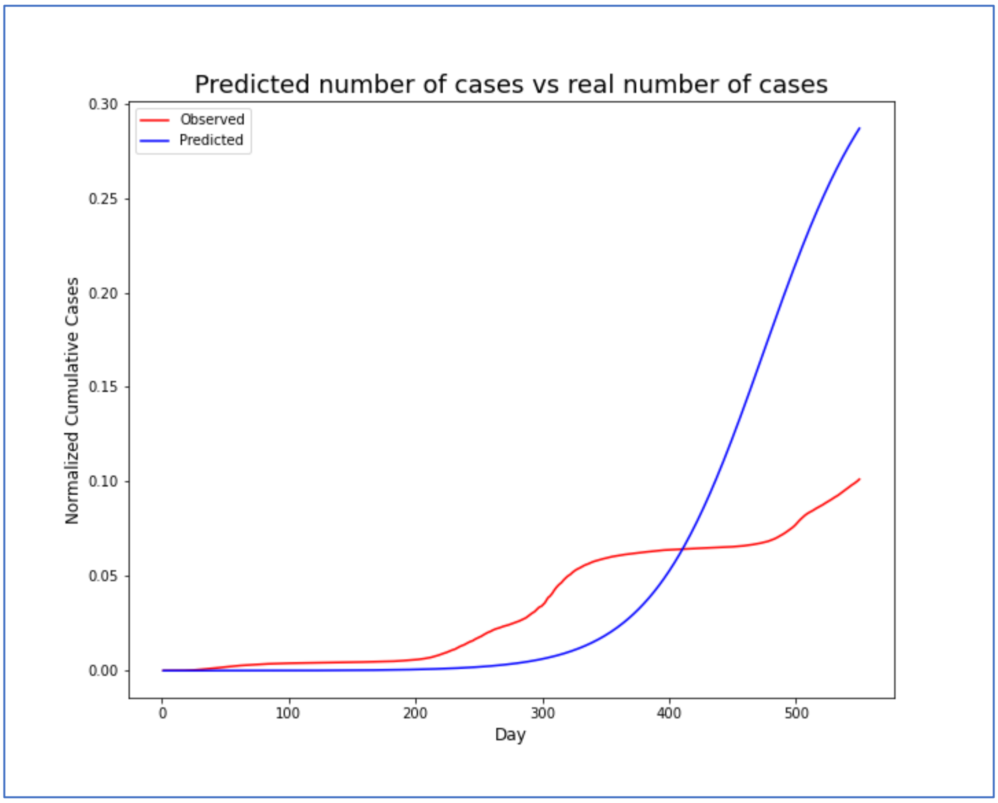

Figure 18: UK prediction (blue) and observed cumulative case (red)

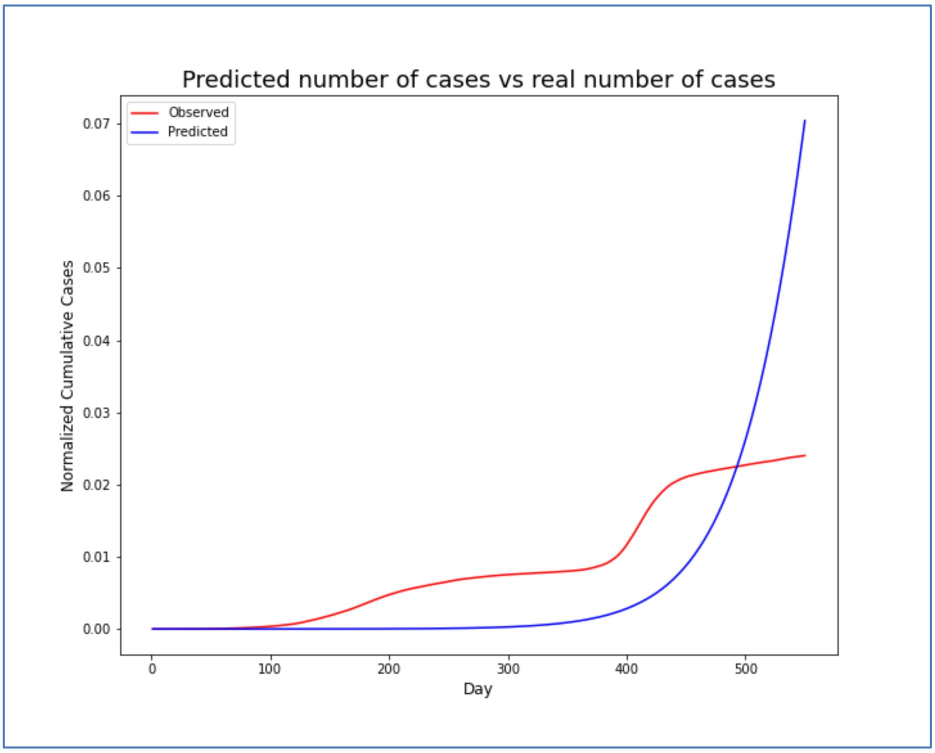

Figure 19: India prediction (blue) and observed cumulative case (red)

The plots obtained from the SIR model is pictured with the observed cumulative cases. The MSE between the observed and prediction shows that the curve can be fitted more appropriately by playing the SIR parameters.

Playing with parameters

Finding an optimum $\mathrm{I(I(0))}$ is crucial to find an appropriate fitting curve to the observed cumulative cases. Here, using trail and error technique we are arbitrarily mixing the value of $\mathrm{I(I(0))}$ and trying to visualize it to see how the curve fits the observed cumulative cases. As we can see that by varying the $\mathrm{I(I(0))}$ we are unable to fit the cumulative cases curve perfectly, but there is a slight improvement of finding a perfect curve approximation for the observed cumulative cases. In addition to the $\mathrm{I(I(0))}$, varying the 'a' value could yield a better curve approximation for all the three countries. Recall, $a = r + b$, where r is the coefficient obtained from the exponential growth fit and b is the reciprocal incubation period of covid which is taken as 0.1. Following are the visualization of observed vs predicted cumulative cases curve by varying $I(I(0))$ and 'a'.

Varying $I(I(0))$:

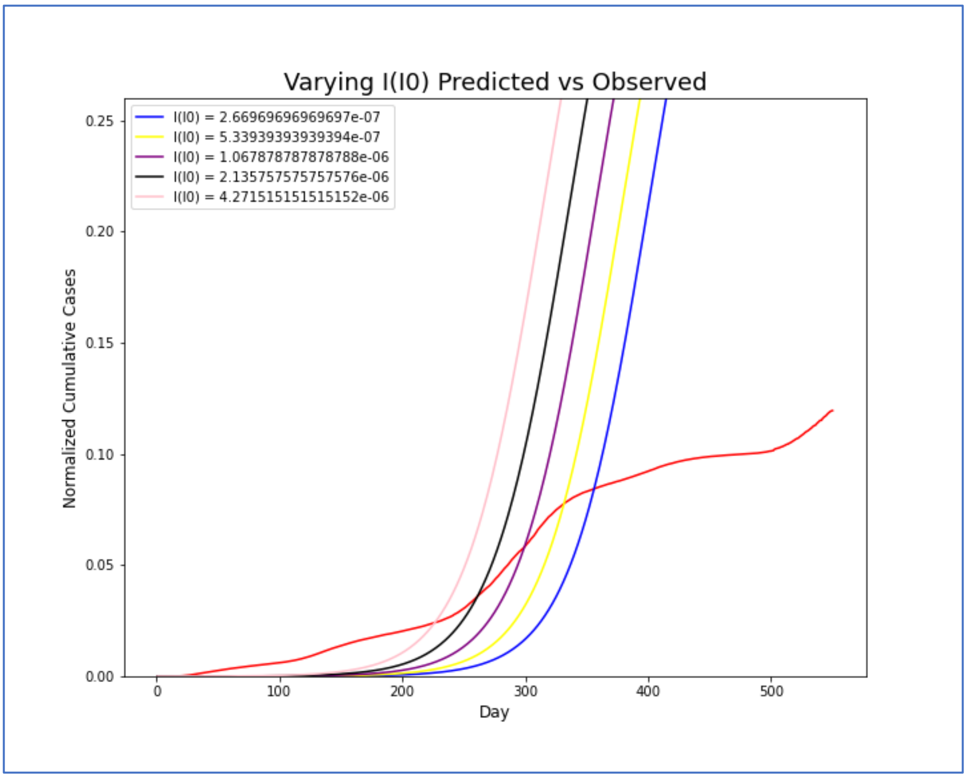

Figure 20: US predicted vs observed by varying I_0

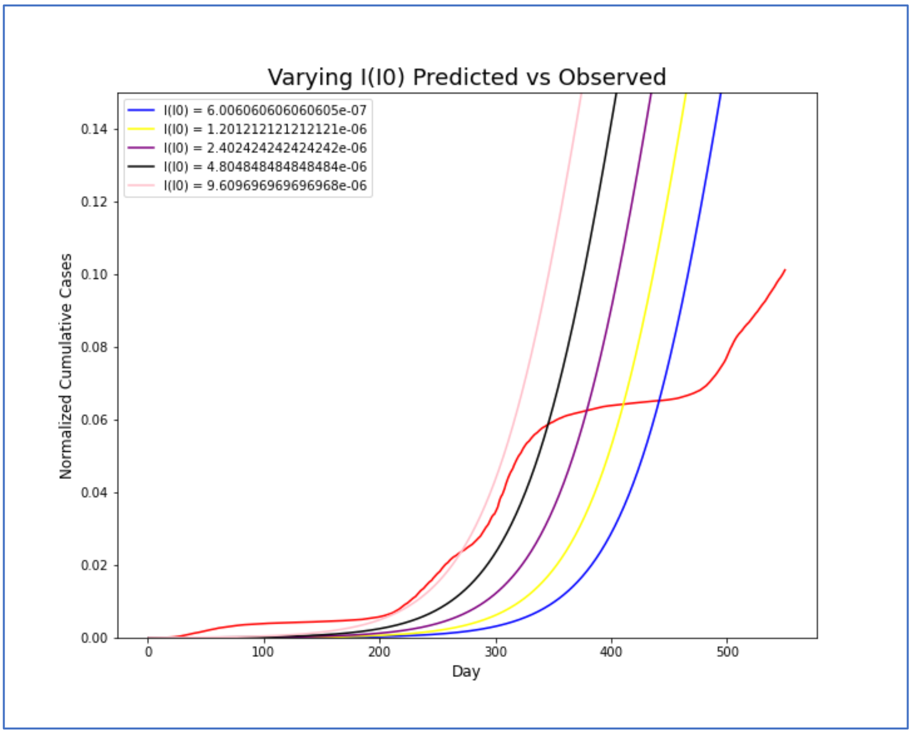

Figure 21: UK predicted vs observed by varying I_0

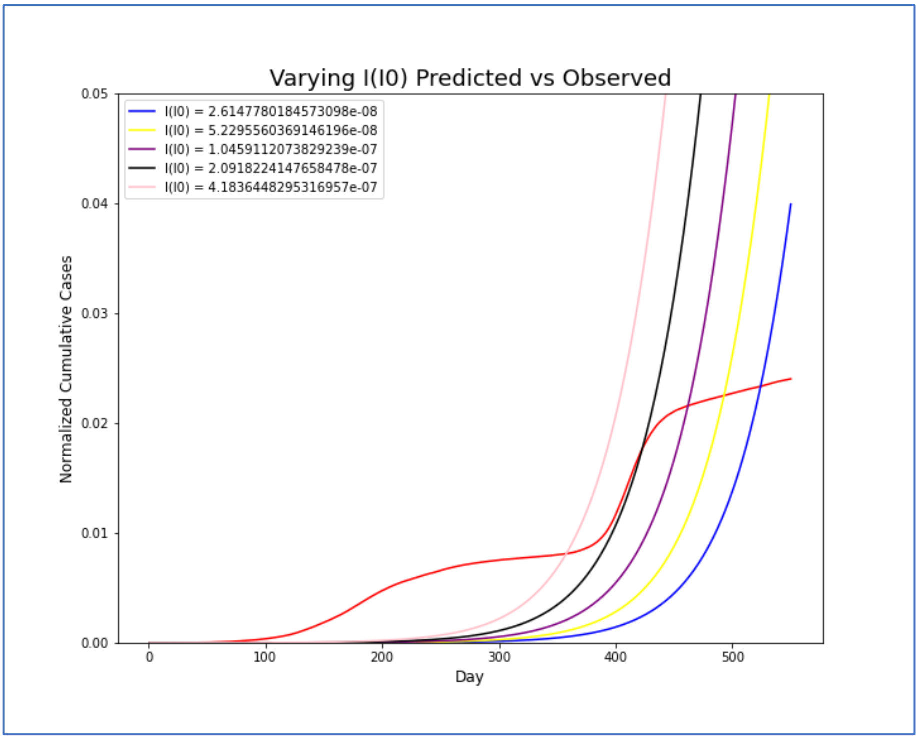

Figure 22: IN predicted vs observed by varying I_0

Above visualizations are obtained by varying the I(I(0)) of the previously obtained SIR predicted cumulative cases and they are colour coded as mentioned in the visualization's legend.

Varying $I(I(0))$ *and* a:

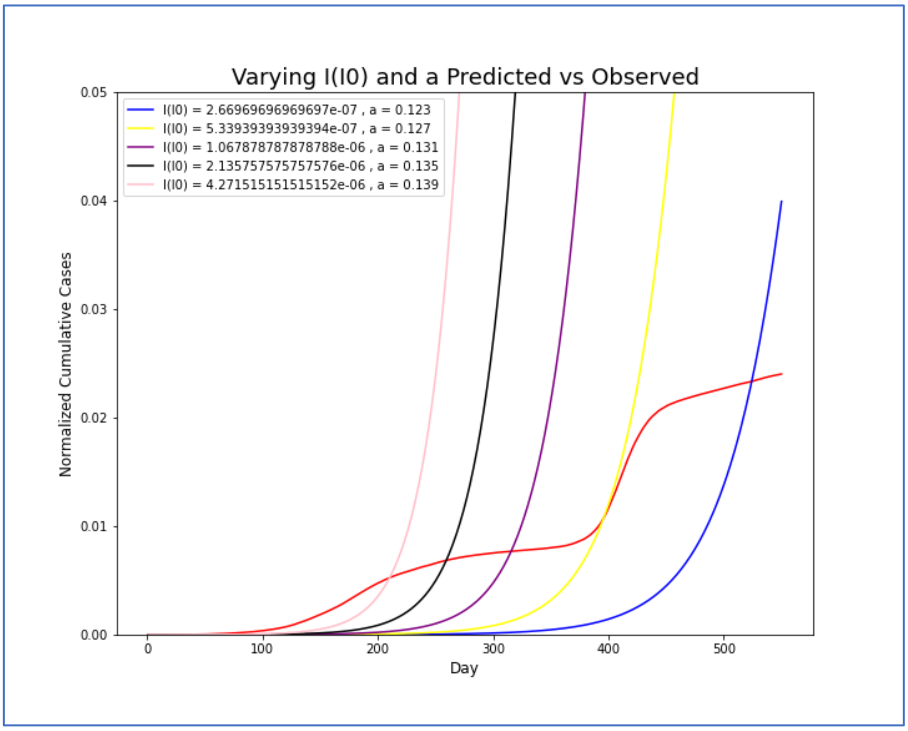

Figure 23: US predicted vs observed by varying I_0 and a

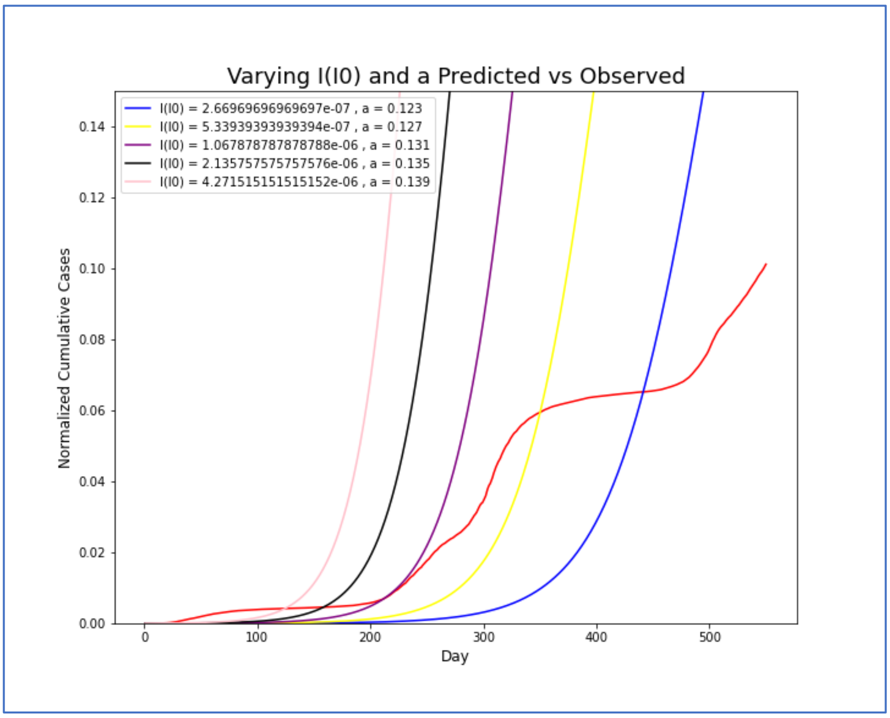

Figure 24: UK predicted vs observed by varying I_0 and a

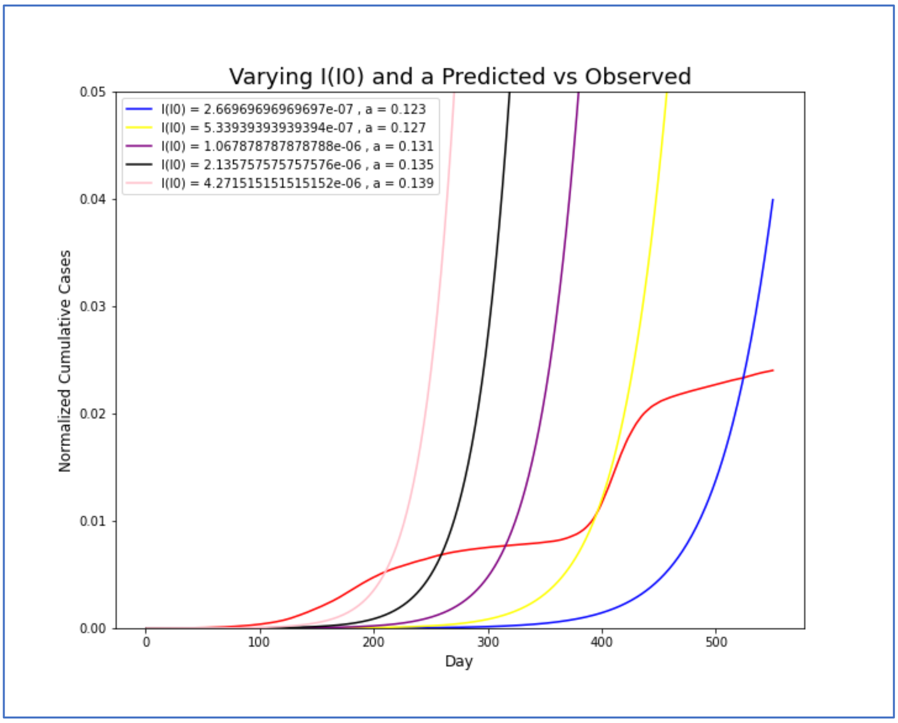

Figure 25: IN predicted vs observed by varying I_0 and a

Above visualizations are obtained by varying the I(I(0)) and a of the previously obtained SIR predicted cumulative cases and they are colour coded as mentioned in the visualization's legend. These results are comparatively better than that of the results obtained from varying only the I(I(0)). Recall that, the I in SIR model denotes number of cumulative people infected by covid19.

SIR model is compared against the cumulative case for all three countries. There are positive and negative effects of modelling SIR. SIR model assumes that the population remains constant for the analysing period from day 1 to day 500. In practical, the population varies every day and there should be parameter which compensates this instability of population. SIR model is very generic and is not capable of modelling the multi wave nature of covid19. Mutation of covid19 makes the disease to have multi wave impact across all the countries in the world and the spread of the disease varies a lot when compared to the original virus originated at Wuhan city, China in November 2019. Also, the incubation period varies among the mutated variants of covid19, thus the b value used in the SIR model varies a lot resulting in a different predicted curve.

SIR model is simple and easy to model with most of the countries affected by covid19 because it considers only three basic components such as susceptible, inflected and recovered. But it fails to address the psychological factors associated with human beings and the government's preventive measure to control the spread of covid19 in their country.

Chapter 3

Introduction

So far in the study, the normalized cumulative cases are modelled with exponential grows in the chapter 1 and with standard SIR model in the chapter 2. In this chapter, introducing more reasonable factors and parameters to the SIR model. Previously used differential equations of SIR model are as follows.

$$\frac{dS}{dt} = -aSI$$

$$\frac{dI}{dt} = aSI - bI$$

$$\frac{dR}{dt} = bI$$

Introduction of general adaptation syndrome (GAS) to the classical SIR model might be an appropriate factor to model the normalized cumulative cases more efficiently. GAS is a three-stage process that gives reason to the physiological and psychological changes happen in humans under stress. Those three phases are: Alarm, Resistance, and Exhaustion. To summarize these influencing factors into the standard SIR model, the S (susceptible) is divided into these four divisions.

S_{ign} – Ignorant people that do not know anything worrying about the epidemic

S_{al} – people in “Alarm phase”

S_{res} – people in “Resistance” state, with very rational and save behaviour

S_{exh} – people in “Exhaustion” state. They are tired of the epidemic, behave unsafe and do not react on alarm stimuli.

$$S = S_{ign} + S_{al} + S_{res} + S_{exh}$$

At the initial state S(0) = S_{ign}(0) and all other components have zero population,

$S_{al}(0) = S_{res}(0) = S_{exh}(0) = 0$

The alarm phase partially included in S_{ign} and S_{res}. The reduced S equation is,

$$S = S_{ign} + S_{res} + S_{exh}$$

The transition of the S variable is,

$$S_{ign} \rightarrow S_{res} \rightarrow S_{exh} \rightarrow S_{ign}$$

With the mass action law formalism, SIR with these additions is:

$$S_{ign} + I \rightarrow 2I; \quad I \rightarrow R$$

Stress reaction:

$$S_{ign} + I \rightarrow S_{res} + I; \ S_{res} \rightarrow S_{exh}; \ S_{exh} \rightarrow S_{ign}$$

Kinetic equations

List of substances

S_{ign} is susceptible healthy ignorant population. For conservation law is b_1.

S_{res} is susceptible healthy resistant population. For conservation law is b_2.

S_{exh} is susceptible healthy exhausted population. For conservation law is b_3.

I is infector population. For conservation law is b_4.

R is removed (recovered and immune or dead) population. For conservation law is b_5.

Table 10: Kinetic equation table

Reactions	Reaction rate	Stoichiometric vector
$\boldsymbol{S_{ign} + I \rightarrow 2I}$	$r_1 = aS_{ign}I$	$(-1, 0, 0, 1, 0)^T$
$\boldsymbol{S_{ign} + I \rightarrow S_{res} + I}$	$r_2 = k_2S_{ign}I$	$(-1, 1, 0, 0, 0)^T$
$\boldsymbol{S_{res} \rightarrow S_{exh}}$	$r_3 = k_3S_{res}$	$(0, -1, 1, 0, 0)^T$
$\boldsymbol{S_{exh} + I \rightarrow 2I}$	$r_4 = aS_{exh}I$	$(0, 0, -1, 1, 0)^T$
$\boldsymbol{I \rightarrow R}$	$r_5 = bI$	$(0, 0, 0, -1, 1)^T$
$\boldsymbol{S_{exh} \rightarrow S_{ign}}$	$r_6 = k_6S_{exh}$	$(1, 0, -1, 0, 0)^T$

ODE:

$$\frac{dc}{dt} = \sum_{\rho=1}^{6} \gamma_\rho r_\rho$$

$$\frac{dc}{dt} = \begin{matrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{matrix} r_1 + \begin{matrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{matrix} r_2 + \begin{matrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{matrix} r_3 + \begin{matrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \end{matrix} r_4 + \begin{matrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{matrix} r_5 + \begin{matrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \end{matrix} r_6$$

$$\frac{dc}{dt} = \begin{matrix} -aS_{ign}I & -k_2S_{ign}I & 0 & 0 & 0 & k_6S_{exh} \\ 0 & k_2S_{ign}I & -k_3S_{res} & 0 & 0 & 0 \\ 0 & 0 & k_3S_{res} & -aS_{exh}I & 0 & k_6S_{exh} \\ aS_{ign}I & 0 & 0 & aS_{exh}I & -bI & 0 \\ 0 & 0 & 0 & 0 & bI & 0 \end{matrix}$$

$$\frac{dS_{ign}}{dt} = -aS_{ign}I - k_2S_{ign}I + k_6S_{exh},$$

$$\frac{dS_{res}}{dt} = k_2S_{ign}I - k_3S_{res},$$

$$\frac{dS_{exh}}{dt} = k_3S_{res} - aS_{exh}I + k_6S_{exh},$$

$$\frac{I}{dt} = aS_{ign}I + aS_{exh}I - bI$$

$$\frac{R}{dt} = bI$$

$$Given, \qquad K_3 = \frac{1}{50}$$

$$K_6 = \frac{1}{100}$$

$$K_2 = 1$$

After substituting the K_2, K_3, K_6

$$\frac{dS_{ign}}{dt} = -aS_{ign}I - S_{ign}I + 0.01 * S_{exh},$$

$$\frac{dS_{res}}{dt} = S_{ign}I - 0.02 * S_{res},$$

$$\frac{dS_{exh}}{dt} = 0.02 * S_{res} - aS_{exh}I + 0.01 * S_{exh},$$

$$\frac{I}{dt} = aS_{ign}I + aS_{exh}I - bI$$

$$\frac{R}{dt} = bI$$

From the chapter two, getting a and b values. Substituting it in the above equation,

For India, b=0.1 and a=0.10612

$$\frac{dS_{ign}}{dt} = (-0.10612) * S_{ign}I - S_{ign}I + 0.01 * S_{exh},$$

$$\frac{dS_{res}}{dt} = S_{ign}I - 0.02 * S_{res},$$

$$\frac{dS_{exh}}{dt} = 0.02 * S_{res} - (-0.10612) * S_{exh}I + 0.01 * S_{exh},$$

$$\frac{I}{dt} = (0.10612) * S_{ign}I + (0.10612) * S_{exh}I - (0.1) * I$$

$$\frac{R}{dt} = (0.1) * I$$

For United Kingdom, b=0.1 and a=0.11143

$$\frac{dS_{ign}}{dt} = (-0.11143) * S_{ign}I - S_{ign}I + 0.01 * S_{exh},$$

$$\frac{dS_{res}}{dt} = S_{ign}I - 0.02 * S_{res},$$

$$\frac{dS_{exh}}{dt} = 0.02 * S_{res} - (-0.11143) * S_{exh}I + 0.01 * S_{exh},$$

$$\frac{I}{dt} = (0.11143) * S_{ign}I + (0.11143) * S_{exh}I - (0.1) * I$$

$$\frac{R}{dt} = (0.1) * I$$

For United States of America,

$$\frac{dS_{ign}}{dt} = (-0.11492\,) * S_{ign}I - S_{ign}I + 0.01 * S_{exh},$$

$$\frac{dS_{res}}{dt} = S_{ign}I - 0.02 * S_{res},$$

$$\frac{dS_{exh}}{dt} = 0.02 * S_{res} - (-0.11492\,) * S_{exh}I + 0.01 * S_{exh},$$

$$\frac{I}{dt} = (0.11492\,) * S_{ign}I + (0.11492\,) * S_{exh}I - (0.1) * I$$

$$\frac{R}{dt} = (0.1) * I$$

Integrating system of ODE mentioned above, by using the function 'odeint' from the library 'scipy.integrate', the differential equation is solved and the results are stored in a array. Comparing this differential equation to standard SIR model, it addresses the psychological state of mind of people by adding alarm, resistance, and exhaustion phase to the SIR model. SIR components remain the same for the second draft and this draft. We saw in the previous study that SIR model is incapable of modelling multi wave dynamics of covid 19. The reason is that SIR model fails to address psychological issue with the pandemic. Playing with I_0 and a does appeal to give a better MSE between observed and predicted cumulative case of covid for each country in the previous. Here we are extending it by modifying the K_3 and K_6 in the differential equation to get the best fit between predicted and observed cumulative case of covid.

Following are the findings derived by solving the ODE and playing with the plays. Results are presented in plots and tabulation for each country.

Optimized MSE while playing around the coefficients, the results are documented and presented in the tabulation for each country. By tweaking the K_2, K_6, I_0 and the a, better approximated model has been found and it is plotted below for each country. By involving additional parameters to SIR model has made the MSE to go further down and a better approximation has been found.

India

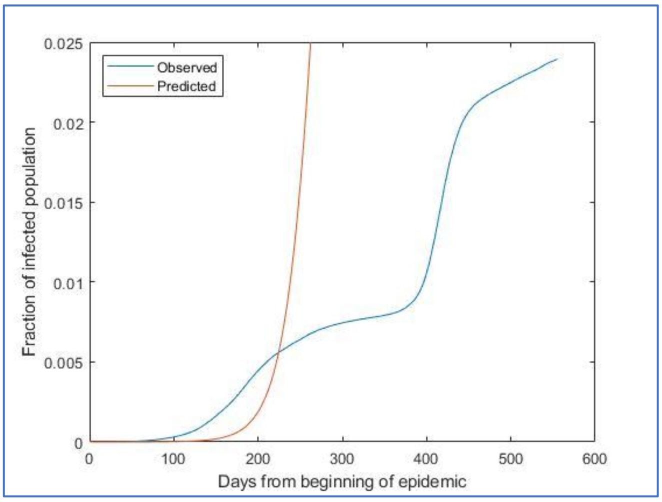

Figure 26: Observed vs Predicted for India

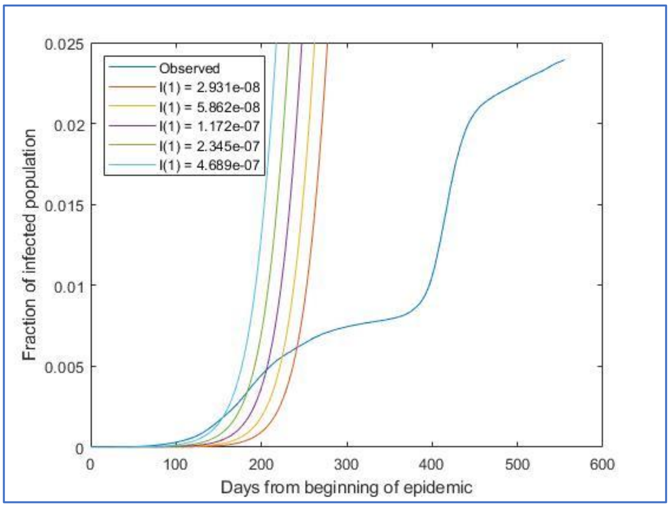

Figure 27: Observed vs Predicted for India by varying I(1)

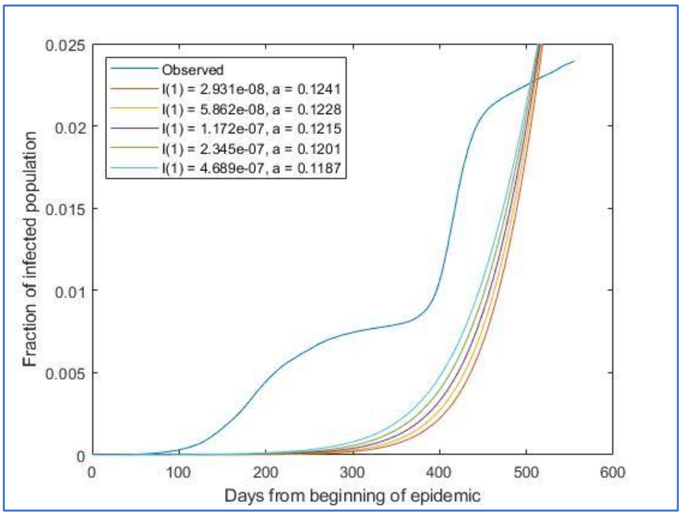

Figure 28: Observed vs Predicted for India by varying I(1) and a

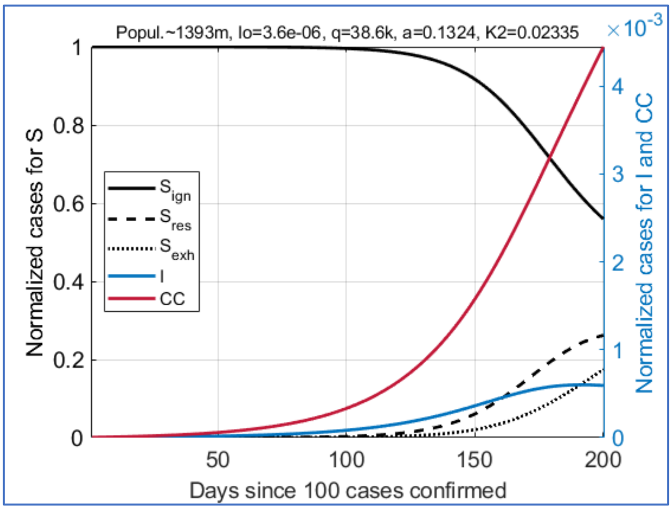

Figure 29: S_{ign}, S_{res}, S_{exh}, Infected (I) and cumulative case (CC) plot for India

S_{ign}, S_{res}, S_{exh}, I (Infected), and CC (Normalized cumulative cases) is visualized for the first 200 days since the start of pandemic (from 100 cases confirmed) for India.

Table 11: Parameters of India (SIR improvised)

INDIA	
Parameters	Value
a	0.1324
b	0.1000
I_0	$3.6*10^{-6}$
c	-13.4497
K2	0.0233
S_0	1.0000
Population	1393000000

United Kingdom

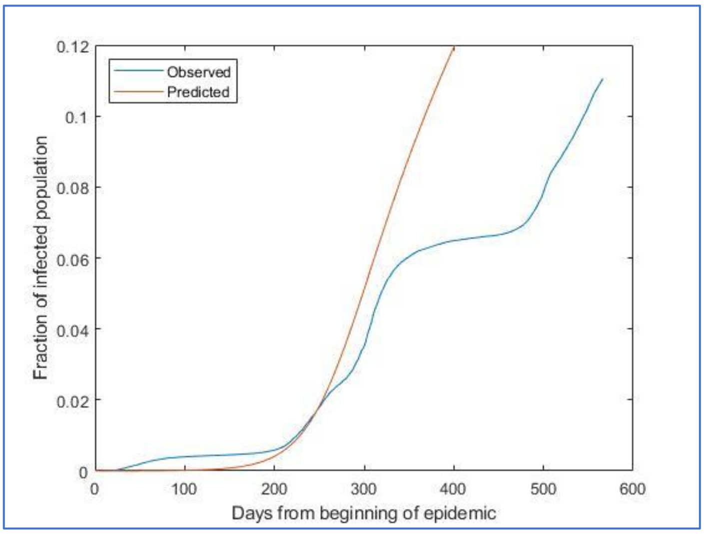

Figure 30: Observed vs Predicted for United Kingdom

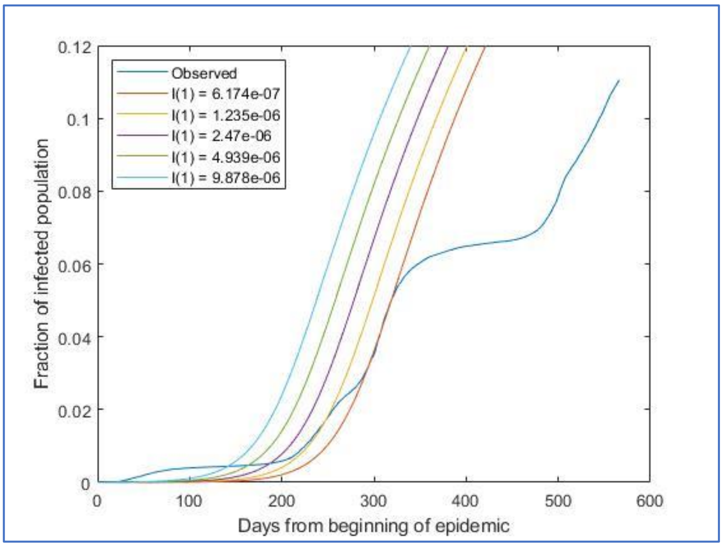

Figure 31: Observed vs Predicted for United Kingdom by varying I(1)

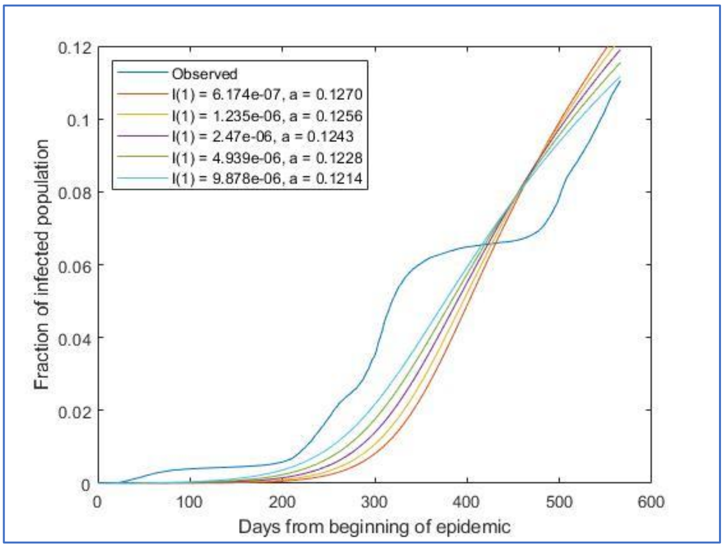

Figure 32: Observed vs Predicted for United Kingdom by varying I(1) and a

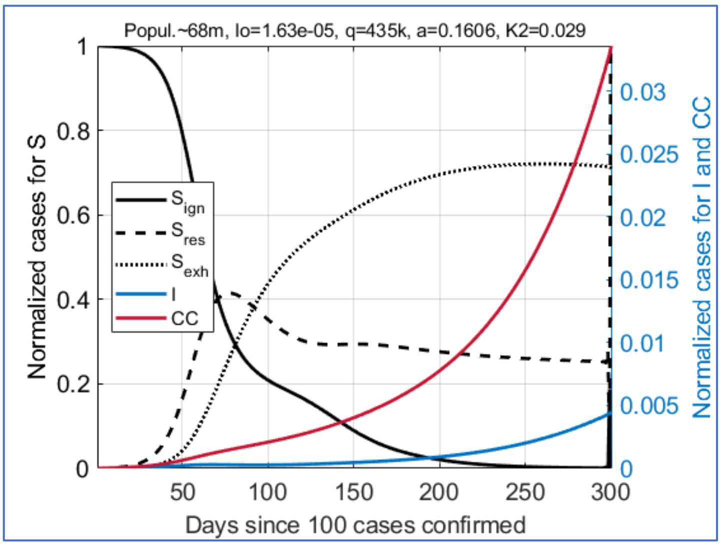

Figure 33: Sign, Sres, Sexh, Infected (I) and cumulative case (CC) plot for United Kingdom

Table 12: Parameters of United Kingdom (SIR improvised)

UNITED KINGDOM	
Parameters	**Value**
a	0.16060
b	0.10000
I_0	$1.63*10^{-5}$
c	-13.44970
K2	0.02900
S_0	1.00000
Population	68000000

United States of America

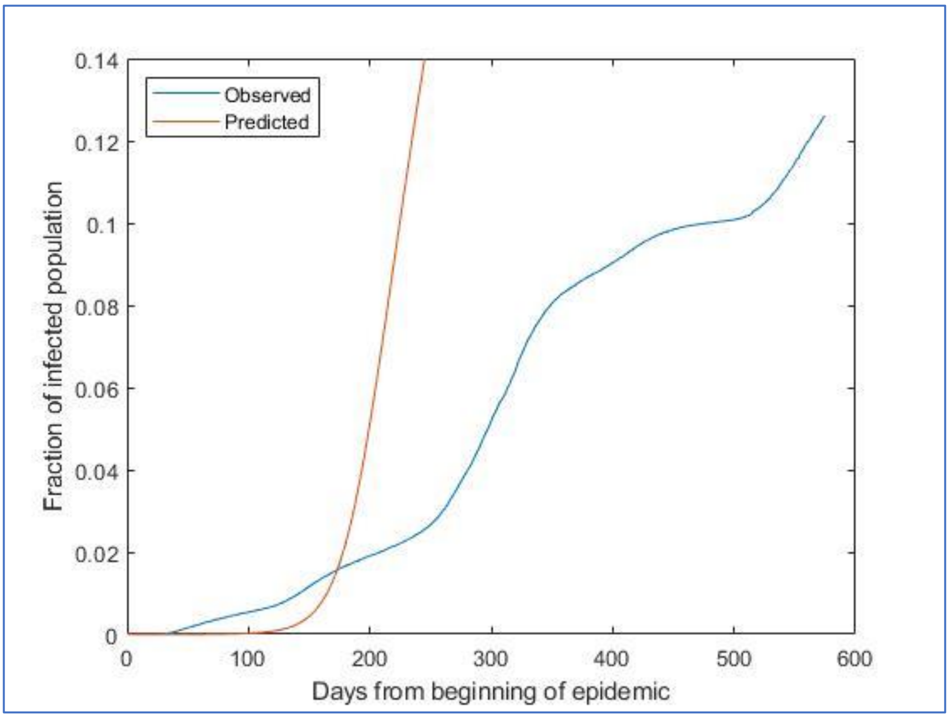

Figure 34: Observed vs Predicted for United States

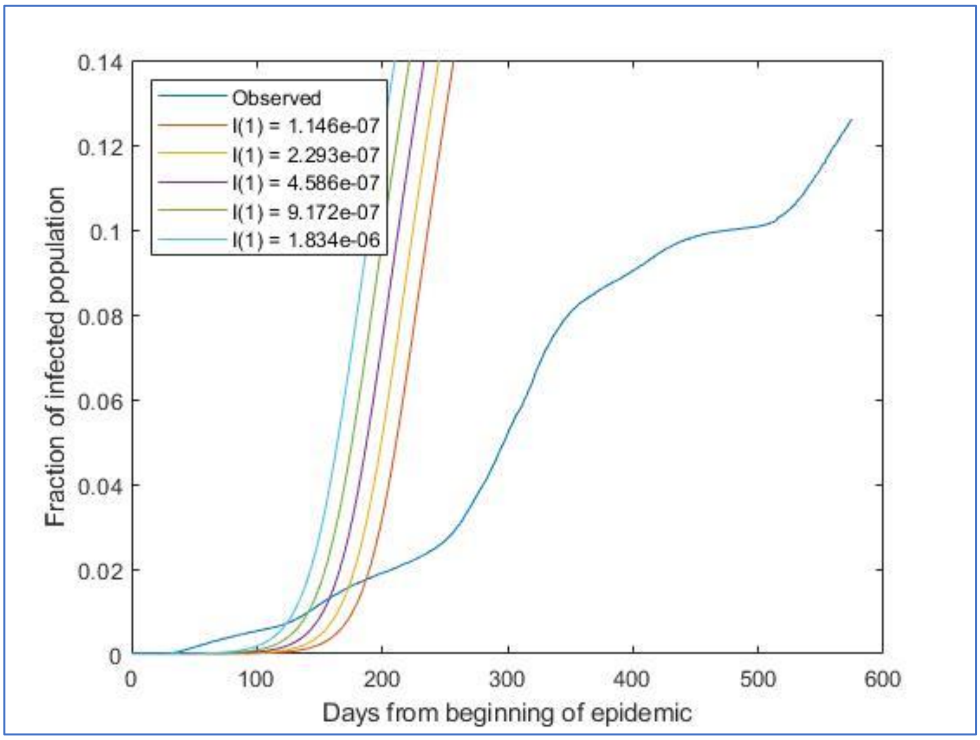

Figure 35: Observed vs Predicted for United States by varying I(1)

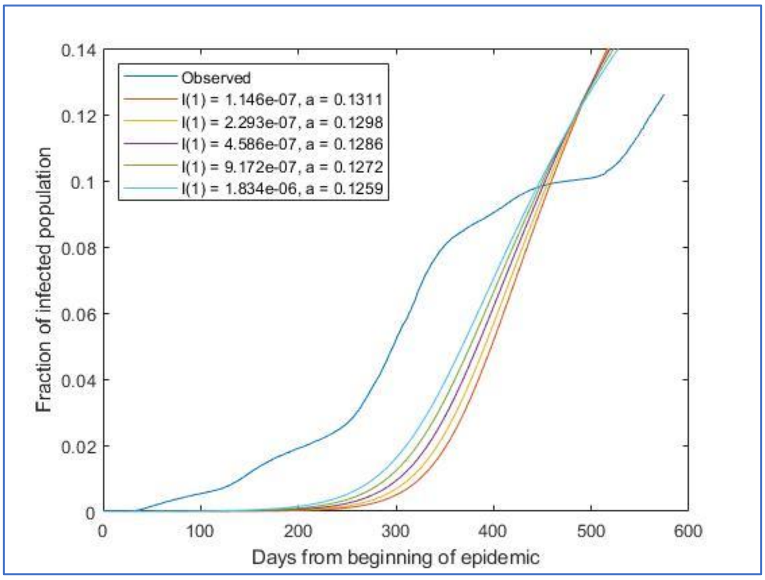

Figure 36: Observed vs Predicted for United States by varying I(1) and a

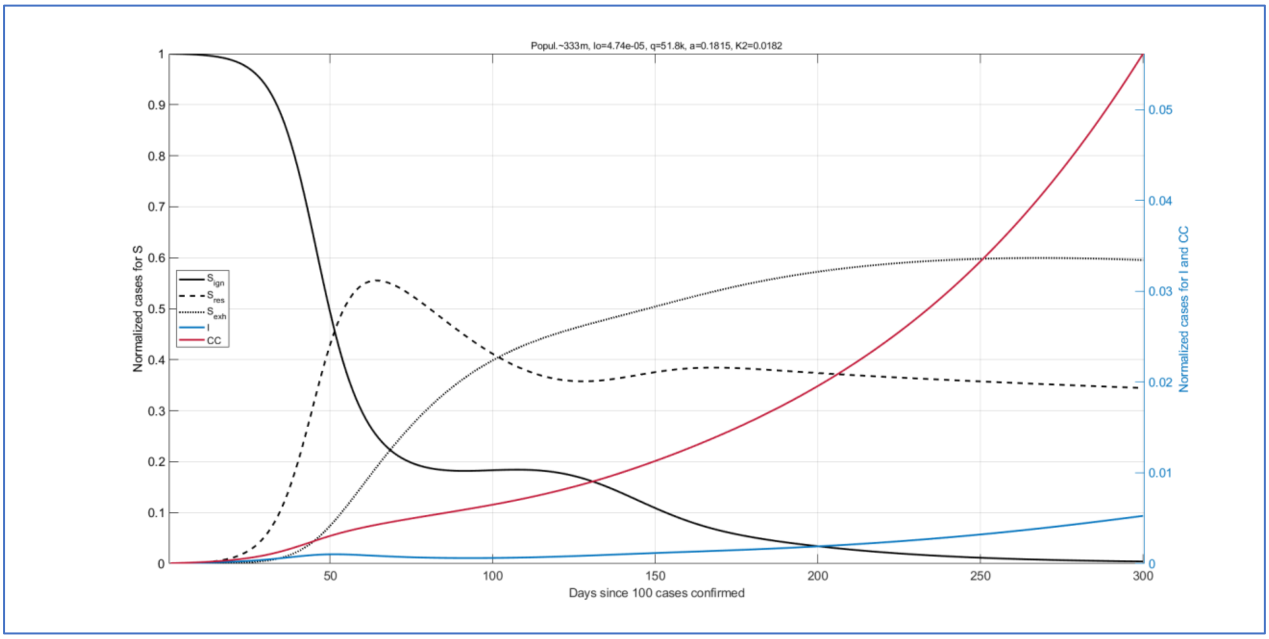

Figure 37: Sign, Sres, Sexh, Infected (I) and cumulative case (CC) plot for United States

Table 13: Parameters of United States of America (SIR improvised)

UNITED STATES OF AMERICA	
Parameters	**Value**
a	0.1815
b	0.1000
I_0	$3.6*10^{-6}$
c	-13.4497
K2	0.0182
S_0	1.0000
Population	333000000

Table 14: Summarized table of Improvised SIR model

Countries	Population (In thousands)	Days Modelled	a	K_2	$q * 10^3$	$I_0 * 10^{-6}$	R^2
IND	1,380,004	200	0.1324	0.02330	21.0	2.64	0.905
US	331,003	300	0.1815	0.01820	73.4	19.20	0.875
UK	67,886	200	0.16060	0.02900	80.1	28.50	0.891

Table 15: SIR Model Improvised using kinetic equations

SIR Model Improvised using kinetic equations	
Country	**MSE**
United States	5. 7123441
United Kingdom	3. 1314290
India	0. 3289083

MSE between the observed and predicted obtained using the SIR model (with above kinetic equation) is shown above. The results are better compared to the chapter 2 SIR model's MSE for all the countries.

Crowd Effect

Modified reaction rate for the below reactions is given and asked to solve the kinetic equations for the crowd effect. Then observed vs predicted plot has been captured for the country's cumulative cases and corresponding MSE value is obtained.

$$S_{ign} + I \rightarrow S_{res} + I$$

Table 16: Crowd effect kinetic equation

Reactions	Reaction rate	Stoichiometric vector
$S_{ign} + I \rightarrow 2I$	$r_1 = aS_{ign}I$	$(-1, 0, 0, 1, 0)^T$
$S_{ign} + I \rightarrow S_{res} + I$	$r_2 = k_2 S_{ign} I$	$(-1, 1, 0, 0, 0)^T$
$S_{res} \rightarrow S_{exh}$	$r_3 = k_3 S_{res}$	$(0, -1, 1, 0, 0)^T$
$S_{exh} + I \rightarrow 2I$	$r_4 = aS_{exh}I$	$(0, 0, -1, 1, 0)^T$
$I \rightarrow R$	$r_5 = bI$	$(0, 0, 0, -1, 1)^T$
$S_{exh} \rightarrow S_{ign}$	$r_6 = k_6 S_{exh}$	$(1, 0, -1, 0, 0)^T$
$S_{ign} + 2I \rightarrow S_{res} + 2I$	$r_2' = qS_{ign}I^2$	$(-1, 1, 0, 0, 0)^T$

Remove all r_2

Add r_2'

Find q

$$\frac{dc}{dt} = \begin{matrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{matrix} r_1 + \begin{matrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{matrix} r_2' + \begin{matrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{matrix} r_3 + \begin{matrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \end{matrix} r_4 + \begin{matrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{matrix} r_5 + \begin{matrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \end{matrix} r_6$$

$$= \begin{matrix} -aS_{ign}I & -qS_{ign}I^2 & 0 & 0 & 0 & k_6S_{exh} \\ 0 & qS_{ign}I^2 & -k_3S_{res} & 0 & 0 & 0 \\ 0 & 0 & k_3S_{res} & -aS_{exh}I & 0 & k_6S_{exh} \\ aS_{ign}I & 0 & 0 & aS_{exh}I & -bI & 0 \\ 0 & 0 & 0 & 0 & bI & 0 \end{matrix}$$

What is q? For $I = I_{panic}$ we have $r_2 = r_2'$.

$$I_p = 2\%$$

$$k_2S_{ign}I_p = qS_{ign}I_p^2$$

$$k_2 = qI_p$$

$$k_2 = q * (0.02)$$

$$q = 50$$

For India, b=0.1 and a=0.1324

$$\frac{dS_{ign}}{dt} = (-0.1324) * S_{ign}I - 50 * S_{ign}I + 0.01 * S_{exh},$$

$$\frac{dS_{res}}{dt} = 50 * S_{ign}I - 0.02 * S_{res},$$

$$\frac{dS_{exh}}{dt} = 0.02 * S_{res} - (-0.1324) * S_{exh}I + 0.01 * S_{exh},$$

$$\frac{I}{dt} = (0.1324) * S_{ign}I + (0.1324) * S_{exh}I - (0.1) * I$$

$$\frac{R}{dt} = (0.1) * I$$

For United Kingdom, b=0.1 and a=0.16060

$$\frac{dS_{ign}}{dt} = (-0.16060) * S_{ign}I - 50 * S_{ign}I + 0.01 * S_{exh},$$

$$\frac{dS_{res}}{dt} = 50 * S_{ign}I - 0.02 * S_{res},$$

$$\frac{dS_{exh}}{dt} = 0.02 * S_{res} - (-0.16060) * S_{exh}I + 0.01 * S_{exh},$$

$$\frac{I}{dt} = (0.16060) * S_{ign}I + (0.16060) * S_{exh}I - (0.1) * I$$

$$\frac{R}{dt} = (0.1) * I$$

For United States of America a=0.1815, b=0.1,

$$\frac{dS_{ign}}{dt} = (-0.1815) * S_{ign}I - 50 * S_{ign}I + 0.01 * S_{exh},$$

$$\frac{dS_{res}}{dt} = 50 * S_{ign}I - 0.02 * S_{res},$$

$$\frac{dS_{exh}}{dt} = 0.02 * S_{res} - (-0.1815) * S_{exh}I + 0.01 * S_{exh},$$

$$\frac{I}{dt} = (0.11492) * S_{ign}I + (0.1815) * S_{exh}I - (0.1) * I$$

$$\frac{R}{dt} = (0.1) * I$$

Table 17: SIR Model Improvised using kinetic equations

SIR Model Improvised using kinetic equations	
Country	**MSE**
United States	1. 0534234
United Kingdom	0.8131429
India	0. 1289083

MSE between observed and predicted cumulative cases for all the countries are mentioned in the above table. The MSE is improved for all the countries and this improvised model is performing exceptionally well.

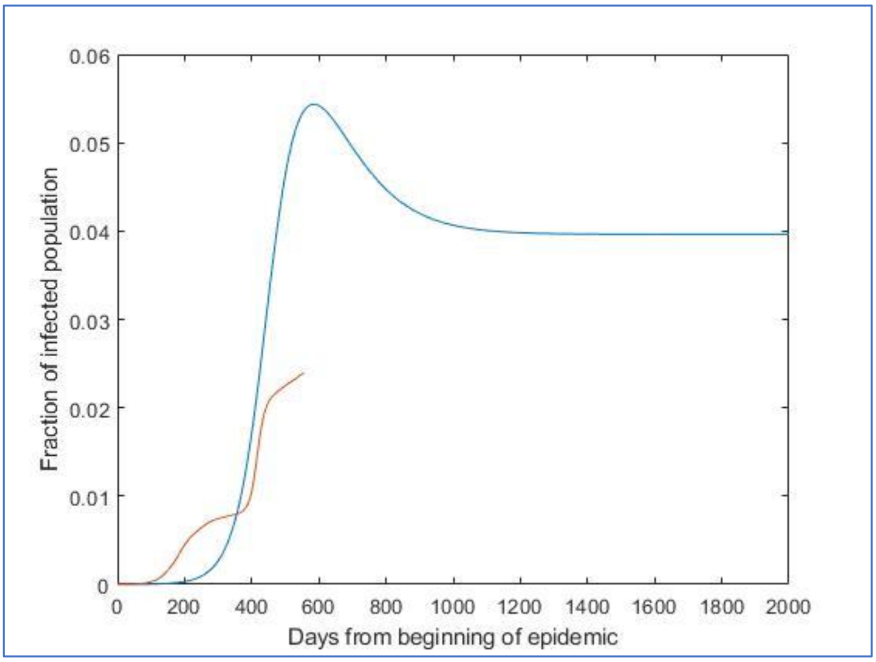

Figure 38: India Observed (Red) vs Predicted Cumulative case (Blue) using crowd effect

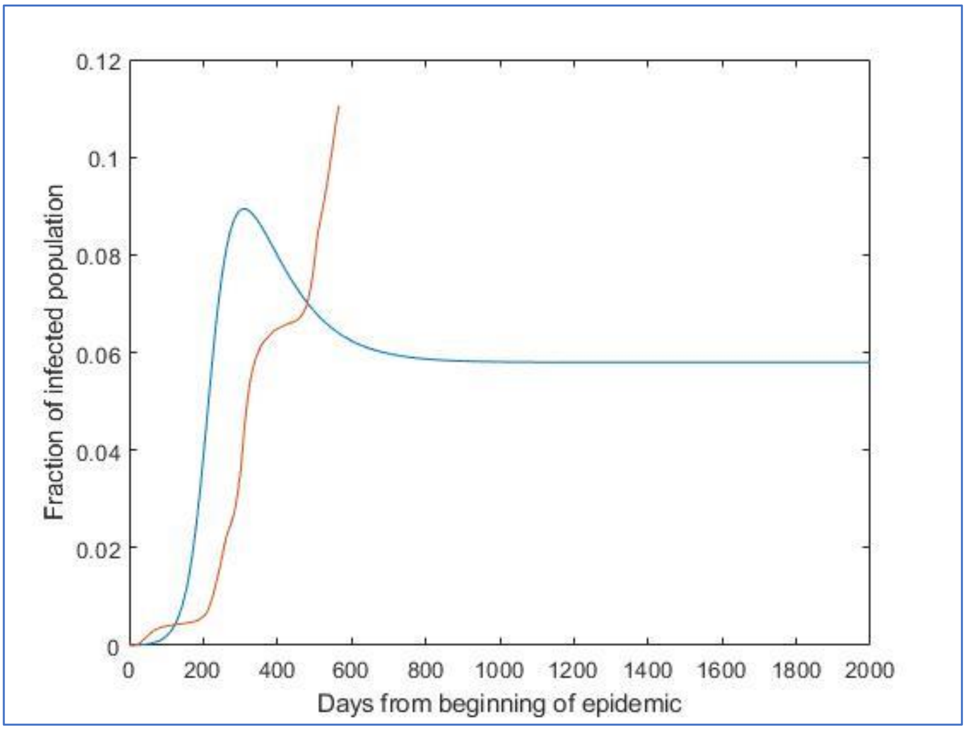

Figure 39: United Kingdom Observed (Red) vs Predicted Cumulative case (Blue) using crowd effect

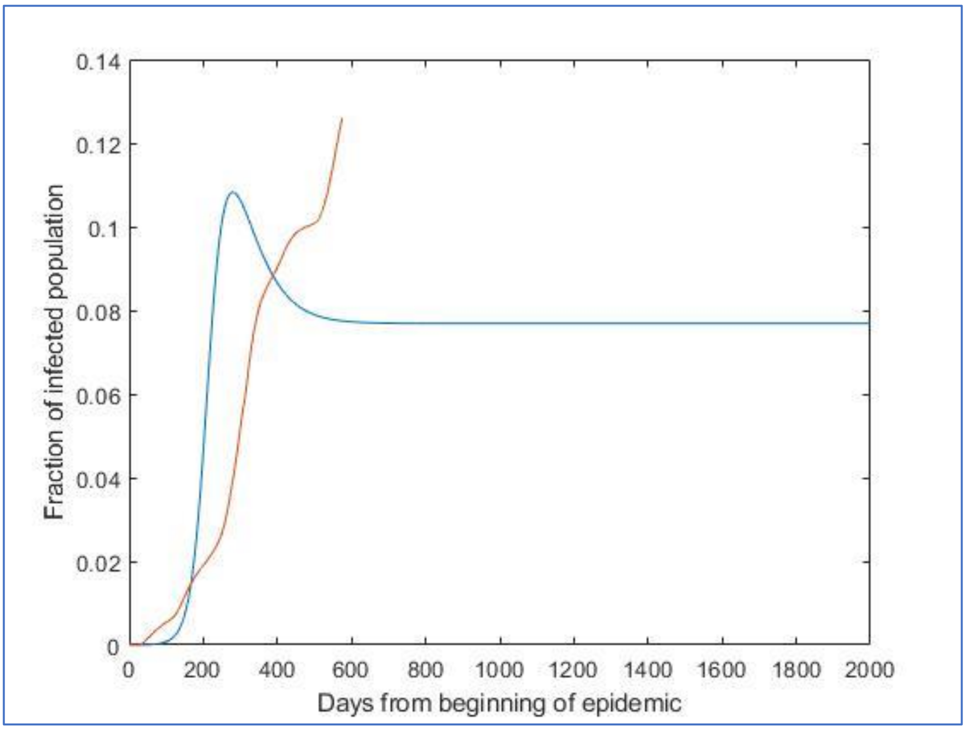

Figure 40: United States of America Observed (Red) vs Predicted Cumulative case (Blue) using crowd effect

Experience of modelling covid19

In draft 2, the cumulative corona virus case of each country was modelled with SIR model. SIR model is very generic and did not perform well as we recorded in the draft 2. As the SIR is itself is very simple and only concentrates on 3 parameters. S is the susceptible, I is Infected and R is recovered or dead. Since modelling an epidemic event like covid 19 is far more complex, introducing more parameters in addition to SIR is necessary to reduce the MSE (mean squared error) between predicted and observed. These parameters should be introduced to balance out the psychological behaviour of the people of each country and the implementation of restrictions during the peak cases. Multi wave dynamics of covid 19 is hard to model with a generic model like SIR, as it doesn't consider the psychological behaviour of people and the government of each country.

There are three phases of Stress: Alarm, Resistance, and Exhaustion. Each country has its own stress behaviour, and it depends on the people's mentality and how strict the government implements the restrictions to control the further spread of covid 19. Alarm is a state where people are in alarm phase i.e., when the population comes to know about the disease. Resistance is with save behaviour i.e., when the population obeys the restriction imposed by the government of each country. Since the incubation stage of covid 19 is 10 days i.e., b = 0.1, spread of covid was controlled at the end of the resistance phase. During Exhaustion phase, the population runs out of the supplies and savings. So, the restrictions are taken away by the government and people tend to behave unsafe during this period. So, once again the country expects second wave or peak. This is a closed loop as discussed.

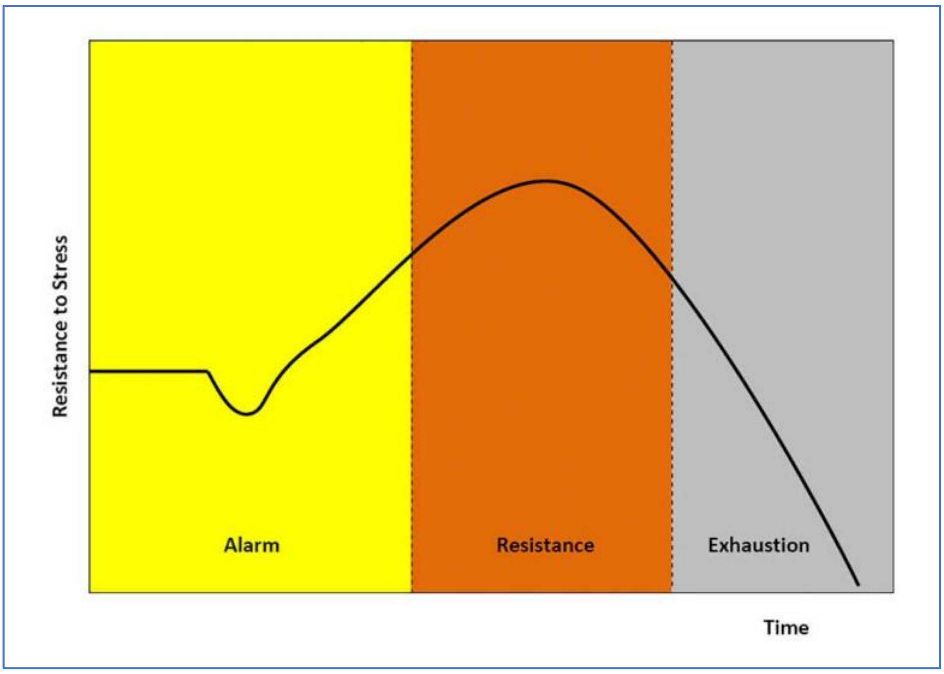

Figure 41: Psychological factor of SIR Model

Considering three states of mind (Alarm, Resistance, Exhaustion) of people has shown improvement to the MSE (Mean squared error) between the predicted and observed cumulative case for each country as recorded above in the document. Introducing few customized parameters of each country might improve the model further. For example, introducing a parameter for festive seasons where people tend to gather and celebrate might help the model further decrease the MSE between the predicted and observed cumulative case for each country. For instance, we can consider Diwali season in India, and can consider Christmas for countries like United States and United Kingdom.

Conclusion

In Chapter 1, the basic data exploration is done and wave analysis for all the countries are defined using straight line and median filter technique. Further analysis is done by implementing exponential grows and logistic growth. For this analysis we have separated all the waves separately and implemented exponential and logistic grows. Obviously, the result should be better in this case as we are implementing the model for each waves separately. But the main objective is to model the cumulative case of each country all together and get an acceptable MSE value between the observed cumulative case and the predicted cumulative case. Concluding the Chapter 1, the wave synchronization between the countries is discussion. The wave synchronization between United Kingdom and United States of America does go together and Indian covid wave has a sort of delay especially with the second wave.

In Chapter 2, a generalized SIR model (S-susceptible, I-Infected, R-Recovered) is modelled with the observed cumulative cases for all the countries. The initialization and SIR parameters are defined using the mentioned equations. By varying these parameters, visualizations of observed and predicted cumulative cases are pictured and discussed the impact of varying the parameters. The good and bad factors to be considered while implementing SIR model is also discussed in the Chapter 2.

In Chapter 3, consideration of psychological factors in the human beings are included to the traditional SIR model. By splitting the S into three different psychological components of human beings, it is possible to implement an improvised version of the tradition SIR model used in the chapter 2. Solving the kinetic equations as given and solving the differential equations is done to compare the results with the other model used in previous chapters. MSE is documented for all the countries using this improvised version of the SIR model. Crowd effect is considered and further tweaking to the kinetic equation is made. Comparison between the with crowd effect and without is done by comparing the MSE values of corresponding countries.

The combination of a dynamic SIR-type model with the classical triad of stages of the general adaptation syndrome, alarm-resistance-exhaustion, makes it possible to describe with high accuracy the available statistical data for the three countries (India, United Kingdom, and the United States of America). The sets of kinetic constants corresponding to optimal fit of model to data were found. These constants characterize the ability of society to mobilize efforts against epidemics and maintain this concentration over time and can further help in the development of management strategies specific to a particular society.

Reference

[1]

"Frontiers | COVID-19 in Wuhan, China: Pressing Realities and City Management | Public Health." https://www.frontiersin.org/articles/10.3389/fpubh.2020.596913/full (accessed Dec. 17, 2021).

[2]

"Modeling Logistic Growth. Modeling the Logistic Growth of the… | by Joos Korstanje | Towards Data Science." https://towardsdatascience.com/modeling-logistic-growth-1367dc971de2?gi=2577576af4e7 (accessed Dec. 17, 2021).