371 Final Project

The Dynamics and Public Health Measures of Tuberculosis in South Africa

How has public health measures impacted the spread of Tuberculosis in South Africa?

Lawrence Martinez ID:1571734 (coding), Joshua George ID:1665548 (analysis), Aram Alyousef ID:1553709 (background information)

University of Alberta

April 14, 2022

Abstract

In this paper we explore how public health measures implemented have affected the growth of tuberculosis in South Africa. To this day, tuberculosis is the leading cause of death due to infection, despite a cure in the form of drugs and medications existing. In many third world countries like South Africa, tuberculosis has greatly impacted their way of life. To better understand tuberculosis, we model data presented from the World Health Organization and simulate the number of TB cases in South Africa. Using a modified SIR model that considers the influx and deaths of individuals, we first model what the current trends of tuberculosis are. Then we modify the parameters with respect to public health measures and interpret what these results mean biologically. The intention of this paper is to gain insight on an infectious killer that is not publicized on mainstream media and to educate people about the dangers that still lurk all around the world.

1 Introduction

The quantitative measurements of the life sciences are extremely complicated and yield complex data. To disentangle and understand data, mathematicians have relied on statistics, mathematical methods, and modern technology (Kranz 13). Thus mathematical modelling is crucial to multiple areas of research and our understanding of the world. Perhaps one of the most paramount accomplishments of science is the utilization of mathematics to model deadly diseases. Data modelling has enhanced our understanding of diseases far more than individually analyzing each case. Though personal hygiene gave humans the upper hand on threatening microorganisms, field experiments, growth chamber experiments, and the combination of the aforementioned have gifted us with essential knowledge for combating current diseases and future pandemics. Modelling diseases has become the accepted form of understanding how infections of certain diseases grow and decay. Though perfectly predicting the number of infections that will occur is impossible, mathematicians show that we can model behaviours and trends of infections (Kabunga 1).

One of the most documented diseases throughout history is tuberculosis (TB). TB is an infectious disease, caused by the bacteria *Mycobacterium tuberculosis*, that most prominently affects the lungs. This organism was discovered in the late 19th century, affecting 1 in 7 people (World Health Organization). The disease can easily be spread through the air by methods of coughing, sneezing, or spitting. Doctors have attempted to treat TB for many years, recommending treatment that consisted of only rest, good food, and warmth. However, in 1943 Selman Waksman found a cure for Tuberculosis in the form of antibiotics. This scientific discovery earned Selman a Nobel prize (Selman). The ability to spread through the air makes TB the top infectious killer, spread-

ing through 10 million people every year and killing 1.5 million of them (World Health Organization). Though this deadly disease affects millions every year, it is easily treatable and curable. Approximately 1 in 4 people are estimated to have been infected with Tuberculosis, while 5-15% are estimated to have fallen ill with Tuberculosis (World Health Organization). Furthermore, TB affects those who are combating HIV and diabetes at a higher rate than those who are relatively healthy (CDC).

As knowledge of the human body is expanded, it became easier to develop new drugs to combat such infections. Though first-world countries have combated TB, less developed nations struggle heavily with these infections. South Africa is an example of one such country. TB is a severely dangerous public health issue in South Africa. Approximately half a million people developing the disease every year, over half of which have pre-existing health conditions such as HIV (K. Selain 3). Though effective treatment is available, and the country is making considerable progress in combating the disease, more could be done to control TB in South Africa and other developing nations.

With all this in mind we focus our study on South Africa and research how public health measures, both at the national level and international level, have impacted the spread of TB. As well, we look into goals for the future and how they can be achieved. To aid our research, we first collect data of TB infection in South Africa. We then develop a mathematical model and determine parameters that best fit the data. Developing a model helps answer our biological question mathematically. Specifically, we are interested in the equilibrium points, what the values and when they are approximately reached. We also explore the shape of the solution curves. Afterwards, we manipulate the parameters according to public health policies and interpret the results biologically.

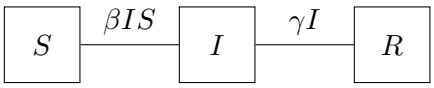
2 Base SIR Model

Many mathematical models of epidemics make the premise that the population may be split into discrete compartments. These compartments are classified based on the current "status" of the disease. The most known compartment model is the susceptible (S), infected (I), and recovered (R), also known as the SIR model. **Susceptible** individuals are defined to have never been infected but are able to catch the disease. **Infected** individuals have the disease and can spread it to susceptible individuals. After the infected period, individuals move on to the **Recovered** compartment.

2.1 Assumptions

The compartments can be modeled according to the following assumptions: the total population is closed and that encounters between infected and susceptible individuals occur at a rate proportional to their respective numbers in the population. This rate of new infections is defined as βIS, where β is a parameter for infectivity. Infected individuals are assumed to all be given the same treatment possible. This treatment is also assumed to be the best available treatment. Individuals are assumed to recover at a constant rate denoted by γ. We also assume that once recovered, they have full immunity and cannot be infected again. Vaccines and vaccinated individuals are assumed to not exist. Other factors such as gender, age, pre-existing health conditions and socioeconomic status are assumed to not be compounding factors.

2.2 Model



The model can be turned into a scheme of Ordinary Differential Equations:

$$\frac{dS}{dt} = -\beta IS$$
$$\frac{dI}{dt} = \beta IS - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

with initial conditions, $S(0) = S_0 > 0$, $I(0) = I_0 > 0$, $R(0) \geq 0$ and parameters defined in Table 1. In this model $S(t), I(t), R(t)$ are the variables that describe the population size with the function letter denoting the compartment. $t = 0$ denotes the beginning of the epidemic. Note that the total population: $N = S(t) + I(t) + R(t)$. As this model is assumed to be closed, we do not consider immigration or emigration. Moreover, we don't consider birth

and natural death, assuming that all infections end with recovery. Therefore the total population size N is constant. In mathematical terms this means $N_0 = S_0 + I_0 + R_0$.

Table 1: Table 1: Description of Parameters used in the SIR model.

Parameter	Description
β	infection rate per year
γ	recovery rate per year

2.3 Mathematical Analysis

We perform a qualitative analysis of the model using the system of differential equations obtained.

Suppose we have a system of differential equations:

$$x_1' = f_1(x_1, x_2)$$
$$x_2' = f_2(x_1, x_2)$$

The x_1-*nullcline*, n_1, is the set of points (x_1, x_2) such that $x_1' = f(x_1, x_2) = 0$ that is, $n_1 := \{(x_1, x_2) \mid f_1(x_1, x_2) = 0\}$. Similarly, the x_2-*nullcline*, $n_2 := \{(x_1, x_2) \mid f_2(x_1, x_2) = 0\}$

Now consider the system of differential equations for our SIR model. Since $R(t)$ has no dependency on $S(t)$ or $I(t)$ we can omit it producing,

$$\begin{pmatrix} S \\ I \end{pmatrix}' = \begin{pmatrix} -\beta IS \\ \beta IS - \gamma I \end{pmatrix}$$

which is the same as,

$$\begin{pmatrix} S \\ I \end{pmatrix}' = \begin{pmatrix} -\beta I & 0 \\ \beta I & -\gamma \end{pmatrix} \begin{pmatrix} S \\ I \end{pmatrix}$$

Evaluating for the nullclines we get, $S' = 0 \implies \beta IS = 0$. This gives us S- nullclines $= \{(S, I) \mid S = 0 \text{ or } I = 0\}$(the S and I axes) and $I' = \beta IS - \gamma I \implies \beta IS = \gamma I \implies I$-nullclines $= \{(S, I) \mid I = 0 \text{ or } S = \frac{\gamma}{\beta}\}$ ($I = 0$ and the vertical line $S = \frac{\gamma}{\beta}$).
The following observations can be made:

- We get a number of steady points on $S^+ := \{(S, 0) \mid S \geq 0\}$.
- From the model we get that $\dfrac{dS}{dt}$ is always decreasing and bounded below by 0, this means $S(t) \leq S_0$
- For $I' = \beta IS - \gamma I$ we get,
  - if $S_0 < \frac{\gamma}{\beta}$ then $I' < 0$, but this means $I(t)$ strictly decreases from $t = 0$. Also since $S(t) \leq S_0 < \frac{\gamma}{\beta}$ we get that no epidemic can occur from this.
  - if $S_0 > \frac{\gamma}{\beta}$ then $I' > 0$, $I(t)$ strictly increases from $t = 0$ to $t = t'$. This also means $S(t) > \frac{\gamma}{\beta}$. From this we get that an epidemic occurs.

To find the stability of each steady point $(S^*, 0)$, by linearization, let $f = (f_1, f_2)$ such that,

$$f_1(S, I) = -\beta IS$$
$$f_2(S, I) = \beta IS - \gamma I$$

Taking the partial derivatives we get the Jacobian matrix J_f,

$$J_f(S, I) = \begin{pmatrix} -\beta I & -\beta S \\ \beta I & \beta S - \gamma \end{pmatrix}$$

with stability points,

$$J_f(S^*, 0) = \begin{pmatrix} 0 & -\beta S^* \\ 0 & \beta S^* - \gamma \end{pmatrix}$$

The eigenvalues of this matrix,

$$\lambda_{1,2} = \lambda^2 - \text{tr}(J_f(S^*, 0)) + \det(J_f(S^*, 0)) = 0$$

which gives us $\lambda_1 = 0$ and $\lambda_2 = \beta S^* - \gamma$. The first eigenvalue $\lambda_1 = 0$ corresponds to the neutrally stable direction along the ray of steady states. The second eigenvalue is positive if $S^* > \frac{\gamma}{\beta}$ and negative otherwise. To construct a phase portrait we write one unknown function I as a function of another S. We can do so by chain rule $I = I(S(t))$.

$$\frac{dI}{dt} = \frac{dI}{dS} . \frac{dS}{dt}$$

This gives us,

$$\frac{dI}{dS} = \frac{\beta IS - \gamma I}{-\beta IS} = -1 + \frac{\gamma}{\beta S}$$

Since we are considering I as a function of S , integrating the above equation from S_0 to S we obtain,

$$I(S) = \frac{\gamma}{\beta} \ln S - S + \text{ constant },$$

where the constant is determined by the initial conditions. It then follows that there is a unique solution curve connecting the equilibrium points in the interval $\gamma/\beta < S < \infty$ to

one in the interval $0 < S < \gamma/\beta$. This justifies the observation as to why if $S_0 > \frac{\gamma}{\beta}$ then $I(t)$ strictly increases in the interval $[0, t')$. This means $I(t)$ increases till a value $S = \theta$ and then it decreases to 0 past this values. In other words, $\lim_{t\to\infty} I(t) = 0$, $\lim_{t\to\infty} S(t) = C_1$ and $\lim_{t\to\infty} R(t) = C_2$ where $C_{1,2}$ are physical constants.

2.4 Description of Data and Model Fitting

With an understanding of the analytics behind the model, we now plot and fit the data to determine the parameters, β, and γ. Data was collected from the World Health Organizations' (WHO) Global Tuberculosis Programme (World Health Organization). The database we accessed presented data starting from 2000. While TB data before 2000 exists, we choose to use the data presented in the main database to understand the modern-day trends. Every year, members of the WHO submit tuberculosis profiles which is then compiled together. These profiles include data on estimated cases, detection rate and mortality rate. As well, many data sets are divided into smaller groups such as children, sex, and other health conditions. Because of the abundance of data, we decided to focus on the general case of new TB incidents and did not look at specific cases. Data for new TB incident data became the value for the infected population with the values 2000 being the initial condition. Although WHO has data on treatment success rate, which could be used to extrapolate the recovered population, we decided against it, focusing on just the infected population. This was to ensure that parameter fitting would fit the infected population. However, we did use the treatment success rate one time to determine the initial condition for recovered. The initial condition for susceptible was then calculated by subtracting the infected and recovered populations from the estimated total population of 2000. This was to keep consistent with the data.

Parameters were then determined using MatLab. Ode45 was used to solve the differential equations and the function fminsearch was used to determine the parameters. The value of the parameters and the initial conditions can be seen in Table 2.

Table 2: Values of Parameters and Initial Conditions.

Parameters	
Parameter/Initial Condition	Value
β	2.47×10^{-8}
γ	0.97
N	44967713
$S(0)$	44408623
$I(0)$	343000
$R(0)$	216090

Biologically, these parameters make sense as β is inversely related to the size of the population. The recovery rate is related to a individuals infectious period. This can last from a least 6 months ($\gamma = 2$) to the worse case scenario of years. A recovery rate of 0.97 suggests that an individual is infected for around 13 months. Applying these values to our ODE system, we plot the solution curves of the system of differential equations and also the phase plane.

Figure 1-2 are consistent with the analysis determine prior. Figure 1 shows the solution curves with the x-axis being the years. The Infected population increases until around x = 11 when it begins to decrease. This continues to decrease until around x = 30 where the population of infected is very small. The susceptible population steadily decreases from x = 0 until x = 30 where it begins to plateau. Conversely, the recovered population increases from x = 0 until x = 30 where it also plateaus. The patterns of these solutions curves suggest that an epidemic occurs and will eventually end at around x = 30. We can then infer that at x = 30, the equilibrium or steady state is reached. The shape of the solution curve is also what we expected with no notable remarks. Figure 2 shows the phase plane with the susceptible population as the x-axis and the infected population as the y-axis. The point is the initial condition and produced a unique curve. Because our initial condition for susceptible was greater than $\frac{\gamma}{\beta}$, an epidemic occurred. This condition also satisfies that $\lambda_2 > 0$, making this region unstable. This instability causes the population to change, reaching $I_{max} = 39271255$ before decreasing to 0 and entering the stable region with $\lambda_1 = 0$ While more can be extrapolated from this model, we do not continue with this model. We do not believe that it is adequate in modelling TB in South Africa and

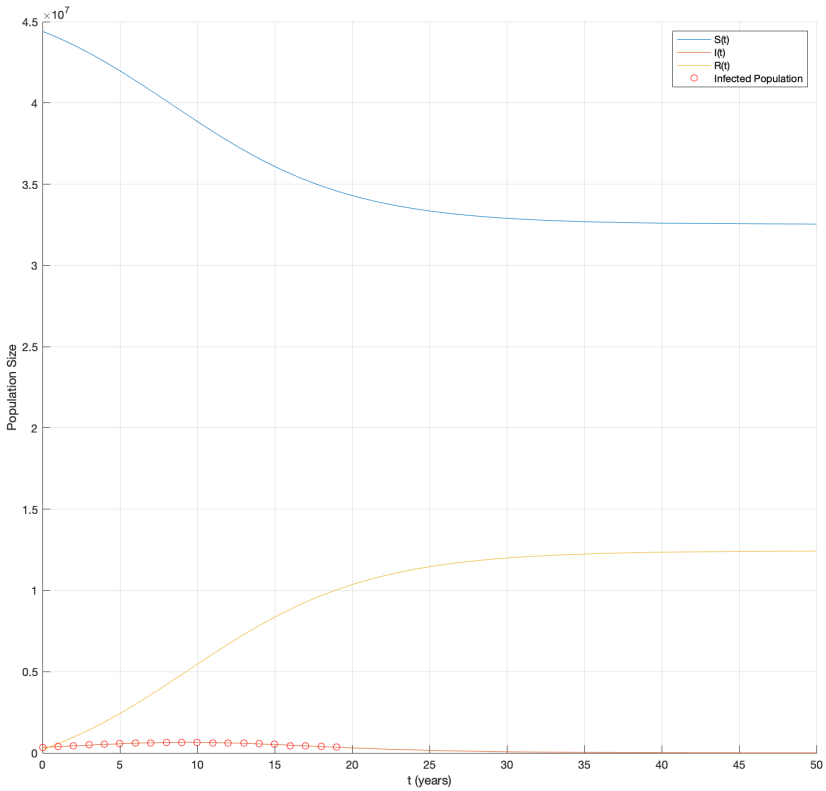


Figure 1: Solution curve of SIR model with $\beta = 2.47\times\ 10^{-8}$ and $\gamma = 0.97$

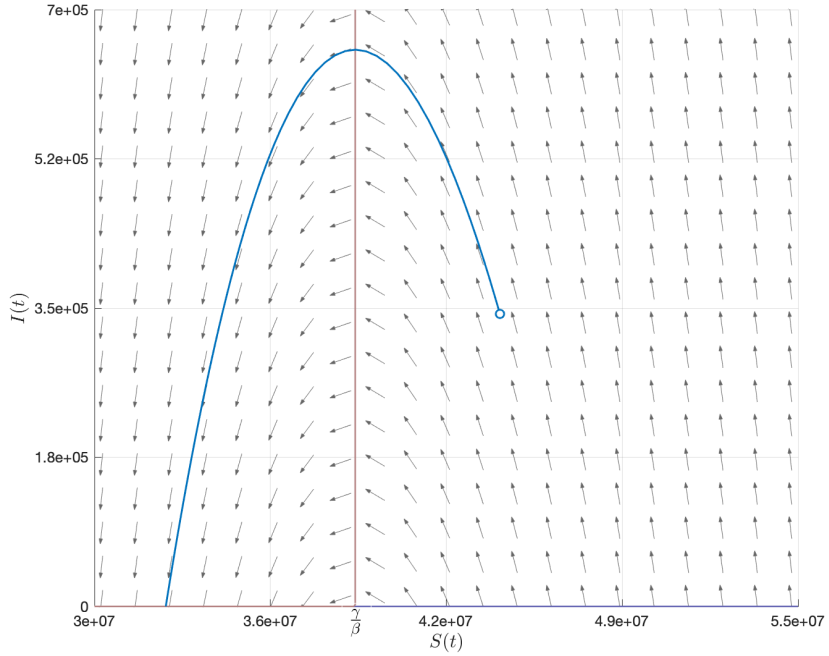


Figure 2: Phase plane with the initial point (44408623,343000). The red vertical is the nullcline $\frac{\gamma}{\beta} = 39271255$.

helps answer our biological question. One of the main issues is that TB is one of the leading causes of death in South Africa. However, a model assumption for SIR is that every infected individual would recover. It also does not consider the natural death rate of the susceptible and recovered populations. Another issue we had was with the closed population assumption. While we already mentioned death, births and the influx of individuals are also not considered. From the data, the total population rises every year which contradicts the closed popula-

tion assumption. We believe that the influx of individuals is a major contributor that has allowed for TB to survive as long as it has. These two main factors are our motivation to produce a different model to answer our biological question.

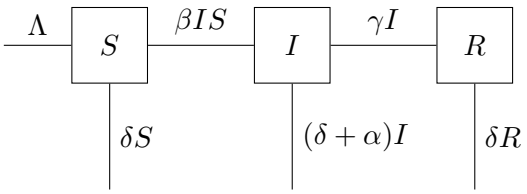
3 Modified SIR Model

We now modify the SIR model to include the influx of individuals and also the deaths due to natural causes and TB. The definition of each compartment, susceptible, infected and recovered remain unchanged, however, additional arrows are added to include the new parameters.

3.1 Assumptions

As we are building on the previous model, many of the same assumptions hold. This includes, encounters between infected and susceptible individuals occur at a rate proportional to their respective numbers, infected individuals are assumed to all be given the same treatment possible and that recovery occurs at a constant rate. Vaccines are assumed to not exist and other factors such as age, gender, pre-existing health conditions and socioeconomic status are not compounding factors. However, because we now consider the influx and deaths of individuals, this population is open. We assume that the population increases by the same constant every year, denoted by Λ. In terms of deaths, we assume that natural death, denoted by δ, to be inversely related to the average life expectancy of an individual. Matching with the previous assumption, this life expectancy is assumed to apply for all individuals regardless of gender, socioeconomic status and pre-existing health conditions. As well, since it is inversely proportional, $0 < \delta < 1$. α is the parameter for death due to TB which we again assume to be exist between (0, 1). α and δ are also assumed to be additive. Finally, we assume the $\Lambda >> \alpha$ or δ to keep consistent with the biological interpretation.

3.2 Model



This modified model be represented as a system of Ordinary Differential Equations with parameters as defined in Table 3:

$$\frac{dS}{dt} = \Lambda - \beta IS - \delta S$$
$$\frac{dI}{dt} = \beta IS - (\delta + \alpha + \gamma)I$$
$$\frac{dR}{dt} = \gamma I - \delta R$$

Table 3: Description of Parameters for the modified SIR model.

Parameters	
Parameters	Description
β	infection rate per year
γ	recovery rate per year
α	TB death rate per year
δ	natural death rate per year
Λ	individual influx

3.3 Mathematical Analysis

Consider the system of ODE, to find the steady points by linearization let $f = (f_1, f_2, f_3)$ where

$$f_1(S, I, R) = \Lambda - \beta IS - \delta S$$
$$f_2(S, I, R) = \beta IS - (\delta + \alpha + \gamma)I$$
$$f_3(S, T, R) = \gamma I - \delta R$$

Adding all three equations together we get, $N' = \Lambda - \delta N - \alpha I$, since in comparison to Λ, the term αI is significantly smaller, we can ignore it and the equation becomes $N' = \Lambda - \delta N$. It is easy to see that the solution to the differential equation is

$$N(t) = N_0 e^{-\delta t} + \frac{\Lambda}{\delta}\left(1 - e^{-\delta t}\right)$$

Therefore as $t \to \infty$, $N \to \frac{\Lambda}{\delta}$ and this is called the *limit population size*.

3.3.1 $\mathcal{R}_0$ and its Derivation

We now introduce the parameter $\mathcal{R}_0$. Also known the basic reproduction number, $\mathcal{R}_0$ measures how transmissible a disease is in a population. The parameter is not a fixed value for all diseases but dependent on population size and model parameters. The larger the value of $\mathcal{R}_0$, the more likely that a disease can spread from host to susceptible individuals. As such, if $\mathcal{R}_0 < 1$, we say that a disease is unable to spread effectively, producing less than 1 secondary case (Milligan et al. 311). Therefore, the population of infected individuals will slowly deplete on its own. Conversely if $\mathcal{R}_0 > 1$, the disease is able to spread effectively, producing more than 1 secondary case, causing an epidemic.

We know that $\mathcal{R}_0$ is a dimensionless quantity and be numerically solved by non-dimensionalizing the system. Consider f_1, f_2. Since δ, α, γ have units (year^{-1} let $\tau = (\alpha+\gamma+\delta)t$, (t time). Then τ is a dimensionless quantity. By this change $I(t) = I(\frac{\tau}{\alpha+\gamma+\delta}) = \hat{I}(\tau)$ and $S(t) = \hat{S}(\tau)$. By the chain rule we get,

$$\frac{d\hat{S}}{d\tau} = \frac{1}{\alpha+\gamma+\delta}\frac{dS}{dt},$$
$$\frac{d\hat{I}}{d\tau} = \frac{1}{\alpha+\gamma+\delta}\frac{dI}{dt}$$

We rescale $\hat{S}, \hat{I}$ variables with limiting population size. Therefore we get $s(t) = \dfrac{\delta\hat{S}}{\Lambda}$ and $i(t) = \dfrac{\delta\hat{S}}{\Lambda}$, and these two are dimensionless quantities. Consider,

$$s(t) = \frac{\delta\hat{S}}{\Lambda} = \frac{\delta(\Lambda - \beta IS - \delta S)}{\Lambda(\alpha+\gamma+\delta)}$$

We get

$$s' = \rho(1-s) - \mathcal{R}_0 si$$
$$i' = (\mathcal{R}_0 s - 1)\, i$$

such that

$$\rho = \frac{\delta}{(\alpha+\gamma+\delta)}, \quad \mathcal{R}_0 = \frac{\Lambda\beta}{\delta(\alpha+\gamma+\delta)}$$

where $\mathcal{R}_0$ is our reproduction number.

3.3.2 Phase-plane analysis

To get the steady points of the system let $(f_1, f_2, f_3) = (0, 0, 0)$. Solving this we get the following steady points,
When $I = 0$ (Disease free equilibrium DFE),

$$(S^*, I^*, R^*) = \left(\frac{\Lambda}{\delta}, 0, 0\right)$$

When $I \neq 0$ (Endemic equilibrium EE),

$$(S^*, I^*, R^*) = \left(\frac{\alpha+\gamma+\delta}{\beta}, \frac{\Lambda}{\alpha+\gamma+\delta} - \frac{\delta}{\beta}, \frac{\gamma}{\delta}\left(\frac{\Lambda}{\alpha+\gamma+\delta} - \frac{\delta}{\beta}\right)\right)$$

To get the eigenvalues, consider the Jacobian, J_f

$$J_f = \begin{pmatrix} -\beta I - \delta & -\beta S & 0 \\ \beta I & \beta S - (\alpha+\delta+\gamma) & 0 \\ 0 & \gamma & -\delta \end{pmatrix}$$

The stability analysis of the DFE is,

$$J_f\left(\frac{\Lambda}{\delta}, 0, 0\right) = \begin{pmatrix} -\delta & -\beta\frac{\Lambda}{\delta} & 0 \\ 0 & \beta\frac{\Lambda}{\delta} - (\alpha+\delta+\gamma) & 0 \\ 0 & \gamma & -\delta \end{pmatrix}$$

with eigenvalues,

$$\lambda_1 = -\delta$$
$$\lambda_2 = \beta\frac{\Lambda}{\delta} - (\alpha+\delta+\gamma)$$
$$\lambda_3 = -\delta$$

We get that $\lambda_{1,3}$ is always negative (as $\delta > 0$) always. Now rearranging λ_2 we get ,

$$\lambda_2 = (\alpha+\delta+\gamma)\left(\frac{\Lambda\beta}{\delta(\alpha+\delta+\gamma)} - 1\right)$$

which is just,

$$\lambda_2 = (\alpha+\delta+\gamma)(\mathcal{R}_0 - 1)$$

When $\mathcal{R}_0 < 1$, all three eigenvalues are negative and the DFE acts as a sink which is stable.

Now consider the EE,

$$J_f\left(\frac{\alpha+\gamma+\delta}{\beta}, \frac{\Lambda}{\alpha+\gamma+\delta} - \frac{\delta}{\beta}, \frac{\gamma}{\delta}\left(\frac{\Lambda}{\alpha+\gamma+\delta} - \frac{\delta}{\beta}\right)\right) =$$

$$\begin{pmatrix} -\beta\left(\frac{\Lambda}{\alpha+\gamma+\delta} - \frac{\delta}{\beta}\right) - \delta & (\alpha+\delta+\gamma) & 0 \\ \beta\left(\frac{\Lambda}{\alpha+\gamma+\delta} - \frac{\delta}{\beta}\right) & 0 & 0 \\ 0 & \gamma & -\delta \end{pmatrix}$$

The eigenvalues are,

$$\lambda_1 = \frac{-\delta\mathcal{R}_0 - \sqrt{(\delta\mathcal{R}_0)^2 - 4\delta(\alpha+\delta+\gamma)(\mathcal{R}_0-1)}}{2}$$

$$\lambda_2 = \frac{-\delta\mathcal{R}_0 + \sqrt{(\delta\mathcal{R}_0)^2 - 4\delta(\alpha+\delta+\gamma)(\mathcal{R}_0-1)}}{2}$$

$$\lambda_3 = -\delta$$

If $(\delta\mathcal{R}_0)^2 < 4\delta(\alpha+\delta+\gamma)(\mathcal{R}_0-1)$ then we obtain complex conjugate eigenvalues. At some value called the bifurcation point we see that the eigenvalues are strictly imaginary. This observation is an indicator for Hopf bifurcation and that. If $\mathcal{R}_0 >$ that the bifurcation value, we see that two branches of solutions exist. Therefore, we should expect that a limits of the solutions approach a stable periodic orbit.

3.4 Description of Data and Model Fitting

As in section 2.4, we now fit the data to our new model to determine the parameters and see how the model aligns with our analysis.

Many of the same controls were kept such as using the data starting from 2000, focusing on general new cases, and not looking at specific demographics, furthermore we kept the same initial conditions. However, we now consider the changing total population size and deaths, both natural and induced by TB. All parameter values and initial conditions can be found on Table 4. Unlike before, not all parameters needed to be fitted, and some were assumed using general statistics.

δ was assumed to be the inverse of South Africa's average life expectancy. From the World Bank, life expectancy varied from 2000 to 2020 (The World Bank). Taking the average of these values we determined that the average life expectancy of South Africa was 58 years old. Therefore $\delta = \frac{1}{58}$.

α was determined by looking at the WHO data. The database has data on estimated mortality due to TB, dividing deaths per year by the number infected of the same year, we get the rate of deaths from the Infected population. However, from our model this was assumed to be $\delta + \alpha$. Therefore, by subtracting δ we acquired α.

Λ was fitted by setting $I(0)$ and $R(0)$ to 0. This meant that $S(t) = N(t)$ and so the population was allowed to grow without disturbance. Using Matlab and the function fminsearch, this growth was determined and equated to Λ.

Finally, β and γ were determined similarly as before. Using the initial conditions to solve the system of ODE's and the Matlab function fminsearch, the infected population was fitted with parameters that minimized the sum of square errors.

Table 4: Values of Parameters and Initial Conditions.

Parameters	
Parameter/Initial Condition	Value
β	4.27×10^{-8}
γ	1.54
Λ	1540918.33
δ	0.017
α	0.24
$S(0)$	44408623
$I(0)$	343000
$R(0)$	216090

Biologically, these values make sense. β is once again inversely related to the total population size. γ suggests that recover takes around 8 months. As well, the influx of individuals is much greater than the proportion of people that die so the total population will increase.

Figure 3-5 are consistent with the mathematical analysis. Figure 3 shows the long term solutions of population compartments. The x-axis is once again time in years. The top line is the total population. From our analysis we found that $\lim_{t\to\infty} N(t) = \frac{\Lambda}{\delta} = 90642254$. While our curve appears to not reach this value, the curve of $N(t)$ considers αI while the mathematical analysis does not. Overall, the difference is minimal and does not impact the quality of the curve and our interpretation.

Answering mathematical questions, we observe the patterns of the solution curves. Unlike before where the solutions followed a more linear pattern before plateauing, we see in this model that the solution oscillate before reaching some limit. A closer inspection of this limit, as seen in Figure 4, shows that it too oscillates. From Figure 3 and Figure 4, we see that the limit for all the solution curves converge to a positive constant which is the EE. This means that our $\mathcal{R}_0 > 1$. The oscillating pattern also matches with the stability and Hopf bifurcation. Since the eigenvalues have negative real parts they will converge to the second

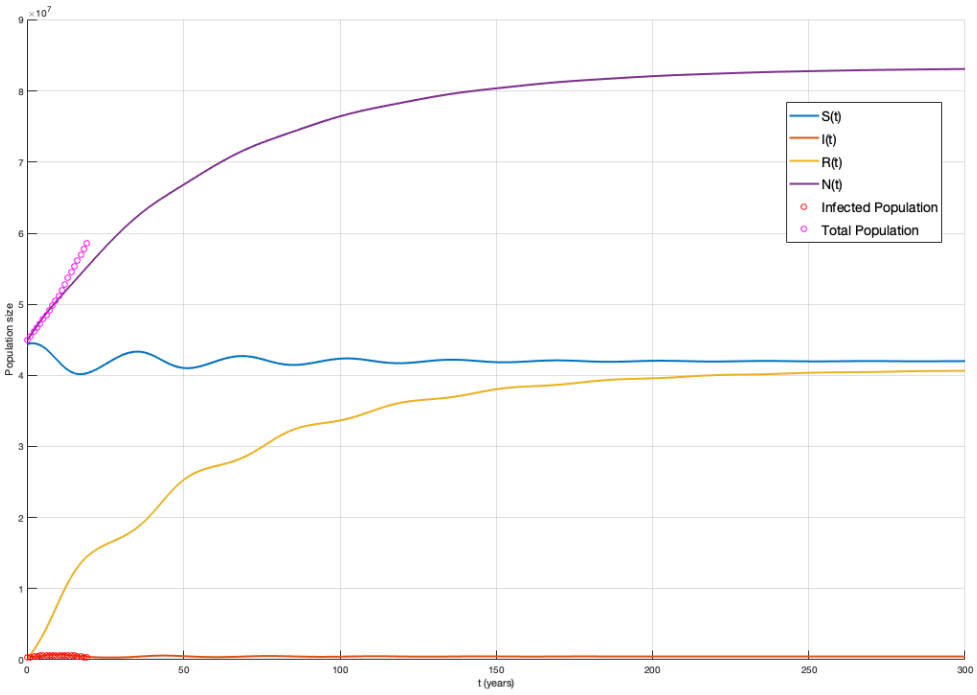


Figure 3: Solution curve of our modified SIR model with $\beta = 4.27\times10^{-8}$, $\gamma = 1.54$, $\Lambda = 1540918$, $\delta = 0.017$ and $\alpha = 0.24$.

steady state. However, from Hopf bifurcation, as $\mathcal{R}_0 > 1 > 0$, this point loses stability and creates two stable periodic orbits. Interestingly, this stability takes a very long to be reached, occurring well beyond our lifetime. This suggests that TB will continue to have an affect on South Africa for a long time.

3.4.1 Modifying Parameters to Answer Biological Question

With a better model that addresses the two main issues we had and an understanding of the dynamics of TB, we explore public health measures in place to see their effect on TB. We also look into future goals and how changing parameters to align with the goals affects the population sizes. South Africa has taken many strides in combating TB. In 2007-2011, South Africa released a National Strategic Plan (NSP) to combat HIV and AIDS (Kapp 1590). While the focus was on HIV and AIDS, many TB patients also have HIV so the measures in place are still relevant. Some of the main goals included increased voluntary testing and counselling to combat stigma and increasing human resources. This is seen in the data as during this time period infected population grew as people got tested. In 2012-2016, following the success of the previous NSP, improvements were made to testing and overall better treatment and medication (Hopkins et al. 1). Better testing allows for early diagnosis and earlier access to medication. Because of this we are interested to see what would have happened if the 2012-2016 measures were not implemented. What would have happened to the cases and limits? Keeping all parameter values the same except for β, we refit using 2000-2010 infected data as this is when infected cases increased.

Figure 5 simulates this scenario. The infected population continues to grow past 2010 when in the original data it begins to decrease. The susceptible and recovered populations have also changed with the susceptible population decreasing and the recovering increasing. Mathematically, this result shows that limits have changed for all compartments. The infected and recovered population limit increases while the susceptible decreases. As well, the curves still approach and oscillate around the equilibrium points as t $\to \infty$. Biologically, this is scenario may be very harmful to South Africa. A large number of infections in a relatively short period of time may put a significant burden on the health care system. More people may not

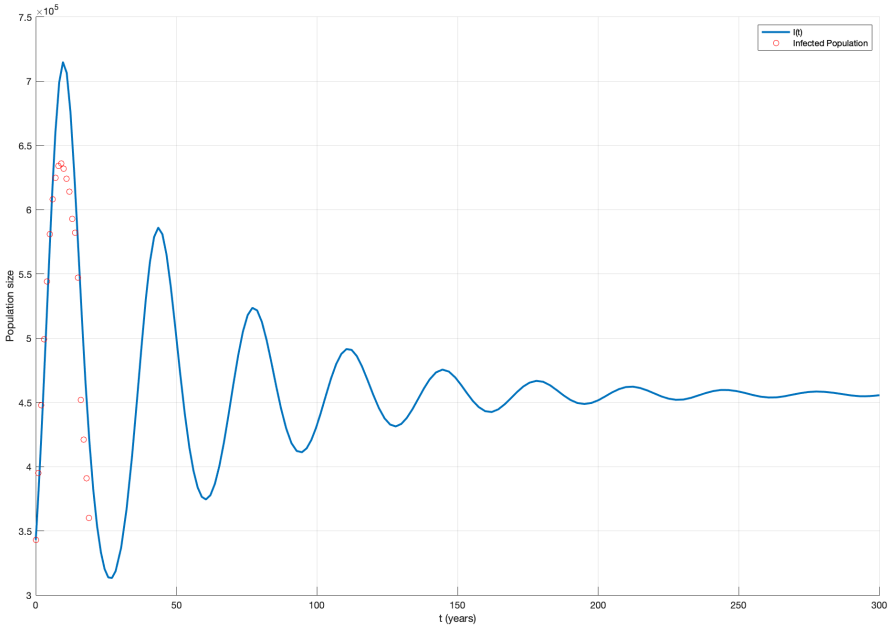


Figure 4: Solution curve of just the infected population with the same parameters as Figure 3.

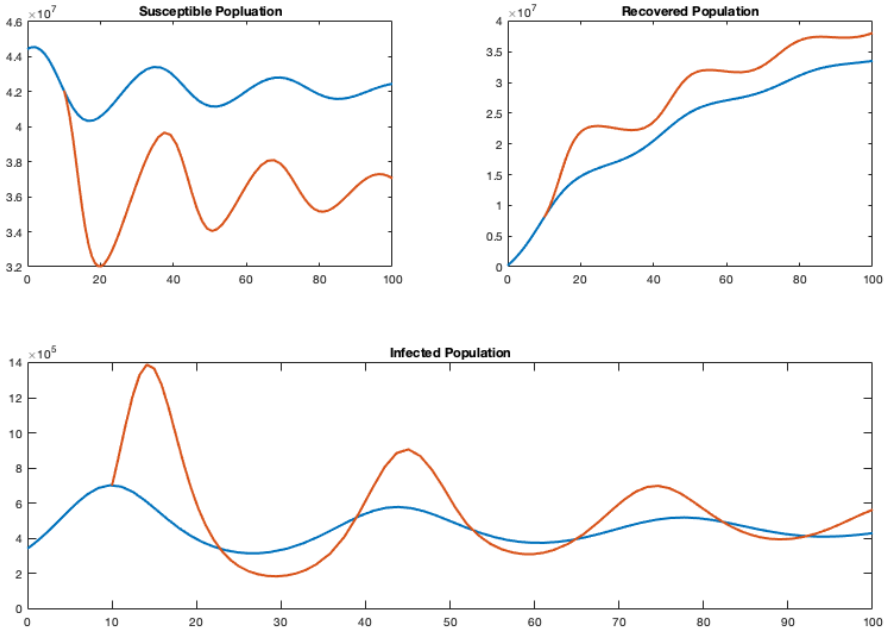


Figure 5: Simulation of TB growth in South Africa before 2012 measures using the modified SIR model. The blue line has no parameter modification. The orange line is when $\beta = 4.944\times\ 10^{-8}$, other parameter values were kept the same. This value was fitted using infected data from 2000 to 2010 in Matlab. he initial condition for the simulation is (S_{10}, I_{10}, R_{10}) = (42043181, 701326, 8047870).

have access to adequate medication or medication may be in short amount. One could also infer that more deaths would have occurred. Therefore, the measures input in 2012-2016 are very beneficial as they helped bring TB infection rates to a manageable level. This also

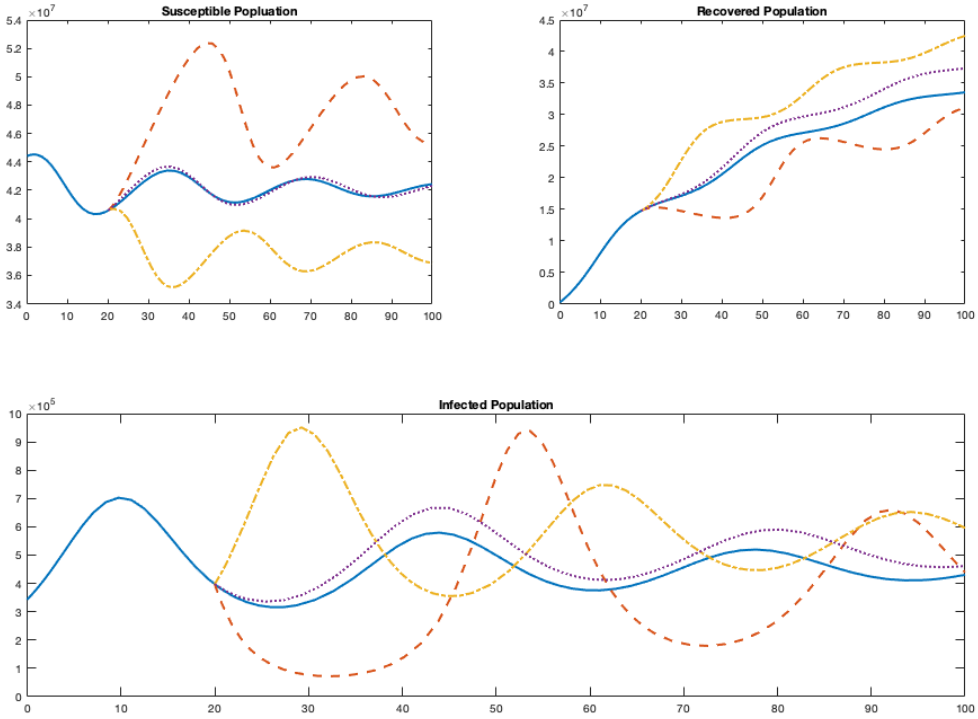


Figure 6: Simulations of the modified SIR model using the goals outlined in the World Health Organization. The solid blue line has no modification to the parameters. The orange dashed line is when β is reduced by 10% ($\beta = 3.8\times\ 10^{-8}$). The yellow dash dotted line is when $\alpha = 0.05$. The purple dotted line is when both parameters are modified. The initial condition for the simulations are $(S_{20}, I_{20}, R_{20}) = (40571098, 397766, 14716652)$.

.

shows the importance of early diagnosis leading to early access to improved treatment.

Internationally, we look at goals WHO has laid out. In their publication, *The End TB Strategy*, they list out targets they want to achieve (World Health Organization, "End TB"). Remarkable targets include decreasing the percentage of people that die from TB to 5% in 2025 from 15% in 2015 and reduce the incidence rate by 10% in 2025 from 2% in 2015. Again, we simulate these scenarios keeping every parameter constant and modifying α and β respectively. A simulation is also done when both parameters are changed. Because we now simulate future events, we change the initial condition to when t = 20.

All scenarios produce differing results. When β is reduced by 10%, we see that the susceptible population remains larger. Since less people overall are being infected, we see that the infected and recovered populations are smaller. This results in the infected and recovered population approaching smaller limits and the susceptible approaching a larger limit compared to the when no parameters were changed. When $\alpha = 0.05$, we see that the susceptible population actually decreases. This is because more people are infected and surviving, passing along the disease to others. However, because less people are dying we see that the recovered population increases. Therefore, the recovered population approaches a larger limit while the susceptible and infected population approach smaller limits. In the case when both modifications are made, we see that solution is very similar to when no modifications to parameters are made. With less people dying there are more people recovering but also more people that can infect others. However, since the β was also reduced the susceptible population does not change as drastically compared to the other two scenarios. What these results mean for WHO is that in order to effectively control TB both measures need to be implemented.

One is not good enough and may actually increase the burden on the countries. Interestingly, while WHO document title suggests that TB can be eradicated, from all the simulations and graphs we see that it not possibly. The constantly changing population size will always provide a host for TB. Therefore, controlling TB to a manageable is a more realistic goal that WHO should focus on.

4 Discussion and Conclusion

While we ultimately found our study to be helpful in modelling TB and answering our biological question, there were some factors we did not consider. These factors are directly related to our model assumptions. As mentioned, vaccines was one measure we did not consider. A vaccinated person can be considered immune to the disease which reduces the susceptible population and decreasing the overall transmission rate. In terms of our model, this would be another arrow from the susceptible population to the recovered population. Another factor we did not consider was preexisting health conditions. Mentioned in South Africa's NSP, HIV and AIDS are very apparent in South Africa. With a weakened immune system, individual treatment may be longer or not as effective. The database we used had data on this but we ultimately choose not to use it. Should we redo the research and look at specific case that include HIV and AIDS, we may see different results. Lastly, the rising number of TB cases that are drug resistant will prove to be a problem. We assume that the treatment given to all individuals will be the same and is of the best quality. However, if TB has mutated to be resistant to these medications, the number of deaths may increase or the period of infection will be longer. This would then require further research into the new variants and the creation of new treatments drugs which may be costly.

Ultimately, we found that mathematical modelling and analysis helped answer our biological question. By first deriving a model that considered the influx and deaths of individuals, we were able to better understand how TB has been able to survive so long in South Africa. The solution curves show that because of an overall increasing population, there are more susceptible people that can be infected. As the infection spreads and increases the infected population, the susceptible population decreases until it reaches a local minimum and then increases again. This produces a wave pattern showing that TB cannot truly be eliminated but rather experiences periods of low and high infections. This understanding was essential before we could modify the parameters and look into how public health measures have impacted TB spread. National measures as seen in South Africa have shown to be very effective. While more testing leads to a larger infected population, it also allows for infected people to begin treatment early. Without improvements to TB treatment, simulations show that more people become infected in a short period of time which can negatively impact the quality of health care. On the international level, we see that both decreasing infection contacts and decreasing TB deaths is essential to keep TB controlled. If only one target is met, the infection cases actually rises which defeats the purpose of the goals. These results further emphasis that TB can never be truly eradicated despite the WHO document title. Overall, while the model is not perfect, it has provided great insight into the dynamics of TB and helps explain how public health measure have helped and will continue to help in controlling TB.

5 Appendix

Code for our the modified SIR model

```
%Creating the modified SIR model
function f =newMSIRmodel(t,y,par)
% Label the parameters and variables
    beta = par(1);
    gamma = par(2);
    S=y(1);
    I=y(2);
    R=y(3);
    N=S+I+R;
% Input the differential equations
    Sdot=1540918.33-beta*I*S-0.0172*S;
    Idot=beta*I*S-gamma*I-0.257*I;
    Rdot=gamma*I-0.0172*R;
    f =[Sdot Idot Rdot]';
end

%Solving the modified SIR model
function sol=newMSIRSol(par,IC,t)
%disp(num2str(par))
    DeHandle=@(t,y) newMSIRmodel(t,y,par);
    [~, Y]=ode45(DeHandle,t,IC);
    sol=Y';
end

%Determining the parameters for the modified SIR model

InfP = [343000  395000  448000  499000  544000  581000  608000  625000  634000
    636000  632000  624000  614000  593000  582000  547000  452000  421000
    391000  360000];
Total = [44967713 45571272  46150913  46719203  47291610  47880595  48489464
    49119766  49779472  50477013  51216967  52003759  52832659  53687125
    54544184  55386369  56207649  57009751  57792520  58558267];

MSIRData = [InfP]

t = [0:9];

IC = [44624713, 343000, 216090]

newMSIRparSol = @(par,t) [0 1 0]*newMSIRSol([par(1) par(2)], IC, t);

SumSquareSIR = @(par) sum(sum((newMSIRparSol(par,t)-MSIRData).^2));

[SIRtheta, fval, exitflag] = fminsearch(SumSquareSIR, [1e-7; 1.67]);
MSIRsol=newMSIRparSol(SIRtheta, t)

figure ;
scatter(t,MSIRData,'.' );
hold on;
plot(t,MSIRsol,'--');
hold off

%Plotting the modified SIR model
function Msir (beta, gamma) %Msir(4.27e-08, 1.54)


format longG
%close all

% x(1) = S, x(2) = I, x(3) = R

```

```
odesys = @(t,x) [1540918.33-beta*x(2)*x(1)-0.0172*x(1); (beta*x(2)*x(1)-gamma*x(2)-0.257*x(2)); gamma*x(2)-0.0172*x(3)];
x0   = [42043181 701326 8047870];          % Initial Conditions
tspan = [10 100];                          % Time Of Integration
[t, x] = ode45(odesys, tspan, x0);       % Integrate


%SusP = [44408623 44719232  44941233  45124193  45298000  45449095  45605364  45774916  45962802  46225863  46634857  46949169  47315289  47728215  48142314  48576429  49122069  49631001  50166160  50678507];
InfP = [343000  395000  448000  499000  544000  581000  608000  625000  634000  636000  632000  624000  614000  593000  582000  547000  452000  421000  391000  360000];
%RecP = [216090 457040  761680  1096010 1449610 1850500 2276100 2719850 3182670 3615150 3950110 4430590 4903370 5365910 5819870 6262940 6633580 6957750 7235360 7519760];
Total = [44967713 45571272  46150913  46719203  47291610  47880595  48489464  49119766  49779472  50477013  51216967  52003759  52832659  53687125  54544184  55386369  56207649  57009751  57792520  58558267];

hold on
subplot(2,2,1)
plot(t, x(:,1), 'LineWidth',2, 'LineStyle','-')
title('Susceptible Popluation')
hold off

hold on
subplot(2,2,[3 4])
plot(t, x(:,2), 'LineWidth',2, 'LineStyle','-')
title('Infected Population')
hold off

hold on
subplot(2,2,2)
plot(t, x(:,3), 'LineWidth',2, 'LineStyle','-')
title('Recovered Population')
hold off

plot(t, x)
plot3(x(:,1),x(:,2),x(:,3))


%tP = [0:19]
%scatter(tP, SusP,  'blue')
%scatter(tP, InfP, 'or')
%scatter(tP, Total, "magenta")
%legend('I(t)',  'Infected Population', 'Location', 'NE')

end
```

Code for the base SIR model

```
%Setting up the system
function f =newSIRmodel(t,y,par)
% Label the parameters and variables
    lambda = par(1);
    gamma = par(2);
    S=y(1);
    I=y(2);
    R=y(3);
    N=S+I+R;
% Input the differential equations
    Sdot=-lambda*I*S;
    Idot=lambda*I*S-gamma*I;
    Rdot=gamma*I;
    f =[Sdot Idot Rdot]';
end
```

```
16
17 %Solving the system
18 function sol=newSIRSol(par,IC,t)
19 %disp(num2str(par))
20     DeHandle=@(t,y) newSIRmodel(t,y,par);
21     [~, Y]=ode45(DeHandle,t,IC);
22     sol=Y';
23 end
24
25 %Fitting for parameters
26
27 InfP = [343000  395000  448000  499000  544000  581000  608000  625000  634000
       636000  632000  624000  614000  593000  582000  547000  452000  421000
       391000  360000];
28
29
30 SIRData = [InfP]
31
32 t = [0:9];
33
34 IC = [44624713, 343000, 216090]
35
36 newSIRparSol = @(par,t) [0 1 0]*newSIRSol([par(1) par(2)], IC, t);
37
38 SumSquareSIR = @(par) sum(sum((newSIRparSol(par,t)-SIRData).^2));
39
40 [SIRtheta, fval, exitflag] = fminsearch(SumSquareSIR, [1e-7; 1.67]);
41 SIRsol=newSIRparSol(SIRtheta, t)
42
43 figure ;
44 scatter(t,SIRData,'.' );
45 hold on;
46 plot(t,SIRsol,'--');
47 hold off
48
49 %Plotting
50 function sir (beta, gamma) %sir(2.46961008565369e-08,0.969706230691677)
51
52 close all
53
54 odesys = @(t,x) [(-beta.*x(1).*x(2)); (beta.*x(1).*x(2)-gamma.*x(2)); gamma.*x
       (2)];
55 x0   = [44408623; 343000; 216090];                          % Initial Conditions
56 tspan = [0 50];                          % Time Of Integration
57 [t, x] = ode45(odesys, tspan, x0);       % Integrate
58
59 tP = [0: 19]
60 %SusP = [44408623 44115673  43758033  43372703  42974103  42536213  42083613
       41622863  41151043  40716563  40385603  39913123  39450343  39008803
       38565843  38157773  37882133  37588963  37341353  37087953];
61 InfP = [343000  395000  448000  499000  544000  581000  608000  625000  634000
       636000  632000  624000  614000  593000  582000  547000  452000  421000
       391000  360000];
62 %RecP = [216090 457040  761680  1096010 1449610 1850500 2276100 2719850 3182670
       3615150 3950110 4430590 4903370 5365910 5819870 6262940 6633580 6957750
       7235360 7519760];
63 hold on
64
65
66 plot(t, x)
67 grid
68
69 %scatter(tP, SusP,  'blue')
70 scatter(tP, InfP, 'or')
71 %scatter(tP, RecP, ['black'])
72
```

```
73 legend('S(t)', 'I(t)', 'R(t)', 'Infected Population', 'Location', 'NE')
74 xlabel('t (years)')
75 ylabel('Population Size')
76 hold off
77
78 end
```

References

[1] Arcari, Christine M. "Understanding and Measuring the Dynamics of Infectious Disease Transmission." Vaccinology: An Essential Guide, by Gregg N. Milligan and Barrett A D T., Wiley Blackwell, 2015, pp. 304–318.

[2] CDC. "Fact Sheets." Centers for Disease Control and Prevention, 28 Oct. 2011, www.cdc.gov/tb/publications/factsheets/general/tb.htm#:%7E:text=Tuberculosis%20(TB)%20is%20a%20disease,they%20do%20not%20get%20treatment.

[3] CDC. "The Facts, Stats, and Impacts of Diabetes." Centers for Disease Control and Prevention, 24 Jan. 2022, www.cdc.gov/diabetes/library/spotlights/diabetes-facts-stats.html#:%7E:text=In%202019%2C%20about%201.4%20million,minority%20groups%2C%20especially%20Black%20teens.

[4] CDC. "TB and Diabetes." Centers for Disease Control and Prevention, 25 Oct. 2021, www.cdc.gov/tb/topic/basics/tb-and-diabetes.html#:%7E:text=People%20living%20with%20diabetes%20who,and%20become%20sick%20with%20TB.&text=Someone%20with%20untreated%20latent%20TB,disease%20than%20someone%20without%20diabetes

[5] CDC. "World TB Day History." Centers for Disease Control and Prevention, 14 Jan. 2022, www.cdc.gov/tb/worldtbday/history.html:%7E:text=TB%20in%20humans%20can%20be,China%20(2%2C300%20years%20ago)

[6] Hau, B., and J. Kranz. "Mathematics and Statistics for Analyses in Epidemiology." Springer.Com, springer, 4 May 1996, link.springer.com/content/pdf/10.1007/978-3-642-75398-5_2.pdf

[7] Hopkins, Kathryn L., et al. "Will the Current National Strategic Plan Enable South Africa to End AIDS, Tuberculosis and Sexually Transmitted Infections by 2022?" Southern African Journal of HIV Medicine, vol. 19, no. 1, 2018, pp. 1–6., doi:10.4102/sajhivmed.v19i1.796.

[8] Kabunga, Kasereka Selain. "Analysis and Simulation of a Mathematical Model of Tuberculosis Transmission in Democratic Republic of the Congo - Advances in Continuous and Discrete Models." SpringerOpen, 16 Nov. 2020,advancesindifferenceequations.springeropen.com/articles/10.1186/s13662-020-03091-0.

[9] Kapp, Clare. "South Africa Unveils New 5-Year HIV/AIDS Plan." The Lancet, vol. 369, no. 9573, 2007, pp. 1589–1590., doi:10.1016/s0140-6736(07)60727-2.

[10] Selman A. Waksman – Facts. NobelPrize.org. Nobel Prize Outreach AB 2022. Sat. 9 Apr 2022. https://www.nobelprize.org/prizes/medicine/1952/waksman/facts/

[11] The World Bank. Life Expectancy at Birth, Total (Years) - South Africa, 2020, data.worldbank.org/indicator/SP.DYN.LE00.IN?end=2020&locations=ZA&name_desc=true&start=2000

[12] World Health Organization. "WHO TB Burden Estimates." Global Tuberculosis Programme, 2021, https://extranet.who.int/tme/generateCSV.asp?ds=estimates.

[13] World Health Organization. "The End TB Strategy." World Health Organization, World Health Organization, 1 Jan. 1970, apps.who.int/iris/handle/10665/331326?locale-attribute=en