Stochastic,Meta-population models,and Lattice-based models

Qinghe Gao, Wenlin Chen

***Abstract*—Basically, SIR model can be categorized as deterministic model and stochastic model. In this report, we discussed the stochastic model first and explore the five hallmarks by Gillespie's direct algorithm. Variability, Variance and Covariance, Increased transient, Stochastic resonance and Extinction.For the spatial model, firstly we used Meta-population model to analyze the delay time and big ρ can reduce the delay time. For the lattice model,the speed of infection is linear. The basic reproductive rate R_0 and interaction rate ρ controls the speed.**

I. Introduction and Background

EPIDEMIOLOGY modelling's research is a difficult road.[1] Basically, the epidemic model can be categorized as deterministic model and stochastic model. At report 1 we discussed the deterministic model in details. Now we will discuss the stochastic model in this report.

Deterministic outcome is always determined when the initial condition was given, while stochastic model's output is always different and random. Especially when the population size is small, there are larger possibility for uncertain events, like extinction. It's important to incorporate the stochastic situation into the model.

Basically, the way to establish the stochastic model can be categorized by adding noise to deterministic model and event drive(Gillespie's algorithm). In the first part of this report, we use Gillespie's algorithm to explore stochastic model's five key hallmarks: Variability, Variance and Covariance, Increased transients, Stochastic resonance, Extinction, based on naive SIR model, SIR model with demography and SIR model with demography and import.

In real life, most of disease transmission is a localized process. So we should use spatial model to describe the process of transmission. In the second part, We use Meta-population to explore the dynamic and the delay phenomenon in disease spread between subpopulations.

Lattice model is a special Meta-population. In the third part, We analyze this model to explore the speed of infection, which can provide the fundamental process how the disease spread and become a epidemic. for further, we study how the speed change with different value of some parameters.

II. Methodology and Discussion

A. Stochastic model and Gillespie's algorithm

Stochastic model have five hallmarks which are distinguish from the deterministic model:

1) Variability. Different simulation of stochastic models produce different outcomes. So it is hard to predict the spread of the disease.
2) Variance and Covariance. Basically, there is negative covariance between the susceptible and infectious people because the interaction between the stochastic and deterministic model. The variance is caused by natural perturbation.
3) Increased transient. When the stochastic model is far way from the equilibrium, there is a strong and powerful restorative force to let its transient-like return. This dynamic of stochastic model is countered by deterministic force, which pull the perturbation back.
4) Stochastic resonance. Deterministic model will decay when approaches the equilibrium. Stochastic model will remain the oscillation behaviour because of the transient-like dynamics. And this oscillation's frequency is close to the natural frequency and it is sustained.
5) Extinction. The fluctuation of infectious people will lead to a extinction in a small group of population. To keep the disease spread for continuous observation, the import of infectious outside is needed.

How to analyze these five hallmarks? Basically, there are two ways: Incorporating noises into deterministic model or using Event-driven method. In this report, we use event-driven method. Basic methodology of event-driven is Gillespie's direct algorithm:

1) List all the events. E_1,E_2 $\cdots E_n$.
2) List all the corresponding rates. R_1,R_2 $\cdots R_n$.
3) R_{total}= $\sum_{i=1}^{N} R_i$.
4) Generate a new time interval δ t=$\frac{-1}{R_{total}}$ $\log(RAND1)$
5) Generate a new incident probability P from random number: RAND2. P=R_{total}* RAND2.
6) Event p occur:
$$\sum_{m=1}^{p-1} R_m < P \leq \sum_{m=1}^{p} R_m$$
7) Return to step2

We applied this algorithm into the naive SIR model. As the equation 1,2,3 shows, the events here are transmission and recovery. And the rates are transmission rate:βXY/N and recovery rate:γY. So the total rate is the sum of βXY/N and γY.

$$\frac{dX}{dt} = -\beta XY/N \tag{1}$$

$$\frac{dY}{dt} = \beta XY/N - \gamma Y \tag{2}$$

$$\frac{dZ}{dt} = \gamma Y \tag{3}$$

If the event occurs randomly, it is reasonably to assume that they are Possion distributed. Since the occur of the events are Possion distribution, if there is no event in (0,t), the probability $P(A_1> t)=e^{-\lambda t}$ Thus the probability that the event happens in (0,t) is $P(A_1< t)=1-e^{-\lambda t}$. In this way we can get the cumulative distribution function is:

$$F(X) = P(X < t) \begin{cases} 1 - e^{\lambda t} \\ 0 \end{cases} \tag{4}$$

In this way, the probability distribution function is exactly the exponential distribution. It is reasonably to assume the tiny t obey this probability distribution function. let's assume $t=F^{-1}(\upsilon)$. In this way, every t can be produced with a random uniform number υ, we can rewrite this equation as below.

$$t = F^{-1}(\upsilon) = \frac{1}{\lambda} ln(1 - \upsilon) \tag{5}$$

In this way, we can plot the naive stochastic SIR model as the figure 1 shows.

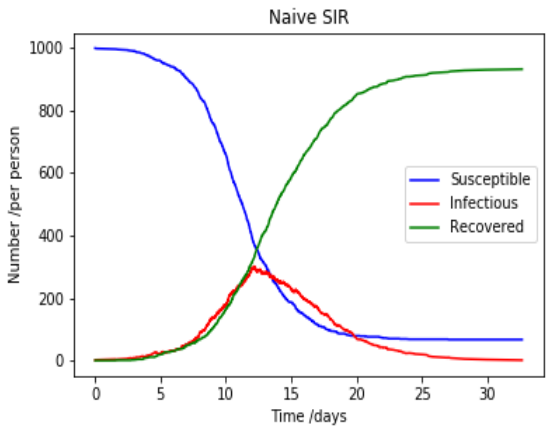

Fig. 1. Naive stochastic SIR model. β =1 per day, 1/γ =3 per day, X_0 =999, Y_0 =1

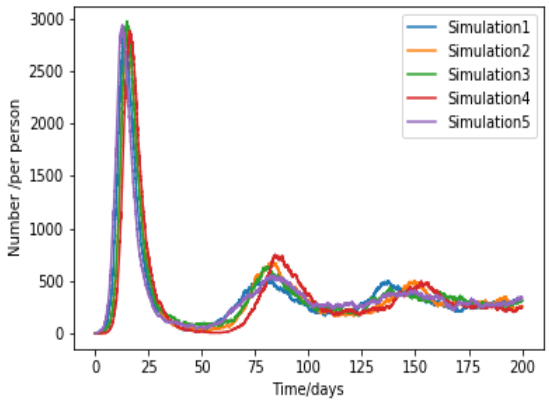

Fig. 2. Infectious individuals of SIR model with demography. four different simulation. X_0=9999,Y_0=1

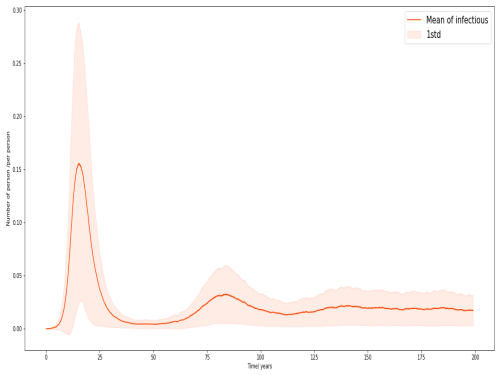

Fig. 3. Infectious individuals of SIR model with demography. Five different simulation.Purple curve is mean, purple shaded is 1 standard deviation

B. *Five hallmarks*

*1) **Variability** :* Different simulation can produce different outcomes, which is the distinguish feature of stochastic model compared with deterministic model. As the Fig.2 shows, we used the same steps of Gillespie's direct algorithm to model stochastic SIR with demography (Equation 6,7,8). We run the simulations five times and different outcome has been arrived.

$$\frac{dX}{dt} = \nu - \beta XY/N - \mu X \tag{6}$$

$$\frac{dY}{dt} = \beta XY/N - \gamma Y - \mu Y \tag{7}$$

$$\frac{dZ}{dt} = \gamma Y - \mu Z \tag{8}$$

Then we run the simulation for five times. With the help of interpolation method, the fitted curve of each simulation is obtained. Then the mean and variance of five simulations are calculated out with the data recorded at the same time plots. As the Fig.3 shows, The red line is the mean curve and the light red shade means 1 standard deviation away from the average value, which also shows the variability of stochastic model.

*2) **Variance and Covariance** :* First we want to know what affects the variance of noise. When we changed the size of population, the variance of noise increase. Just like the Fig.4 shows.

Since this is stochastic model, the random process will lead to variance. It is clearly to see that the variance of infectious are increased as the variance of the noise increased.(Fig.5) Because when there are stronger perturbations, the number of infectious people will fluctuate greatly, which means that the variance of infectious people will increase.

Besides, we can observe the negative covariacne between infectious and susceptible(Fig.6). This is because the stochastic model is also a single-direction spread model like deterministic model, from susceptible to infectious. And the noise sometimes performs better on infectious than the susceptible. Hence, increasing noise make amplitude fluctuations with a strong negative covariance between X and Y.

Since this negative covariance of X and Y, the transmission rate, which is the production of β, X and Y, will decay. Then the mean number of susceptible people will increase as the variance of noise increase(Fig.7)

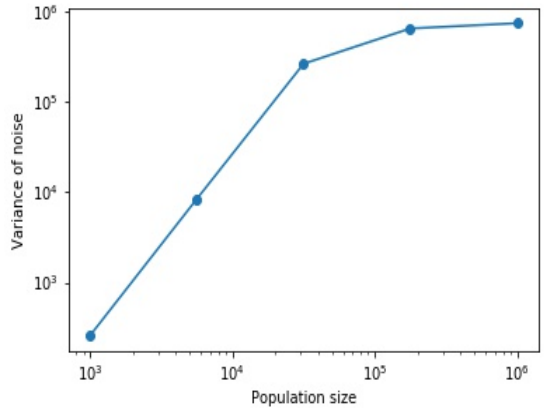

Fig. 4. The correlation of size of population and variance of noise

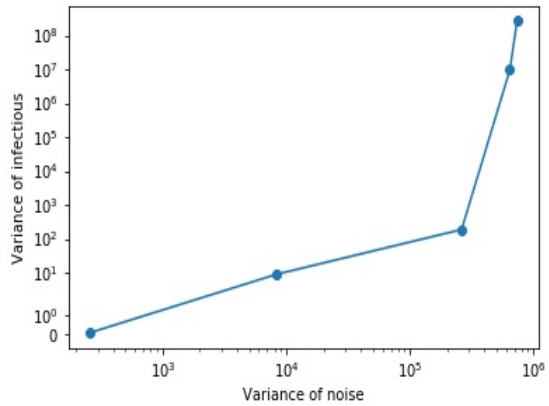

Fig. 5. The correlation of infectious' variance and variance of noise

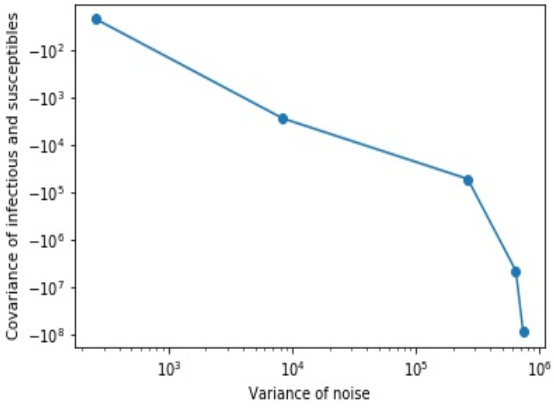

Fig. 6. The correlation of Covariance and variance of noise

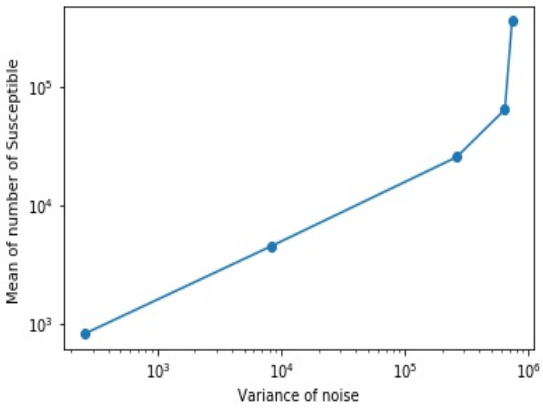

Fig. 7. The correlation of susceptibles variance and variance of noise

*3) **Increased Transient** :* The stochastic model is fully random. But when the curve is greatly away from the equilibrium, there is a restorative force to pull it back. As the Fig.8 shows, while the deterministic model will coverage into a fixed point, the stochastic model will fluctuate around the fixed point. And the stochastic values will not be far away from the fixed point. Fig.9 also show the same increased transient dynamic. The amount of deterministic infectious people remain constant, while the amount of stochastic infectious people doesn't remain constant. It will also deviate from equilibrium continuously.

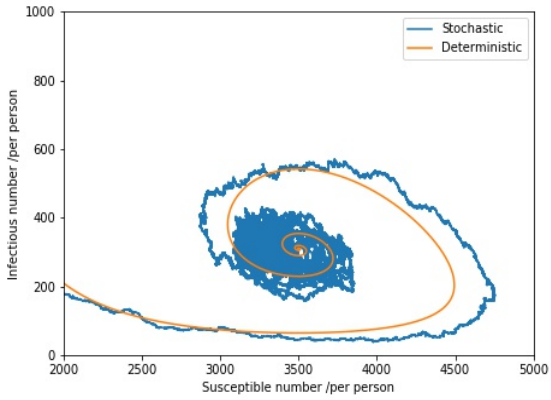

Fig. 8. The correlation of susceptible and infectious. Yellow line is deterministic and blue is stochastic. β =1 per day, $1/\gamma$ =3 per day,$1/\mu$ =60 per day

*4) **Stochastic Resonance** :* Since the SIR model is weakly stable, when the stochastic model approach the equilibrium, the transient-like will deviate the wave from the equilibrium. If the time is long enough, the oscillatory dynamics may be excited by noise. These noises cause the wave continuously deviate from the equilibrium and hold wave back if it is far away(Fig.9). This is called the stochastic resonance, and its oscillation's frequency is close to the natural frequency. We used Fourier analysis to explore this frequency. The deterministic frequency is 0.0155 and the amplitude is 0.00154(Fig.10).

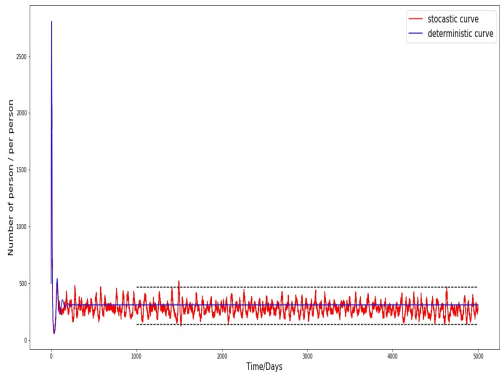

Fig. 9. Transient situation. Blue line is deterministic and red is stochastic.β =1 per day, 1/γ =3 per day,1/μ =60 per day

The stochastic model's frequency is 0.0165 and amplitude is 0.00241(Fig.11). The frequency are pretty close to each other. And the natural frequency is 0.01656 (given by 1/T in chapter 2).

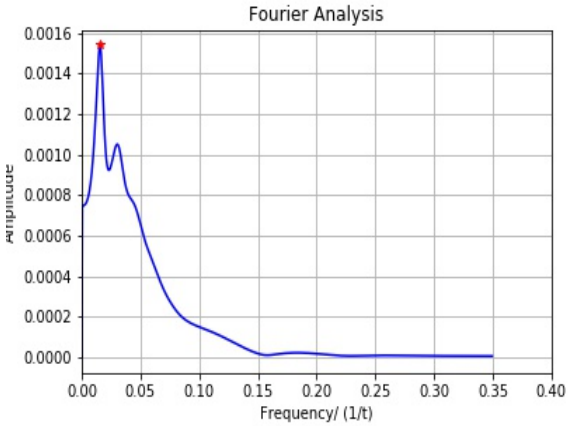

Fig. 10. Deterministic frequency, frequency is 0.0155 and the amplitude is 0.00154

*5) **Extinction** :* Since the event-driven model is a discrete model, the number of infectious sometimes drops from one to zero suddenly, unlike the continuously deterministic model. In closed population, an extinction happens when infectious number becomes zero, which makes the transmission rate become zero. Then this zero-state is remained. It is clear that these extinction is related to the population size and we want to explore the correlation between them.

Though diseases in real life may frequently undergo extinction, the imports of infectious people from the outside population can prevent the extinction. The import rate of infectious people is proportional to the square of the population size. Besides, according to the data that, Rate of measles imports = 5.5*$10^{-5}\sqrt{N}$, we can get a general import rate: δ= 0.0625μ(R_0-1)$\sqrt{N}$. So we use this rate to analyse the relation between the extinction and population size.

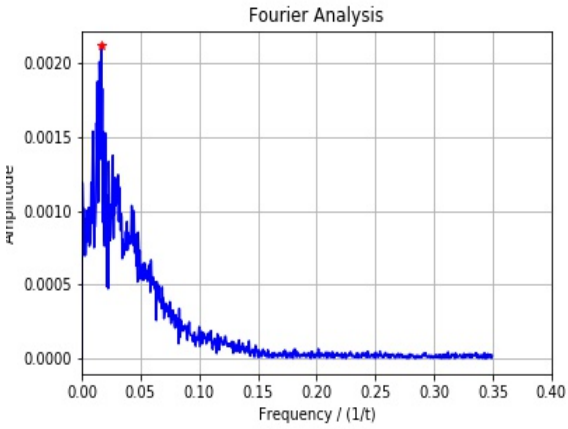

Fig. 11. Stochastic frequency, frequency is 0.0165 and the amplitude is 0.00241

As the Fig.11,12 shows, as the population size increases, the times of extinction (per year) and the extinction time (per year) will drop significantly. Small population will easily become extinction because of the large amplitude oscillation. While large population is much stabler than the small population.

Besides, taking population size is 5000, we used different values of R_0 to explore the correlation. It turns out that the big R_0 will suffer less extinction because the big transmission leads to more infectious people.(Fig.13)

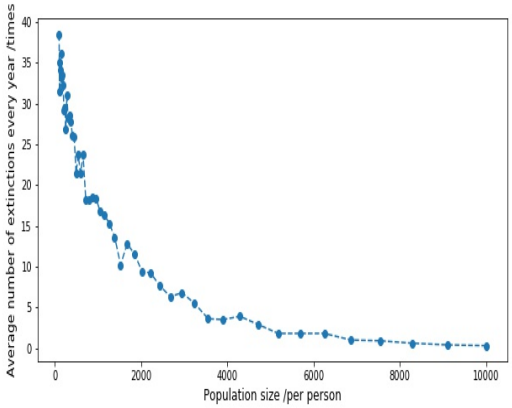

Fig. 12. Correlation between population size and extinction per year.β =1 per day, 1/γ =3 per day,1/μ =60*365 per day

*6) **Advanced Topic:First passage time** :* First passage time is one way to measure the persistence. First passage time is the average time spent until the extinction happens. Besides, this process doesn't need import infectious outside because we measure the first time of extinction. As the Fig.14 shows, when the population is small, its first passage time is above zero. Because there are big fluctuations which lead to extinction frequently in small population. And when the population is small, its first passage time approximately closes to zero.

We find that, when the population is bigger than 3652, no extinction happens. And the population with 1082 people has

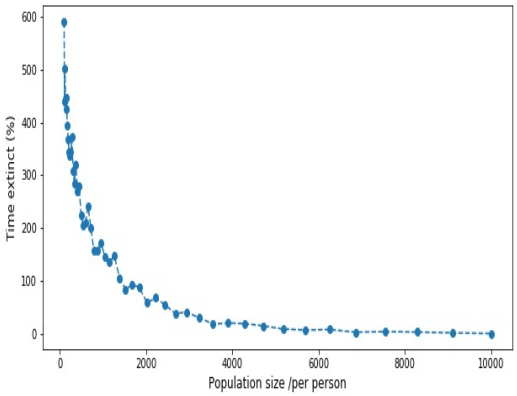

Fig. 13. Correlation between population size and time of extinction per year.β =1 per day, 1/γ =3 per day,1/μ =60*365 per day

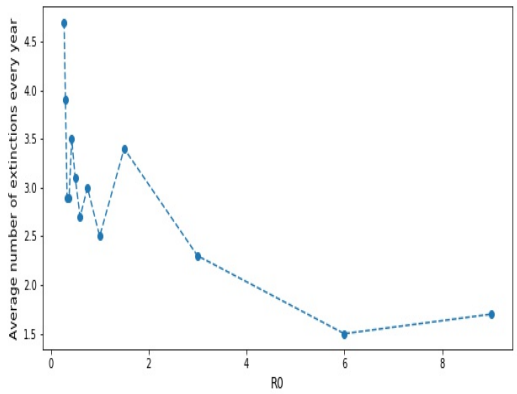

Fig. 14. Correlation between R_0 and extinction per year, population size=5000.β =1/11 1 per day, 1/γ =3 per day,1/μ =60*365 per day

the biggest passage time. These discoveries shows that the critical community size is 3652.

C. *Spatial model*

In real life, transmission may be a local process and population will spread unevenly, individuals move around between areas. So in order to capture this transmission's dynamics we should consider the spatial model. There are many models we can choose depended on the host organism. They are Meta-population, individual based models,lattice model,continuous PDE based models.

Before we choose the model, we should make some concepts clear.

1) Heterogeneity. Different social structures has different transmission rate and have different dynamics.
2) Interaction. Basically, the interactions between individuals are different because of different kinds of movements of people in different models.
3) Isolation. It is obvious that if the population has been isolated, the disease can not spread.

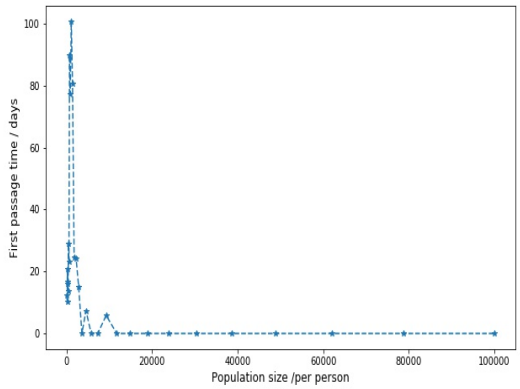

Fig. 15. First Passage time.529 population size has the big passage time. And CCS is 3652.β =1 per day, 1/γ =3 per day,1/μ =60 per day

4) Local extinction. Small population size is prone to be extincted. And when it has interaction with larger population size, new invasions of pathogen will happen.
5) Scales. This scales mainly include the interactions and simulations.

*1) **Meta-populations model** :* In this report, we use Meta-population model to analyze the dynamics because it is suitable to human diseases. Meta-population has three concepts: 1) Subpopulation. The entire observed population is divided into subpopulations. 2) Each population has its own dynamics. 3) Interaction between subpopulations. In order to describe the dynamics: we should import a variable which represents the strength of interaction:ρ_{ij} Then we can get:

$$\lambda_i = \beta_i \sum_{j=1}^{n} \rho_{ij} \frac{Y_j}{N_i} \tag{9}$$

This equation ρ_{ij} means the strength of interaction of transmission to population i from population j. Besides, we use N_i because this infection happened in i population. We assume that two subpopulation are fully susceptible populations. Besides ρ_{ii} = 1 and $\rho_{ij} < 1$. and the populations are of the same size, and ignore demography. Then for the j population, we have two rate $\beta_j X_j$ Y_j/N_j and $\gamma_j Y_j$. For the i population, we have $\beta_i X_i$ Y_i/N_i, $\gamma_i Y_i$ and $\beta_i \rho_{ij} Y_j X_i/N_i$. Thus, we use Gillespie's direct algorithm to simulate. As the Fig.15 shows, it is clearly to see the delay time of epidemic between the i and j population. And the lag time is 19.29 days, because the infectious people in j population need a chance to have interactions with susceptible people in i population.

Besides, we want to know whether the different ρ will influence the delay. As Fig.16 shows, the lag time drop significantly as the ρ increasing. Because when the ρ increase, the interaction between the i and j will increase. The epidemic delay will decrease significantly.

Then, we increase the number of subpopulations to explore the dynamics further. We use a one unit of lattice model, which has one subpopulation in central with four adjacent ones. To simplify it, we assume that only the central subpopulation j can infect the four neighborhoods (in four di-

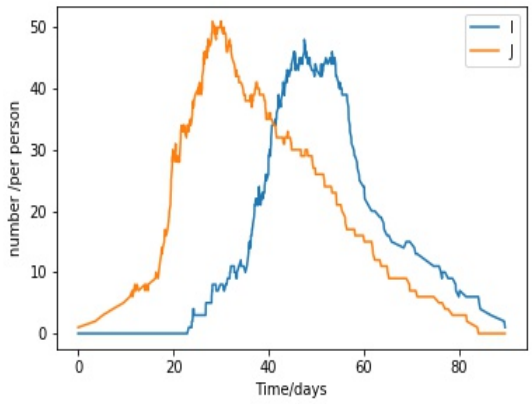

Fig. 16. i and j infectious people. The delay time is 22.30 days.X_i = 100, Y_i = 0,Z_i = 0,X_j= 100.Y_j = 1.Z_j= 0.β = 0.3745days,1/γ =14 days,ρ=0.01

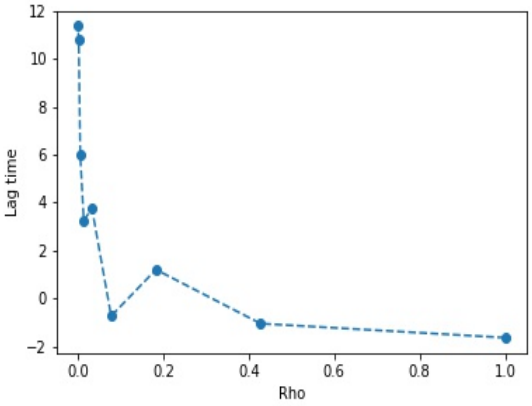

Fig. 17. The correlation between ρ and lag time.X_j= 100 Y_i = 0,Z_i = 0,X_j= 90.Y_j = 8.Z_j= 2.β = 1 days ,1/γ = 14days

rections). Andρ_{1j}=0.001,ρ_{2j}=0.002,ρ_{3j}=0.01,ρ_{4j}=0.02.ρ_{nn}=1, andρ_{jn}=0. It is clearly to see that the big ρ has less delay with the epidemic of j. And subpopulations with similar ρ have close epidemic time.

Here the conclusion can be arrived that, there is a delay between the two subpopulation, and The delay will decrease when the ρ increase.

*2) **Lattice-based model** :* Based on the one unit Lattice-based model discussed in Fig.19, we expand the one unit to a 9x9 grid spatial model, which has 81 subpopulations. But we use deterministic SIR model with demography to simulation this dynamic in this part for simplification and efficiency. The interaction rules between subpopulations are 1) each subpopulation interacts with the four neighborhoods in up,down,left and right direction. 2) each subpopulation can infect others, while being infected at the same time.

For each subpopulation j, it has 9 transmission rate; four rates are infected by others: $\beta X_i rho_{j1} Y_j/N_i$, $\beta X_i \rho_{j2} Y_j/N_i$, $\beta X_i \rho_{j3} Y_j/N_i$, $\beta X_i \rho_{j4} Y_j/N_i$; four rates are infect others:$\beta X_i \rho_{1j} Y_j/N_i$,$\beta X_i \rho_{2j} Y_j/N_i$,$\beta X_i \rho_{3j} Y_j/N_i$,$\beta X_i \rho_{4j} Y_j/N_i$; one transmission rate among itself: $\beta X_i Y_i$. It also has the

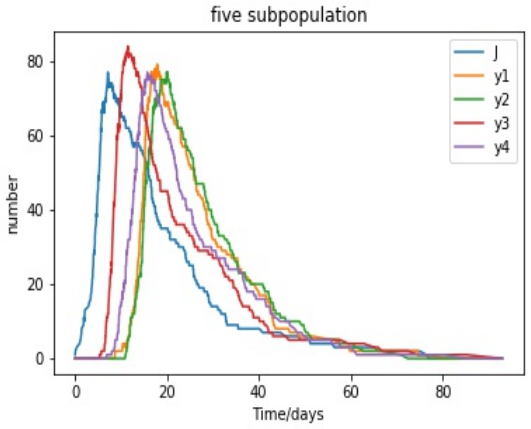

Fig. 18. Five subpopulation:One unit of lattice model.ρ_{1j}=0.001,ρ_{2j}=0.002,ρ_{3j}=0.01,ρ_{4j}=0.02.ρ_{nn}=1, andρ_{jn}=0.

natural birth rate, death rate and recovery rate like basic demography model. The odes of Lattice-based model are written as below.

$$\frac{dX_i}{dt} = \mu - \beta X_i \frac{(1-\sum_j \rho_{ij})Y_i + \sum_j \rho_{ij} Y_j}{(1-\sum_j \rho_{ij})N_i + \sum_j \rho_{ij} N_j} - \nu X_i \quad (10)$$

$$\frac{dY_i}{dt} = \beta X_i \frac{(1-\sum_j \rho_{ij})Y_i + \sum_j \rho_{ij} Y_j}{(1-\sum_j \rho_{ij})N_i + \sum_j \rho_{ij} N_j} - \nu Y_i - \gamma Y_i \quad (11)$$

$$\rho_{ij}=\rho_{ji}\begin{cases}\rho & \text{if i and j are neighbors}\\ 0 & \text{otherwise}\end{cases}$$

We assume that only the central lattice has 1 percent infectious people at the start, and the rest subpopulations are full susceptible. We initialize a 25x25 grid to observe the dynamic. The simulation result is shown in Fig.19. The spread of infection after 4 years like water waves. At this time, infectious epidemic move from central to four direction. We can see that the lattices in the blue circle have more infectious.

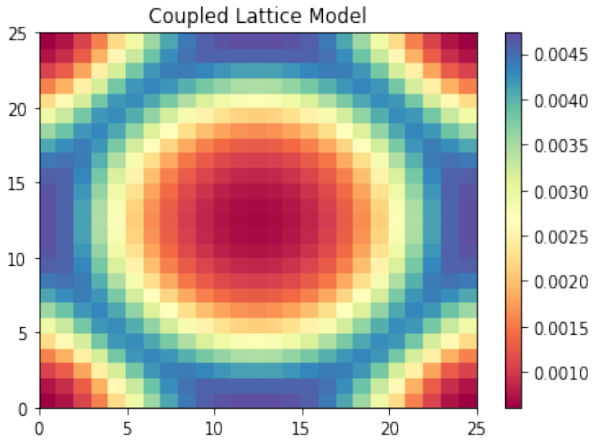

Fig. 19. The spread of disease from the central lattice after 4 years from the start. The degree of infectious density is shown as the right bar. β = 2 days,1/γ =100 days,μ=1e-4 days ,ν=1e-4 days,ρ=0.1,X_0=0.1

Since the spread of disease is like water waves from near to far. We want to explore the speed of infection, which is defined

as how fast the epidemic happens in each grid from central to edge. This time, we simplify the model to a 9x9 grid to explore the speed, due to the limit of computer performance During our simulation, we find that not only the speeds in horizontal and vertical direction same, but also the speeds in opposite direction are same. Here, we will talk about the speed in horizontal direction from central to right. AS shown in Fig.20, X_i (i=0,1,2,3,4) records the infection number at epidemic state, with corresponding recording time t_i (i=0,1,2,3,4).

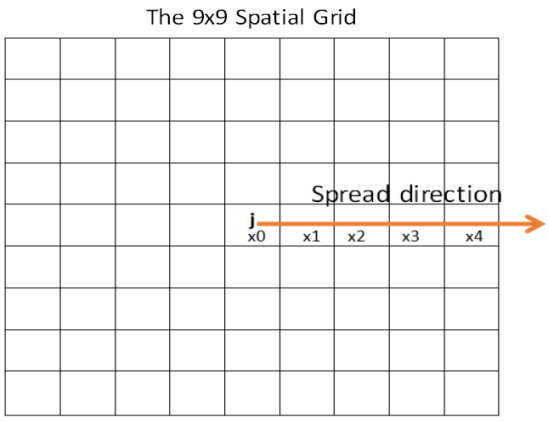

Fig. 20. Schematic diagram of studying the speed of disease transmission. x0,1,2,3,4 represent the number of infectious people on epidemic state at time t0,t1,t2,t3,t4

With the time going on, each subpopulation undergoes the epidemic state at different time. The dynamic of infectious people of the five observed subpoplution in Fig.20 is shown in Fig.21.

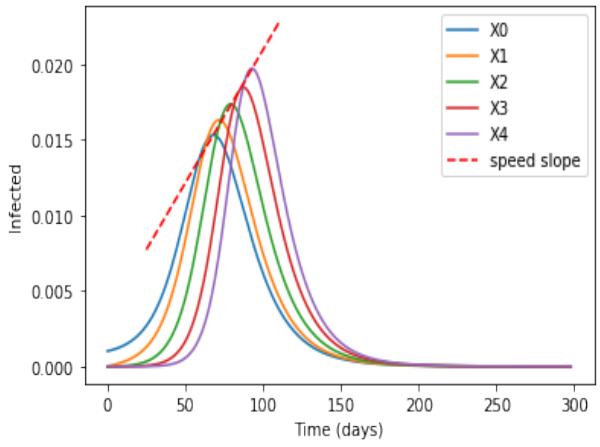

Fig. 21. The dynamics of infectious people in the five grids from x0 to x4. β = 2days,1/γ = 100days,μ=1e-4days,ν=1e-4days,ρ=0.1 ,X_0=0.1

As the speed slope in Fig.21, it is clearly that the epidemic point of each subpopulation locates at the same line. This indicates that they have same speed approximately, When speed is calculated as (x_n-x_{n-1}) / (t_n-t_{n-1}), which exactly is the slope between epidemic points. This also verifies that the spread of infection is linear.

To explore the speed further, we want to have a look of the connection between the initial parameters value and speed. At the first, we simulates with different beta, which contributes to the basic reproductive rate R_0. It reasonably to guess that the R_0 will influence the speed as it will influence the transmission.

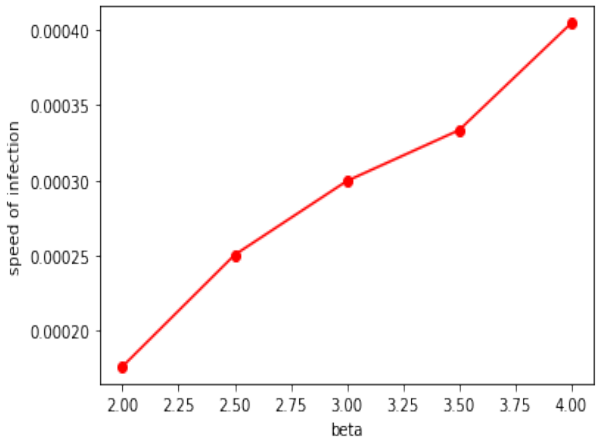

Fig. 22. The increase of speed when β increase. 1/γ = 100days,μ=1e-4days,ν=1e-4days,ρ=0.1,X_0=0.1.

As shown in Fig.22, when the β increase, the speed also increase. What's more, the growth in speed seems like a linear growth against the beta changes.
Beside the R_0, we also explore the relationship between speed and interaction rate ρ_{ji} in Fig.22.

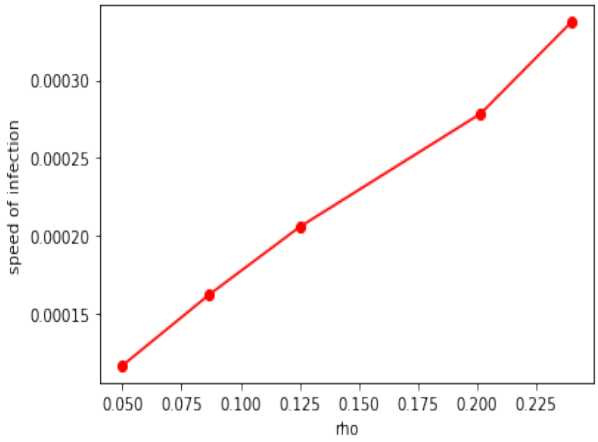

Fig. 23. The increase of speed when ρ increase. β=2 days, 1/γ = 100 days,μ=1e-4 days,ν=1e-4 days, X_0=0.1

It it seems like, when ρ increase, the speed also increase due to the big chance for disease spreading.
Here the conclusion can be arrived that, the speed of infection is linear. The basic reproductive rate R_0 and interaction rate ρ controls the speed.

*

a.Interpolation

Since we used Gillespie's algorithm to simulate the stochastic SIR with demography model, and we also want to get the variance of noise at second hallmark. Our method is using

deterministic outcome minus the stochastic outcome. However, the stochastic model's time interval is random, so we need a function in order to let both deterministic and stochastic model have same time interval. In this report, we used interpolation(1D)[2] to fit the stochastic model. The theory behind the linear interpolation is simple. Let take two points in stochastic model as example.[3] (x_0,y_0) and (x_1,y_1). We assumed a function:

$$y = \varphi(x) = ax + b$$

This function:

$$y_0 = \varphi(x_0), y_1 = \varphi(x_1)$$

Then we can get the equation expression

$$y = \varphi(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0)$$

We called $\frac{y_j - y_i}{x_j - y_i}$ is f(X)'s first order mean difference at x_i and x_j point. then the function can be rewrite like this:

$$\varphi(x) = f(x_0) + f(x_0, x_1)(x - x_0)$$

Finally, we used y_0 and y_i to substitute the function:

$$\varphi(x) = \frac{x - x_1}{x_0 - x_1}y_0 + \frac{x - x_0}{x_1 - x_0}y_1$$

We used this theory to simulate the whole points of stochastic model, and apply the same time interval of deterministic model. Then we used deterministic model minus the stochastic model.

b.Runge – kuttaMethod

In numerical integration, we integrate the deterministic odes with the help of python function odeint(), which based on the Runge-kutta Method. This method is an advanced Euler Method, with a higher orders in Taylor Expansion. It picks more points on t-axis between (t,t+σt) to predict the next value f(t+σ) by integrating on discrete time.

III. Results And Conclusion

In conclusion, there are three parts introduced here.

1). For the stochastic model, Each simulation has different outcome, which shows the variability. Increasing the population size can increase the variance of noise.

And the number of susceptible and variance of infectious increase as the increase of noise.

And there is negative covariance between the susceptible and infectious. And we can see the transient-like return in the plot. Besides, The frequency of stochastic oscillation is similar to the natural frequency, which is called stochastic resonance.

Furthermore, small population is prone to suffer the extinction and the big R_0 can reduce the probability of suffering the extinction.

For the first passage time, in our model the population is bigger than 5736, no extinction happens. And the population with 529 people has the biggest passage time. These discoveries shows that the critical community size is 5736

2). For the Meta-population model, there is a dealy between the two subpopulation, and The delay will decrease when the ρ increase. Besides, increase the number of subpopulation, the lag time also depend on the ρ

3). For the lattice model, we found the speed of infection is linear. The basic reproductive rate R_0 and interaction rate ρ controls the speed.

References

[1] Matt J Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, 2011.

[2] ALEXANDER L Read. Linear interpolation of histograms. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 425(1-2):357–360, 1999.

[3] Philip J Davis. *Interpolation and approximation*. Courier Corporation, 1975.