SIR Model(ODE)

QinghеGao, 12589896

***Abstract*—SIR model is an epidemiological model that computes the theoretical number of people infected in a constant population over time. In this report we started with the simple SIR model and to analyze the dynamic and equilibrium state. Then we add demography and mortality into the SIR model. Damp oscillation has been found in the curve and we used both theoretical and Fourier analysis to analyze it. It turn outs that R_0 plays an important role in every model. So finally we got the advanced model, which added seasonal forcing into it. The model has become robust.**

I. Introduction and Background

THE process of modeling in epidemiology is a long and tough road. Until 1900s dynamical systems approaches were increasingly applied to epidemiology [1]. And theoretical epidemiology has developed significantly in concept and technology. In this report, we will discuss the simplest epidemiological modelsSIR model.

SIR model is one of the simplest model and many models are derivation form this. This model was studied in depth by Kermack and McKendrick[2], and there are three compartments-Susceptible(S), Infected(I), and Recovered(R). And the population will have the transitions $S \to I \to R$.

Threshold phenomenon and epidemic burnout of SIR model without demography has been analyzed and it turned out that the R_0(basic reproductive ratio) is the most important parameter to control the epidemic burnout.

And the SIR model without demography was applied to a case. A influenza in a British boarding school. And we got the R_0 is 3.67 from the best fit parameters. And we proposed a vaccination plan for this case to prevent the outbreak of influenza.

Then, we added the demography to SIR model. Equilibrium state and stability properties was analyzed in the report. And we also used Fourier analysis to interpret the oscillatory dynamics.

And we introduced the infection-induced mortality to the SIR model with demography. the parameter ρ means the probability of an individual in the I class dying from the infection. we analyzed the the situation if ρ approached the 1.

Then, we used the approaches to analyze the SEIR model, which adds a new compartment Exposed(E) to represent the category for these individuals who are infected but not yet infectious. And we add seasonal forcing into SEIR model, each compartment has period.

II. Methodology and Discussion

A. SIR model without demography

1) Threshold phenomenon and epidemic : The model was initially studied in depth by Kermack and McKendrick. And they categorized hosts within a population as Susceptible, Infected, and Recovered. Because we ignore the population demography, we only have the transitions $S \to I \to R$

For the first $S \to I$ process, we define the force of infectionγ, which is defined as the per capita rate at which susceptible individuals contract the infection. Then we got $\lambda = \beta Y/N$ and $\lambda = \beta Y$ (where Y is the number of infectious disease, β is the product of contact rates and transmission probability and N is total number of the population). The distinction between these two transmission mechanisms becomes pronounced when host population size varies, otherwise the 1/N can be absorbed in $1/\beta$.

For the second $I \to R$ process. We define the recovery rate γ(which is inverse of the infectious period). So we use SIR instead of XY and got the SIR model without demography differential equations:

$$\frac{dS}{dt} = -\beta SI \tag{1}$$

$$\frac{dI}{dt} = \beta SI - \gamma I \tag{2}$$

$$\frac{dR}{dt} = \gamma I \tag{3}$$

Then we used the python odeint function to integrate the differential equations. First we used different parameters to plot two picture, and it seems that figure1 has the epidemiology but figure2 hasn't. So what exactly determines epidemiology?

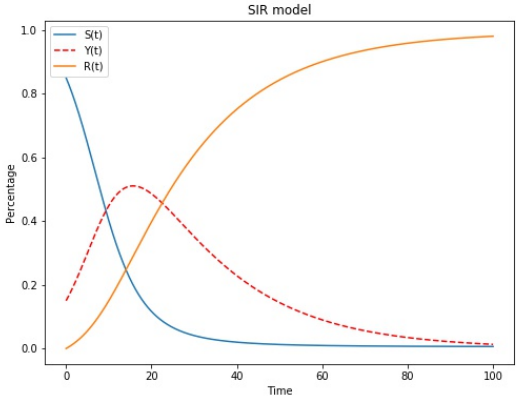

Fig. 1. SIR model without demography. β =1/4 per day, 1/γ =20 per day, S_0 =0.85, I_0 =0.15

Firstly, we started with the equation 3 in this form

$$\frac{dI}{dt} = I(\beta S - \gamma) \tag{4}$$

We could conclude that if γ/β is greater than the initial susceptible S(0), then $\frac{dI}{dt}$<0, so we can see the disease dies

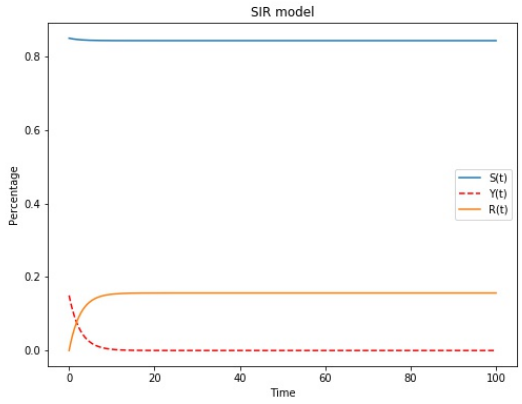

Fig. 2. SIR model without demography. β =1/50 per day, 1/γ =2.5 per day, S_0 =0.85, I_0 =0.15

out like Fig.2. And if γ/β is less than the initial susceptible S(0). We got the epidemiology just like the Fig.1. So we let $\frac{dI}{dt}$=0, then S(0)=γ/β. we defined the inverse of γ/β as R_0(basic reproductive ratio), which is defined as average number of infectious compartment produced by one infection in susceptible compartment. Then S(0)=1/R_0. We assumed an extreme situation that all the initial population is susceptible(S(0)=1). Finally, we can conclude that if R_0<1 and there is no epidemiology, the disease will die out. Only if R_0>1, the infectious disease can invade and become an epidemiology.

In order to verify our analysis, we want to have the change of infectious people with respect to R_0. We used equation1 and 2 to get $\frac{dS}{dR}$=- R_0 S. Then we integrated this equation to get: S(t)=S(0) $e^{-R(t)R_0}$. Thus, we can get S(∞)=1-R(∞)=S(0) $e^{-R(\infty)R_0}$ (I(∞)=0). Thus we can get the final equation: 1-R(∞)-S(0) $e^{-R(0)R(\infty)}$=0 (R(∞) equals to the total number of infective)

Using this euqation, Fig.3 clearly shows our conclusion, it is clearly to see that the green area is R_0<1 and there is no epidemiology. Only if R_0>1, the infectious disease can invade and become an epidemiology.

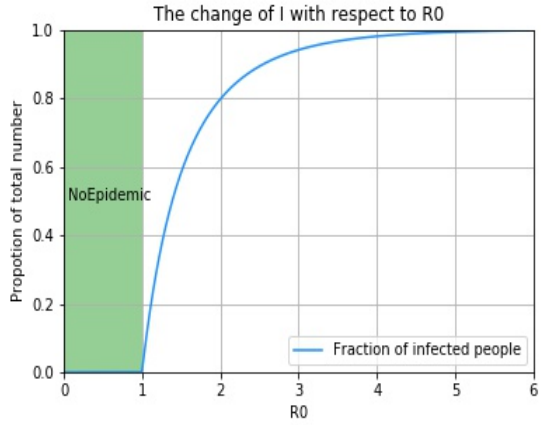

Fig. 3. The change of I with respect to R_0. The curve is fitting by equation function in Box1.2 and function "fsolve" in python

TABLE I
BOARDING SCHOOL DATA FOR INDIVIDUALS CONFINED TO BED

Day	1	2	3	4	5	6	7
Nmber	3	8	28	75	221	291	255
Day	8	9	10	11	12	13	14
Number	235	190	125	70	28	12	5

The phenomenon which whether the epidemiology happens or not is called threshold phenomenon. This threshold effect can be also illustrated by the phase plate(Fig.4). When S(0) is smaller than γ /β, the curve of the infectious population directly towards 0. When S(0) is bigger than γ /β, the curve increase firstly then decrease towards 0, which means that there is an epidemiology.**Only if R_0>1, the infectious disease can invade and become an epidemiology.**

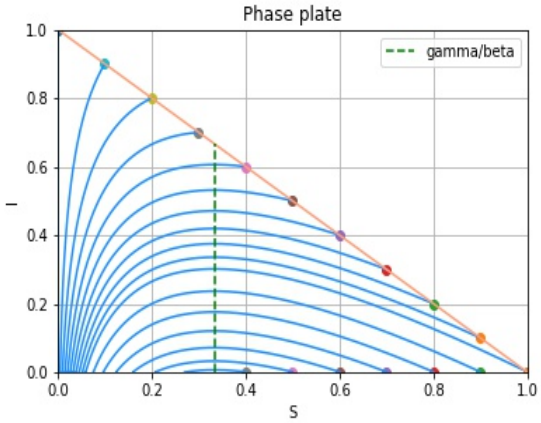

Fig. 4. Phase plane of SIR model without demography. The vertical dash line is γ /β. The epidemiology only occurs when S(0) is above the γ /β. The pink line means S(0)+I(0)=1 since R(0)=0

*2) **Application of SIR model without demography** :* Since we have already familiar with the SIR model without demography, we analyzed the case in English boarding school. An influenza outbreak in this school, and total number of boys is 763. The influenza started by one boy and the table1 show the number of individuals confined to bed as time changed. It seems that we can use python to fit the curve to find the best parameters(least suqares). Thus, we used python curvefit function to fit the data. The infectious period(1/γ) is 2.2 days, and the transmission rate(β) is 1.67 per day. Thus the R_0 is 3.67.

*3) **Propose a vaccination plan** :* In order to propose a vaccination, we need a new parameter called P , which is an outcome of vaccination coverage and vaccine efficacy. Then we can easily get the X'(0)=X(0)(1-P), so we used X'(0) to substitute the the original susceptible. Then we could rewrite the equation of SIR differential equations:

$$\frac{dX}{dt} = -\beta X'Y/N \tag{5}$$

$$\frac{dY}{dt} = \beta X'Y/N - \gamma Y \tag{6}$$

And we have discussed only if $\frac{dY}{dt}$<0, we could prevent disease spread. So we used $\frac{dY}{dt}$=0 and wrote a loop in python

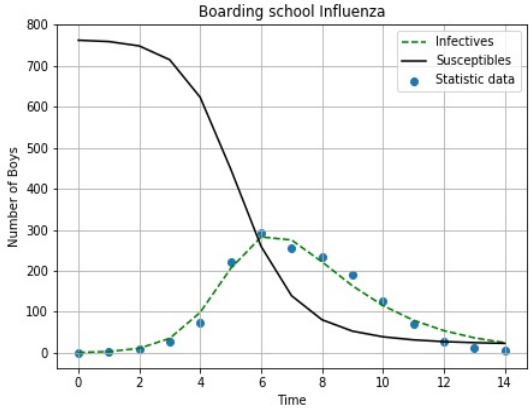

Fig. 5. Curvefitting of the Data,1/γ=2.2 days, β=1.67 per day, R_0=3.67

to find the threshold point. As the Fig.6 shows the max initial susceptibleX'$_{max}$(0)=205, there is no epidemiology. Thus, we used X'(0)=X(0)(1-P) and got the P=557/763. Thus, the vaccination plan is at least 557 boys should be vaccinated. Besides, we used mass vaccination strategy to verify our method, the final result was exactly same as our result.

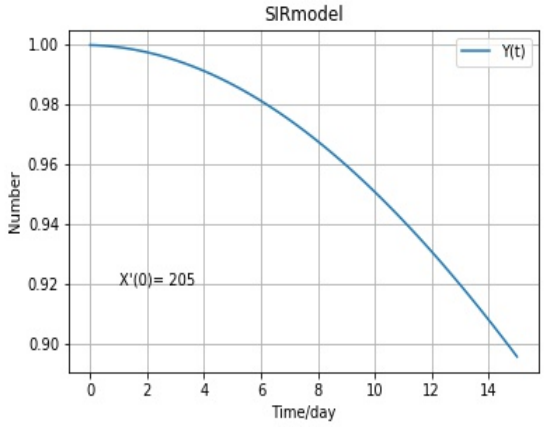

Fig. 6. No epidemic of Y(t) , the max initial susceptible should be less than 205 people

——————————————————————————————

a.NumericalIntegration

Since all of them are differential equations and it is not possible to directly get the exact equation, it is necessary to use the numerical integration. So we used the simplest method-"Euler Method" to integrate. Since we talked about the SIR model, so we used I(t). We divided the infective's domain[0,1] into N equal parts. Each step size h=1/N. Since we already knew the initial value of infective:I(t_0).Then we used Taylor's formula:

$$\begin{aligned} I(t_1) &= I(t_0+h) \\ &= I(t_0) + hI'(t_0) + \frac{h^2}{2}I''(t_0) + hI'(t_0) + \frac{h^3}{6}I'''(\zeta) \\ &= I_0 + hI(t_0, I_0) + H_0 \end{aligned}$$

Besides, $\zeta\in(t_0,t_1)$, and we ignored second order H_0, now we get:

$$I(t_1) = I_0 + hI(t_0, I_0)$$

Using the same way to get the recurrence formula:

$$I_{n+1} = I_n + hI(t_n, I_n), n = 0, 1\cdots N-1$$

Besides, its geometric meaning is quite obvious.

b.Fsolvefunction

Fsolve function is to find numerical solution of linear and nonlinear equations. Fsolve(fun,x_0) where fun is a function of calculating the error of the system of equations, whose parameter x is an array whose value is a set of possible solutions of the system of equations, and fun returns the variance of each variance obtained after bringing x into the system of equations Error. x_0 is a set of initial values of unknowns.

When solving, it automatically calculates the partial derivatives of the equations at a certain point to each unknown variable. These partial derivatives form a two-dimensional array, which is mathematically a Jacobian matrix. If there are too many unknowns.

And the number of unknowns associated with each equation is small, that is, when the Jacobian matrix is sparse, the arithmetic speed can be improved by passing the Jacobian matrix.

c.Curvefit

Python's curve fit is a function for people to fit a curve, which is based on Least squares. We used a non-linear example.

When we have a set of data (x_1,y_1. x_2,y_2. x_3,y_3. $\cdots$) ,and the data near a straight line y=b+ax, So we used the residual to write a equation:

$$R = \sum(ax_i + b - y_i)^2$$

And only when the gradient is 0, we can obtain the minimum value. So we can obtain the equation:

$$\frac{\partial}{\partial a} = 2\sum(a - b - y_i)x_i = 0$$
$$\frac{\partial}{\partial b} = \sum(a - b - y_i) = 0$$

Then we get the answer:

$$a = (\sum y_i - b\sum x_i)/n$$
$$b = [n\sum(x_i y_i) - (\sum x_i \sum y_i)]/(n\sum x_i^2 - \sum x_i \sum x_i)$$

——————————————————————————————

B. SIR model with demography

*1) **Dynamic of SIR model with demography** :* In this part, we introduce the demography to the SIR model. We assumed a natural lifespan is 1/μ years, then the mortality(natural) is μ. Since this assumption is natural lifespan, the total population is still stable as the SIR model without demography(S+I+R=1). Then, we can rewrite the differential equations:

$$\frac{dS}{dt} = \mu - \beta SI - \mu S \tag{7}$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I \tag{8}$$

$$\frac{dR}{dt} = \gamma I - \mu R \tag{9}$$

Then as the Fig.7 and Fig.8 show, we again used two groups of different parameters to fit.

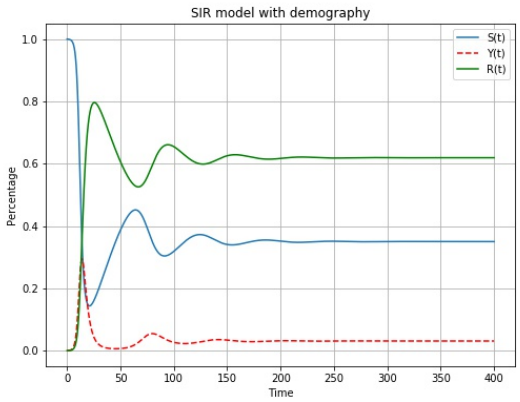

Fig. 7. SIR model with demography, β=1 day, 1/γ=3days, 1/μ=60 years

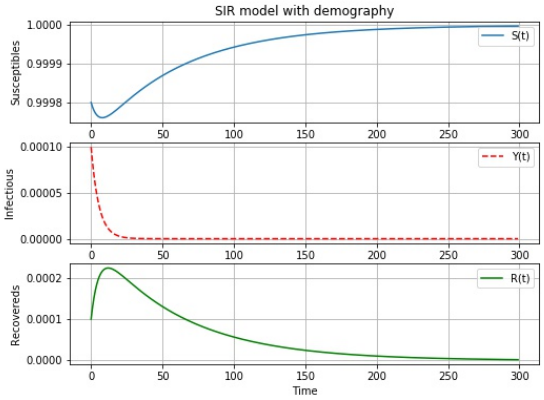

Fig. 8. SIR model with demography, β=1/6 day, 1/γ=3days, 1/μ=60 years

It shows that Fig.7 has clear epidemiology and in Fig.8 the disease dies out soon. So compared to the original model, how the dynamic change in SIR model with demography? Just like the way we analyzed before, we used $\frac{dI}{dt}$=0, then S(0)=(γ+μ)/β. Again we introduced β/(γ+μ)=R_0(basic reproductive ratio) And assumed an extreme situation that all the people are susceptible(S(0)=1). Then only if R_0>1, the disease can be transmitted like Fig.7. Otherwise the disease will die out soon like Fig.8.

In order to analyzed the dynamics further, we should discuss the equilibrium state and its porperities. To find the equilibrium, we should let $\frac{dS}{dt}$=$\frac{dI}{dt}$=$\frac{dR}{dt}$=0. It is obvious that we can find two fixed points. The first is disease free point: (S*, I*, R*)=(1,0,0). To find second point, we should use R_0 to rewrite the equation7. Then we got: I*=$\frac{\mu}{\beta}$*(R_0-1). Since S*+I*+R*=1, we can get the second fixed point (S*, I*, R*)=(1/R_0, $\frac{\mu}{\beta}$*(R_0-1),1-1/R_0-$\frac{\mu}{\beta}$*(R_0-1)). Fig.9 clearly shows that the disease-free equilibrium since R_0<1. Fig.10 shows an endemic equilibrium and the line will converge to the fixed point(red*) because R_0>1.

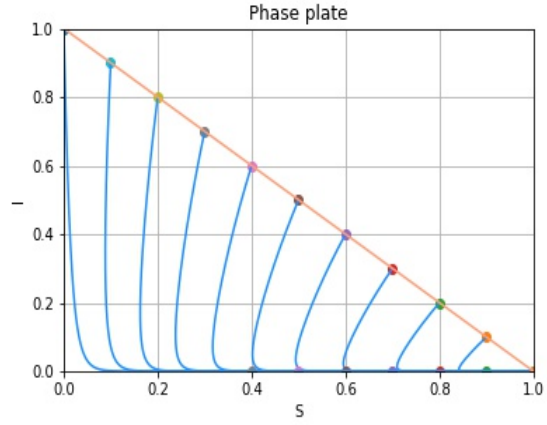

Fig. 9. Phase of plate. Disease-free equilibrium of SIR model with demography, β=1/2 day, 1/γ=3days, 1/μ=60 years. R_0=20/7.

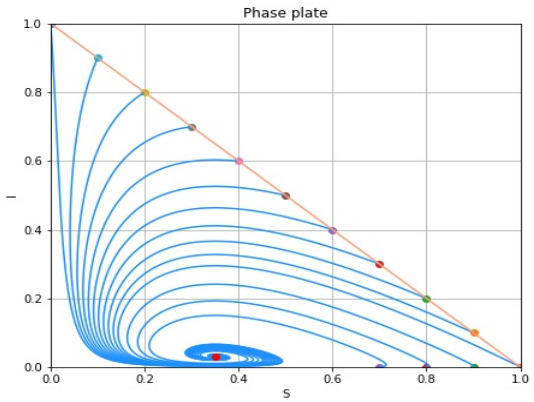

Fig. 10. Phase of plate. Endemic equilibrium of SIR model with demography, β=1 day, 1/γ=3days, 1/μ=60 years. R_0=20/7. Red dot is fixed point

It is clearly to see that the infectious curve has a damp oscillations, why this oscillation happens and What are requirements to observe this behavior? We should analyze the stability analysis. To Analyze the system dynamic we should use linearization approximation. Since the total population will not change, we just used two differential equations(7), and (8). We already got the fixed point (S*, I*)=(1/R_0, $\frac{\mu}{\beta}$*(R_0-1)). Then we can get:[3]

$$S(t) = S* + \eta(t) \tag{10}$$

$$I(t) = I* + \theta(t) \tag{11}$$

$\eta(t)$ and $\theta(t)$ are the part which deviate from equilibrium. That's what we should talk about. So we rewrite the equation(10),(11):

$$\frac{d\eta}{dt} = -[\mu(R_0 - 1) + \mu]\eta - \beta\theta/R_0 + NL(\eta, \theta) \tag{12}$$

$$\frac{d\theta}{dt} = \mu(R_0 - 1)\eta + NL(\eta, \theta) \tag{13}$$

In equations, NL(η,θ) is all nonlinear term. Now we can write as matrix way.

$$\begin{bmatrix} -[\mu(R_0 - 1) + \mu] & -\beta/R_0 \\ \mu(R_0 - 1) & 0 \end{bmatrix}$$

We called this $\vec{J}$ and its eigenvalues represented the dynamic of system. We substituted parameter in equation 14 and then the final equation was 15.

$$\Lambda^2 - tr\vec{J}\Lambda + det\vec{J} = 0 \tag{14}$$

$$\Lambda^2 - \mu R_0 \Lambda + \mu(\mu + \gamma)(R_0 - 1) = 0 \tag{15}$$

So if we want a damp oscillation, we should let R_0>1, then we compared $(\mu R_0)^2$ and $4\mu(\mu+\gamma)(R_0\text{-}1)$. Because μ is quite small number(its reciprocal is lifespan), which is smaller than the γ. And R_0 is bigger than 1, so it is safe to assume $(\mu R_0)^2$ is smaller than $4\mu(\mu+\gamma)(R_0\text{-}1)$. Then we got the stable damped oscillations. **So if it is endemic state and R_0>1, we have damped oscillations.**

Then the amplitude and frequency should be analyzed. Since $(\mu R_0)^2$ can be ignored, so we could get the answer from the equation 15.

$$\begin{aligned} \lambda_{1,2} &= \frac{1}{2}(\sqrt{-4\mu(\gamma + \mu)(R_0 - 1)}) \\ &= \frac{\mu R_0}{2} \pm \frac{i}{\sqrt{AG}} \end{aligned} \tag{16}$$

A=1/($\mu(R_0$-1)), means the mean age at infection and G=1/(μ+γ), means infected period. Then the circle frequency $\omega=\frac{i}{\sqrt{AG}}$, and the T= $2\pi\sqrt{AG}$. Then we use Fourier analysis to analyze the infectious data from Fig.7. And after Fourier analysis we got two peaks, the first frequency and amplitude: 0.015 0.00764 and second frequency and amplitude: 0.03,0.00515. Then we used parameters from Fig.7 to calculate the frequency(1/T) and the frequency is 0.0165, which is close to the second peak's frequency.

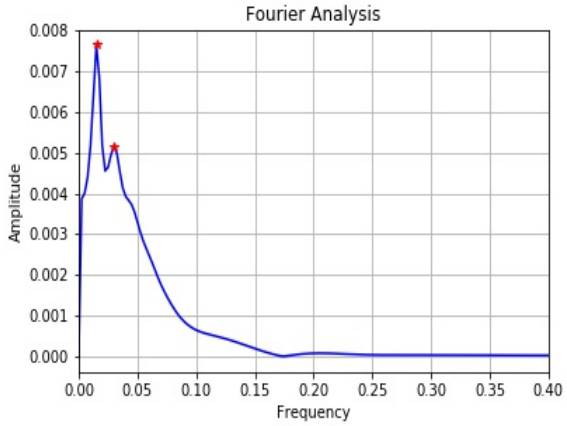

Fig. 11. Fourier analysis of infectious data from Fig.7. Red * are the peaks.First frequency and amplitude: 0.015 0.00764. Second frequency and amplitude: 0.03,0.00515

*2) **SIR model with disease induced mortality** :* Infectious disease actually can cause the death, so we can't assume the total population is constant. And we introduce ρ, which is the probability that a people dies because of infection. And a fixed birth rate υ then we can get the equation:

$$\frac{dS}{dt} = \upsilon - \beta SI - \mu S \tag{17}$$

$$\frac{dI}{dt} = \beta SI - \frac{(\gamma + \mu)}{(1 - \rho)} I \tag{18}$$

$$R_0 = \frac{\beta\upsilon(1 - \rho)}{\mu(\mu + \gamma)} \tag{19}$$

The density and frequency transmission should discuss separately because N is changing now. For the density-dependent transmission, we want to find the equilibrium state. We used the method discussed before. For the disease free(υ / μ,0,0), and for the (X*,Y*,Z*)=(υ/(μR_0),μ(R_0-1)/β,γ(R_0-1)/β). Then we got N* equation:

$$N* = \frac{\upsilon}{\mu R_0}[1 + (1 - \rho)(R_0 - 1)] \tag{20}$$

Then for the frequency-dependent transmission, We got the equations:

$$N* = \frac{\upsilon}{\mu}(\frac{R_0(1 - \rho)}{(R_0 - \rho)}) \tag{21}$$

As the figure 12, 13 and equation 18 show, both percentage of infective drop quickly when ρ approaches 1. Meanwhile, the R_0 will drop quickly to the zero.Besides, we used equation 20,21 to find out with the ρ changes, how will the N* change. Just as Fig.12 shows, when the ρ and R_0 are large, total population will decrease significantly. Besides, when ρ approach to 1, N* of frequency-dependent transmission drop quickly than density-dependent transmission. Because frequency-dependent transmission's contact rate will drop as populations decreases.

Thus, when ρ approaches to the 1, R_0 drop to 0. Proportion of infective and total population drop significantly. And N* of frequency-dependent transmission drop quickly than density-dependent transmission.

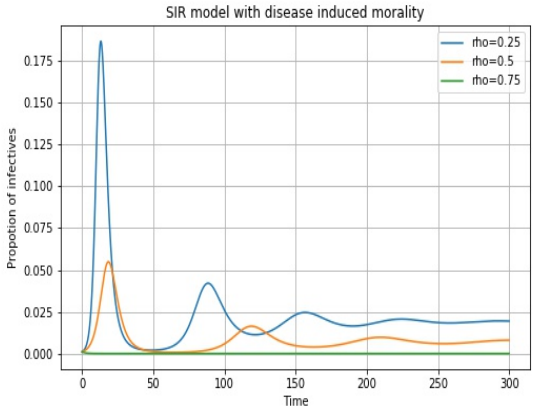

Fig. 12. SIR model with mortality(density), three different ρ shows that when ρ approach 1, indicatives die almost instantaneously,and R_0 drops to zero

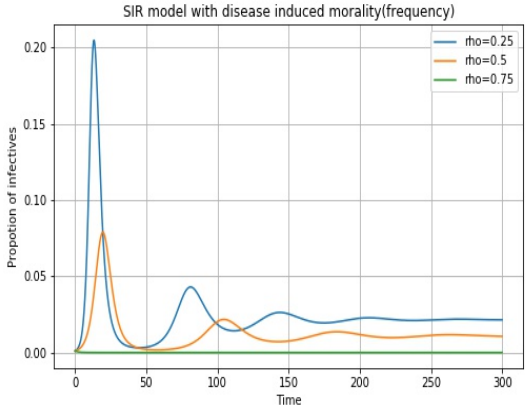

Fig. 13. SIR model with mortality(frequency), three different ρ shows that when ρ approach 1, indicatives die almost instantaneously, and R_0 drops to zero

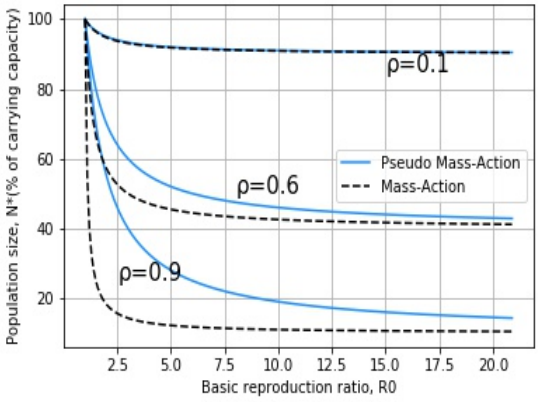

Fig. 14. The population size, N*, for diseases that are associated with mortality. Three different disease mortality probabilities are considered, ρ= 0.1,ρ=0.6 ρ= 0.9, and the two mixing assumptions of density dependence (βXY, solid line) and frequency dependence (βXY/N, dashed line) are shown.

FourierAnalysis

Fourier transform[4], which means that a certain function that satisfies certain conditions can be expressed as a trigonometric function (sine and/or cosine function) or a linear combination of their integrals. The equation can be written like this:

$$F(\omega) = \Gamma[f(t)] = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$$

FFT (Fast Fourier Transformation) is a fast algorithm of discrete Fourier transform (DFT). Because FFT decline the time complexity to the $n\log_2 n$. The FFT can separate the wave(take infective wave) into different frequency's wave, in this way we can analyze the frequency and amplitude after FFT.

C. *SEIR model*

In this part, we introduce an "Exposed" compartment into the SIR model, which means that the individuals in this compartment are infected but not infectious.The average duration of period is given by $1/\sigma$. In this way, we get the new models: $S \to E \to I \to R$. Then we get the equations:

$$\frac{dS}{dt} = \mu - (\beta I + \mu)S \quad (22)$$

$$\frac{dE}{dt} = \beta SI - (\mu + \sigma)E \quad (23)$$

$$\frac{dI}{dt} = \sigma E - (\mu + \gamma)I \quad (24)$$

$$\frac{dR}{dt} = \gamma I - \mu R \quad (25)$$

We used numerical integration to integrate the differential equations, The Fig.12 shows the curve of different compartments. Then we want to analyze the dynamic of this model. Similar to the SIR model, the disease free point(S*,E*,I*,R*)=(1,0,0,0). And the endemic fixed point is (S*,E*,I*,R*)=(1/R_0, $\mu(\mu+\gamma)(R_0\text{-}1)/\beta\sigma$, $\mu(R_0\text{-}1)/\beta$, 1-S*-E*-I*). It is clearly to see that SEIR model's R_0 is similar to the SIR model because μ can be ignored. And only if $R_0>1$, the disease can invade which is the same as SIR model. After exploring stability, SEIR model's oscillations period is also T=2πAG, where G is slightly different:G=1/(μ+γ)+1/(μ+σ). Then we used the Fourier analysis to verify. The frequency of the peak is 0.00857, which is close to the 1/T =0.00667. **Though SIR model and SEIR have many similarities, but SEIR model's pathogen have much slower invasion. Because the pathogen has to pass through exposed compartment.**

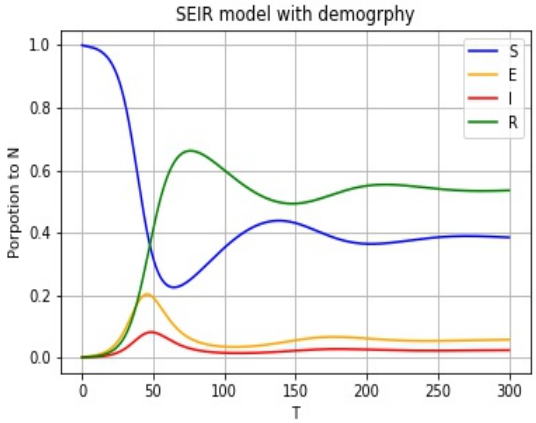

Fig. 15. SEIR model, β=1 day, 1/γ=3days, 1/μ=70years. 1/σ=7days.

Besides, we plot SEIR model's phase plane to verify the dynamic. As the Fig.17 shows,we used many different S,E,I,R and it seems all the curve will converge to the fixed point(red dot in Fig) if $R_0>1$, which is exactly similar to the SIR model.

Then we add seasonal forcing to the SEIR model, the the transmission rate β is variable now, and the equations:

$$\beta(t) = \beta_0(1 + \beta_1 cos(\omega t))(26)$$

Then we used this equations to rewrite the equations 22-25 and plot the curve. As the Fig.16 shows, all the compartments

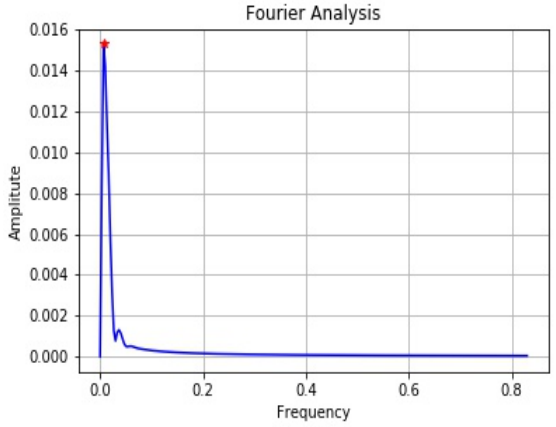

Fig. 16. Fourier analysis of infectious data. Red * is the peak.Frequency and amplitude: 0.0085 0.0153.

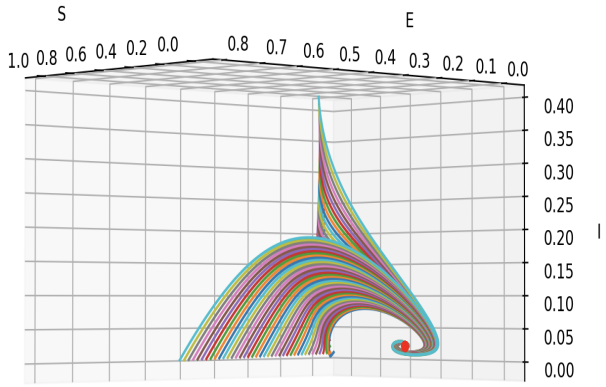

Fig. 17. SEIR phase plane,β=1 day, 1/γ=3days, 1/μ=70years. 1/σ=7days. Red dot is fixed pointnt of endemic

have the periodic signal. Fourier analysis says it have many peaks, one of peak's frequency(0.011) is similar to the frequency(0.0189) calculated by equation.

III. Results And Conclusion

In conclusion, there are four models introduced here.

1). For the basic SIR model, R_0 control the epidemic occurrence. Only if $R_0>1$, the epidemiology will show. And we find the perfect fit for the data from a school, which the 1/γ=2.2 days, β=1.67 per day, R_0=3.67. Besides, for the vaccination at least 557 person should be vaccinated.

2). For SIR model with demography, we can see the damp oscillatory which require $R_0>1$. And for the For SIR model induced mortality, f the ρ approaches to one, R_0 drop to 0. Proportion of infective and total population drop significantly.

3). For SEIR model, it has so much similarities to the SIR model, but SEIR model's pathogen have much slower invasion because it has to pass through exposed compartment.

4). For SEIR model with seasonal forcing, each compartment has period. By using Fourier analysis we find one of the peak's frequency is close to the frequency of infective curve.

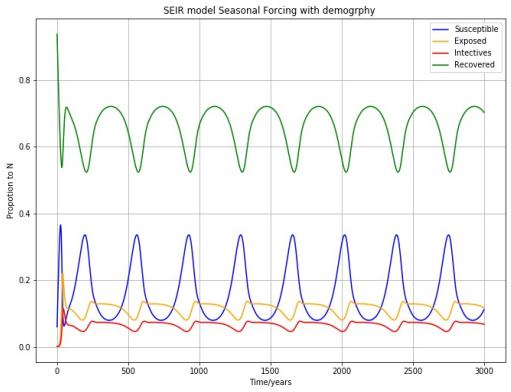

Fig. 18. SEIR model with seasonal forcing, β_0=2 days, β_1=0.6 day, 1/μ=60days, γ=1/5 day,1/σ=8 days,ω=2*π/365

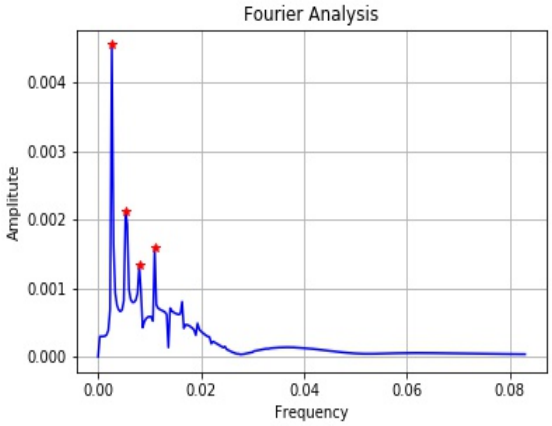

Fig. 19. Fourier analysis of infectious data. Red * are the peaks.First frequency and amplitude: 0.00267 0.0045,second frequency and amplitude: 0.00533 0.00134, third frequency and amplitude: 0.008 0.0015,fourth frequency and amplitude: 0.011 0.00159

References

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

[2] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[3] Marc Choisy, Jean-François Guégan, and P Rohani. Mathematical modeling of infectious diseases dynamics. *Encyclopedia of infectious diseases: modern methodologies*, pages 379–404, 2007.

[4] David H Bailey and Paul N Swarztrauber. A fast method for the numerical evaluation of continuous fourier and laplace transforms. *SIAM Journal on Scientific Computing*, 15(5):1105–1110, 1994.