Assignment 4: Final Project

FIT3139: Computational Modelling and Simulation

Shang-Fu Tsou 32444583

Due Date: 16 June 2023

Extended Due Date: 21 June 2023

Table of Contents

1 Specification Table

Base Model	A simulation of a COVID-19 Epidemic Based on a the SEIR Model
Extension Assumptions	The main difference between the initial assumptions for the SEIR model discussed in class and those that were used for this project, is that the population is assumed to be constant for the purpose of this project. Many new parameters were introduced to the extended SEIR models, so removing the fluctuation in the population is a way to reduce the complexity of a still-complex model. The extensions that were implemented on top of the base SEIR model takes things like local isolation policies, personal interests and strategic decision-making into consideration to better understand the epidemiological model.
Techniques Showcased	Game Theory
	Genetic Algorithm
Modelling Question 1	How does human behavior affect the containment and spread of an epidemic?
Modelling Question 2	Can the SEIR model accurately model the uniqueness of a local municipal based?

2 Introduction

The Susceptible-Exposed-Infectious-Recovered (SEIR) model is a coupled dynamical model that is commonly used to study the spread of an epidemic. The model aims to describe the social dynamics between agents. However, the base SEIR model is oversimplified in many ways. It assumes that a given population is homogenous and that everyone is equally susceptible without consideration of individuals' or groups of individuals' behavioral patterns. To target this limitation, we use game theory to model the strategic decision-making performed by rational agents who aim to act in their own best interest, or rather, what they perceive to be in their interest based on the predetermined payoffs for each strategy. The aim of this extension is to model and understand how human behaviors affect the containment and spread of an epidemic. As a result, we will be able to take into account the heterogeneity of a population and consider various groups of agents categorized by their behaviors, namely the cooperators and the defectors.

Another limitation posed by the base SEIR model is the unconditionally predetermined and fixed parameters without taking into consideration the nature of a given population, the area which the population resides, and the local policies that were implemented to combat the spread of the disease or virus. Often, these parameters are roughly estimated and defined as constants. However, with epidemiological models like the SEIR model, as the dynamics between variables become more complex from the consideration of local policies, the model becomes increasingly sensitive to initial inputs and parameters. To account for this limitation, we will be extending the SEIR model with terms that consider the local policies established, as well as an implementation of the genetic algorithm to optimize the initial parameters so that the simulated data can accurately model the real collected data of a local area and therefore be used for the analysis of future trends. This extension aims to produce a more detailed and accurate SEIR model that can account for the unique characteristics of a local area and predict the local epidemic trend.

3 Model Description

The base SEIR model is as follows:

$$\begin{cases} S_{t+1} = S_t - fS_tI_t \\ E_{t+1} = fS_tI_t \\ I_{t+1} = aE_t \\ R_{t+1} = R_t + (1-a)E_t + I_t \end{cases} \quad (1)$$

The basic idea of the model essentially describes how susceptibles become exposed a fraction at a time. Another fraction of the exposed then become infected. Finally, the infected becomes recovered.

3.1 Assumptions

- Every agent in the population is equally susceptible to the disease
- Incubation period is exactly 1 week
- Infective period is exactly 1 week
- An infected individual infects a constant fraction of susceptible individuals
- No deaths occur

3.2 Game Theory

3.2.1 Extended Assumptions

- Cooperators are those that choose to self-isolate, and defectors are those that choose to live without taking any safety measures
- Defectors have a greater probability of getting infected than cooperators
- Cooperators receive a constant payoff that reduces their probability of getting infected
- The perceived payoff for defectors is dependent on the number of exposed individuals
- Cooperators can become defectors and defectors can become cooperators
- The population remains constant, that is, the rate of natural births and natural deaths are equal and the movement of people across borders is not considered

3.2.3 Definition of Parameters

Parameter	Definition
S_C	Susceptible cooperators
S_D	Susceptible defectors
E_C	Exposed cooperators
E_D	Exposed defectors
I_C	Infected cooperators
I_D	Infected defectors
R	Recovered individuals
π_C	Cooperator perceived payoff
π_D	Defector perceived payoff
Ω	Cooperator perceived cost
δ	Perceived disease cost
β_C	Cooperator infection probability
β_D	Defector infection probability

β_a	Average infection probability which cooperator and defector interact with
a	Control parameter that determines the likelihood of cross-infection
Θ	Probability of a given agent adopting the strategy of another agent
k	Irrationality constant of the Fermi rule
Φ_S	Rate at which S_C agents convert to S_D agents
Φ_I	Rate at which I_C agents convert to I_D agents
τ	Rate at which one adopts a new strategy
γ	Probability of recovery

3.2.4 Model Description

We start by defining the perceived costs and payoffs for the cooperators and the defectors. As assumed, cooperators have a constant perceived cost while defectors have a perceived cost based on their infection probability multiplied by the perceived disease cost. Due to the existence of the exposed population, the perceived payoffs for the defectors can be dependent on X, where X is either E or I, depending on whether the agent is currently a susceptible or exposed respectively. Infected defectors' payoffs will still be dependent on I.

$$\pi_C = -\Omega \tag{2.1}$$

$$\pi_D = -\delta\beta_D X \tag{2.2}$$

To estimate the probability of a given agent adopting the strategy of another agent, we use Fermi's rule. Fermi's rule describes the transition rate from one system to another due to some perturbation. This formula is a variation of the rule. The probability is dependent on the perceived cost of each agent.

$$\Theta(\pi_i, \pi_j) = \frac{1}{1 + e^{-\frac{\pi_j - \pi_i}{k}}} \tag{2.3}$$

The value of k is set as 0.1 to account for the possibility of an agent making a mistake.

Next, we need to consider the rate at which agents convert from one strategy to another. We do so by considering the number of encounters between any cooperators and defectors. The rate at which cooperators become defectors in a given compartment would therefore be the number of interactions that caused cooperators to become defectors minus the number interactions that caused defectors to become cooperators.

$$\Phi_S = S_C(S_D + E_D + I_D)\Theta(\pi_C, \pi_D) - S_D(S_C + E_C + I_C)\Theta(\pi_D, \pi_C) \tag{2.4}$$

$$\Phi_E = E_C(S_D + E_D + I_D)\Theta(\pi_C, \pi_D) - E_D(S_C + E_C + I_C)\Theta(\pi_D, \pi_C) \tag{2.5}$$

$$\Phi_I = I_C(S_D + E_D + I_D)\Theta(\pi_C, \pi_D) - I_D(S_C + E_C + I_C)\Theta(\pi_D, \pi_C) \tag{2.6}$$

Combining the payoffs and conversion rates with the base SEIR model, we get the following:

$$\begin{cases} \dfrac{dS_C}{dt} = -S_C(\beta_a E_D + \beta_C E_C) - \tau\Phi_S \\ \dfrac{dS_D}{dt} = -S_D(\beta_D E_D + \beta_a E_C) + \tau\Phi_S \\ \dfrac{dE_C}{dt} = S_C(\beta_a E_D + \beta_C E_C) - \gamma E_C - \tau\Phi_E \\ \dfrac{dE_D}{dt} = S_D(\beta_D E_D + \beta_a E_C) - \gamma E_D + \tau\Phi_E \\ \dfrac{dI_C}{dt} = E_C - \gamma I_C - \tau\Phi_I \\ \dfrac{dI_D}{dt} = E_D - \gamma I_D + \tau\Phi_I \\ \dfrac{dR}{dt} = \gamma(I_C + I_D) \end{cases} \quad (2.7)$$

At each stage, S, E, and I, the model accounts for the possibility of a cooperator becoming a defector and the possibility of a defector becoming a cooperator. The rate at which each agent adopts a new strategy is given by the terms with $\tau\Phi$. The terms with $S_C(\beta_a E_D + \beta_C E_C)$ essentially calculates how many interactions between susceptible cooperators crossed with exposed defectors and susceptible cooperators crossed with exposed cooperators caused a susceptible member to become exposed.

Similarly for the terms $S_D(\beta_D E_D + \beta_a E_C)$, the number of interactions between susceptible defectors crossed with exposed cooperators and susceptible defectors crossed with exposed defectors that caused a susceptible defector to become exposed is given.

The final S is the sum of S_C and S_D, the final E is the sum of E_C and E_D, the final I is the sum of I_C and I_D, and the R is R.

The result is a continuous and deterministic model of SEIR.

3.3 Genetic Algorithm

3.3.1 Extended Assumptions

- COVID-19 is infectious during the incubation period, that is, the exposed can be infectious
- Susceptible, exposed, and infected individuals can all be isolated
    - Susceptible individuals get isolated if in contact with exposed
    - Exposed individuals get isolated if they get a positive PCR test with no symptoms
    - Infected individuals get isolated if they are symptomatic
- Isolated individuals are not infectious
- Instead of infective period, the isolation period is 2 weeks
- The population remains constant, that is, the rate of natural births and natural deaths are equal and the movement of people across borders is not considered

3.3.2 Definition of Parameters

Parameter	Definition	Requires Optimization
S	Susceptible individuals	N
S_i	Isolated susceptible individuals	N
E	Exposed individuals	Y
E_i	Isolated exposed individuals	N
I	Infected individuals	N
I_i	Isolated infected individuals	N
R	Recovered individuals	N
N	Total Population	N
α	Conversion rate of the exposed to the infected (inverse of incubation period)	N
μ	Isolation period	N
q_E	Isolation rate of the exposed	N
q_I	Isolation rate of the infected	N
q_S	Proportion of close contacts isolated	Y
r_1	Number of effective contacts of the infected	Y
r_2	Number of effective contacts of the exposed	Y
β_E	Probability of transmission from the exposed to susceptible	Y
β_I	Probability of transmission from the infected to susceptible	Y
γ_I	Recovery rate of an infected person	Y
γ_{II}	Recovery rate of isolated infected person	Y

3.3.3 Wuhan, China

In Wuhan, the population in 2019-2020 was reported to be 11,081,000. By definition α is the inverse of the incubation period and μ we simply set to be 14. Since the exposed only have to be isolated if they test positive with a PCR test, q_E would be the accuracy of a PCR test which would be 0.7. Infected individuals must be isolated so the probability q_I would be 1.0.

Parameter	$\boldsymbol{N}$	$\boldsymbol{\alpha}$	$\boldsymbol{\mu}$	$\boldsymbol{q_E}$	$\boldsymbol{q_I}$
Value	11,081,000	1/7	14	0.7	1.0

3.3.4 Model Description

Extending from the base model shown in equation Eq.(1), we need replace the parameters f and a with their equivalent using our model's parameters. Since we want our model to be in iterative form, we also need to make sure we always use current terms to calculate the next terms. The parameter f is just the fraction of susceptible individuals that can be infected, so it is equivalent to $\frac{r_1 \beta_I I_t}{N}$.

Parameter a is essentially our α. From there, we just need to add in the parameters that do not account for isolation (for now).

$$\begin{cases} S_{t+1} = S_t - r_1\beta_I I_t \dfrac{S_t}{N} \\ E_{t+1} = E_t + r_1\beta_I I_t \dfrac{S_t}{N} - \alpha E_t \\ I_{t+1} = I_t + \alpha E_t - \gamma_I I_t \\ R_{t+1} = R_t + \gamma_I I_t \end{cases}$$

(3.1)

We then also need to consider the possibility of transmission from an exposed to a susceptible individual as we are now assuming that COVID-19 is infectious during the incubation period. Hence, we arrive at the following system:

$$\begin{cases} S_{t+1} = S_t - \dfrac{S_t}{N}(r_1\beta_I I_t + r_2\beta_E E_t) \\ E_{t+1} = \dfrac{S_t}{N}(E_t + r_1\beta_I I_t + r_2\beta_E E_t) - \alpha E_t \\ I_{t+1} = I_t + \alpha E_t - \gamma_I I_t \\ R_{t+1} = R_t + \gamma_I I_t \end{cases}$$

(3.2)

Now to account for the policy where susceptible, exposed, and infected individuals can all be isolated, we add in our new parameters S_i, E_i, and I_i. Here, we also need to consider how N is the total population. However, when we use it for the fraction of the population that can be infected, we must now remove the isolated population, as we assume they are no longer infectious.

$$\begin{cases} S_{t+1} = S_t - \dfrac{S_t}{N - S_{i_t} - E_{i_t} - I_{i_t}}[r_1 I_t(\beta_I + q_S - q_S\beta_I) + r_2 E_t(\beta_E + q_S - q_S\beta_E)] \\ S_{i_{t+1}} = S_{i_t} + \dfrac{S_t}{N - S_{i_t} - E_{i_t} - I_{i_t}} q_S[r_1 I_t(1-\beta_I) + r_2 E_t(1-\beta_E)] \\ E_{t+1} = E_t + \dfrac{S_t}{N - S_{i_t} - E_{i_t} - I_{i_t}}(r_1\beta_I I_t + r_2\beta_E E_t) - \alpha E_t(1-q_E) - q_E E_t \\ E_{i_{t+1}} = E_{i_t} + q_E E_t - \alpha E_{i_t} \\ I_{t+1} = I_t + \alpha E_t(1-q_E) - q_I I_t - \gamma_I I_t(1-q_I) \\ I_{i_{t+1}} = I_{i_t} + \alpha E_{i_t} + q_I I_t - \gamma_{II} I_{i_t} \\ R_{t+1} = R_t + \gamma_I I_t(1-q_I) + \gamma_{II} I_{i_t} \end{cases}$$

(3.3)

At this stage, a lot of seemingly new parameters have been introduced to the system. Firstly, we replaced $\frac{S_t}{N}$ with $\frac{S_t}{N - S_{i_t} - E_{i_t} - I_{i_t}}$ to remove the isolated individuals from contributing to further infection or exposure. Then, wherever there was a possibility of infection or exposure, that is, we now need to consider the fact that they wont necessarily be infected or exposed, instead, they will enter isolation. For example, for the equation of $S_{i_{t+1}}$, we multiplied the susceptible population by q_S, to get the fraction of close-contact susceptible individuals that are isolated. Instead of $r_1\beta_I I_t + r_2\beta_E E_t$, we now have $r_1 I_t(1-\beta_I) + r_2 E_t(1-\beta_E)$ because we want the susceptible individuals who will not become infected to go into isolation, hence the terms $1-\beta_I$ and $1-\beta_E$.

Similarly, we further subtract $\alpha E_t(1-q_E)$ from the exposed as these are the people that do not go into isolation but become part of the infected population. The same concept is applied to the transition from the infected to isolated infected or the recovered.

There is one last thing we need to consider, which is the timeline in which the disease spread. The first 14 days would be the initial infection period. Since the incubation period is 7 days, people are unlikely to be infected until at least day 8. We then give the virus a week to spread. These first two weeks will be modelled by Eq.(3.2). This is also to model the actual practices local authorities have conducted in Wuhan, China. After these two weeks, the danger is recognized, and isolation measures are then put in place. Since the isolation period is 14 days, the next two weeks will be modelled by Eq.(3.3). However, after these two weeks, the isolated susceptible individuals that come out of isolation without testing positive will now return to the susceptible population. The model after the first 28 days will therefore be as such:

$$\begin{cases} S_{t+1} = S_t - \dfrac{S_t}{N - S_{i_t} - E_{i_t} - I_{i_t}} [r_1 I_t(\beta_I + q_S - q_S\beta_I) + r_2 E_t(\beta_E + q_S - q_S\beta_E)] + S_{i_{t-\mu}} \\ S_{i_{t+1}} = S_{i_t} + \dfrac{S_t}{N - S_{i_t} - E_{i_t} - I_{i_t}} q_S[r_1 I_t(1-\beta_I) + r_2 E_t(1-\beta_E)] - S_{i_{t-\mu}} \\ E_{t+1} = E_t + \dfrac{S_t}{N - S_{i_t} - E_{i_t} - I_{i_t}} (r_1\beta_I I_t + r_2\beta_E E_t) - \alpha E_t(1-q_E) - q_E E_t \\ E_{i_{t+1}} = E_{i_t} + q_E E_t - \alpha E_{i_t} \\ I_{t+1} = I_t + \alpha E_t(1-q_E) - q_I I_t - \gamma_I I_t(1-q_I) \\ I_{i_{t+1}} = I_{i_t} + \alpha E_{i_t} + q_I I_t - \gamma_{II} I_{i_t} \\ R_{t+1} = R_t + \gamma_I I_t(1-q_I) + \gamma_{II} I_{i_t} \end{cases}$$

(3.4)

Now, the model is ready. In the end, the susceptible population will be the sum of S and S_i, the exposed population will be the sum of E and E_i, the infected population will be the sum of I and I_i, and the recovered population will be R.

3.3.5 Fitness and Selection

To measure the performance of the simulation, the total simulated infected population will be plotted against the real infected population of Wuhan across 66 days, from 26th January 2020 to 31st March 2020. This data was complied by going through the daily summaries published by Wuhan Municipal Health Commission on their official website and is in the python source code. The raw data is collected in the forms of daily new COVID cases (labelled 'Iraw') and the sum of daily recovered and deaths (labeled 'RREAL_WH'). We subtract the list of 'removed' from the list of new cases pairwise, which gives us the net increase or decrease in COVID patients each day. A new list is then created where each index is the sum of the net COVID patients up to that day (labelled 'IREAL_WH').

The fitness of the result is determined by calculating the root square mean error from the simulated infected population versus the actual data IREAL_WH.

$$Fitness = \sqrt{average(I_{sim} - I_{real})^2}$$

(3.5)

The lower the fitness value, the closer the simulated data is to the real data. In terms of implementation, this means as long as the fitness value is greater than a certain threshold, we keep modifying the initial parameters until we eventually find a set of values that are good enough and can accurately model the real data. If that is the case, it would be that we can use our model to predict the future trends of the epidemic locally.

4 Results

4.1 Game Theory

4.1.1 Constant Parameters Used

Parameter	k	a	Ω	τ	γ	β_C
Value	0.1	0.1	1	1	1	1

4.1.2 Variable Parameters Initial Values

Parameter	β_D	β_a	δ
Value	10	$\frac{a(\beta_C + \beta_D)}{2}$	10

4.1.3 Results

As an initial condition, we set the population so that only a very small fraction of the fraction have been exposed or infected. We let $S_0 = 0.99$, $E_0 = 0.9(1 - S_0)$, and $I_0 = 0.1(1 - S_0)$. We assume that the population starts with an equal number of cooperators and defectors. The system of differential equations is solved using the SciPy function ODEINT. The resulting plot shown in Figure 1.

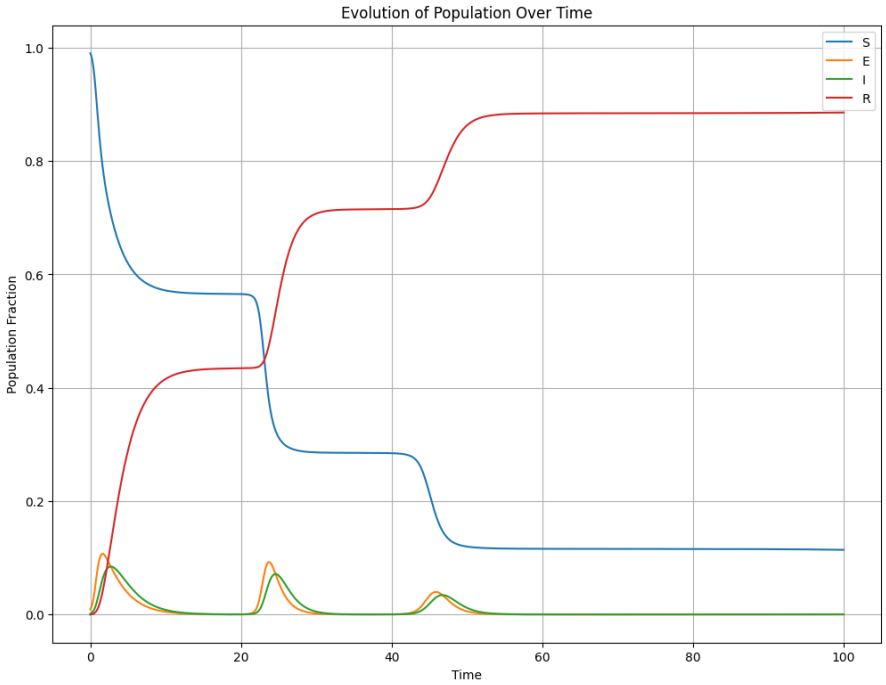

Figure 1

Some key trends that we can immediately identify is that the local maximums of the exposed always corresponds to the exponential decrement of the susceptibles and when the susceptible population starts to plateau, the infected reaches a local maximum. Furthermore, when the infected population decreases, the recovered population increases. These observations are homogenous to the design of the model any agent would go from susceptible to exposed to infected then to recovery, in that exact order. The decrement of any preceding compartment would naturally correlate to the increment of the subsequent compartment.

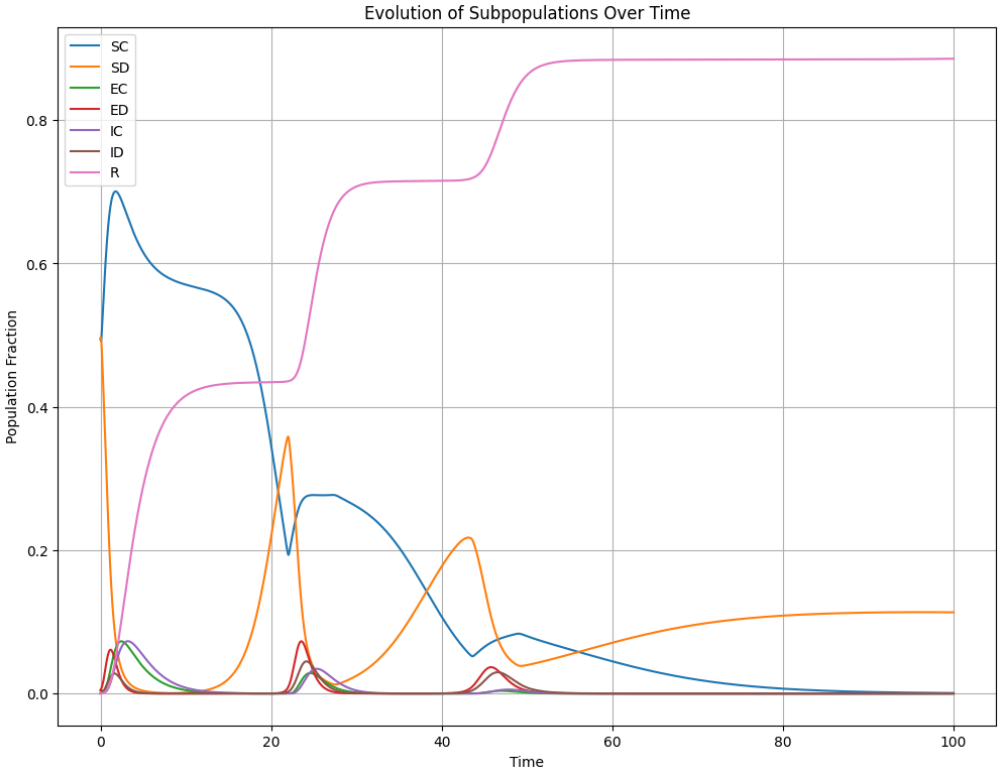

Figure 2

Figure 2 illustrates the evolution of each subpopulation over the course of the simulation. If we look at the exposed cooperators and defectors, we see that at the first local maximum, there are more exposed cooperators than defectors, and the same goes for the infected. However, by the second local maximum, the number exposed cooperators quickly dropped, while the number of exposed defectors increased and was greater than the first local maximum. Similarly, the number of infected cooperators was less than the number of infected defectors by the second local maximum. This shows that the model was successful in modelling the higher infection possibility of the defectors compared with the cooperators at the same timeframe.

If we look at the susceptible subpopulations, we can also see that the local maximum of one corresponds to the local minimum of the other, and vice versa. This shows demonstrates the conversions between susceptible cooperators and defectors along with the possibility that an agent takes on the strategy of another. From day 50 and onwards, we can also observe a trend that shows the number of susceptible defectors continue to increase while the number of susceptible cooperators decrease and approach 0. This could be a sign that the remaining population have either been infected and recovered, or the susceptible cooperators that have yet to be infected are starting to convert to the defectors. However, it can be expected that eventually all subpopulations will plateau and reach an equilibrium. If we plot this over 200 days, we get the following graph.

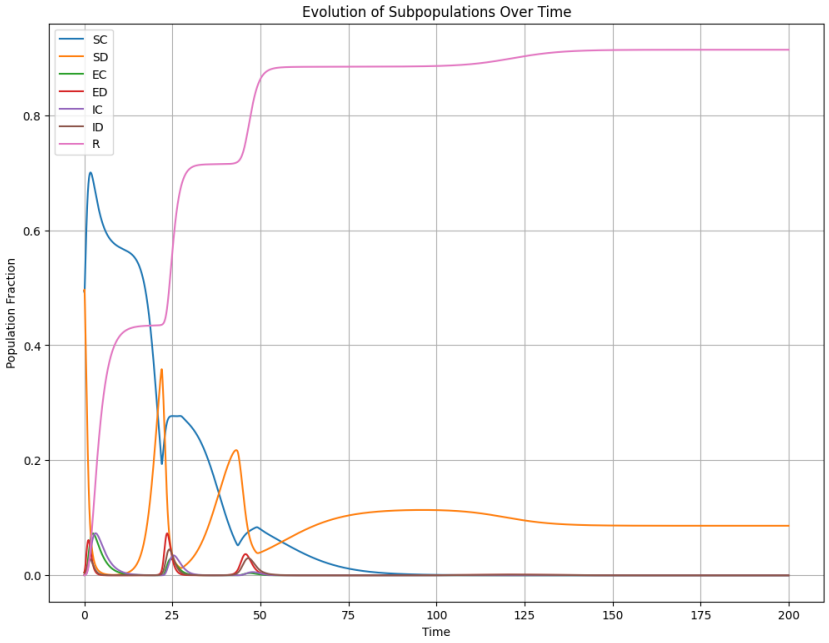

Figure 3

We can see that in Figure 3, recovered and susceptible defectors start to plateau at around day 125, while the other populations approach 0. If we consider one of the methods of cooperation as getting vaccinated, for example, this would be an example of crowd immunity, where those that have been cooperative are either no longer susceptible as they developed immunity through vaccination or they have been infected previously. Since the majority of the population is now healthy, the susceptible defectors are also unlikely to become exposed or infected. However, based on real-world cases, we know this is extremely difficult to achieve, which shows the limitation of this model, that is, the lack of consideration for viral mutation, as well as population density or environment factor that could reinitiate another cycle of infection.

From the above results, we can see that the existence of defectors causes a perturbation in a population that increases the spread of a disease. However, as long as a majority of the population is or becomes cooperative, in theory the spread of an epidemic can be drastically reduced.

4.2 Genetic Algorithm

4.2.1 Optimized Parameters

By implementing the genetic algorithm and fitness function, eventually we will be able to estimate the following parameter values that produce a model that is accurate enough to simulate the real-world data of a local area.

Parameter	q_S	r_1	r_2	β_E	β_I	γ_I	γ_{II}	E
Value	0.0046	4.6701	4.1396	0.0244	0.4896	0.0752	0.7522	606

4.2.2 Results

The results from parameter optimization were substituted into the system of equations in Eq.(3.2), Eq.(3.3), and Eq.(3.4). The correlation between the simulated results and real values was strong and proved that genetic algorithm in combination with the more detailed SEIR model is capable of restoring the unknown parameters listed in *4.2.1 Optimized Parameters*.

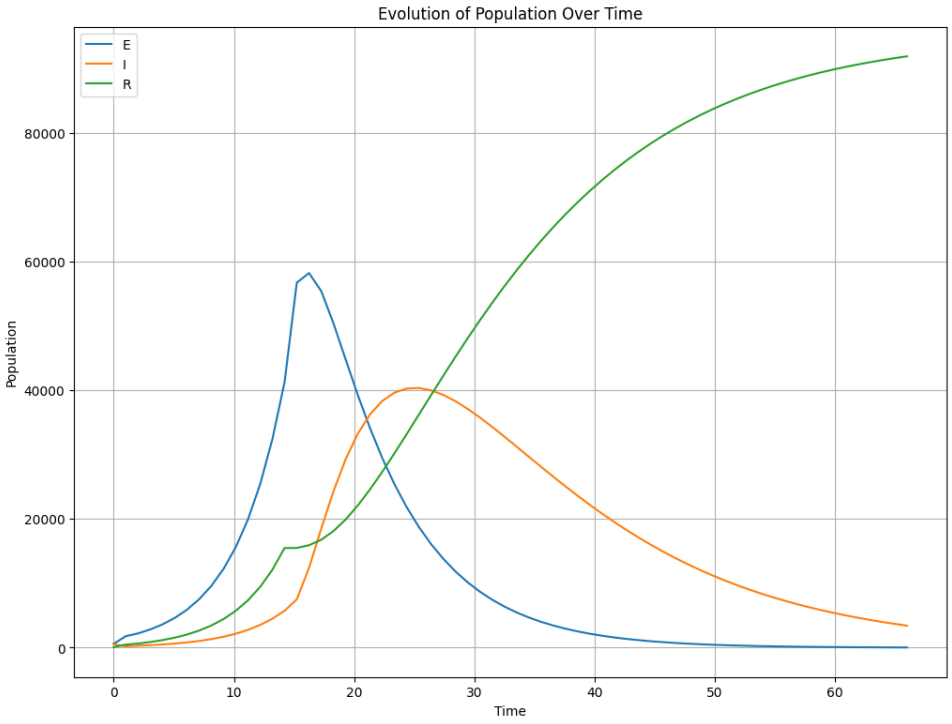

Figure 4

From Figure 4, we can see that the simulated data shows a clear difference between before and after day 14. As the first 14 days do not take into account the isolation policies, we can observe that the number exposed and increases exponentially. We see that the number of exposed peaks at day 16, after which, the population starts to drop, while the infected population rises at a much steeper rate, peaking at around day 25. We can note here that day 28 and onwards, we start bringing the isolated susceptible individuals back into the susceptible population. However, the infected population and the exposed population had both reached a local maximum before day 28 and there the number of exposed and infected was on a steady decline ever since. Bringing the isolated susceptible individuals back does not seem to increase the number of people exposed or infected, as we also have to consider how the infected and the exposed can be isolated and removed from the infectious population.

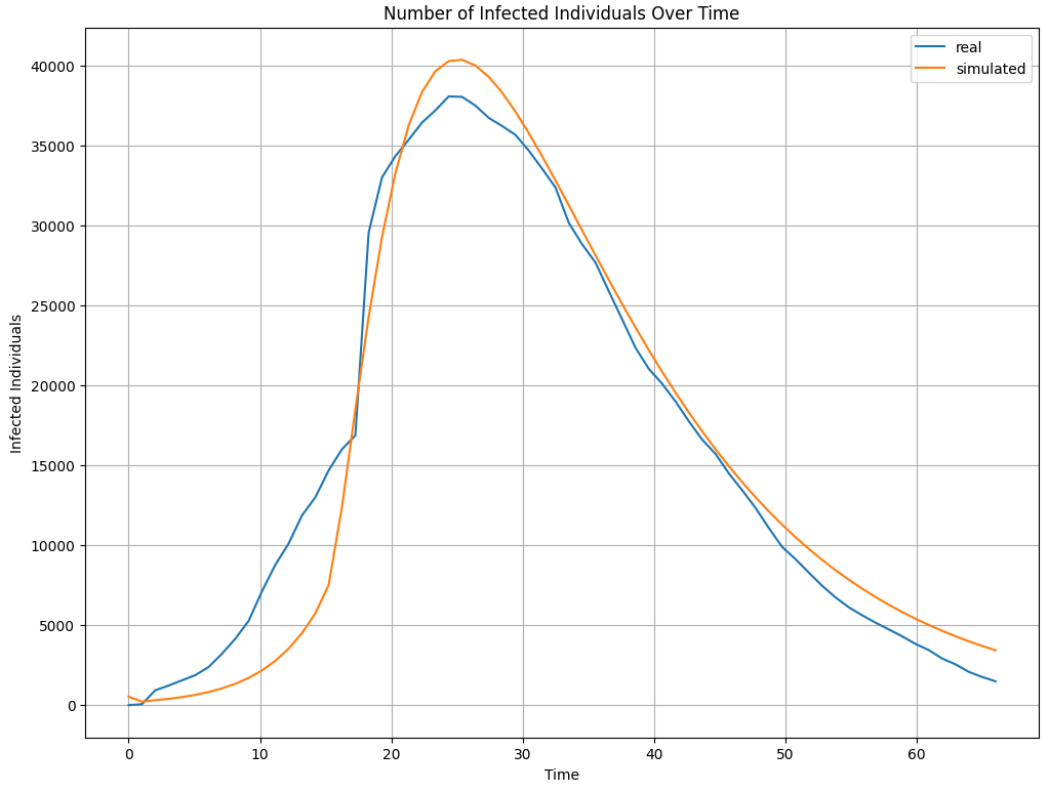

Figure 5

If we graph the simulated data against the actual data collected for Wuhan in 2020, we get Figure 5. As the figure shows, the simulated data adheres quite closely to the real data, particularly when the number of infected individuals is drastically increasing or decreasing. This set of data yielded a root mean square error of around 68, which, given an actual maximum number of infected individuals of around 37500, the RMSE is relatively small.

5 List of Algorithms and Concepts

Key Concepts

- Coupled Models and SEIR (Week 5 Content)
- Evolutionary Computation & Genetic Algorithm (Week 9 Content)
- Game Theory (Week 11 Content)

Key Algorithms

- SEIR (Week 5 Live Coding)

Further Credit & Source for Algorithms and Derivations

- Amaral, M. A., Oliveira, M. M., & Javarone, M. A. (2021). An epidemiological model with voluntary quarantine strategies governed by evolutionary game dynamics. Chaos, solitons, and fractals, 143, 110616. https://doi.org/10.1016/j.chaos.2020.110616
- Qiu, Z., Sun, Y., He, X., Wei, J., Zhou, R., Bai, J., & Du, S. (2022). Application of genetic algorithm combined with improved SEIR model in predicting the epidemic trend of COVID-19, China. Scientific reports, 12(1), 8910. https://doi.org/10.1038/s41598-022-12958-z