Epidemiology: The Spread of Influenza in the South Atlantic

Zoe Cai, Laura Nolt, Paul Scanlon, Zach Valadez

Colorado State University, Math 435

Abstract

The purpose of this paper is to analyze the spread of Influenza virus in the South Atlantic. The aim is to identify the rate Influenza is spreading, in an attempt to come up with solutions of how slow the spread. Using a simple SIR model, there seems to be evidence for a few different public health solutions to further slow the spread of Influenza.

1. Introduction/Motivation

Influenza or the flu is one of the most common illnesses today and can be contracted in many different fashions. Influenza causes anywhere from 5 to 20 percent of the US population to have the virus and can cost hospitals more than 10 billion dollars a year [1]. The motivation is that the virus is fairly easy to prevent in most cases. With a good past data set, we can create a model that will observe how Influenza spreads in order to prevent more spreading or future outbreaks. This would not only help people avoid catching the virus, but also save hospitals billions of dollars a year.

1.1. About the Disease

Influenza viruses cause a very contagious respiratory virus that targets one's nose, throat, and lungs. Influenza can be characterized by symptoms such as cough, sore throat, runny or stuffy nose, body aches, fatigue, and sometimes even fever or vomiting. These symptoms usually appear two days after the infection, but appearance of these symptoms can range between one to four days after the virus enters the body. Once contracting the virus, a person is the most contagious 3-4 days after experiencing major symptoms. However, people infected with Influenza can start infecting others as early as one to four days before their symptoms develop and up to a week after the symptoms present themselves. People with weaker immune systems can pass the virus for more than one week [8]. Due to the highly contagious nature of the Influenza viruses, there are numerous ways a person can contract the virus. It is believed that the virus is spread more commonly through person-to-person contact with one who is contagious, usually when the infected person coughs, sneezes, or talks. These actions cause droplets of the virus to land on nearby people, infecting them through the mouth, nose, or inhaling of the droplets. Less common, the Influenza virus can be contracted when an infected person leaves behind traces of the virus on surfaces or objects. In this case, one may contract the disease when other people touch these surfaces and proceed to touch their mouth, nose, or eyes [7]. Influenza is a virus that can be contracted all year round, but it does see an increase in activity in October, followed by peak activity from December to February before activity slows down again in April. While it is uncommon,

Influenza activity can last as late as May [6]. Although the Influenza viruses range from mild to severe symptoms, they can also lead to more complications, such as bacterial pneumonia and sinus infections. Influenza can be deadly to those who are already sick, age 5 or younger, age 65 or older, pregnant women, and anyone with chronic medical conditions such as asthma, diabetes or heart disease. It is important to take precautionary measures to avoid contracting any of the Influenza viruses. Such measures include getting the Influenza vaccine and being aware of covering coughs, covering sneezes, and frequently washing hands [8].

1.2. Annotation

k: number of contacts an infected individual has with susceptible individuals per unit time
τ: rate of influenza transmission
b: number of secondary infections an infected individual makes per unit time
α: normalized number of secondary infections an infected individual makes per unit time
β: recovery rate
$\frac{1}{\beta}$: duration of infection

1.3. SIR Model

Also known as a "Compartment Model", SIR models are widely used for the study of Epidemiology. There are three main groups in this model: those who are susceptible (S), infected/infectious (I), or recovered/removed (R). The SIR model is especially good for predicting the spread of diseases and illnesses, but it also offers a way to see how these illnesses and diseases would be affected under different scenarios, such as vaccine distribution. The way this model works is by compartmentalizing, as its other name gives away. In the model S, I, and R are each variables that represent the number of people in each category at any given time. In Figure 1 we can see how the flow of the SIR model goes from susceptible, to infected, and finally to recovered. In a more complicated model, we can have recovered going back to susceptible, as well. Some may even be SIS models, where each individual will bounce back and fourth between susceptible and infectious.

Figure 1: SIR flow from "The SIR model" by Sinead Morris

As mentioned above, the model variables S, I, and R are all in respect to time. So, our variables are all functions as time such that $S = S(t)$, $I = I(t)$, and $R = R(t)$ with t being time. These functions allow dynamic fluctuation and change over time, which is why the SIR model is an excellent model for keeping track of the spread of diseases and illnesses. Figure 1 shows the clear transition from S to I and from I to R. Within this model, one has the basic assumption of the Law of Mass Action. This assumption states that the rate of new infection is proportional to the product of susceptible and infectious populations. From this it is easy to say there is some αI that transitions from compartment S to I and some

βI that transitions from compartment I to R [3]. Utilizing differential equations, one can represent the model of a certain population N as follows:

$$dS/dt = -\alpha SI \tag{1}$$

$$dI/dt = \alpha SI - \beta I \tag{2}$$

$$dR/dt = \beta I \tag{3}$$

$$dS/dt + dI/dt + dR/dt = 0 \tag{4}$$

$$S(t) + I(t) + R(t) = N \tag{5}$$

$$b = k\tau \tag{6}$$

$$\alpha = b/N \tag{7}$$

1.4. Assumptions

In order to create a model, we have to make some basic assumptions. All these assumptions are to simplify the model so it is easier to use and interpret. The first assumption is that the total population is large in size and closed. A large population is needed in order for the SIR model to be accurate. We assume a closed population because if we did not, we would have to account for deaths and births and all the other variables that change the overall population we are examining. Similar to what is discussed above, we are assuming only an SIR model. This means that once a subject compartmentalizes to R, they will not be susceptible again and cannot be infected again. We also assume our Influenza has zero latency and people who become infected are immediately infectious. Lastly, we assume the members of the sample population are homogeneously distributed, which means each person encounters each other with equal probability.

2. Data

This study focuses on the 2017-2018 Influenza season in the South Atlantic area. First, we found data from the CDC with the number of people in the South Atlantic area hospitalized due to influenza viruses. We then narrowed our population to those hospitalized from October 1st, 2017 to September 30th, 2018 for the variable of Infected people. Taking the average of hospitalized patients each week, the total population of the studied area is around 196,418 people. [9]

3. Statistical Features of Infected Patients Number

3.1. Statistical Distribution Functions

By utilizing the number of total infected patients, we can model a simple distribution that shows the behavior of the Flu virus for the 2017-2018 year. By first looking at our data from a statistical lens, we can visually see and obtain an general idea of the frequency and spread of the Flu viruses. We can see that this data distribution reiterates the fact that the Flu has peak activity during and around the winter season. These spikes in virus activity are visually apparent on the distributions shown in Figure 2.

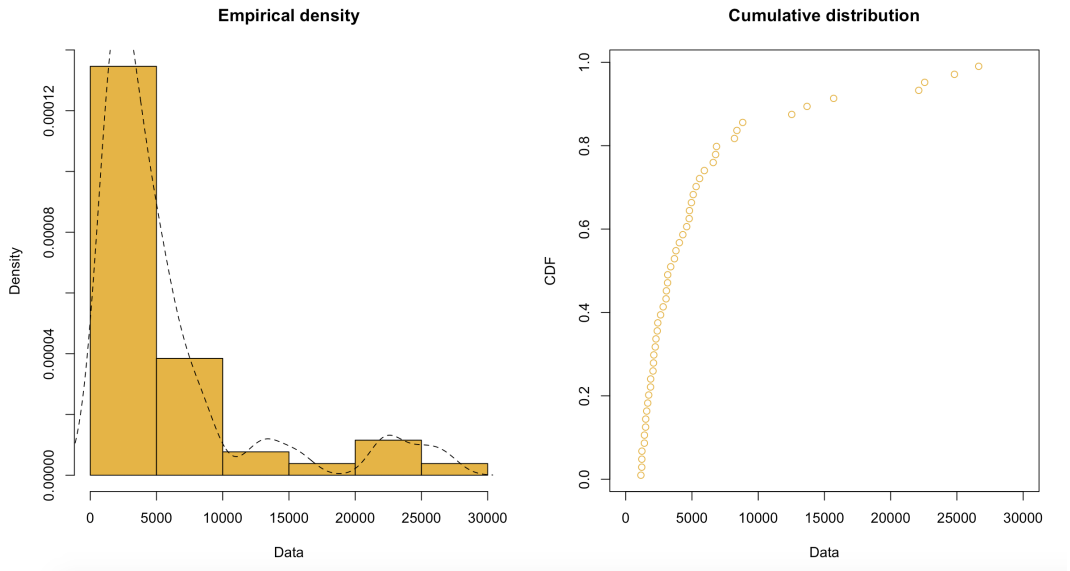

Figure 2: Density of Infected Population

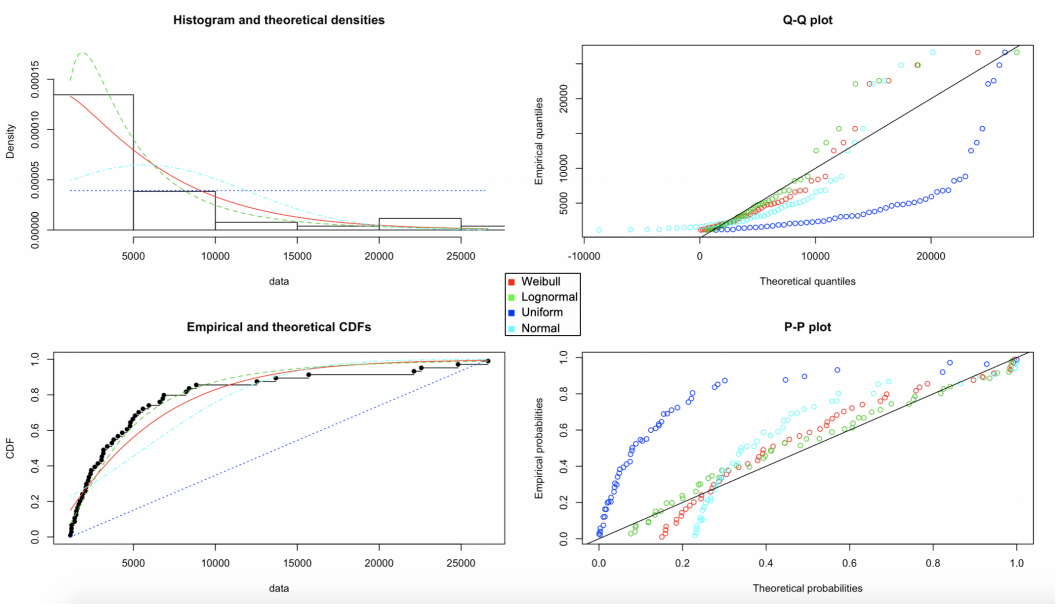

Figure 3: Distribution Fitting

Distribution functions that may be intuitively easier to compare data to include the Weibull, log normal, uniform, and normal distributions. By plotting the data against these common distribution functions, we can better assess our data and draw further conclusions. Visually analyzing the numerous graphs in Figure 3, it is evident that the Influenza density more closely resembles a log normal distribution. By definition, the log normal distribution is "a continuous probability distribution of a random variable whose logarithm is normally distributed" [4]. The log normal probability density function (PDF) and cumulative distribution function (CDF) are represented as the following equations:

$$Lognormal(\mu, \sigma^2) \tag{8}$$

$$PDF : f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\frac{-(lnx-\mu)^2}{2\sigma^2}} \tag{9}$$

$$CDF : F_X(x) = \Phi\left(\frac{(lnx) - \mu}{\sigma}\right) \tag{10}$$

$$\mu \in (-\infty, +\infty), \sigma > 0, x \in (0, +\infty)$$

where Φ is the CDF of the standard normal N(0,1) represented as $\Phi(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-t^2/2}dt$ [4]. We can also calculate the expected value ($E[X]$), variance ($Var[X]$), and standard deviation of the log normal distribution utilizing the following equations:

$$E[X] = e^{\mu+\sigma^2/2} \tag{11}$$

$$Var[X] = E[X^2] - (E[X])^2 = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1) \tag{12}$$

$$SD[X] = \sqrt{Var[X]} = e^{\mu+\sigma^2/2}\sqrt{e^{\sigma^2} - 1} \tag{13}$$

3.2. Results

Using the given CDC data, we can calculate each of these values:

$$E[X] = \mu = 5709.519,\ Var[X] = \sigma^2 = 38762663,\ SD[X] = \sigma = 6225.967$$

Examining these statistical values, a high variance indicates the data points are more spread out from each other and from the mean. Due to the high nature of these numbers and the dramatic changes of infected individuals each week, this high variance is expected. After visually deciding the log normal distribution is the better choice, we next applied goodness-of-fit tests to confirm our decision. Figure 9 shows the results of the tests.

```
Goodness-of-fit statistics
                              weibull  lognormal   uniform    normal
Kolmogorov-Smirnov statistic 0.1505099 0.09043922 0.5839013 0.2394722
Cramer-von Mises statistic   0.2995777 0.09750059 7.0317518 1.0540464
Anderson-Darling statistic   1.9939600 0.77193031       Inf 5.8879894

Goodness-of-fit criteria
                               weibull lognormal uniform   normal
Akaike's Information Criterion 1006.693  991.4341      NA 1059.154
Bayesian Information Criterion 1010.595  995.3365      NA 1063.057
```

Figure 4: Goodness-of-Fit Tests

The results of the tests show that the log normal distribution is our preferred choice over the other distributions, due to the smaller nature of the values. These small values mean that the distance between our data distribution and the log normal distribution is less than the others. Research has shown that log normal distributions are "important in the description of natural phenomena...because many natural growth processes are driven by the

accumulation of many small percentage changes" [4]. This distribution is widely used for modeling and exploring numerous epidemics.

Statistically and with the flu data that we obtained from the CDC, we can conclude that the virus spreads and grows at a logarithmic rate.

4. Simulation of 2017-18 Influenza in South Atlantic by SIR model

4.1. Preliminary Modeling - SIR model

We can now tell how the flu virus grows and spreads statistically, so we need some way to model this phenomenon. The first approach in an attempt to create this model is to work with a differential equation concept of the SIR model. We began by creating the function infections.py to model the basic step-wise behavior of the differential equation, with fixed values of alpha (α) and beta (β). A small, fake data set was then created, to ensure our model worked. We started with 100 individuals, 1 initial infected and 99 susceptible, and fixed the values of α and β, as produced in Figure 5.

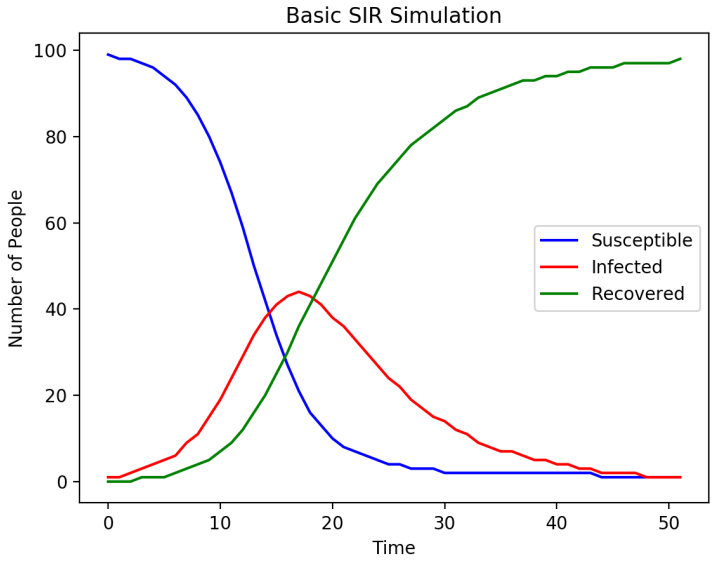

Figure 6: "Basic SIR Simulation"

The basic behavior of the model transfers all susceptible individuals to the recovered stage in an exponential manner. The individuals must become infected at some point along the distribution to make this transition. For this basic simulation, α represents an 8 day recovery period with the logic that a 0.125 rate of recovery per day models period. β represents a 50% chance of contraction, modeled as a function of the current number of infected and susceptible individuals. After creating and examining this preliminary model, we were then ready to introduce real data. Utilizing the Influenza data from the CDC, we can see how our model fits. We uploaded the CDC data, guessed and checked a set of initial values, and prepared to optimize them.

```python
import numpy as np
import scipy.optimize as opt
def infection(S, I, R, A, B, t):
    '''
    This function predicts the number of infected
    individuals at a given time based on initial value data
    '''
    t += 1
    Sdata = [0 for i in range(t)]
    Idata = [0 for i in range(t)]
    Rdata = [0 for i in range(t)]
    for i in range(t):
        Sdata[i] = round(S)
        Idata[i] = round(I)
        Rdata[i] = round(R)
        Snext = S - B*S*I
        Inext = I + B*S*I - A*I
        Rnext = R + A*I
        S = Snext
        I = Inext
        R = Rnext
    return I
if __name__ == "__main__": # Execution
    # File and variable declaration omitted for clarity.
    def obj(x): # The Objective Function
        a,b = x
        o = 0
        tv = np.arange(len(I)) # I = data
        Ipred = [0 for t in tv]
        Ipred = np.array(Ipred)
        for t in tv:
            Ipred[t] = infection(np.int(Savg), I[0], 0, a, b, t)
        o = np.linalg.norm(I - Ipred)**2
        return o
    # The following line preforms the full optimization with
    # the 2nd argument as the initial guess.
    opt.minimize(obj, (0.2, 0.0000019), method = 'Nelder-Mead')
```

Figure 5: Implementation of our Python prediction function, Python Sum of Least Squares Objective Function implementation, and SciPy variable optimization

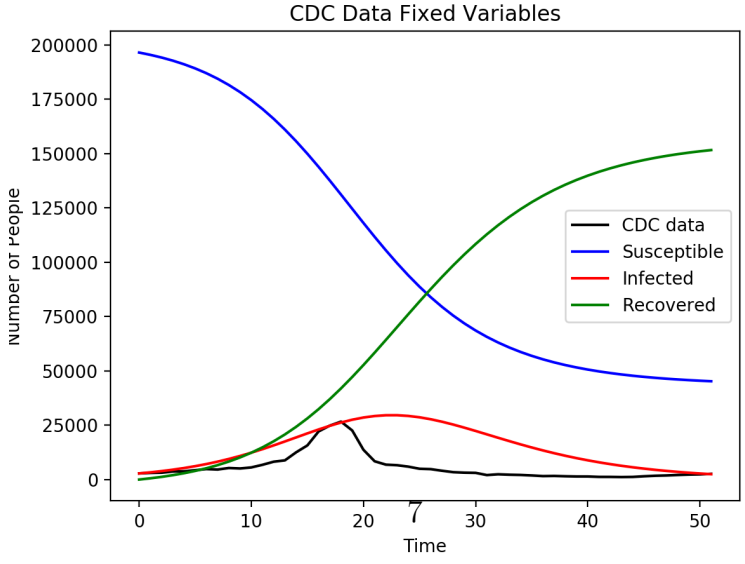

Figure 7: CDC Data SIR Simulation

In a data set of this size, one can easily see the affects of perturbations on α and β. The simulation of the hypothetical set cannot oscillate significantly, due to its very simple initial values. As a result we can see that minor perturbations on α affect the infected model's peak height, and perturbations on β affect the length of the distribution - i.e. the duration of the epidemic.

Once we found initial values for an optimization function that did not cause a null epidemic result, we were able to reasonably model a least squares objective function. From Figure 6 above, it is clear the initial guess is not optimal because the Sum of Squares error is 6.9×10^9. After some more research on how to optimize α and β, we found "On-the-fly Modelling and Prediction of Epidemic Phenomena" by Ionela Roxana Danila from Imperial College London. This study included the best algorithms for minimizing the objective function, such as the scipy.optomize.minimize function that allows specification of the Nelder-Mead algorithm [5]. The optimized results changed our initial guesses of α and β and the error reduced to 7.4×10^8, a difference of 6.2×10^9 from the original guess.

4.2. Simulation of 2017-18 Influenza in South Atlantic

To confirm the preliminary modeling done in the previous section, we used the least square method to determine the value of α and β. Subsequently, a 4^{th} order Runge-Kutta method was being used to solve the simulation numerically, with a time step of 0.0001. As a result the SSE is minimized, giving $\alpha = 2.8 \times 10^{-6}$ and $\beta = 0.3667$. Figure 8 represents the SIR model simulation of 2017-18 Influenza in South Atlantic, while Figure 9 compared the estimated number of infected people to the actual number.

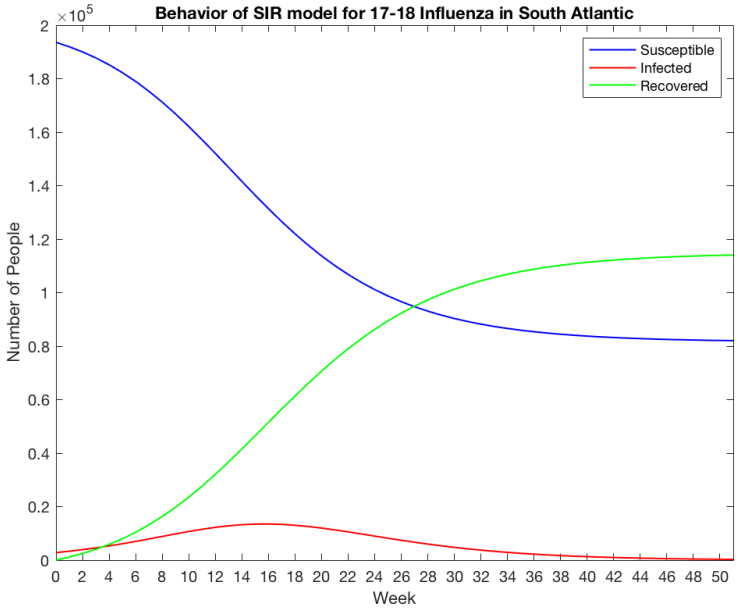

Figure 8: SIR model prediction of 2017-18 Influenza in South Atlantic

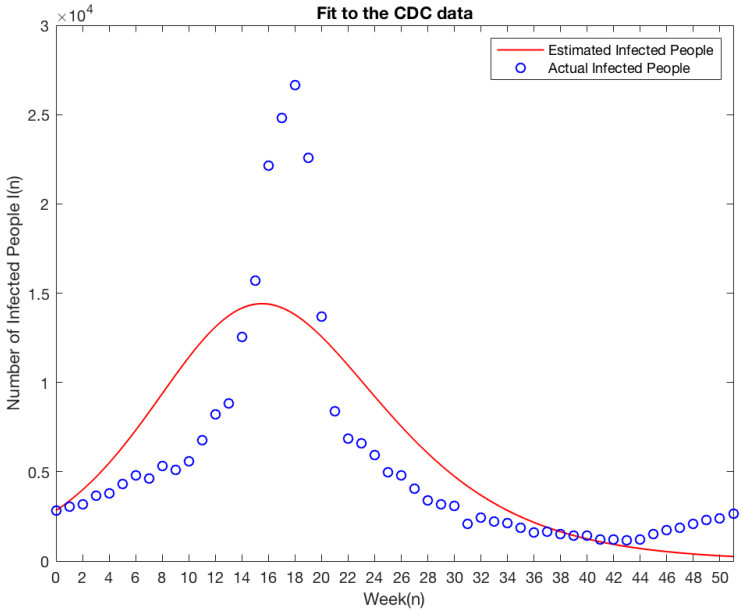

Figure 9: Estimated Number of Infected People versus Actual Number

4.2.1. Analysis of the result

Long term limits exist. From Figure 8, we can easily notice that the numbers of people in all three groups will reach long term limits at the end of the Influenza Season. This can be proved by SIR model equations. Since the right hand side of Equation (1) is negative and right hand side of Equation (3) is positive, we know that $dS/dt \leq 0$ and $dR/dt \geq 0$. This means that $S(t)$ is monotonically decreasing and $R(t)$ is monotonically increasing. Therefore, due to the boundary conditions, we can get $0 \leq S(t) \leq S(0) \leq N$ and $0 \leq R(0) \leq R(t) \leq N$, which implies that the limits of S and R exist. In other words, $S(\infty) = lim_{t\to\infty}S(t)$ and $R(\infty) = lim_{t\to\infty}R(t)$ exist, thus $I(\infty) = lim_{t\to\infty}I(t) = N - S(\infty) - R(\infty)$ exists.

The disease finally dies out. Figure 8 shows that the number of Infected people decrease close to zero at the end of Influenza Season. This can be proved by SIR model equations with contradiction. From Equation (3), we can derive that $dR/dt \geq \beta I(\infty)/2 \geq 0$. But if $S(\infty) = lim_{t\to\infty}S(t) \neq 0$, then $R(\infty) = \infty$, which is a contradiction.

The existence of an Influenza Pandemic. From Figure 8 and 9, the number of infected people increases at the beginning, reaches its maximum, and then decreases to zero as $t \to \infty$. It reaches the peak at around week 16, which is the end of January and the beginning of February. However, this is not always the case with the simulation of the SIR model. For example, in Figure 10 is the case in which $b = 0.4$ and $\beta = 0.6$. It is evident that the number of infected people decreases monotonically and quickly to zero.

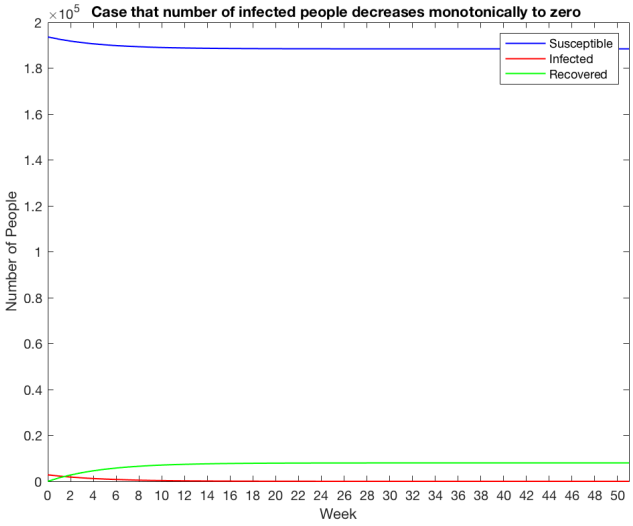

Figure 10: Special case: $\alpha = 0.4$, $\beta = 0.6$

In epidemiology, the basic reproduction rate is one of the most important quantities, denoted by $R_0 = b/\beta$. This represents the number of secondary infections one infected person generates on average over the course of their infectious period amongst a population that is totally susceptible [12]. However, not all contacts will be susceptible to infection. For example, a person could have conferred life-long immunity due to previous infection. To solve this problem, the number of new infections produced by an infected person can be better measured by Effective Reproduction Rate (R_e) [18].

$$R_e = \frac{b}{\beta}\frac{S(0)}{N} = Db\frac{S(0)}{N} = Dk\tau\frac{S(0)}{N} \tag{14}$$

R_e is commonly considered as a threshold value for an epidemic. If $R_e \leq 1$, this means that each infected individual produces on average less then one new infected individual. So, the infection can be predicted to be cleared from the initial individual. Figure 10 shows an example in the case of $R_e \leq 1$. If $R_e > 1$, then $I(t)$ starts increasing, reaches its maximum, and then decreases to zero as $t \to \infty$. This indicates an incidence of epidemic. [14]

In our study, effective reproduction rate is $R_e = \frac{\alpha}{\beta}S(0) = \frac{2.8*10^{-6}}{0.3667}193593 = 1.478$. Since this is greater than 1, it indicates that the Influenza will evolve into a Pandemic in 2017-2018, as shown in Figure 8 and Figure 9.

5. Public Health Interventions

Effective Reproduction Rate Threshold theorem provides us with ways to prevent an epidemic, including reducing R_e to less than one. According to Equation (14), we can reduce the duration of infection D, the contact rate k, the transmission rate τ and the initial susceptible population to make the Effective Reproduction Rate as small as possible. Examples of specific actions are shown below.

5.1. Reduce the duration of infection D by taking antivirals such as Tamiflu within 24 hours.

Tamiflu(oseltamivir) is a prescription medication used to treat influenza, or flu, virus. It can shorten the duration of the flu if taken as soon as symptoms start. A 2012 analysis compared the effect of oral oseltamivir with no treatment and found that oseltamivir reduced fever duration by 33 hours.[15] Similarily, two 2014 systematic reviews found that oral oseltamivir reduced time to alleviation of symptoms in adults and children by about 17 hours to 29 hours compared with placebo.[16] [13] In 2016 study conducted by Blanco, et al, antiviral treatment accounts for 4%-14% decrease of infected people.[2]

5.2. Reduce the contact rate k by isolation of susceptible individuals

Voluntary home confinement of infected individuals can reduce contact between infectious people and susceptible people, so voluntary self-isolation is considered as an intervention capable of limiting the transmission of pandemic influenza by CDC.[11] Voluntary self-isolation can reduce transmission to some extent, and is extremely critical when antiviral drugs are not immediately available.[20] The study conducted by Qingxia and Dingcheng shows the reproduction rate decreases when the proportion of isolation increases. [19] More specifically, patient isolation will introduce a 5%-16% of infected cases reduction.[2]

5.3. Reduce S(0) by offering some percentage of population Flu Vaccine

Assumes that vaccine is 100% effective, then although vaccinating an entire population can prevent the epidemic of Influenza, it would be super expensive and not practical. Therefore, one question arose that what fraction of the susceptible class should be vaccinated to prevent an pandemic.

Using the idea of Effective Reproduction Rate Threshold theorem, we know that to prevent an epidemic, we require $R_e \leq 1$. Denote x the fraction of the susceptible persons who gets vaccinated. Then $(1-x)S(0)$ represents the population of susceptible people. To prevent an epidemic, we want $(1-x)S(0)\frac{\alpha}{\beta} \leq 1$, and the result is $x \geq 1-\frac{1}{R_e} = 1-\frac{1}{1.476} = 32\%$ in our case. The adjusted vaccine effectiveness against medically attended influenza is around 65% [10], therefore, it is reasonable to estimate that around $\frac{32\%}{65\%} = 49.2\%$ of the population in South Atlantic should be vaccinated to prevent an Influenza pandemic.

5.4. Reduce the rate of transmission τ by washing hands frequently or avoiding crowds during Influenza Season.

The likelihood to be infected by influenza can also be reduced by washing with soap and water or avoid crowds such as wearing face masks. [11] Hand-washing or use of face masks will introduce a 11%-27% or 3%-10% of infected cases reduction.[2] A case-control study to test whether the risk of influenza transmission associated with self-reported hand-washing shows that odds ratios of influenza infection decreased around 26% by hand-washing. [17]

In our study, we introduce three specific controls and simulate SIR Influenza model respectively to show the efficiency of introduced controls. The first method is that 49.2% of population get vaccinated as calculated above. The second control is decreasing 10% of the duration of infectiousness by taking antiviral drugs within 24 hours after the appearance of symptom, which is reasonable since a flu will typically last one or two weeks and we estimate Tamiflu can reduce the duration of the flu by around 1 day. The third control is washing hand frequently or self-isolation, and the combined effect is estimated to decrease the contact

rate by around 10%. Figure 11 shows the curve of Infected people for three interventions and actual number.

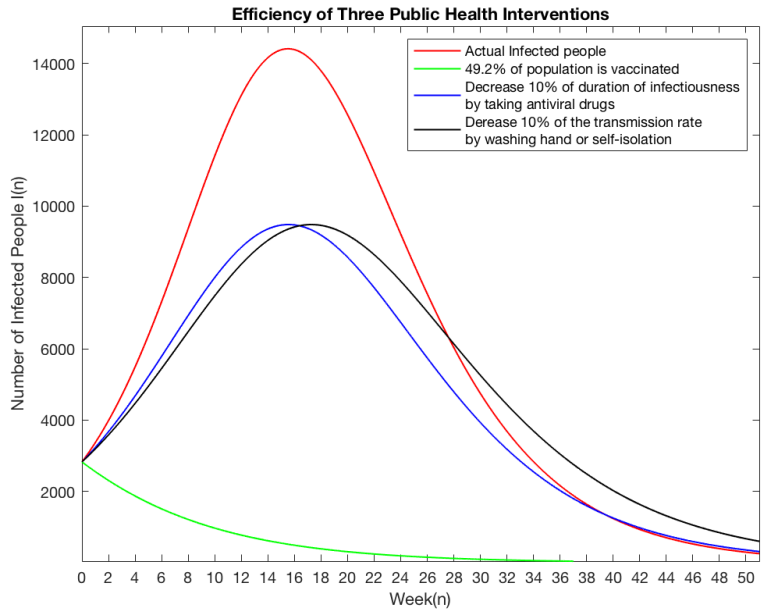

Figure 11: Effect of Three Kinds of Controls on Dynamic of 2017-18 Influenza in South Atlantic

From the figure above, we can notice that the intervention of getting 49.2% population vaccinated would not lead to a pandemic, and the number of infected people would decrease monotonically and quickly to zero instead, indicating this is an efficient public health intervention to prevent a outbreak of Influenza. The black curve and blue curve have similar shape with the same peak number of infected people, which is much lower than the actual peak number during 2017-18 Influenza Season in South Atlantic. The only difference is that the time of the peak of blue curve is a little bit earlier than that of black curve. Therefore, the interventions of taking antiviral drugs and washing hand or self-isolation decreases the extent of the pandemic a lot, and have similar effectiveness. Both of them fail to avoid a outbreak of Influenza. However, we can expect that such a pandemic can be avoided as long as more public health propaganda will be conducted. The control method of getting a fraction of people vaccinated is considered to be the most efficient way to prevent a pandemic, but the related department needs to care about those people who are not suitable to be given a vaccine.

6. Future Work

In this paper, we researched and expanded on the spread of the Flu among patients in one population. For the future directions of this study, more research can be done regarding different populations, such as the whole South Atlantic. By only studying the population of patients at doctors' offices, we miss a huge chunk of the entire population. This would affect the total number of susceptible individuals, thus changing the results of our SIR model. One could also explore different strands of the Influenza virus. For example, there are

differences between the H1N1 strands and the H3 strands, thus resulting in differing data. Further research could also include vaccines, preventative care, and treatment options. This paper could be taken in a different direction when accounting for different treatment options. Finding which ones are more or less successful could create a future study about possible eradication of the Flu viruses. There could also be a study on a more complicated model so that people are not immune after reaching recovery state as well as having a dynamic population that can increase and decrease.

6.1. Different Models

From Figure 9, we can notice that our simulation of Influenza by SIR model didn't catch all the outlier point of Infected people. This actually can be solved by adding some additional variable to the model such as in the SEIR model where there is a fourth partial differential equation added to the system for the period of "Exposure" which accounts for the time it takes to show symptoms and infect others with the E variable. This would significantly change the shape of the graph, and the step-wise prediction algorithm. Another is the MSIR model where the portion of the population immune or vaccinated, the "Maternally-Derived Immunity" variable, is taken into account with M. The same result occurs with the differential equation system and allows for more accurate comparison of the infection curve to the recovered and susceptible curves. Most commonly attributed to Influenza is the SIS model, mentioned previously, where, likely due to mutation, individuals will bounce back and fourth between susceptible and infectious.

Based on the behavior of most epidemics, a combination of all of these models would best fit real data. A model such as the MSEIRS model takes into account vaccination, immunity, mutation, exposure and incubation period, and seasonal trends would be most realistic. Fitting an SIR model to real data efficiently compartmentalizes a limited number of variables, but fails to account for multiple real life factors. Our simulation effectively applied probability, statistics, difference modeling, optimization, and error accountability to real life data despite the lack of desired fit. This analysis can be used for modeling on different demographics as well as future seasons of Influenza.

7. Conclusions

By simulating the 2017-18 Influenza in South Atlantic by SIR model, it shows that the number of infected people increases at the beginning around October of 2017, reaches its maximum at around the end of January and the beginning of February, and then decreases to zero till the end of September 2018. The Effective Reproduction Rate is 1.478 which is larger than 1, indicating that the Influenza will evolve into a pandemic. Several possible public health interventions can be introduced to slow the spread of Influenza such as reducing the duration of infection, the contact rate between susceptible and infected, the initial susceptible people and the rate of transmission. As a result, the intervention of getting 49.2% of population vaccinated to reduce the susceptible people performs best in reliving the seriousness of Influenza or even preventing a pandemic.

8. Distribution of Work

Paul: Finding data, coding/diff eq, Future Work

Zoe: Data/Assumption, coding in Matlab, simulation and analysis of model result, public health interventions

Zach: Introduction, description of SIR, slides

Laura: Introduction, description of SIR, distributions, Future Work

Appendices

A. Matlab Code

A.1. 4^{th} order Runge-Kutta method

```
%4th order Runge–Kutta method algorithm
function [T,x] = RK4(f,t0, x0, dt, tt)
%t0 is the beginning time, dt is time step size and tt is the
   ending time
%x0 is the initial value, f is the objective function
    t = t0;
    x{1} = x0;
    T(1) = t;
    i = 1;
    while (t + dt < tt)
        F1 = dt*f(t,x{i});
        F2 = dt*f(t+dt/2, x{i}+F1/2);
        F3 = dt*f(t+dt/2, x{i}+F2/2);
        F4 = dt*f(t+dt, x{i}+F3);
        x{i+1} = x{i} + 1/6*F1 + 2/6*F2 + 2/6*F3 + 1/6*F4;
        t = t + dt;
        T(i+1) = t;
        i = i + 1;
    end
    dt = tt-t;
     F1 = dt*f(t,x{i});
     F2 = dt*f(t+dt/2, x{i}+F1/2);
     F3 = dt*f(t+dt/2, x{i}+F2/2);
     F4 = dt*f(t+dt, x{i}+F3);
    x{i+1} = x{i} + 1/6*F1 + 2/6*F2 + 2/6*F3 + 1/6*F4;
    T(i+1) = t;
end
```

A.2. Least Square Method

```
%Main: Least Sqaure method
p0 = [3.5552,3.1129]; %initial guess of \alpha and \beta
f = @(p) leastsquare(p);
p = fminsearch(f, p0);
[p(1),p(2)] %give you the value of parameter \alpha and \beta


function err = leastsquare(p)
data = xlsread('data.xlsx','Sheet1','c2:c52');
```

```
b = p(1);
r = p(2);
F = @(t,X) [-b*X(1)*X(2)/196418;b*X(1)*X(2)/196418-r*X(2);r*X(2)];
    %Total population: 196418
dt = 0.001; %step size for 4th order Runge-Kutta method
tt = 51;
[T,x] = RK4(F,0,[193593;2825;0],dt,tt);

%Objective function for least square method
err = zeros(1,51);
for i = 1:51
    j = i*1000+1;
    err(i) = x{j}(2)-data(i);
end
err = err*err'
end
```

A.3. SIR model simulation

```
tic;
b=0.5468; %the estimated value from least square method
k=0.37;  %the estimated value from least square method
N=196418; %the total population
f = @(t,X) [-b*X(1)*X(2)/N;b*X(1)*X(2)/N-k*X(2);k*X(2)]; %SIR
   model
delta_t = 1 %step size which can be changed
stop_t = 51;  %in weeks
x0 = [193593;2825;0]; %for initial value: S+I=N
all_x = RK4(f,0,x0,delta_t,stop_t); %simulate SIR model by Runge-
   Kutta method
n = size(all_x)
plot([0:n(2)-1]*delta_t,all_x(1,:),'b','linewidth',1);hold on;  %
   Suscepitble people curve
plot([0:n(2)-1]*delta_t,all_x(2,:),'r','linewidth',1);hold on;  %
   Infected people curve
plot([0:n(2)-1]*delta_t,all_x(3,:),'g','linewidth',1);hold on;  %
   Recorvered people curve
xlim([0,51])
set(gca,'XTick',[0:2:51])
title('Behavior of SIR model for 17-18 Influenza in South Atlantic
   ')
xlabel('Week'), ylabel('Number of People')
legend("Susceptible","Infected","Recovered")
toc;
```

A.4. Compared Estimated Infection People Number with Actual Number

```
b=0.5515;
k=0.3667;
N=196418;
f = @(t,X) [-b*X(1)*X(2)/N;b*X(1)*X(2)/N-k*X(2);k*X(2)]; %SIR model
delta_t = 0.1;
stop_t = 51;  %in weeks
x0 = [193593;2825;0]; %initial condition
all_x = RK4(f,0,x0,delta_t,stop_t);
n = size(all_x)
plot([0:n(2)-1]*delta_t,all_x(2,:),'r','linewidth',1);hold on; % the estimated number from SIR model
data = xlsread('Data.xlsx','Sheet1','c1:c52');
plot([0:stop_t],data,'bo','linewidth',0.3);hold on; %the actual number
xlim([0,51])
set(gca,'XTick',[0:2:51])
title('Fit to the CDC data')
xlabel('Week(n)'), ylabel('Number of Infected People I(n)')
legend("Estimated Infected People","Actual Infected People")
```

A.5. Test the Efficiency of Three Public Health Interventions

```
%%Test the efficiency of Three Public Health Intervention%%
%First intervention: 49.2% of the population get vaccinated.
b=0.5515;
k=0.3667;
N=196418;
s_0 = 0.492*(N-2825); %place changed
i_0 = 2825;
r_0 = N-0.492*(N-2825)-2825; %place changed
x0 = [s_0;i_0;r_0];
%Second intervention:increase the recovery rate or decrease the infectious period
b=0.5515;
k=0.4074; %place changed
N=196418;
x0 = [193593;2825;0];
%Thrid intervention: decrease the contact rate
b=0.4964; %place changed
k=0.3667;
N=196418;
x0 = [193593;2825;0];


f = @(t,X) [-b*X(1)*X(2)/N;b*X(1)*X(2)/N-k*X(2);k*X(2)];
```

```
23 delta_t = 0.01;
24 stop_t = 51;
25 all_x = RK4(f,0,x0,delta_t,stop_t);
26 n = size(all_x)
27 plot([0:n(2)-1]*delta_t,all_x(2,:),'g','linewidth',1);hold on;
```

Bibliography

[1] What are your odds of getting the flu?, 2017.

[2] Natalia Blanco, Marisa C. Eisenberg, Terri Stillwell, and Betsy Foxman. What transmission precautions best control influenza spread in a hospital? *American Journal of Epidemiology*, 183(11):1045–1054, 2016.

[3] Wikipedia contributors. Compartmental models in epidemiology, 2019.

[4] Wikipedia contributors. Log-normal distribution, 2019.

[5] Ionela Roxana Danila. On-the-fly modelling and prediction of epidemic phenomena, 2014.

[6] CDC: Centers for Disease Control and Prevention. The flu season, 2018.

[7] CDC: Centers for Disease Control and Prevention. How flu spreads, 2018.

[8] CDC: Centers for Disease Control and Prevention. Key facts about influenza (flu), 2018.

[9] CDC: Centers for Disease Control and Prevention. National, regional, and state level outpatient illness and viral surveillance, 2018.

[10] CDC: Centers for Disease Control and Prevention. Preliminzary estimates of 2017-18 seasonal influenza vaccine effectiveness against laboratory-confirmed influenza from the us flu ve and haiven networks, 2018.

[11] CDC: Centers for Disease Control and Prevention. Prevention strategies for seasonal influenza in healthcare settings, 2018.

[12] Christophe Fraser, Christl A. Donnelly, and Simon Cauchemez. Pandemic potential of a strain of influenza a (h1n1): Early findings. *Science*, 324(5934):1557–1561, 2009.

[13] A.M Fry, D. Goswami, and et al K. Nahar. Efficacy of oseltamivir treatment started within 5 days of symptom onset to reduce influenza illness duration and virus shedding in an urban setting in bangladesh: a randomised placebo-controlled trial. *Lancet Infect Dis*, 14(2):109–118, 2014.

[14] J.M Heffernan, R.J Smith, and L.M Wahl. Perspectives on the basic reproductive ratio. *Journal of the Royal Society*, 2(4):281–93, 2005.

[15] J. Hsu and et al R. Maustafa. Antivirals for treatment of influenza: a systematic review and meta-analysis of observational studies. *Ann Intern Med*, 156(7):512–524, 2012.

[16] T. Jefferson, M.A Jones, and P. Doshi. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. *Cochrane Database Syst Rev*, 10(4), 2014.

[17] Mingbin Liu, Jianming Ou, and et al. Lijie Zhang. Protective effect of hand-washing and good hygienic habits against seasonal influenza - a case-control study? *Medicine (Baltimore)*, 95(11), 2016.

[18] Kenneth J. Rothman, Sander Greenland, and Timothy L.Lash. *Modern Epidemiology*. Lippincott Williams Wilkins, 2008.

[19] Qingxia Zhang and Dingcheng Wang. Antiviral prophylaxis and isolation for the control of pandemic influenza. *Int J Environ Res Public Health*, 11(8):7690–7712, 2014.

[20] Qingxia Zhang and Dingcheng Wang. Assessing the role of voluntary self-isolation in the control of pandemic influenza using a household epidemic model. *Int J Environ Res Public Health*, 12(8):9750–9767, 2015.