**Tracking the spread of the COVID-19 virus in India**

**Abstract**

An Extended-Susceptible-Infected-Removed (eSIR) model has been developed to track the time-evolution of the spread of the COVID-19 virus within a community. The model is based upon the well-known SIR model. The main differences between the SIR model and the eSIR model is that in the eSIR model, the total population is not defined and the number of susceptible individuals is a variable that can be increased in surge periods and not decrease monotonically as in the SIR model. In this paper, the eSIR model has been applied to track the time-evolution of the virus in India. As new data becomes available, the eSIR model input parameters are adjusted to give a best fit of the eSIR model predictions with the published data. The eSIR model’s predictions can be used in both the qualitative and the quantitative analysis of the spread of COVID-19 and thus gives an insight into the spread of the virus that the raw data alone does not.

*Keywords*: COVID-19, infectios disease, virus spreading, eSIR model.

**1. Introduction**

A novel strand of Coronavirus (SARS-CoV-2) was identified in Wuhan, Hubei Province, China in December 2019 causing a severe and potentially fatal respiratory syndrome known as COVID-19. On March 11, 2020 the World Health Organization (WHO) declared a world- wide pandemic. By mid-June, 2020 most countries have been affected and the total infections around the globe has exceed seven million and the number of deaths is approaching half-million. At the beginning of April, 2002 there was a just over 1000 infected individuals in India, but my mid-June the number of infections has increased to more than a quarter-million. So far, government actions in India have not bought the spread of the virus under control despite the shuttering down of the economy. So, it is imperative that mathematical models be used to monitor the spread of the virus to give a scientific basis for the control measures implemented by governments and to access their success. Such models can be used to predict the number of infections and deaths in the future, to monitor any changes in the trends of the spread of the infection and provide estimates of time scales involved.. Using mathematical models, one can gain a better qualitative understanding of the spread and control of the virus as well as provide a quantitative framework to analyse the published data. Insights can be drawn from the mathematical models that may be nearly impossible to discern from the data alone.

The standard well-known SIR (susceptible-Infected-Removed) model cannot be applied successfully to the study the evolution of the virus in India. The SIR model assumes a homogeneous mixing of the infected individuals within a constant total population and the susceptible population decreasing monotonically towards zero. The proposed eSIR (Extended-Susceptible-Infected-Removed) model relaxes some of the imposed restrictions of the SIR model. The eSIR model does not consider the total population, nor a homogenous mixing of infected individuals within a community. Rather, as people become infected and move about, more individuals are added to the susceptible population. Hence, the susceptible population is considered as a variable that can be increased at various times to account for newly infected individuals spreading throughout a community. The SIR model and its extended modifications such as eSIR mathematical model in different forms have been used in some previous studies [19-21] to model in current situation due to the spread of COVID-19.

**Methods**:

We used time-series datasets [22,23] of COVID-19 from Marchy, 2020 to June, 2020. An infectious disease dynamical model (eSIR) was introduced, which is applied to understand the effects and estimate the disease trend in India and then compared it with modelling framework. We compare the pandemic development of the disease and predict the characteristics using the eSIR model and assess the effectiveness of modelling approach.

**2. The Susceptible-Infected-Removed (SIR) model**

The SIR (Susceptible-Infected-Removed) model consists of a system of three coupled non-linear ordinary diﬀerential equations (equation 1). A total population *N* is divided into three distinct groups: the susceptible population *S*, the infected population  *I* and the removed population *R*.

***Susceptible population***: the number of individuals who are not infected, however, could become infected .

***Infected individuals***:  the number individuals who have already been attacked by the virus and can transmit it to those individuals who are susceptible.

***Removed individuals***:  is the number of individuals who have been removed from the sufferings of the virus and are assumed to be immune from being infected again.



 (1)



The SIR disease transmission model is derived assuming several strong assumptions. It is assumed that the members of the susceptible and infected populations are homogeneously distributed in space and time. An individual removed from the infected population has a lifetime immunity. The total population is constant where *N* = *S* + *I* + *R* and the number of births and the number of deaths from causes other than the virus are ignored.

The system of equations (equation 1) can be solved using the Euler method or the Runge-Kutta method as shown in figure 1.



Fig. 1. The time variation of the population: susceptible *S*, infected *I*, removed *R*, and effective reproductive number.

Model parameters: *S*(0) = 1.0, *I*(0) = 0.001; *R*(0) = 0, a = 0.20, b = 0.05.

The evolution of the infected population is governed by the second ODE in system 1, where *a* is the transmission rate constant and *b* is the removal rate constant. At the start of an epidemic, where , the number of infections increases exponentially . Then the rate of infection falls to zero at the peak where . When S drops below about 0.2, the infected population decreases exponentially .

We can define an effective reproductive rate *Re*, where *Re* = *a S* / *b*. The evolution of the disease depends upon the value of the reproductive rate *Re*. When the effective reproductive rate *Re* is less than one, then the infected population will decrease monotonically to

zero, when greater than one, it increases and when *Re* = 1, the rate of increase of the infected population is zero and this corresponds in the peak number of active infected individuals. The

reproductive rate *Re* acts as a threshold value that determines whether an infectious disease will quickly die out or whether it will continue to grow. The existence of a threshold for infection is not obvious from the data but can be discerned from a mathematical model.

The rate of increase in the number of infections depends on the product of the number of infected and susceptible individuals. An understanding of the system of equations (1) explains the staggering increase in the infection rate around the world. Only by reducing the number of people who can become infected can the spread of the virus be contained.

One cannot get accurate predictions by applying the SIR model to the data of the pandemic in India. Figure 1 identifies the problem in applying the SIR model to the Indian data starting at the 14 March, 2020. The model and predictions are in agreement only at the start of the pandemic. The number of active infections in India does not peak, it only keeps getting bigger and bigger. However, by modifying the SIR model by resetting the susceptible population at surge times, it is possible to fit the model predictions to all the available data. Therefore, applying the eSIR model to the Indian data, one can make a quantitative analyse of the spread of the virus.



Fig. 1. The model predictions (14 March, 2020 to 30 September, 2020) and Indian data (14 March, 2020 to 9 June, 2020) for the infections, recoveries and deaths.

**3. The Extended-Susceptible-Infected-Removed (eSIR) model**

In a community such as India there is not a homogenous mixing of the infected and susceptible individuals within the population. Only part of the total population will be become susceptible to infected individuals. When infected individuals persist and move about within a community, further individuals can become susceptible to the disease. The susceptible population will not necessarily decrease monotonically with time. As susceptible individuals become infected, these new infections act as a source for more individuals to become susceptible. This gives rise to a positive feedback loop leading to a very rapid rise in the number of active infected cases in a surge period where the number of susceptible individuals increases and not decreases. The SIR model is modify so that any time *ts*, the susceptible population can be reset to *Ss*(*ts*).

The system of equations (1) are solved for the variables *S*, *I* and *R* where *S*, *R* and *I* [0 1]. These three populations are multiplied by a scaling factor *f* to give the number of individuals in each population group to match the model outputs with the actual COVID-19 data for India.

The solution of the system of equations (1) gives only the values for the removed population *R* as a function a function of time *t*. Individuals removed from the infected population *R* have either recovered, *C* or has died, *D* where *R* = *C*+*D*. The number of deaths, *D* is estimated by fitting a non-linear function to a plot of the data for the number of actual deaths verses the data for the number of removals. The non-linear function used to find the number of deaths, *D* from the number of removals *R* is given by equation 2

, (2)

where *D*0 and *k* are constants chosen to give the best fit between the non-linear function and the data as shown in figure 3. The number of recoveries, *C* is then simply *C* = *R* – *D*.

The eSIR model input parameters are the population scaling factor, *f*, the initial infected population, *I*(0), the rate constants *a* and *b*, death constants, *D*0 and *k*, and the surge periods *ts* and *Ss*. This set of constants is chosen to best fit the eSIR model predictions with the dataset for India. The eSIR model is used to track the data. As new data becomes available, slight adjusts to the input parameters can be made to get a better agreement between the model predictions and the data for each population. The eSIR model cannot predict the time and height of the peak in the number of active infections. But, if the number of active infections keeps increasing, the variable *S* can be increased which has the effect of increasing the peak value for active infections and delaying the time at which the peak value will occur. The adjustments that are necessary in resetting the value of *S* are an important indicator of the success or failure of government actions in controlling the spread of the virus. High values of the reset value of *S* and a high reset frequency means that the spread of the virus is not contained. Only when the susceptible population decreases towards zero will the number of active infections also decline to zero. Either though the peak for active infections in India has not been reached, the eSIR model gives estimates of the duration of the pandemic and minimum estimates for the total cumulative number of infections and deaths.

For the virus to be controlled and to limit its impact on society, governments must take actions to prevent people to become susceptible to the disease. Some actions that are proving to be successful include: testing, contract tracing of people who may have been in contact with an infected individual and then isolating them, isolating suspected individuals, restricting movement of people, restrictions on large gatherings, hygiene measures such as washing hands, and wearing face masks particularly on public transport or in crowed indoor spaces masks. By using the predictions of mathematical model and comparing it with published datasets, one is more able to access actions that have been taken to control the spread of the virus.

**4. Results**

Data has been collected for India from 14 March, 2020 (day zero) until mid-June, 2020. The results of the modelling for the spread of the COVID-19 virus in India are displayed in figures 2 and 3 for a period of 200 days. The crosses are for the data and the smooth lines for the model predictions (14 March, 2002 to 30 September, 2020). The model input parameters are : *f* = 1.300 x105, *I*(0) = 1.000x10-4, *a* = 0.200, *b* = 0.040, *D*0 = 1.00x104, *k* = 1.200x10-5.



Fig. 2. Model predictions and data as functions of time for the cumulative total number of infections, the active number of infections, the number of recoveries and the number of deaths / 1000.

The model predictions in figure 1 show that the infected population is still increasing and a peak in the active infections has not yet been attained. Although the selection of model’s input parameters and surge times are subjective, the data can be tracked quite successful, but does not predict the peak for the active infected population. However, the eSIR model’s predictions are useful. If the peak does occur in mid-June, then by the end of September the active number of infections will be very low and the virus is no longer of a great concern. The minimum number of cumulative infections will be in the order of 350 000 and the minimum number of deaths in the order of 10 000.



Fig. 3. Model predictions and data. The plot of the data for number of deaths verses removals is use to estimate the number of deaths and recoveries.

Figure 3 shows that the peak in the number of actively infected may not occur for some time. The plot of the susceptible variable, *S* shows that the susceptible population has to be repeatedly reset to higher values. This is a strong indicter that the spread of the virus is not contained during May and the early part of June. The plots for the removals against infection or total infections show that the data has not reached the knee section of the model’s prediction again providing evidence that the peak in the active infections may not occur soon.

It is most likely the spread of the virus as of mid-June, 2020 is not controlled and there will be an

|  |  |  |
| --- | --- | --- |
|  | DATA  9 June, 2020 | eSIR Model Predictions  30 September, 2020 |
| active infections *I* | 133729 | 2984 |
| total cum. infections *Itot* | 283896 | 354849 |
| recoveries *C* | 1422417 | 342030 |
| deaths *D* | 7750 | 9835 |
| % active infections | 48.4 | 0.9 |
| % recoveries | 48.8 | 96.3 |
| % deaths | 2.8 | 2.9 |

Table 1. Summary of the data for 9 June, 2020 and the eSIR model predictions for 30 September, 2020.

For the data recorded on 9 June, 2020 9,

**4. Conclusions**

**References**