

# Preliminary Report of the Susceptible-Infected-Removed (SIR) Model for Disease Spread

Yu An

Frieda Han

Yun Lin

---

*Keywords:* SIR Model, disease spread, infectious disease, contagion modeling

---

## 1. Introduction to the SIR model

The SIR (susceptible-infected-removed) model, contributed by Kermack, W. O., McKendrick, A. G., Ronald Ross and others, was formed in the early twenties and made a big breakthrough in modeling the trend of the possible disease outbreak as well as providing key information for the public and authorities regarding severeness and control (1). The establishment of the simplified model is implemented here in this project (4)(9)(8).

Under a fixed homogeneous population ( $N$ ), three distinct subpopulations are divided ( $S+I+R = N$ ):

1. S - those who are susceptible and have not yet caught the disease
2. I - those who are infected and may spread the disease
3. R - those who have recovered without the ability to spread and are immuned to the disease

These groups are divided in a way that during an epidemic, people are interested in the amount of infected patients that can transmit the disease, the number of agents who are immuned against the disease and the portions who are not yet being infected and are susceptible. Denote the sizes of these groups at time  $t$  by  $S(t)$ ,  $I(t)$ , and  $R(t)$ . Moreover, introduce notations involved in this model:

- $b$ : a fixed number of interactions that spread the disease per day per infected individual
- $k$ : a fixed fraction of the infected subpopulation that recovers per day

Altogether, the SIR model consists of a system of three ordinary differential equations (ODEs) that be written by

$$\begin{aligned}\frac{ds}{dt} &= -b * s(t) * i(t) \\ \frac{di}{dt} &= b * s(t) * i(t) - k * i(t) \\ \frac{dr}{dt} &= k * i(t)\end{aligned}$$

where  $s, i$  and  $r$  here are the percentage of each group size over the total population.

- Equation 1 captures the infection of  $S$  by  $I$  through interaction with parameter  $b$ .
- Equation 2 captures the infection of  $S$  and the removal of  $I$ .
- Equation 3 captures the exit of  $I$  and the enter of  $R$  with rate  $k$ .

Besides a closed population with no social structure, the SIR model also assumes that  $R$  subpopulation is permanently immuned and the  $I$  subpopulation is contagious at the moment of infection. Moreover, the model assumes a mass-action mixing, meaning the fraction of interactions between  $S$  and  $I$  is proportional to the product of their population sizes (10). In other words, each agent will have equal probability in interacting with every other agents per unit time. This implies that an increase in the size of the subpopulations will lead to an increase in the infections and that  $I$  and  $S$  are homogeneously distributed in space. This might be problematic in real-world cases, as people tends to encounter within a small community. While these assumptions can sometimes be strong and adversely affecting the prediction in applications, they can be adjusted in various forms to bring complexity and create extended and modified versions of the SIR models. The project explores some variation and modification in section 4.

## 2. sir package

The *sir* package consists of 3 parts. *sir/discrete.py* contains all definitions of class and functions required for the discrete agent-based simulation. *sir/ode.py* contains all definitions of functions used in continues/ ODE-based simulation. *sir/\_\_init\_\_.py* is required to import the directory as a package.

	<b>discrete.py</b>	<b>ode.py</b>
class	Person	
function	count_infectious count_removed count_susceptible sir_model_simulation	ODE_simulation

## 3. Preliminary investigations using the agent-based and continuous models

In this section, the population size is assumed to be 10000. The initial condition is that 0.1% of the population has been infected, while others remain susceptible.  $b$  denotes the number of interactions each day that could spread the disease (per individual) and  $k$  denotes the fraction of the infectious population which recovers each day.

### 3.1. ODE simulation

Here we simulated the  $S, I, R$  populations over time under different  $b$  and  $k$ . It is obvious that if  $k$  is approximate zero, the infectious population gets larger but cannot recover, hence all population will get be infectious eventually. It is also obvious that if  $k$  is approximate one, infectious people are sure to removed, hence no more susceptible people get infected. This two special but trivial cases were not simulated. We simulated for  $b = 1/5, 2/5, 3/5, 4/5, 1, 2$  and  $k = 1/5, 2/5, 3/5, 4/5$ . It can be seen from Figure 1 that

1. If  $k = b$ , there will be a very small surge of infectious population (less than 10% of total population), so the susceptible population decrease a bit. Soon the infectious population will decrease to around zero since they are removed (recovered or dead).
2. If  $k > b$ , there will almost no surge of infectious population, the initial infectious people will soon removed. The susceptible population remains uninfected.
3. If  $k < b$ , there will be a significant surge of infectious population. As  $k$  is significantly larger than zero, the infectious population will eventually recovered. Moreover, when  $k < b$ , the larger the difference between  $b$  and  $k$ , the greater the surge will be. When  $b$  is larger than 1, it can be seen that almost the entire population will have been infected and recovered.

In order to have a closer look at the qualitative behavior of the simulations based on the parameters  $b$  and  $k$ , we drew phase diagrams to look at the each population proportion with respect to  $k$  and  $b$  at several time points ( $T = 20, 40, 60, 80$ ).

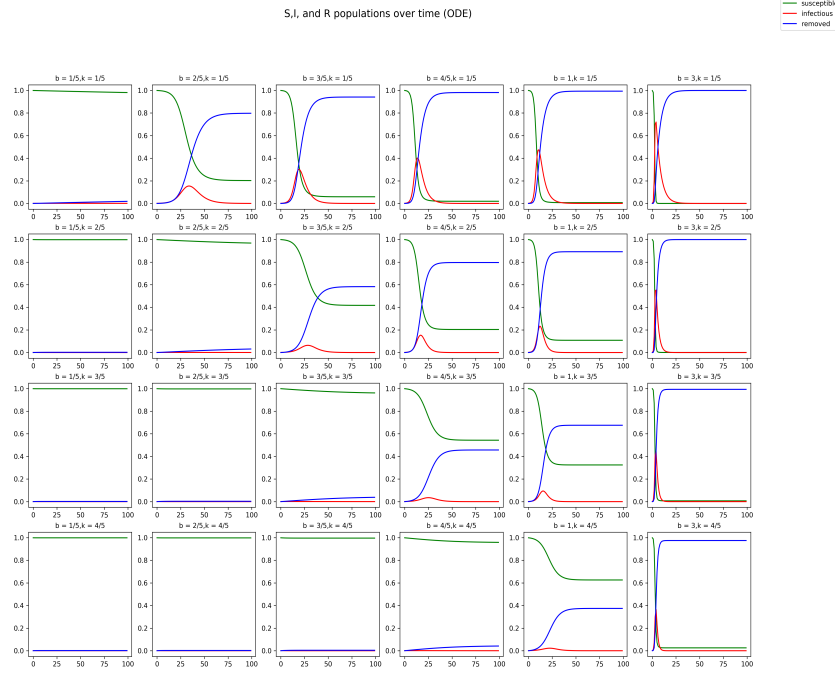


Figure 1: Example plot for ODE simulation under different values of  $b$  and  $k$

### 3.2. ODE simulation

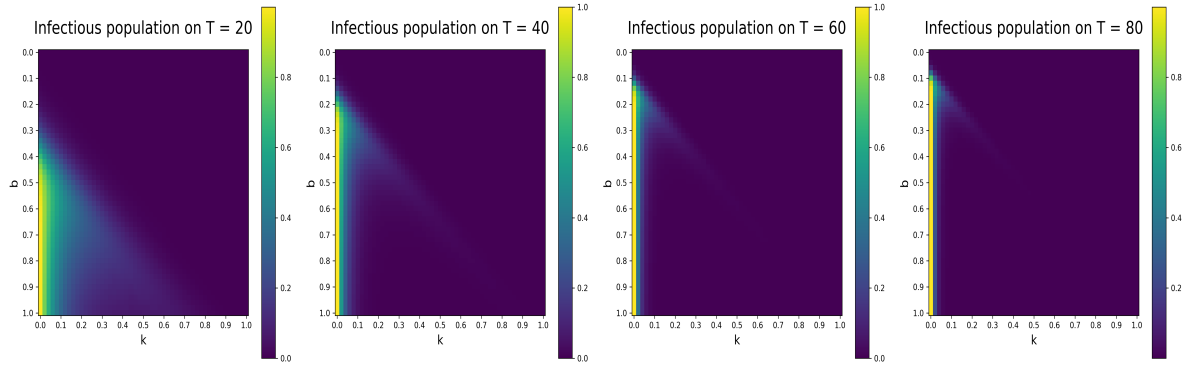


Figure 2: Phase diagram of infectious population for ODE simulation

From the four phase diagrams above (Figure 2) for infectious population, it can be seen that

1. The dark area on top right of each diagram indicates that if  $k > b$ , infectious population size is close to zero all the time, which means that there is almost no surge of infectious population and the initial infectious people are soon removed. It agrees with the result of Image 1.
2. The lighter area first appears in the lower left corner, then moves in the upper right and become darker itself at the same time. It indicates that if  $k < b$ , when the difference between  $k$  and  $b$  is larger, the surge of infectious population appears earlier and greater. It is also earlier and quicker for the infectious population goes to zero.

From the diagrams above (Figure 3) for susceptible and removed population, it can be seen that

1. If  $k > b$ , susceptible population proportion is always close to one and removed population proportion is always close to zero, since there are almost no more infectious people.

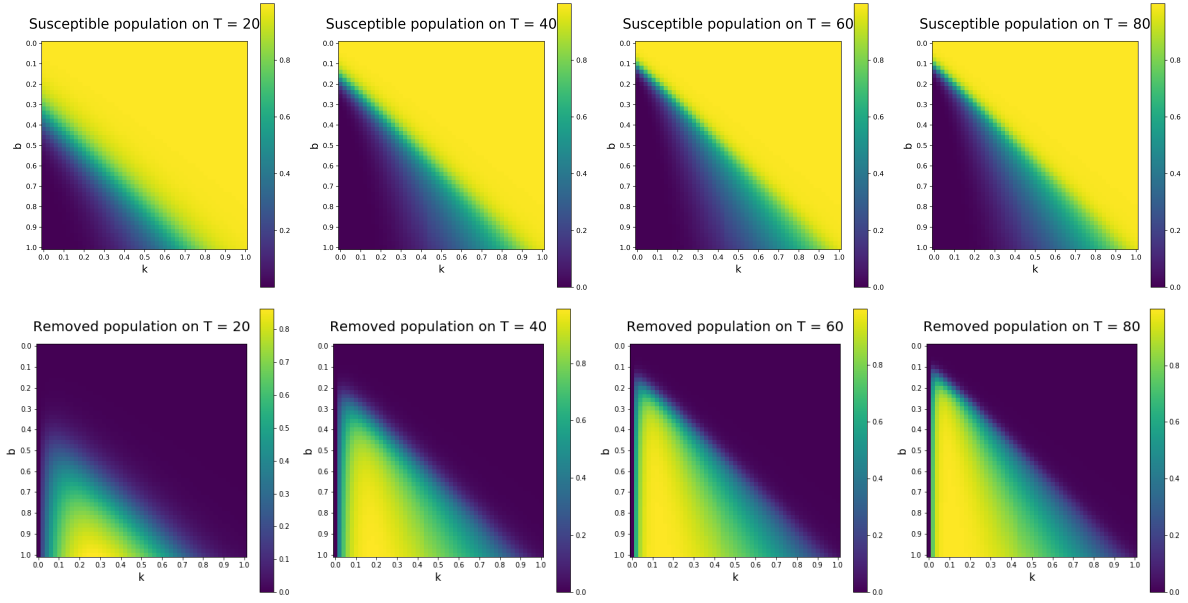


Figure 3: Phase diagrams of susceptible and removed population for ODE simulation

2. IF  $k < b$ , After the surge and recover of infectious population, the susceptible population proportion at the end is smaller when the difference between  $b$  and  $k$  is greater. It means if  $b$  is greater enough than  $b$ , everyone is eventually infected.

### 3.3. Agent-based model simulation

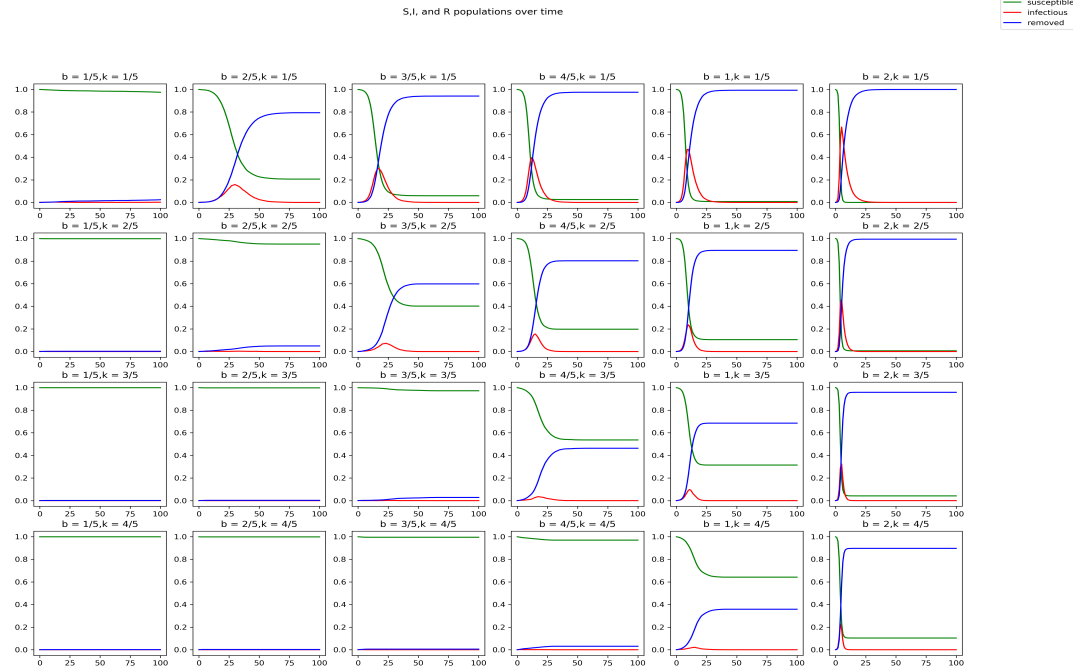


Figure 4: Example plot for agent-based model simulation under different values of  $b$  and  $k$

The example plot for agent-based model simulation (Figure 4) agrees with that for ODE simulation.

We then drew 12 phase diagrams to see the qualitative behavior of the simulations based on the parameters  $b$  and  $k$ :

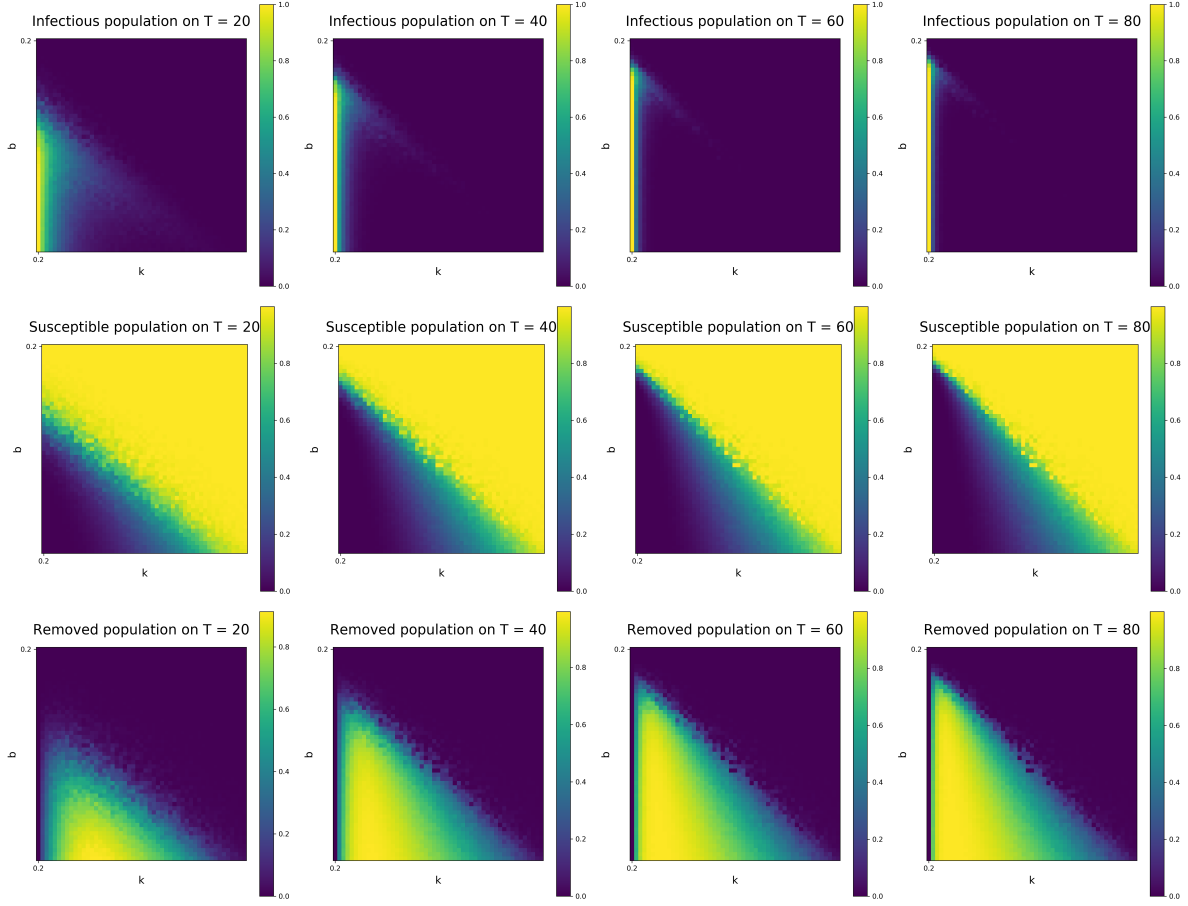


Figure 5: Phase diagrams of susceptible and removed population for agent-based model simulation

The phase diagrams (Figure 5) also agrees with that for ODE simulations, though they are not as smooth as plots for ODE simulations.

#### 4. Variations/improvements proposal

##### 4.1. Modeling the use of quarantine

The SIR model assumes the frequency of interaction between one and the other is kept stable even when one of them is infected. According to actual situation, we can see that many countries/regions try to use quarantine to reduce the interactions between people, which is the parameter  $b$  in the SIR model. Moreover, under the case of COVID-19, there's a time delay between one being infection and being confirmed as sick. Therefore isolating infected patients is an effective way to reduce the rate of disease transmission. Based on this idea, we propose the following SUQC (11) model to describe this situation.

In this model, the total population  $N$  is divided into 4 parts:

1.  $S$  (Susceptible) - the individual who has not yet caught the disease
2.  $U$  (Unquarantined) - the unquarantined infected individual who may spread the disease to susceptible individuals
3.  $Q$  (Quarantined) - the quarantined infected individual who loss ability of infecting susceptible individuals

#### 4. $C$ (Confirmed) - the individual who is confirmed sick

All parts above are functions of time  $t$ .

The number of removed individuals  $R$  is not included. This model assumes that once the infected are quarantined, we assume that the probability of infecting susceptible individuals is zero. So whether they are recovered/removed won't affect the whole dynamic system. This model exerts more focus on the direct impact of quarantine measures on disease spread.

There are 3 parameters involved in this model:

- $\alpha$ : the number of individuals infected by an unquarantined individual per day
- $\beta$ : the rate at which the cases are confirmed
- $\gamma$ : the rate at which unquarantined infected individuals get quarantined

which gives us the following differential equations:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\alpha * U(t) * S(t)}{N} \\ \frac{dQ}{dt} &= \gamma * U(t) - \beta * Q(t) \\ \frac{dU}{dt} &= \frac{\alpha * U(t) * S(t)}{N} - \gamma * U(t) \\ \frac{dC}{dt} &= \beta * Q(t)\end{aligned}$$

With the model above, we can build discrete and continuous simulations to see how these 4 parts of population changes over time with different parameters.

#### 4.2. Modeling the introduction of a new state

In the SIR model,  $R$  subpopulations contains those who have recovered or have died. The real-world case can be a bit more complicated. In previous studies, researchers have found that it seemed important to separate the infectious and the recovering phase (5). From a biological point of view, an infectious period might average out about 4 to 5 days while resolving period average out about 14 to 20 days (2). Hence, it can be problematic when fitting real-world data on the SIR model. The introduction of the Susceptible-Infected-Recovery-Death(SIRD) model helps distinguish the two states. After infection, a person then comes to the 'Resolving' state with two consequences: either recovered or died. The SIRD model clarifies the transmission process and help explore the fatality in evaluating the severeness of the disease.

In this model, the total population  $N$  is divided into 5 groups (5):

1.  $S(t)$  - the number of individuals susceptible
2.  $I(t)$  - the number of individuals infected
3.  $R(t)$  - the number of individuals that are being resolved
4.  $D(t)$  - the number of individuals that died from the disease
5.  $C(t)$  - the number of individuals that recovered from the disease

Similar to the SIR model, some of the assumptions for SIRD model requires a fixed homogeneous population with  $C$  subpopulation permanently immuned and the  $I$  subpopulation contagious at the moment of infection.

There are 4 parameters involved in this model:

- $\beta$ : a fixed number of interactions that spread the disease per day per infected individual
- $\gamma$ : a Poisson rate of infectiousness resolves

- $\theta$ : a fixed fraction of people exiting the resolving state
- $\delta$ : a fixed fraction of the resolved consequences being died per day

which gives us the following differential equations:

$$\begin{aligned}\frac{dS(t)}{dt} &= -\frac{\beta * S(t) * I(t)}{N} \\ \frac{dI(t)}{dt} &= \beta * S(t) * I(t) - \gamma * I(t) \\ \frac{dR(t)}{dt} &= \gamma * I(t) - \theta * R(t) \\ \frac{dD(t)}{dt} &= \delta * \theta * R(t) \\ \frac{dC(t)}{dt} &= (1 - \delta) * \theta * R(t)\end{aligned}$$

- Equation 1 captures the new infection of  $S$  by  $I$  through interaction with parameter  $\beta$ .
- Equation 2 captures the new infection of  $S$  and the removal of resolving infectious  $I$ .
- Equation 3 captures the resolving infectious  $I$  and the removal of resolved cases  $R$ .
- Equation 4 captures resolved cases that died
- Equation 5 captures resolved cases that recovered

Run a simulation by starting with a small number of infectious individuals (or maybe fitting the COVID-19 data in Illinois using JHU CSSE COVID-19 Dataset (6)). Explore the trend of  $S, I, R, D$  and  $C$  under different choices of parameters. In addition, investigate some qualitative behaviors of the simulations such as the regimes of the parameters when  $I$  goes to zero.

Moreover, if we consider  $\beta_t$  to be a time-varying parameter, it can then capture and reflect the changes on interactions due to lockdown, social distancing or other individual behaviors (3).

#### 4.3. Modeling the appearance of a new strain

It happens for diseases to reemerge due to the appearance of a new strain. An original strain may mutate through the antigenic shift or antigenic drift. The resulted newer strain does not exhibit cross-immunity with the original one, which means that people who recovered from the original strain will become susceptible to the new strain (7). This situation can be modelled by performing modifications of the SIR model.

A new strain is often introduced when an equilibrium has been achieved by the original strain. It has transmission parameter  $b'$  and the recover parameter  $k'$ . Instead of three categories  $S$ ,  $I$  and  $R$  in the SIR model, new categories are created:

- $S$ : an individual who has not yet caught the disease, but may be infected by the original strain or the new strain.
- $I_1$ : an individual is infectious of the original strain.
- $I_2$ : an individual is infectious of the new strain.
- $R_1$ : People who have removed after being infected by the original strain. Part of them have recovered from the disease while others have died. People who recovered is susceptible of the new strain. Here we suppose the number of people died is proportional to size of infectious population of constant  $d$ .

- $R_2$ : People who have removed after being infected by the newer strain.

Similar to SIR model, system of differential equations can be expressed as:

$$\begin{aligned}\frac{ds}{dt} &= -bs(t)i_1(t) - b's(t)i_2(t) \\ \frac{di_1}{dt} &= bs(t)i_1(t) - ki_1(t) \\ \frac{di_2}{dt} &= b'i_2(t)[s(t) + (1-d)r_1(t)] - k'i_2(t) \\ \frac{dr_1}{dt} &= ki_1(t) - b'i_2(t)(1-d)r_1(t) \\ \frac{dr_2}{dt} &= k'i_2(t)\end{aligned}$$

Based on the model discussed above, it is natural to perform discrete and continuous simulations of these 5 populations and have a look at how the population sizes change over time. We will also solve the equilibrium points for this model where the sizes of each of 5 populations do not change over time and explore what regime of  $b$ ,  $k$ ,  $b'$ ,  $k'$  and  $d$  affect the equilibrium.

## References

- [1] R. M. Anderson. Discussion: the kermack-mckendrick epidemic threshold theorem. *Bulletin of mathematical biology*, 53(1-2):1, 1991.
- [2] Y. M. Bar-On, A. Flamholz, R. Phillips, and R. Milo. Science forum: Sars-cov-2 (covid-19) by the numbers. *Elife*, 9:e57309, 2020.
- [3] S. Chatterjee, A. Sarkar, S. Chatterjee, M. Karmakar, and R. Paul. Studying the progress of covid-19 outbreak in india using sird model. *medRxiv*, 2020.
- [4] I. Cooper, A. Mondal, and C. G. Antonopoulos. A sir model assumption for the spread of covid-19 in different communities. *Chaos, Solitons & Fractals*, 139:110057, 2020.
- [5] J. Fernández-Villaverde and C. I. Jones. Estimating and simulating a sird model of covid-19 for many countries, states, and cities. Technical report, National Bureau of Economic Research, 2020.
- [6] J. H. U. C. for Systems Science and Engineering. Covid-19 dataset.
- [7] M. Fudolig and R. Howard. The local stability of a modified multi-strain sir model for emerging viral strains. *medRxiv*, 2020.
- [8] W. MathWorld. Kermack-mckendrick model.
- [9] M. A. of America. The sir model for spread of disease - the differential equation model.
- [10] H. H. Weiss. The sir model and the foundations of public health. *Materials matematics*, pages 0001–17, 2013.
- [11] S. Zhao and H. Chen. Modeling the epidemic dynamics and control of covid-19 outbreak in china. *Quantitative Biology*, pages 1–9, 2020.