Some Basic SIR Models with Variations

Hsiang Wang, Shenghan Mei, Xinyu Shi

March 23, 2021

Abstract

This paper introduces SIR models of Covid-19 with some variations. First, we introduce some notations for SIR models. Then we discuss agent-based and ODE/PDE models in one-dimension and two-dimension respectively with some proper simulations. Last but not least, we show our extensions to the SIR models including simulating elementary schools disease spreading, reinfection and masks wearing situations.

Introduction

SIR model shows how disease spreads through a population. People are in one of three states at time t:

$s(t)$: the number of susceptible population in the total population at time t.

$i(t)$: the number of infectious population in the total population at time t.

$r(t)$: the number of removed population in the total population at time t.

The model is characterized by the following system of differential equations:

$\frac{ds}{dt} = -b \times s(t) \times i(t)$

$\frac{dr}{dt} = k \times i(t)$

$\frac{di}{dt} = b \times s(t) \times i(t) - k \times i(t)$

b represents the number of interactions for each individual per day

k represents the recovery rate from the diesease

There are some more parameters:

N : Total number of people in a population

T : Time duration

ii : initial fraction of population been infected

The ultimate purpose for a model is to serve for real world application, and coronavirus is the best example we can use and extremely care recently. One extension model we propose is to apply our basic SIR model into simulating elementary schools situation to decide whether the school should reopen or not. What I propose is to input the N(population) as the average elementary school students, I(Infectious people) and R(recovered people) to the model, as well as the average k(recover rates) from research as a baseline. After that, we can change the b(contact people) to observe what is the trend for number of infectious people. At last, we will give our conclusions and some improvements to our model.

Secondly, it is likely for people to be reinfected if people are not immune to the disease after they are recovered. Therefore, we could implement reinfection into the SIR model. For Agent-Based model, this could be implemented by adding probability of reinfection, where reinfected people would be infected

again. Alternatively, we could let people who are recovered becomes susceptible with certain probability. In addition, we could add a new variable immunization. This variable stores the probability of been immunized from the disease. For example, immunization = 0.2 means 80% chance of been reinfection. The value increases if individuals recover from the disease. For ODE model, this could be implement by adding probability of reinfection r. This has the form:

$$\frac{ds}{dt} = -b \times s(t) \times i(t)$$

$$\frac{dr}{dt} = k \times i(t) - r \times r(t)$$

$$\frac{di}{dt} = b \times s(t) \times i(t) - k \times i(t) + r \times r(t)$$

It is obvious that people wearing masks can slow down the disease spread. We can know that most of Chinese people wore face masks and the actual numbers of infections have decreased at a greater rate (Cooper et al). Therefore, we could implement the effect of the use of masks into the SIR model. For ODE model, we can implement it by reducing the parameter b. We can obtain data from to calculate the efficiency of the use of masks on the number of interactions each day that could spread the disease. Thus, we can judge whether the use of masks can slow down the spread of the virus.

Some SIR models

Basic SIR models

Agent-based model:

The key idea is that we have total number of N people and some of them are infected initially. These N people interact with b other people per day. If the person is infectious, the person who interacts with him will get infected, which can be done by the discrete method. Later, we introduce k, which is the probability of a person gets recovered. We can choose the duration we want to simulate in the model.
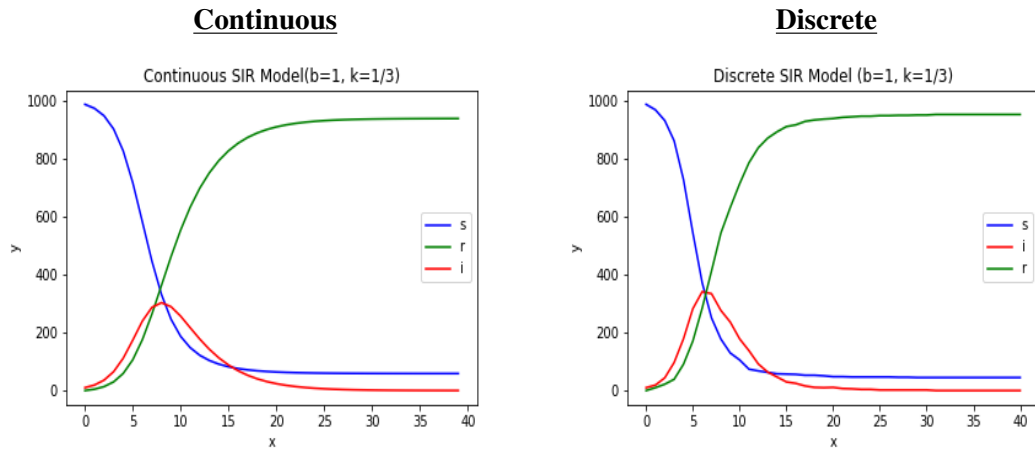
ODE model:

This model treats SIR model as initial value problem. s(t), i(t), r(t) as fraction of population with $s(t) + i(t) + r(t) = 1$ for this model. We could apply $solve_{ivp}$ from scipy to solve. The parameters are b,k, N and ii.

We fixed parameters for these simulation:
- number of days = 40
- percentage of initial infected people = 0.01
- number of people = 1000

Figure 1: First Simulation models with b=1, k=1/3

Continuous   **Discrete**

For these models, the number of susceptible people decreases exponentially until day 18 at about 70 and then keeps no change. The number of infectious people peaks at about day 9 and decrease to 0 after day 26, which is also the time that recovery people stays at about 930 after an exponential raise.
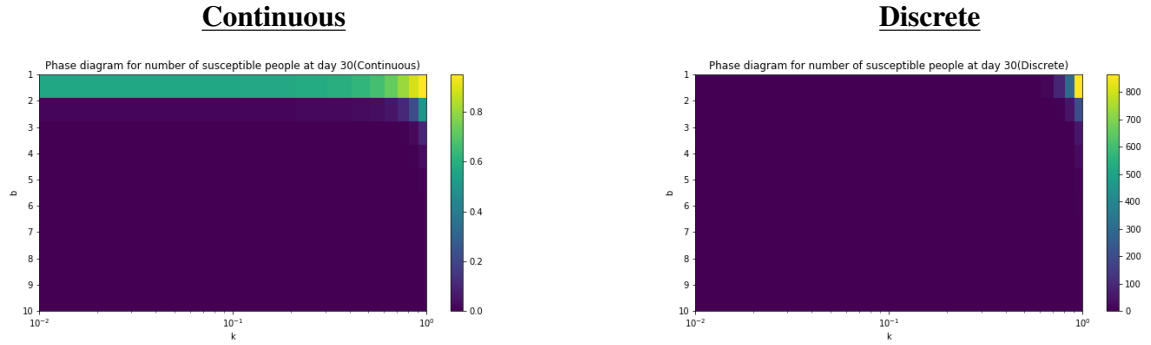
Figure 2: Second Simulation models with b=5, k=0.2

Continuous   **Discrete**
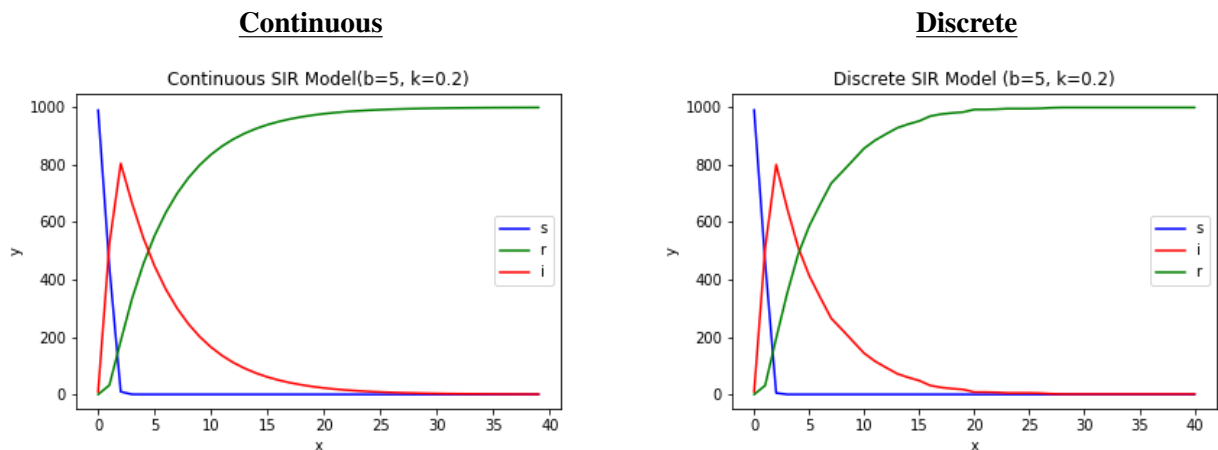
In this case, because of number of contact people increase to 5, the number of infectious people reaches its peak at the first few days with approximately 820 people. This leads to number of susceptible people drops to 0 within 3 days. The rate of recovery is relatively slow compare with the rate of infected. It takes around 24 days to recover every patient.
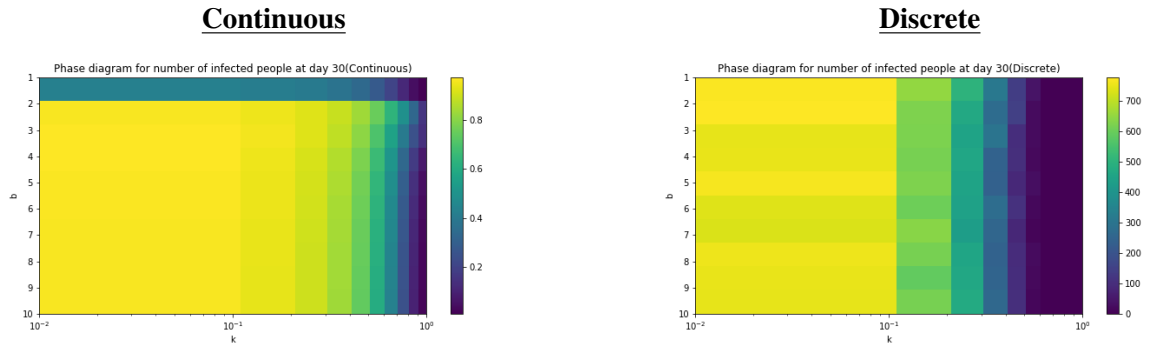
These simulations demonstrate discrete and continuous models are very similar. There are few points in the graph are not smooth is due to calculation errors from the randomness in the discrete model.

Figure 3: Phase for Susceptible

Continuous

Phase diagram for number of susceptible people at day 30(Continuous)

Discrete

Phase diagram for number of susceptible people at day 30(Discrete)
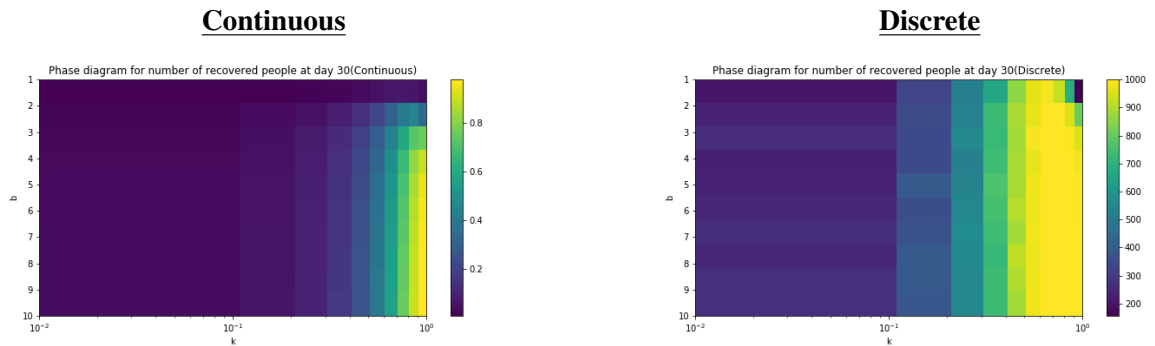
It is no hard to observe that at day 30, both models illustrate that the susceptible people are about zero in most situations with some exceptions, mainly case b = 1 in continuous model and b = k = 1 in discrete model. This should not be surprising since more people they interact with (b increases), it takes shorter time to spread the disease. At day 30, almost the entire population has already gotten.

Figure 4: Phase for Infectious

Continuous

Phase diagram for number of infected people at day 30(Continuous)

Discrete

Phase diagram for number of infected people at day 30(Discrete)

From the diagram for number of infected people at day 30, as recovery rate (k) increases, number of infected people drops. This result makes common sense because people will be removed from disease if they have higher recovery rates.

Figure 5: Phase for Recovery

Continuous

Phase diagram for number of recovered people at day 30(Continuous)

Discrete

Phase diagram for number of recovered people at day 30(Discrete)

4

In terms of the recovery people at day 30, the diagram should be nearly the complement of the infected people diagram since people are only on this three stages in our model. From the first phase diagram, almost the no people are susceptible except the some cases mentioned.

Spatial SIR model

The idea for introducing spatial models is that in some cases, disease spreading is more localized, which means people usually interact with a specific group of people (in the work or family communities). Proposing spatial models can resolve this problem to some extent.

Agent-based model:

Based on the basic agent-based model we had above, we did some slight modification to get the spatial models. We first create a 2D position vector, then a person with random position can take a step of length p in a random direction. After that, the people they interact with is all people surrounding by him with radius q. We used KDTree from *Scipy* package to achieve our goal. This way, we replaced parameters b from our previous model since the use of q describes numbers of interactions.

PDE model: Alternatively, we can tackle spatial component in SIR model through partial differential equation(PDE). Solving PDE first we create the 2-dimensional finite difference Laplacian matrix L, and we discretize the square into an M * M grid. Then we turn this into a system of ODEs by vectorising the 2-dimensional arrays s, i and r. And we can get the following equations:

$\partial s(x,t)/\partial t = -b \times s(x,t) \times i(x,t) + p \times L \times s(x,t)$
$\partial r(x,t)/\partial t = k \times i(x,t) + p \times L \times r(x,t)$
$\partial i(x,t)/\partial t = b \times s(x,t) \times i(x,t) - k \times i(x,t) + p \times L \times i(x,t)$

Figure 6: Relation between p and infected people(rate)

Discrete   **Continuous**
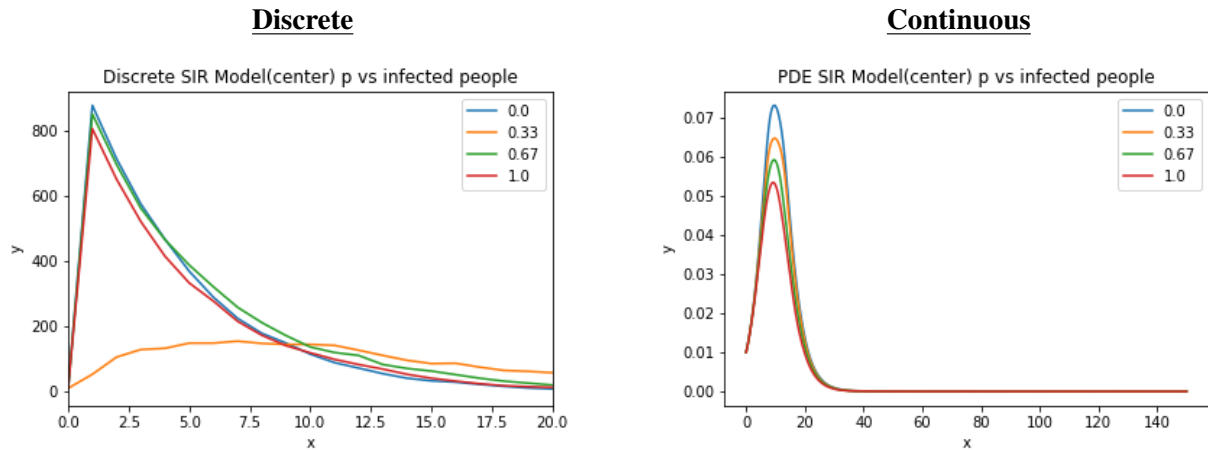
The above figure shows the evolution of infected people(rate) changes as time increases, when $b = 1, k = \frac{2}{3}$ and people initial position is at center. Both Continuous and Discrete model shows as p increases, the maximum infected people in a day decreases. This makes sense because as p increases, people are more likely to spread out, which lead to lower infected probability.

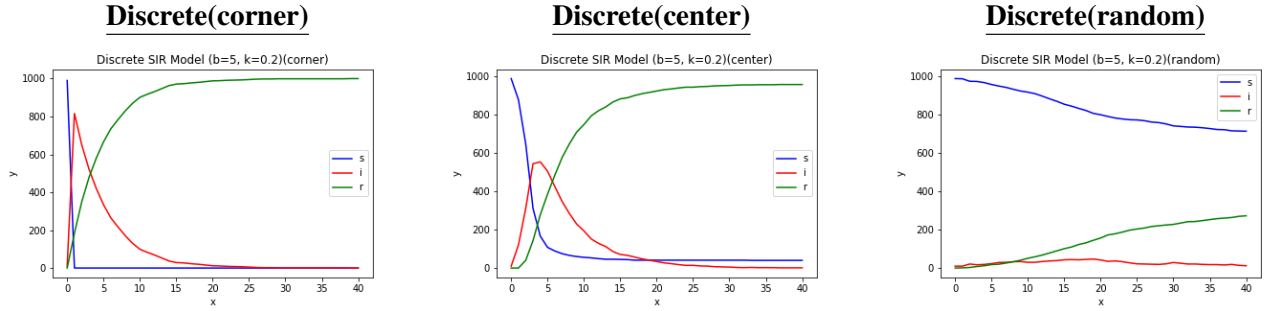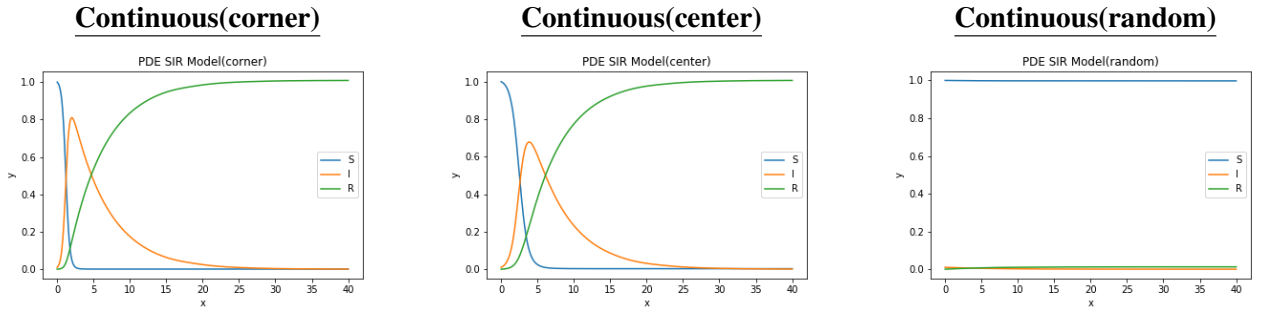Figure 7: Discrete spatial SIR with different initial position

Discrete(corner)   **Discrete(center)**   **Discrete(random)**

Figure 8: Continuous spatial SIR with different initial position

Continuous(corner)   **Continuous(center)**   **Continuous(random)**

Both continuous and discrete model shows the peak of infected people(rate) is highest when the initial position is corner, then it follows by center and lastly random. This is because when people cluster at corner, they are more likely to be infected since the area is more dense and movements are restricted. people cluster at center has less restriction on movement, which lead to lower peak for the infected people(rate). The random starting position is not compact. Therefore, it has the lowest peak of infected people(rate).

Variations investigations

Model 1

Our basic SIR model can help us determine whether we should reopen elementary schools. According to the research, the average number of students per school for all elementary schools is 473 (Public Elementary Schools). This is used as our reference for the studies. The recovery rates is about 13.425 days for young people from 0 to 19 years old, which is calculated roughly by average of 13.61 day(male) and 13.24 days(female) (Voinsky et al.). By following a Poisson distribution with $\lambda = 0.074$, we calculate the average daily recover rates k. In terms of number of contact people, this is the variable we are varying.

Consider the real situation that little children have limited self-discipliary so we are not assuming they can obey the CDC guidelines such as wearing masks properly, or keeping the 6-feet away from others. Therefore, we used our basic SIR model to do this simulation. Since both continuous and discrete models are the basically the same, we randomly choose to implement the continuous model:

Figure 9: SIR Model vs Maximum Infected People as b Varies
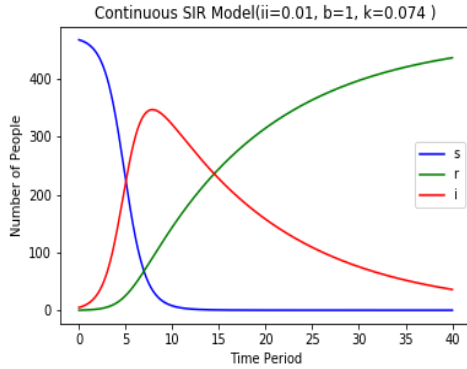
SIR for b = 1, k = 0.074   **Max infected people vs b**
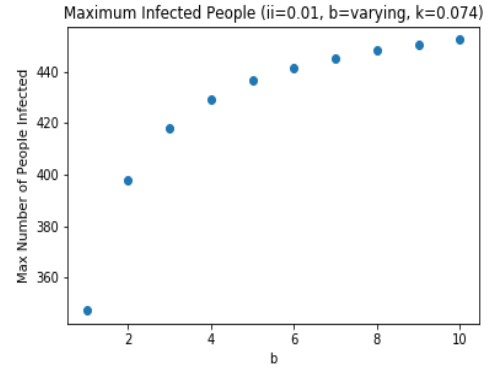
From the left figure, when ii = 0.01, b = 1, k = 0.074, one clear observation is that we have a very tall, skinny curve for infected people with peak at around Day 7 for about 350 students, which is very bad. Why? Because many people get sick simultaneously and we do not have enough time to prevent to spread the disease (Gavin, 2020).

From the figure on the right, we can see that maximum number of infected people has an increasing trend as b increases, but as b gets larger, the curve becomes more stable. This result is not surprising since when b = 1, the maximum infected people is already 350, which is 74% of the total students, therefore, as b raises, the maximum infected people is limited, which shows an increasing trend. This plot is useful in a way to give information to schools about how many people will get infected at maximum if they do not do any actions when reopening the schools.

Model 2

Reinfection is a situation when people are re-infected by a disease. This can occur when people are not immune to the disease. For some viruses, infection can provide lifelong immunity. For the other viruses, infection only provides temporary immunity, which could lead to potential reinfection. *Tillett et. al* states the first confirmed case of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) reinfection in the USA. This shows that reinfection could occur for some diseases. COVID-19 reinfection is still under investigation. *Akiko* shows a evidence of COVID-19 reinfection, while *Osmam et. al* shows previous negative results might be attributed to false-negative laboratory results and prolonged viral shedding, rather than to re-infection. Reinfection for COVID-19 is still debatable.

In this model, we would investigate the possibility of COVID 19 reinfection. We would fit COVID-19 data to SIR ODE model, which includes reinfection component.

After we fit the data to the model, we could estimate the reinfected rate *r*. We would conclude there is evidence of COVID 19 reinfection if *r* is greater than 0.05.

The implementation is completed in two steps: loss function and optimization. The loss function is defined as:

$$l(r) = ||Susceptible_{data} - Susceptible_{simulation}||_2 + ||IandR_{data} - IandR_{simulation}||_2$$

$Susceptible_{data}$ and $IandR_{data}$ is collected from United States Centers for Disease Control and Prevention.

The data only contain total number of cases in United States, which is equivalent to total number of infected

and removed in the SIR model. Therefore, we need to combine I and R for the loss function. Optimization is implemented through *Scipy*. We fix number of days to be 319, initial infected percentage to be 0.01, then we can estimate b, k, r through optimization.

Notice that for unconstrained optimization, the estimated values are $b = -0.00000966, k = -0.985, r = 0.977$, which is impractical as b and k should be non-negative. One solution for this is to fix $k = 0.074$ from model 1 and $b = 0.0001$ for simplification.

Figure 10: Figures for reinfection model

Loss Function

Covid SIR model vs data
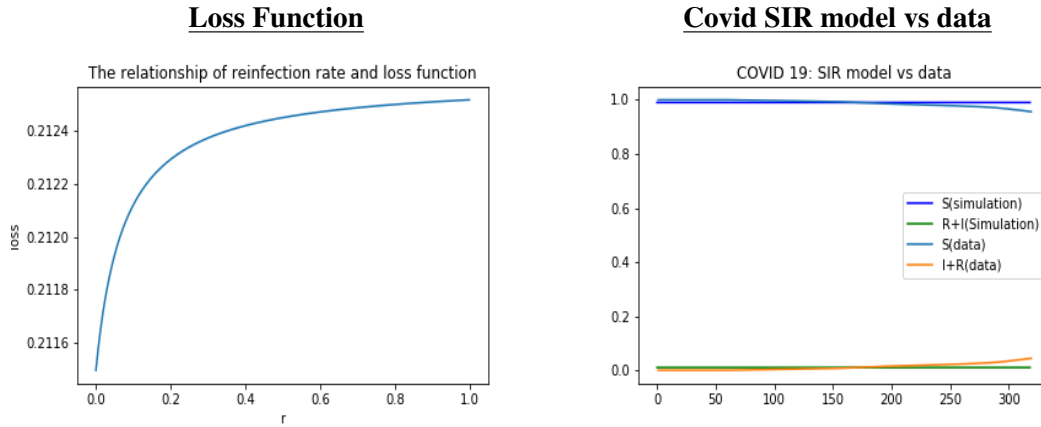
When we fix $b = 0.0001$ and $k = 0.074$, we can get the above plot by obtaining the loss through manipulate r. The figures show as reinfection rate decreases, the loss function also decreases. This implies there is no evidence of COVID 19 reinfection.

This figure shows SIR simulation with $b = 0.0001$, $k = 0.074$ and $r = 0.01$ vs COVID 19 data in the United States. The simulated SIR model fits well with the data before 200 days. The model fails to capture increase trend of infected and removed people in the United States after 200 days. Therefore, there are some room of improvement for SIR model.

Model 3

It is obvious that people wearing masks can slow down the disease spread (Cooper et al). Therefore, we could implement the effect of the use of masks into the SIR model. For ODE model, we can implement it by reducing the parameter b. We can obtain data to calculate the effect of the use of masks on the number of interactions each day that could spread the disease (C, MacIntyre, et al). Thus, we can judge whether the use of masks can slow down the spread of the virus.

We know that in Cooper's paper, the authors conducted a prospective cluster-randomized trial(Cooper et al). During the 2006 and 2007 winter seasons, 286 exposed adults from 143 households were recruited. They found that during a severe pandemic when use of face masks is greater, pandemic transmission in households could be reduced.

We introduce a new parameter q (0<q<1) to reduce the parameter b, then we can get the following equations:
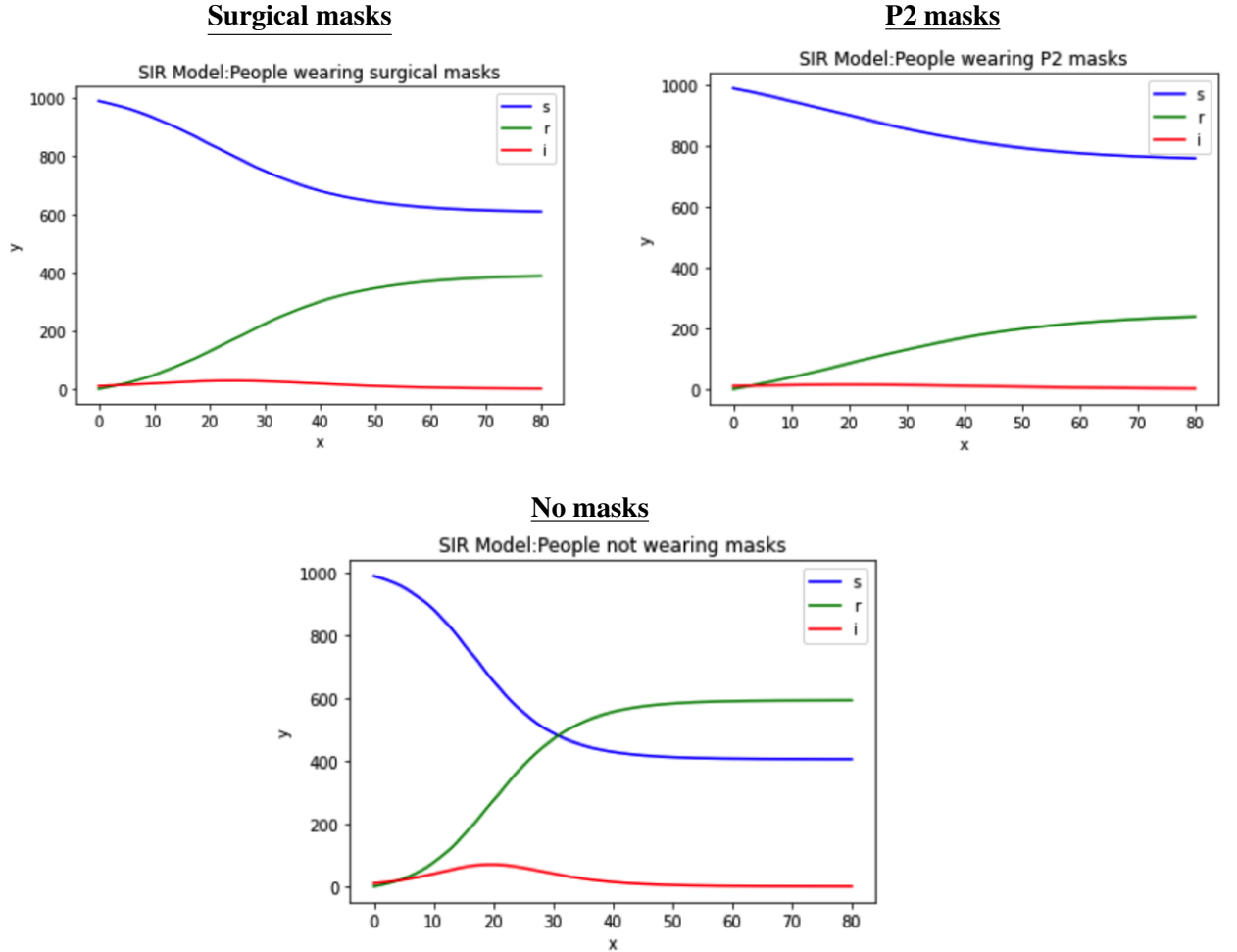
$$\frac{ds}{dt} = -b \times s(t) \times i(t)$$

8

$$\frac{dr}{dt} = k \times i(t)$$

$$\frac{di}{dt} = b \times q \times s(t) \times i(t) - k \times i(t)$$

We have the parameter b = 1/2, k = 1/3 and the initial condition s = 0.99, r = 0, i = 0.01. From the figure shown in Appendix, we know that when we use surgical mask, q ≈ 0.83. And when we use P2 mask, q ≈ 0.74.

Figure 11: Different mask conditions

Surgical masks

P2 masks

No masks

From the plot above, we know that the use of masks can slow down the spread of the virus from seeing there is a small peak at Day 20 while we do not observe that in the cases of wearing mask. In addition, we can compare with the P2 masks and surgical masks and we can get the conclusion that people wearing P2 masks is better than wearing surgical masks since the infectious curve is more flat in P2 mask situation.

Conclusion

To conclude, our basic SIR models in 1D and 2D provide us with clear understanding about how many people will be infected/recovered during a specific period of time by manipulating variables. Both models

indicate that as b increases, the peak for infected people takes place much earlier. When recovery rate raises, infected people drops since more sick people recovered from disease so they are not infectious anymore as well.

Spatial models add more dynamic to the SIR model. We observe from different starting position, the resulting model will be noticeably different. Parameter p also affects the model significantly. As p increases, the maximum infected people(rate) decreases. One limitation of the spatial models is there are some randomness involve in the discrete model, which leads to some fluctuation in the figures. One could address this by implement Monte Carlo method, which can tackle randomness effectively.

According to what we investigated above for model 1, it is dangerous to reopen the school at this point, if students cannot strictly follow the public health guidelines. Even with just the initial infectious people is 0.01, and each student interact with 1 other, 74% of the students in the school will get sick at Day 7, which is a very scary result. One improvement can be considered with this model is that we haven't consider the factor that sick students are not allowed to go back to school. This will slow down our results of infected people since they will not interact with other students.

For SIR model with reinfection component, we show that there is no evidence of COVID 19 reinfection. For continuous model, the reinfection rate is fixed. This neglects that fact that immune system enhance after people are recovered from diseases. Vaccine is am example for this. Discrete model addresses this issue by decreasing reinfection rate when people recover. However, there is no simple solution for continuous model.

In terms of the third model, we can generate the following two conclusions. First, the use of masks can slow down the spread of the virus. Moreover, people wearing P2 masks is better than wearing surgical masks. A limitation here is that this trial only compared the effects of two masks and we can also compare the effects of various masks to determine which mask works best. In addition, we can introduce a new parameter d to increase the social distance between people to slow down the spread of the virus, which can make the model more complete.

Bibliography

1. C, MacIntyre, et al. "Face mask use and control of respiratory virus transmission in households." National Library of Medicine. Published by National Center for Biotechnology Information., Feb. 2009, https://pubmed.ncbi.nlm.nih.gov/19193267/.

2. Cooper, Ian, et al. "A SIR model assumption for the spread of COVID-19 in different communities." Chaos Solitons Fractals. Published by Elsevier Ltd., Oct. 2020, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7321055/.

3. Gavin, K., 2020. Flattening The Curve For COVID-19: What Does It Mean And How Can You Help?. [online] Healthblog.uofmhealth.org. Available at: <https://healthblog.uofmhealth.org/wellness-prevention/flattening-curve-for-covid-19-what-does-it-mean-and-how-can-you-help> [Accessed 8 December 2020].

4. Osman A.A. and Al Daajani M.M. and Alsahafi A.J. (2020). Re-positive coronavirus disease 2019 PCR test: could it be a reinfection? *New Microbes and New Infections* (37)

5. Public Elementary Schools, by Grade Span, Average School Size, and State or Jurisdiction: 2005-06. National Center for Education Statistics, U.S. Department of Education,

6. Tillett R.L., Sevinsky J.R., Hartley P.D., et al. (2020). Genomic evidence for reinfection with SARS-CoV-2: a case study. *Lancet Infect Diseases*;

7. Iwasaki A. (2020). What reinfections mean for COVID-19. *The Lancet Infectious Diseases*.

8. Voinsky, Irena, et al. "Effects of Age and Sex on Recovery from COVID-19: Analysis of 5769 Israeli Patients." The Journal of Infection, The British Infection Association. Published by Elsevier Ltd., Aug. 2020, www.ncbi.nlm.nih.gov/pmc/articles/PMC7229949/.

Appendix

Figure 1: Summary statistics of wearing/no wearing masks

Variable	Control group, no. (%), n = 50	Surgical mask group No. (%), n = 47	p value	P2 mask group No. (%), n = 46	p value
Living arrangement					
Reside in house	38 (76)	32 (68)	0.39	33 (72)	0.64
>4 persons in house	13 (26)	18 (38)	0.20	19 (41)	0.11
≥3 adults in house	8 (16)	11 (23)	0.36	12 (26)	0.23

Citied: C, MacIntyre, et al. "Face mask use and control of respiratory virus transmission in households." National Library of Medicine. Published by National Center for Biotechnology Information., Feb. 2009, https://pubmed.ncbi.nlm.nih.gov/19193267/.