Applied math Project3

20180075 Kim myeong cheol

May 2025

1 Code

All implementations for the simulation of the virus evolution model discussed in this report are available on GitHub:

`https://github.com/dayforday2468/Applied_math_Project3`

This repository contains the following scripts:

- `evolution.py`: Implements the genetic programming framework for simulating virus evolution.
- `feasible.py`: Defines the feasible parameter region for transmission and fatality rates based on theoretical hypotheses.
- `SIR_model.py`: Defines the modified SIR model equations. This script also includes functions for solving the system numerically.

All scripts are located in the `code` folder, and output figures are saved in the `result` folder.

2 Motivation

In 2019, the world experienced a massive COVID-19 pandemic. At the beginning, many predicted that the virus would become a major threat to humanity, bringing dramatic changes to our lives. For a while, this seemed to be true. To prevent the spread of the virus, social distancing was introduced. As a result, social connections weakened, and online activities began to dominate over offline interactions. Daily life became more individualized.

However, we are no longer afraid of COVID-19. Even though news report infection and death statistics, few people pay attention anymore. This made me wonder whether the change is solely due to vaccines. Since immunity wanes over time and the virus has undergone mutations that evade existing immunity, a similar pandemic could have reemerged if vaccines were the only factor. However, as of 2025, no such resurgence has occurred, which makes me question whether the vaccine alone can explain the current stability.

In my opinion, this new stability is largely due to the evolution of COVID-19 itself. In the early stages, the virus had not yet adapted to humans—it had a high fatality rate and a relatively low transmission rate. But such traits are not advantageous for long-term survival. Evolution favors strains with lower fatality and higher transmission, making coexistence with humans more likely.

3 Hypotheses

To explain the evolutionary trend of COVID-19, two representative hypotheses are proposed in [1].

Hypothesis 1 (Avirulence Theory)**.** *This hypothesis assumes that viruses evolve to become less virulent over time. Lower virulence enables infected hosts to remain mobile and socially active, thereby increasing the chance of transmission. In this view, evolutionary pressure selects for strains that cause milder symptoms and promote coexistence with the host population.*

Hypothesis 2 (Virulence–Transmission Trade-off)**.** *This hypothesis suggests that there is a trade-off between virulence and transmission: while higher virulence may increase infectivity, it also shortens the infectious period by killing the host. Evolution thus favors an optimal level of virulence that balances these opposing effects and maximizes the virus's overall fitness.*

Throughout this report, **we interpret virulence as the fatality rate** of the disease. While virulence can encompass broader pathological effects, fatality rate serves as a practical and quantifiable proxy in the context of pandemic modeling and evolutionary dynamics.

4 Problem

We aim to demonstrate the two verbal hypotheses through quantitative analysis. To simulate the evolution of the virus, we use **Genetic Programming**.

4.1 SIR Model with Fatality Rate

To investigate the virulence–transmission trade-off, we consider a modified SIR model where the virus is characterized by two key parameters: **the transmission rate** and **the fatality rate**.

Unlike the standard SIR model, which assumes a fixed population size, we introduce demographic dynamics to account for both virus-induced deaths and natural processes. Specifically, we include a logistic birth term and natural mortality across all compartments. We assume that all living individuals contribute equally to reproduction and that all newborns enter the susceptible class.

The resulting system of differential equations is:

$$\dot{S} = N\left(1 - \frac{N}{K}\right) - \alpha\frac{S}{N}I - \mu S$$
$$\dot{I} = \alpha\frac{S}{N}I - (\beta + \gamma + \mu)I$$
$$\dot{R} = \gamma I - \mu R$$

Symbol	Description	Variable Parameter
N	Total population	–
K	Environmental capacity	No
α	Transmission rate	**Yes**
β	Fatality rate	**Yes**
γ	Recovery rate	No
μ	Natural death rate	No

Table 1: Model parameters and their roles

5 Approach

5.1 Initial Population

To see the spread of the virus, we set the initial population as follow:

$$S_0 = 499 \quad I_0 = 1 \quad R_0 = 0$$

5.2 Non-variable Parameters

To simplify our analysis, we minimize redundancies wherever possible. As part of this effort, we fix the non-variable parameters as follows:

Symbol	Description	Value
K	Environmental capacity	500
γ	Recovery rate	0.07
μ	Natural death rate	0.00002

Table 2: Non-variable parameters and their values

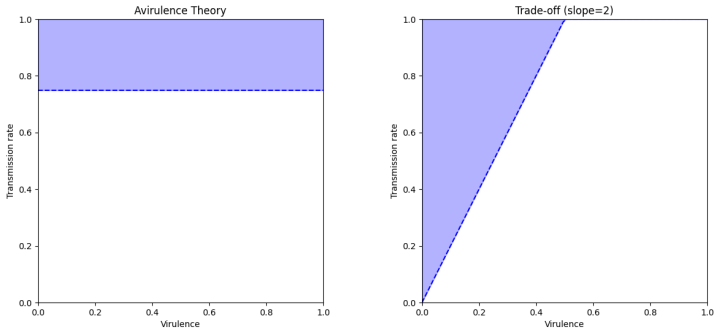
The value of K is set to ensure consistency with previous projects. The recovery rate γ is based on an assumed average recovery duration of 14 days. The natural death rate μ is estimated from an annual mortality rate of 0.01%.

5.3 Variable Parameters

The variable parameters in our model are the transmission rate α and the fatality rate β. Inspired by the theoretical constraints suggested in [1], we define the feasible regions differently for each hypothesis.

For Avirulence Theory, we assume that there is a strict upper bound on the transmission rate, regardless of virulence. For the Virulence–Transmission Trade-off, we assume a linear trade-off between the two parameters, where the slope of the trade-off line defines the relationship.

The following figures illustrate the feasible regions in each scenario:



(a) Avirulence Theory (b) Trade-off (slope=2)

Figure 1: Feasible parameter regions under different hypotheses. Shaded areas indicate the disallowed combinations.

5.4 Genetic Programming

We use Genetic Programming (GP) to simulate the evolution of virus traits. Our GP framework is tailored to reflect the biological characteristics of virus evolution, specifically the asexual reproduction mechanism. Therefore, we do not use crossover.

The procedure is as follows:

1. **Initialization**: We generate an initial population of viruses, each characterized by random values of transmission and fatality rates within the feasible region.

2. **Mutation**: New offspring are generated by introducing small random changes to the traits of selected parent viruses, allowing exploration of the trait space.

3. **Selection**: After simulating each virus's fitness, we select a subset of individuals with higher fitness for survival into the next generation.

4. **Termination**: The process is repeated for a fixed number of generations.

The initial population is sampled from a uniform distribution over the entire feasible region. For mutation, small random perturbations are drawn uniformly from a square neighborhood of side length 0.04 around each parent's traits, constrained within the feasible region.

The GP parameters are given as follow:

Parameter	value
Population size	100
Mutation rate	0.1
Mutation range	0.02
Max generation	50

Table 3: GP parameters

5.5 Fitness

For simplicity, we define fitness as **the number of infected individuals I at the end of the simulation period**. This choice reflects a straightforward measure of a virus's short-term impact on the population.

5.6 ODE Solve

Once model parameters and initial values are fixed, we numerically solve the differential equations:

$$\begin{aligned}
\dot{S} &= N\left(1 - \frac{N}{K}\right) - \frac{S}{N} I\alpha - \mu S \\
\dot{I} &= \frac{S}{N} I\alpha - I(\gamma + \beta + \mu) \\
\dot{R} &= \gamma I - \mu R
\end{aligned}$$

We integrate the system over the time period $[0, T]$, where $T = \frac{2}{\gamma}$, using the `solve_ivp` function in Python. The time period is chosen to approximate a single generation of virus.

6 Result

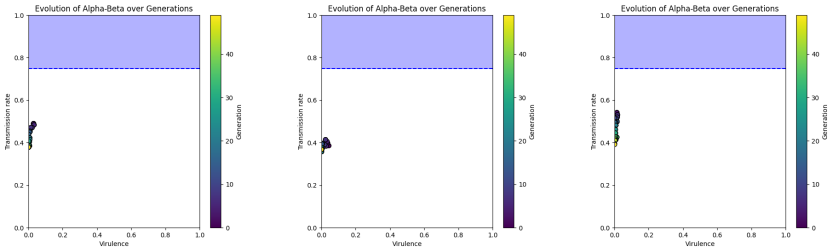
For reproducibility, we fixed the random seed in all simulations.

6.1 Avirulence Theory

The results of the genetic programming simulations under the Avirulence Theory are shown in Figure 2. Each plot represents the evolution trajectories for

different random seeds. Despite differences in initial conditions, the evolution of the virus tends to converge towards a specific region. The virulence parameter β converges to zero, aligning with the Avirulence Theory that predicts lower fatality rates in the long run.

Interestingly, the transmission rate α converges to some positive constant, likely because the fitness function is defined as the number of infected individuals after a fixed period. This means that evolutionary pressure leads to the transmission rate evolving toward the value that maximizes the number of infected individuals at the measurement time, rather than favoring long-term coexistence. If we used a different fitness function, the evolutionary trajectories might shift toward maximizing transmission, as predicted by the Avirulence Theory.



(a) $seed = 1$ (b) $seed = 2$ (c) $seed = 3$

Figure 2: GP result for Avirulence Theory with different seeds

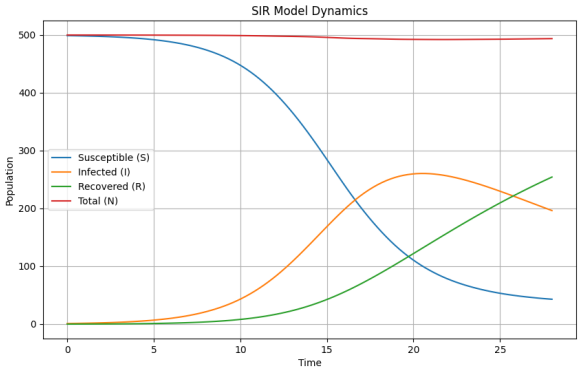
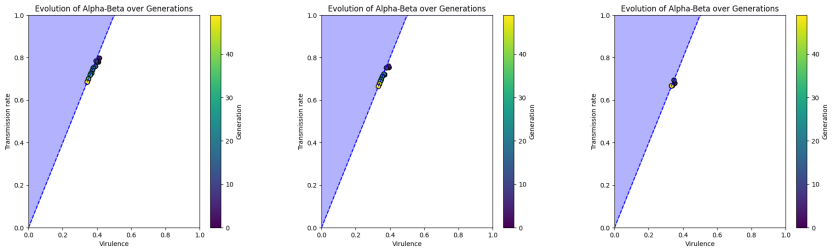


Figure 3: Infection dynamics of the best virus (seed = 1) of final generation under Avirulence Theory

6.2 Virulence Transmission Trade-off

Figure 4 shows the evolutionary dynamics of the virus under the Virulence Transmission Trade-off hypothesis. Across all runs, we observe a consistent pattern: the evolutionary trajectories converge towards a narrow region along the trade-off boundary, balancing transmission rate and fatality rate.

This behavior suggests that the feasible region acts as a constraint, guiding the evolutionary process toward a compromise between virulence and transmission. Different definitions of fitness might shift the final location, but the trade-off constraint is expected to shape the outcome similarly.



(a) $seed = 1$ (b) $seed = 2$ (c) $seed = 3$

Figure 4: GP result for Virulence transmission trade-off with different seeds

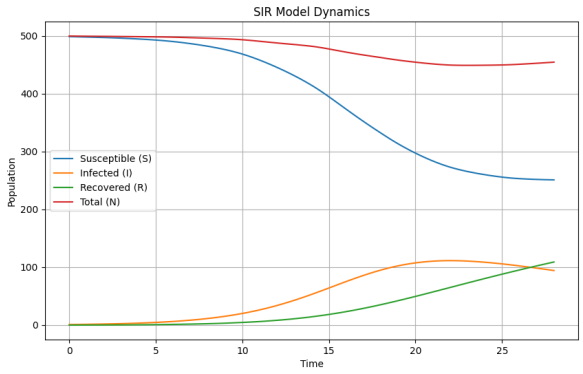


Figure 5: Infection dynamics of the best virus (seed = 1) of final generation under Virulence Transmission trade-off

7 Conclusion

7.1 Are Hypotheses Correct?

For the Avirulence Theory, we observed a discrepancy between the verbal explanation and the experimental results. While the theory suggests that viruses should evolve toward lower fatality and higher transmission rates, our simulation showed a consistent trend of decreasing fatality rates, but the transmission rate converged to a fixed value rather than continuously increasing. Therefore, based on this experiment alone, it is insufficient to confirm the validity of the Avirulence Theory. Further analysis with different fitness definitions would be necessary to draw a more robust conclusion.

In contrast, for the Virulence–Transmission Trade-off, our results closely align with the verbal explanation. The evolutionary process stabilizes at a point that balances the trade-off between transmission and fatality rates, as the theory predicts. The experimental evidence provides strong support for the validity of the Virulence–Transmission Trade-off hypothesis in our model setup.

7.2 Fitness Suggestion

To explore the evolution of virus traits under different selection pressures, we suggest alternative fitness definitions:

- **Max** I: The peak infected population during the simulation period. This measures the virus's potential to cause large outbreaks.
- **Final** $I + R$: The total number of individuals who have been infected. This is proportional to the cumulative viral replication and indicates the virus's long-term reproductive success.
- **Basic Reproduction Number** R_0: The expected number of secondary infections generated by one infected individual in a fully susceptible population.

7.2.1 Deriving R_0

Let us define R_0 as the expected number of secondary infections caused by a single infected individual in a fully susceptible population. Starting from the infection equation:

$$\dot{I} = \alpha \frac{S}{N} I - (\gamma + \beta + \mu) I$$

At the initial stage, we have $S = N - 1$ and $I = 1$. So,

$$\dot{I} \approx \alpha I - (\gamma + \beta + \mu) I$$

Thus, an infected individual produces new infections at rate α per unit time and remains infectious for an average duration of $\frac{1}{\gamma+\beta+\mu}$. Therefore,

$$R_0 = \alpha \times \frac{1}{\gamma + \beta + \mu} = \frac{\alpha}{\gamma + \beta + \mu}$$

References

[1] Chadi M. Saad-Roy et al. "SARS-CoV-2 virulence evolution". In: *Journal of Evolutionary Biology* 34.12 (2021), pp. 1867–1875.