Refined SIR Model with Vaccination and its Application in 2022 NYC Influenza A Activity Prediction

Shizhao Yang

Abstract

In this project, I seek to provide a comparison between the influenza A activities in the past five years and predict the influenza trend this year in New York. I will start by analyzing a refined SIR model with the consideration of vaccination through a series of ODE methods. Then, the refined SIR model will be adopted to analyze and predict the relevant flu data. Finally, the model's limitations and possible extensions will be discussed.

1 Introduction

Seasonal Influenza, more commonly referred to as the flu, is an infectious disease caused by influenza viruses that can afflict individuals with mild to severe symptoms as well as life-threatening complications, including death. Due to virus characteristics and transmission spread by virus-containing respiratory droplets and aerosols, influenza outbreaks often occur from October of each year to May of the following year, peaking in January and February. In addition to vaccine and pharmaceutical development, the monitoring and forecasting of the epidemic by mathematical modeling at the macro level is equally important for it can be used by the government to make policies to slow and reduce outbreak peaks.

The use of compartmental models in infectious diseases began in the early 20th century and has been proven its feasibility and widely applied to different diseases such as Flu, COVID-19, and Ebola[1]. In the basic SIR model, the total population is divided into three compartments: Susceptible (S), Infectious (I), or Removed (R). Then, the model often runs with deterministic ordinary differential equations and gives predictions of the spread of the diseases.

According to The Centers for Disease Control and Prevention (CDC), the seasonal flu outbreak in New York this year is the most rapid and severest in the past six years. Therefore, monitoring and prediction of the disease are especially important. In this project, I aim to use a derivative of the SIR model with the consideration of the impact of vaccination to analyze and predict the influenza A trend and compare it to the past years.

2 Model Description

The classic epidemic SIR model is adopted from Hethcote's paper[2]. On top of that, inspired by a SIR model with vaccination that is used in Covid-19[3], I add two parameters: vaccination rate v and vaccine efficacy rate η. Therefore, the deterministic SIR model with vaccination, given the initial values is:

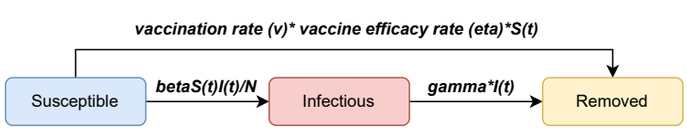


Figure 1: SIRV model diagram

$$\frac{dS}{dt} = -\frac{\beta IS}{N} - v\eta S \qquad S(0) = S_0 \geq 0 \tag{1}$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I \qquad I(0) = I_0 \geq 0 \tag{2}$$

$$\frac{dR}{dt} = \gamma I + v\eta S \qquad R(0) = R_0 \geq 0 \tag{3}$$

$$N = S + I + R. \tag{4}$$

Assuming the population N is closed, i.e. no births, emigration, and immigration, we divide it into three compartments at the time t: the number of susceptible individuals $S(t)$, which can be infected by infectious people; the number of infectious individuals $I(t)$; and the number of removed individuals $R(t)$, which includes the number of dead, recovered people and those who immunized through vaccines.

As Figure 1 shows, the transition from different compartments can be described as the following. First, the susceptible individuals will become (1) infectious at a speed of $\frac{\beta IS}{N}$, where transmission rate β is the average number of adequate contacts per person, per unit time; or become (2) removed individuals through vaccination, where vaccination rate v is the proportion of susceptible people who get vaccinated per unit time (note that we assume v is a constant for simplification), and vaccine efficacy rate η denotes the possibility that an individual is immune to the flu by injection. Then, the infectious individuals will become removed at a speed of γI, where removal rate γ is the inverse average removal time. The parameters and their meaning are shown in Table 1.

notation	meaning
S	number of susceptible individuals
I	number of infectious individuals
R	number of removed individuals
N	closed population size
β	transmission rate
γ	removal rate
v	vaccination rate
η	vaccine efficacy rate
Re_0	basic reproduction number
$Re(t)$	replacement number

Table 1: Notations (note that we assume vaccination rate to be a constant in this project for simplification.)

2.1 Basic Reproduction Number and Replacement Number

Consider the scenario in which a single infectious agent is introduced into a completely susceptible host population,i.e. $I_0 = 1$, $S_0 = N - 1$. Therefore, the growth rate of infectious individuals at $t = 0$ is,

$$\frac{dI(0)}{dt} = \frac{\beta I_0 S_0}{N} - \gamma I_0 = \gamma(Re_0 - 1),$$

where

$$Re_0 \equiv \frac{\beta S_0}{\gamma N} \tag{5}$$

is defined as **Basic Reproduction Number** Re_0[2]. $Re_0 = 1$ is the critical value to determine whether an epidemic would occur, namely, if $Re_0 > 1$, I grows exponentially, otherwise, the epidemic wouldn't occur.

Define the **Replacement Number** $Re(t)$ as the average number of additional infectives per infective at any time during the epidemic, we have,

$$Re(t) \equiv \frac{\beta S(t)}{\gamma N} \tag{6}$$

where $Re(0) = Re_0$ at time $t = 0$.

The requirement for **Herd Immunity** to take place is $Re(t) < 1$, when the epidemic will slowly abate. Since equation (6), we have,

$$\begin{aligned} Re(t) &< 1 \\ \frac{\beta S(t)}{\gamma N} &< 1 \\ \frac{S(t)}{N} &< \frac{1}{Re_0} \end{aligned}$$

Therefore, for herd immunity to happen, the ratio between $S(t)$ and N must smaller than $1/Re_0$.

2.2 Fixed Points and Phase Plane

2.2.1 Fixed Points

To analyze the fixed points, first write the equations as a 2-dimensional system.

$$\dot{S} = -\frac{\beta SI}{N} - v\eta S \tag{7}$$

$$\dot{I} = \frac{\beta SI}{N} - \gamma I \tag{8}$$

therefore, R can be calculated by $R = N - S - I$ and the region T for the system is

$$T = \{(S, I) | S \geq 0, I \geq 0, S + I \leq N\}$$

By solving the null-clines function, i.e. $\dot{S} = 0$ and $\dot{I} = 0$ at the same time, one set of fixed points will be:

$$(S^*, I^*) = (0, 0)$$

The Jacobian Matrix for fixed point (S^*, I^*) is given by:

$$\mathcal{J}(S^*, I^*) = \begin{bmatrix} -\frac{\beta}{N} I - v\eta & -\frac{\beta}{N} S \\ \frac{\beta}{N} I & \frac{\beta}{N} S - \gamma \end{bmatrix}_{(S^*, I^*)}$$

Thus,

$$\mathcal{J}(0, 0) = \begin{bmatrix} -v\eta & 0 \\ 0 & -\gamma \end{bmatrix}. \tag{9}$$

To calculate the eigenvalues, let $\det(J(0,0) - \lambda I) = 0$, we have,

$$\begin{aligned} \det(J(0,0) - \lambda I) &= 0 \\ \begin{vmatrix} -v\eta - \lambda & 0 \\ 0 & -\gamma - \lambda \end{vmatrix} &= 0 \\ (-v\eta - \lambda)(-\gamma - \lambda) &= 0 \\ \Rightarrow \lambda_1 = -v\eta,\ \lambda_2 &= -\gamma \end{aligned}$$

Since the parameters v, η, and γ are positive, we have two negative eigenvalues which implies the fixed point $(0, 0)$ is a stable node. The analysis shows that, at the end of the epidemic, there will be no susceptible or infected individuals. Individuals will either be vaccinated or recovered or dead.

2.2.2 Phase Plane

To better articulate, I used python to draw the phase plane of a model, in which the parameters are set as $N = 1e7$, which is approximate to the population of NYC; $\beta = 0.5$; $\gamma = 0.2$; $v\eta = 0.0005$; and $T = 35$ weeks (average period for a flu season). As Figure 2 shows, in most cases, the susceptible and infectious populations will both converge to the fixed points $(0, 0)$, given 35 weeks. However, if the initial infectious population is too large, then the epidemic may need a longer time to die out.

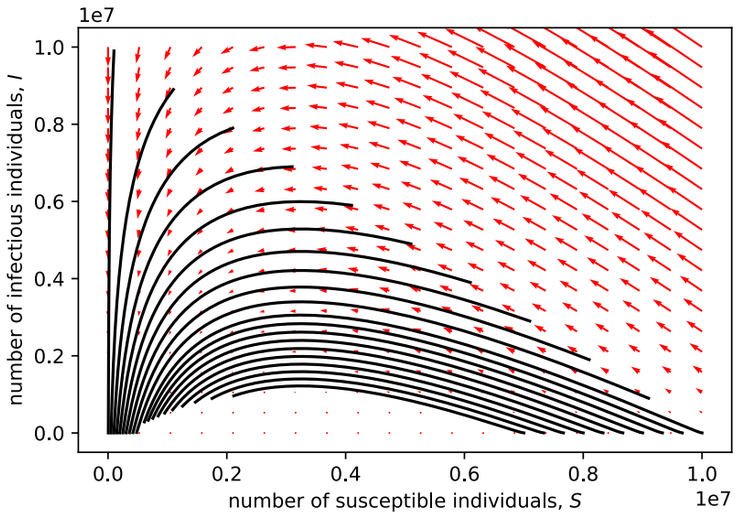


Figure 2: Phase Plane

3 Result

3.1 Dataset

The influenza dataset is downloaded from New York Health Data website[1]. The dataset contains weekly counts of laboratory-confirmed influenza cases, by county and influenza type, since 2009-10 season.

According to The Centers for Disease Control and Prevention (CDC), there are four types of Influenza viruses: A, B, C, and D. Among these, Influenza A and B cause seasonal epidemics in the United States. Since individuals who are immune to type A can still be infected by type B, the confirmed cases caused by different types cannot be counted together. **Therefore, in this project, I will only focus on the analysis of Influenza A, since it takes up 70% of the total confirmed cases.**

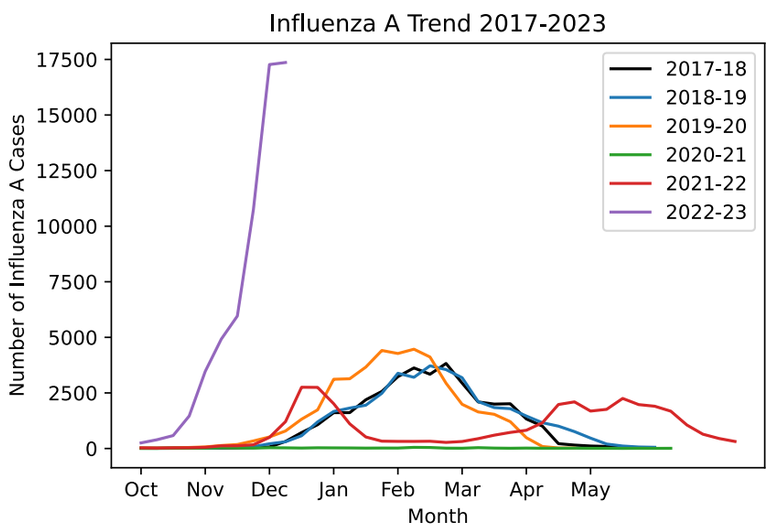


Figure 3: Trend of Confirmed Influenza A Cases in the Past 6 Flu Seasons

Figure 3 depicts the trends of confirmed Influenza A cases in NYC for the past

[1]New York influenza data: `https://health.data.ny.gov/Health/Influenza-Laboratory-Confirmed-Cases-By-County-Beg/jr8b-6gh6/`

6 flu seasons, where a flu season is defined as the 40th week of a year (usually the 1st week of October) to the 20th week of the following year (usually the 3rd week of May) according to CDC. The count data is collected weekly, so we have an average of 35 data points for each flu season from 2017 to 2022, and 10 data points for season 2022-2023 by 12/16/2022. As the figure shows, the flu trends of 2020-2021 and 2021-2022 are very irregular due to the covid-19 pandemic, which cannot provide useful information about the Influenza A disease. **Therefore, we will only focus on analyzing the data of 4 seasons: 2017-18, 2018-19, 2019-20, and 2022-2023.**

3.2 Parameter Estimation

After data munging, I used the python curve_fit function to estimate the transmission rate β and removal rate γ. Before running the optimization function, one important part is to set up other parameters. As Table 2 shows, the total closed population N of NYC is 8468000; the vaccination rate is 0.005 according to health data NY website, meaning that there is an average 0.5% of the total susceptible population that get vaccinated per day; and according to the CDC, the vaccine efficacy rate η of influenza A vaccine is about 40% to 60%, so I set $\eta = 0.5$, meaning that an individual has 50% chance to be fully immune to influenza A after injection.

parameter	value
N	8468000 (NYC)
v	0.005
η	0.5

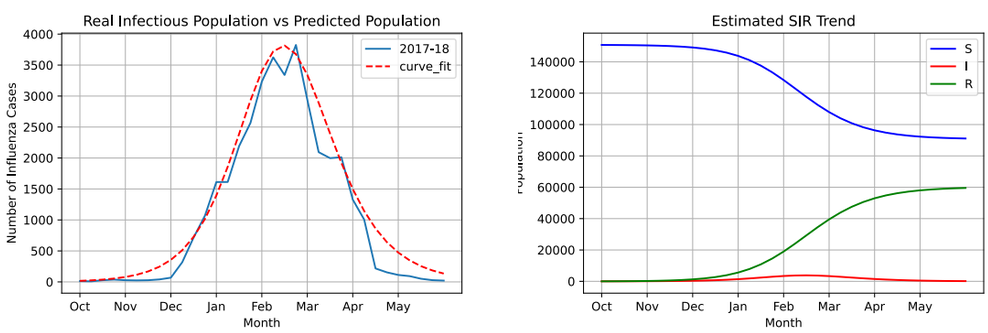
Table 2: Parameter

Figure[4, 5, 6] shows the estimation results for seasons 2017-18, 2018-19, and 2019-20: each figure (a) shows the fitting results of the infectious population with the dashed line representing the estimated model and valid line representing the real data, and figure (b) shows the estimated SIR trend. To evaluate the fitting quality, I adopt the RMSE(root of mean squared error) method:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \parallel y(i) - \hat{y}(i) \parallel^2}{N}}.$$

The RMSE for the estimation results is 300.1, 209.4, and 301.4 respectively. Given a large population base number, the RMSE results show that the estimation can reflect the real situation objectively.

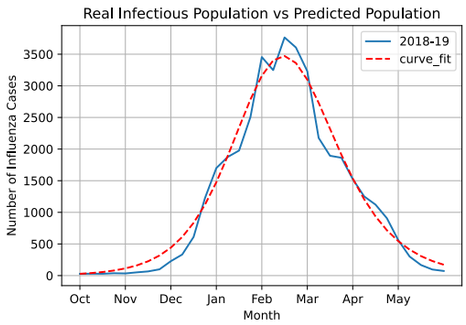
As Table 3 shows, the transmission rate β and the removal rate γ of the seasons from 2017 to 2020 are very stable, with averages of 0.253 and 0.199 respectively. This leads to a stable basic reproduction number Re_0 of those past three years, averaging 1.273. Besides, we can see from the plots that the peak of the infectious population usually happened between mid-February.



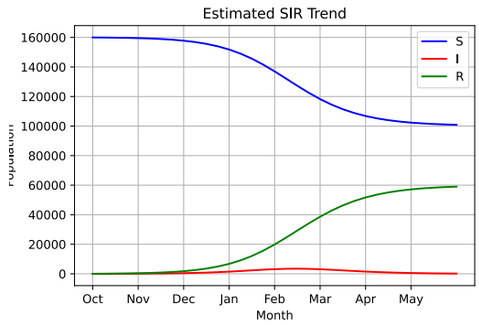
(a) Real Infectious Population vs Predicted Population

(b) Estimated SIR Trend

Figure 4: 2017-18

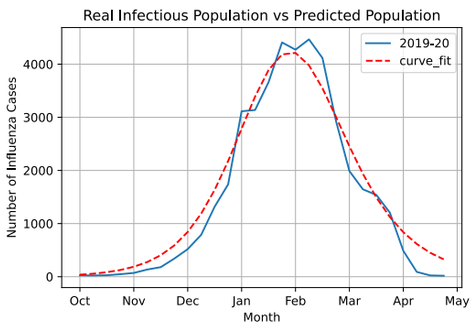


(a) Real Infectious Population vs Predicted Population

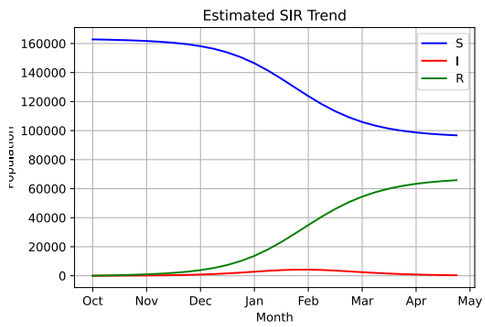


(b) Estimated SIR Trend

Figure 5: 2018-19



(a) Real Infectious Population vs Predicted Population



(b) Estimated SIR Trend

Figure 6: 2019-20

	I_0	β	γ	Re_0
2017-18	17	0.255	0.199	1.276
2018-19	28	0.250	0.199	1.252
2019-20	38	0.256	0.199	1.292
2022-23	290	0.278	0.187	1.484

Table 3: Result

3.3 Prediction of this Year

After investigating the influenza A epidemics of the past seasons, the same method was implemented on the elevating flu epidemic data of this year. According to Table 3, the transmission rate $\beta = 0.278$, 10% larger than the average of the past three years, meanwhile the removal rate γ is less than the average of the past, which give rise to a larger $Re_0 = 1.484$. The statistics show that influenza A this year has a stronger spreading ability and takes a little longer for people to recover. Together with a large initial infectious number I_0, it explains why the epidemic trend of this year is much steeper than usual as it's showed in Figure 7.

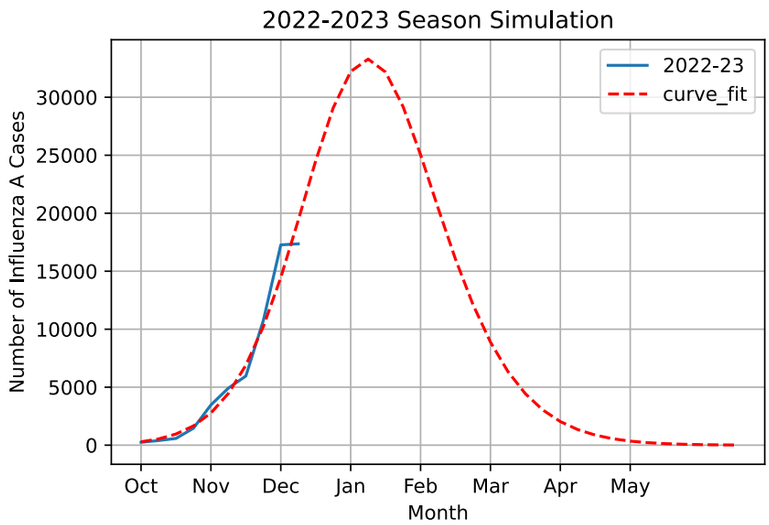


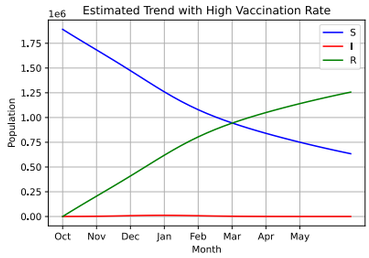
Figure 7: 2022-2023

3.4 Effect of Vaccination in Epidemic

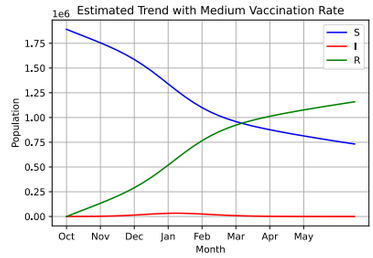
To further investigate the effect of vaccination rate under the condition of this year, I plotted the SIR trend and calculated the infection peak and its related time under three different vaccination rates v: 0.008 (high), 0.005 (medium), and 0.001 (low). Note that the medium rate is the real value mentioned in Table 2.

As shown in figure 8a, 8b, and 8c, with fixed transmission rate $\beta = 0.278$, removal rate $\gamma = 0.187$, and vaccine efficacy rate η, a higher vaccination rate would lead the susceptible population to decrease and a smaller infection peak since more individuals will be vaccinated and moved from susceptible population to removed ones.

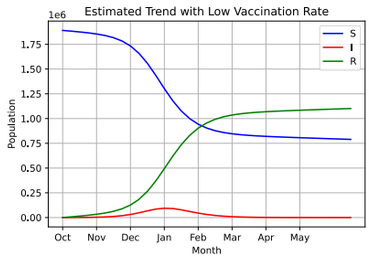
From figure 8d and Table 4, the models show that if we keep the current vaccination rate $v = 0.005$ (medium), then the estimated infection peak would be 33292.29 happening on 2022-12-24. However, if the vaccination rate v is 0.008 (high) at the beginning of the epidemic and remains, the infection peak would be reduced to 12167.38. On the contrary, if the vaccination rate v is low (0.001), then the infection peak would be as 3 times many (946065.95) as the medium v. The analysis demonstrates that uplifting the vaccination rate, will greatly reduce the infection peak, however, it wouldn't postpone or advance the infection peak date, which is in line with equation(2).



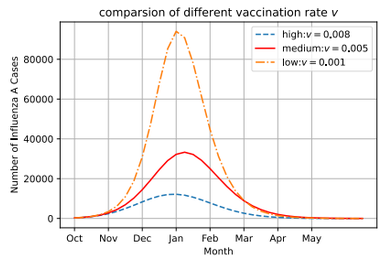
(a) High Vaccination Rate: $v = 0.008$



(b) Medium Vaccination Rate: $v = 0.005$



(c) Low Vaccination Rate: $v = 0.001$



(d) Infectious number under different vaccination rate v

Figure 8: Analysis of different vaccination rates

vaccination rate v	infection peak	infection peak time
high (0.008)	12167.38	2022-12-24(week 51)
medium (0.005)	33292.29	2022-12-31(week 52)
low (0.001)	94065.95	2022-12-24(week 51)

Table 4: infection peak and related time under different vaccination rates

4 Conclusion

In this project, I have successfully analyzed the refined SIR model with vaccination, from aspects of fixed points and its stability, phase plane, and related indexes such as reproduction number. Then, the model is adopted to estimate the transmission rate β, removal rate γ, and SIR trend of the influenza A epidemics of this season and past seasons (see Figure[4, 5, 6, 7] and Table 3). Finally, the project focused on the effect of vaccination rate in varying infection peaks and its date(see Figure[8] and Table 4).

4.1 limitation

Although the project gives us some insights into the influenza A trend of NYC for now and past seasons by using this refined SIR model with vaccination, there are still several limitations to this model.

First, due to the limit time issue, I set the vaccination rate v as a constant for simplification, which is usually not the case in real life. For better modeling, the epidemic, the vaccination rate $v(t)$ should be a function of t, and according to past statistics, $v(t)$ usually grows fast at the beginning of each flu season (October to December) and becomes steady as time passes. Besides, it may be better to take the exposed term $E(t)$ into consideration to make the model more realistic. Second, due to the constrain of the number of data points for each season, the estimation has a rather high RMSE. The situation will be better if the government increases the frequency of flu data collection.

5 Acknowledgement

Special thanks to Professor K Shafer Smith and TA Dave Connelly, for your devotion to this semester and valuable suggestions for this final project.

References

[1] Wikipedia contributors, "Compartmental models in epidemiology — Wikipedia, the free encyclopedia," 2022, [Online; accessed 16-December-2022]. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Compartmental_models_in_epidemiology&oldid=1127422701

[2] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Review*, vol. 42, no. 4, pp. 599–653, 2000. [Online]. Available: https://doi.org/10.1137/S0036144500371907

[3] T. T. Marinov and R. S. Marinova, "Adaptive sir model with vaccination: Simultaneous identification of rates and functions illustrated with covid-19," *Scientific Reports*, vol. 12, no. 1, pp. 1–13, 2022.