Hindawi
Journal of Mathematics
Volume 2023, Article ID 6156749, 12 pages
https://doi.org/10.1155/2023/6156749

Research Article

Modeling the Spread of COVID-19 Using Nonautonomous Dynamical System with Simplex Algorithm-Based Optimization for Time-Varying Parameters

Kevin Yotongyos and Somchai Sriyab

Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand

Correspondence should be addressed to Somchai Sriyab; somchai.sriyab@gmail.com

Received 5 October 2022; Revised 4 July 2023; Accepted 20 July 2023; Published 9 August 2023

Academic Editor: Francisco J. Garcia Pacheco

The SIRDV (Susceptible, Infected, Recovered, Death, Vaccinated) compartmental model along with time-varying parameters is used to model the spread of COVID-19 in the United States. Time-varying parameters account for changes in transmission rates, people's behaviors, safety precautions, government regulations, the rate of vaccinations, and also the probabilities of recovery and death. By using a parameter estimation based on the simplex algorithm, the system of differential equations is able to match real COVID-19 data for infections, deaths, and vaccinations in the United States of America with relatively high precision. Autoregression is used to forecast parameters in order to forecast solutions. Van den Driessche's next-generation approach for basic reproduction number agrees well across the entire time period. Analyses on sensitivity and elasticity are performed on the reproduction number with respect to transmission, exit, and natural death rates in order to observe the changes from a small change in parameter values. Model validation through the Akaike Information Criterion ensures that the model is suitable and optimal for modeling the spread of COVID-19.

1. Introduction

The Kermack–McKendrick SIR-type model and its variants have become one of the most useful tools for modeling the spread of infectious diseases. Much research has been done on exploring the capabilities and extensions of SIR-type models through applications in many infectious diseases [1]. The SIR and SEIR model variants are commonly used to analyze the spread of infectious diseases in population. The SIR model divides the population into three compartments: Susceptible (S), iInfected (I), and Recovered (R). It assumes that individuals can be categorized into one of these compartments, with recovered individuals gaining lifelong immunity. The model tracks the transitions between compartments, such as susceptible individuals becoming infected and infected individuals recovering. The SEIR model expands on this by adding an Exposed (E) compartment to represent individuals in the incubation period. It accounts for the fact that individuals can transmit the infection before showing symptoms [2]. Both models help researchers understand disease dynamics, evaluate interventions, and make predictions. While these models simplify real-world complexities, they provide valuable insights into how diseases spread and can inform public health strategies.

Due to COVID-19, mathematical epidemiology, particularly the use of the SIR model, has gained popularity and is used by many researchers throughout the world who are interested in studying the topic [3–13]. Since COVID-19 is a disease with a latent period, some include the exposed compartment in the SIR model [2]. SEIR models are also suitable and popular amongst epidemiologists to study the outbreak of COVID-19 in different countries and regions [8]. On December 31, 2020, WHO issues its first emergency

use validation for a COVID-19 vaccine [14]. The vaccinated compartment, along with other compartments such as quarantined and death, is added to the SEIR or SIR model [15]. However, while the models previously mentioned are able to simulate and match real COVID-19 data for a relatively short span of time of around a few weeks to months, they fail to compare to the data over a long period of time due to many factors changing and affecting the disease. These models often assume constant, time-independent parameters for transmission, death, and recovery rates, which would be sufficient for a short time span; however, it is not realistic due to the changes in safety regulations set by the government, the behavior of people, the ability to cure COVID-19 patients, and the rate at which people are getting vaccinated.

There are several publications with similar interests in modeling the spread of COVID-19 over a long period of time using varying parameters. An example of a model using time-varying parameters for a longer time span include Nastasi's use of nine different time-dependent transmission frequency parameter values throughout the span of 400 days in order to account for different policy measures in Great Britain and Israel [16]. Similarly, Girardi and Gaetan apply the concept of time-varying parameters to the transmission rate, which is able to qualitatively describe the number of cases over the span of several months in three different countries [17]. Another research explores the same notion of time-varying parameters in the sense of parameter estimation through deep learning on the transmission rate and deceased rate [18]. Furthermore, a study involving the estimation of the time-varying reproduction number has been done based on various cities in China [19]. From these works, it can be seen that the time-varying parameters are what make SIR-type models capable of modeling COVID-19 over a longer period of time, accounting for changes that affect the spread of the disease.

The objective of this research is to use an SIRDV model to match the United State's real COVID-19 data with high precision. The compartments compared with real data are infected (I), death (D), and vaccinated (V). In order to achieve high precision, this model varies the transmission rate β, death probability p, and vaccination rates σ on a daily basis. The method used for parameter estimation is developed based on MATLAB's *fminsearch* function and the simplex algorithm. Specifically, the *fminsearch* function in MATLAB utilizes the Nelder–Mead simplex algorithm, a derivative-free optimization method, to find the minimum value of a given function. It constructs a simplex in the parameter space, iteratively updating its vertices based on function evaluations, aiming to converge towards the minimum. *fminsearch* is particularly useful for optimizing functions without readily available derivative information or in cases where computing derivatives is challenging. Since the simplex algorithm is lightweight and fast, all three parameters with over 800 final values each after 10 attempts of optimization are able to be obtained in under 6 minutes. The solutions are plotted against real active cases data from Worldometer [20], deaths, and fully vaccinated people on Our World in Data [21].

The parameters obtained by the method are used to compute time-varying basic reproduction numbers and are compared to reproduction rates presented by Our World in Data [21]. Sensitivity analysis is performed by day since different parameter values yield different values for sensitivity. Elasticity values are calculated in order to observe the basic reproduction numbers' response to each parameter. Error analysis is performed, and Akaike Information Criterion (AIC) values are also computed on a daily basis in order to show that the model is indeed valid when compared to similar models [22].

2. Model

The constructed SIRDV model is similar to a typical SIR-type model, with the same compartments being susceptible (S), infected (I), and recovered (R); however, with the addition of deaths (D) due to COVID-19 and vaccinated (V), the infected class has the same exit rate for recovery and death, with the difference being the death probability p. It is allowed for recovered people to become susceptible again at the rate of α and for fully vaccinated people to become infected, but with a probability factor of r. The diagram and the nonlinear differential equation for the SIRDV model are represented in Figure 1 and equations (1)–(5), respectively.

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \mu N - \frac{\beta \mathrm{SI}}{N} - (\sigma + \mu)S + \alpha R, \tag{1}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\beta \mathrm{SI}}{N} + r\frac{\beta \mathrm{VI}}{N} - (\gamma + \mu)I, \tag{2}$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = (1 - p)\gamma I - (\alpha + \mu)R, \tag{3}$$

$$\frac{\mathrm{d}D}{\mathrm{d}t} = p\gamma I. \tag{4}$$

When the vaccinated compartment is considered, there are two differential equations used for vaccinations, (5) and (6). When optimizing for compartments other than V, equation (5) is used. Since the vaccination data are only a count of who has received the vaccine (not accounting for people who are infected or died after the vaccination), equation (6) must be used for optimizing the vaccination compartment. Equation (6) is a modification of equation (5), where all negative terms denoting the population exiting the vaccination compartment are omitted.

$$\frac{\mathrm{d}V}{\mathrm{d}t} = -r\frac{\beta \mathrm{VI}}{N} + \sigma S - \mu V, \tag{5}$$

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \sigma S. \tag{6}$$

The term N denotes the current population, in which only the alive population takes into account, thus excluding people in compartment D.

$$N = S + I + R + V. \tag{7}$$

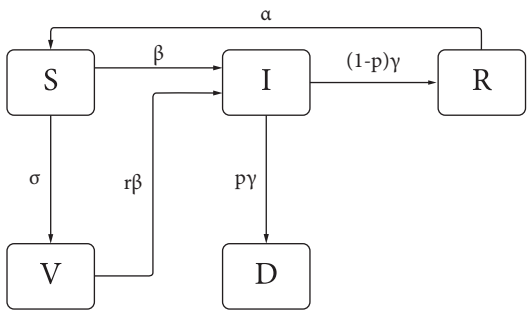

Figure 1: Diagram of the SIRDV model.

3. Parameter Estimation

3.1. Parameters. Constant parameters are kept as a single value throughout the simulation under the assumption that there are no significant changes in the values over the entire time span. It is assumed that the average exit rate from infectious, recovery to susceptible, natural death rate, and probability of infection from being vaccinated do not change over the entire time span. Constant parameter values are listed in Table 1; meanwhile, varying parameters will attain many different values upon optimization, their values will not be presented within Table 1.

3.2. Time-Varying Parameter Estimation. Constant parameter values are not realistic due to the changes in safety regulations set by the government, behavior of people, ability to cure COVID-19 patients, and vaccination rate; hence, there are many previous studies implementing the time-varying parameters in the models [16–19]. Therefore, the time-varying parameter estimation is an essential technique to optimize the parameter in the SIRDV model. Parameters γ, α, μ, and r are kept constant throughout all time steps, as there are no reported significant changes in the rate at which people recover, become susceptible again, die of other causes, or change in the probability of being infected after being fully vaccinated. The transmission rate, probability of death, and vaccination rate are expected to change throughout the course of the pandemic, and will also have the greatest impact when attempting to match the solutions of the model to the real data. Therefore, only parameters β, p, and σ vary with time and undergo the parameter estimation method.

MATLAB's built-in*fminsearch* function which is based on the simplex algorithm is used to minimize the sum of squared errors (SSE) of the solution computed by MATLAB's *ode45* based on the Runge–Kutta numerical scheme and the real data in time steps of a single day by one parameter at a time. In MATLAB, *ode45* is a numerical integration algorithm used to solve ODEs. It employs a variable step-size, Runge–Kutta method of order 4(5), meaning it combines the fourth and fifth-order Runge–Kutta methods to achieve higher accuracy. *ode45* subdivides the integration interval into smaller intervals and approximates the solution by iteratively computing intermediate values using the Runge–Kutta formulae. It dynamically adjusts the step-size based on the estimated error, aiming to maintain accuracy while efficiently integrating the ODE. *ode45* is a popular choice for solving a wide range of ODE problems due to its balance between accuracy and computational efficiency.

The *fminsearch* function in MATLAB is a numerical optimization algorithm that aims to find the minimum value of a given function without requiring derivative information. It utilizes the Nelder–Mead simplex algorithm, which constructs a geometric figure called a simplex in the parameter space. The algorithm iteratively adjusts the simplex's vertices based on function evaluations, such as reflecting, expanding, contracting, or shrinking, in order to converge towards the minimum. By iteratively updating the simplex, *fminsearch* searches for the optimal solution within the parameter space, making it useful for optimizing functions where derivatives are not readily available or difficult to compute.

In the parameter estimation method in Figure 2, the initial parameter value guesses are improved upon via the simplex algorithm, leading to the optimal parameter values by minimizing the error between the real data and the numerical solution. The parameters can be estimated in any order and will yield the same results; however, choosing the order from left to right in the model results in less time and optimization attempts are needed. After performing the optimization algorithm for one time step, the optimized parameter value and its corresponding solution value are then stored in the parameters and solution arrays, respectively. The process is then repeated by taking the endpoint of the solution array as the initial condition for when the dynamical system is to be solved and optimized again, and the next initial guess for the parameter is reset to 0.

The *fminsearch* function is highly efficient, however unconstrained; therefore, it may return values that are invalid for being less than 0 or greater than 1 for certain parameters. In order to avoid negative parameters, a lower bound condition is placed such that the dynamical system is solved with the assumption that the parameter is 0 if the optimal parameter returned by the algorithm is negative. An upper bound condition is also in place for p and σ, since those parameters must be strictly less than 1. Since *fminsearch* operates by searching for a local minimum in the neighborhood of the initial guess, the initial guess for every parameter is 0. The optimization procedure is applied multiple times to yield the best results for all classes.

3.3. Results of Estimated Parameters. The values of β, p, and σ obtained from the parameter estimation technique are plotted starting from January 1st, 2020, in Figure 3, in order to demonstrate the change in people's behavior over the time span, as well as to observe any trends that may be present.

The estimated parameter values appear to generally follow a nonlinear trend with very few outliers, which lead to the ability to analyze people's behavior over the course of the pandemic. The transmission rate β appears to take a drop significantly during the months of COVID-19 lockdown,

Table 1: Parameter interpretation and values.

Parameters	Description	Values (da y^{-1})	Types
β	Transmission rate	—	Varying
γ	Infectious exit rate	1/14	Constant
α	Recovery to susceptible rate	1/30	Constant
p	Probability of death	—	Varying
σ	Vaccination rate	—	Varying
μ	Natural death rate	0.0000248	Constant
r	Probability of infection after vaccinated	0.01	Constant

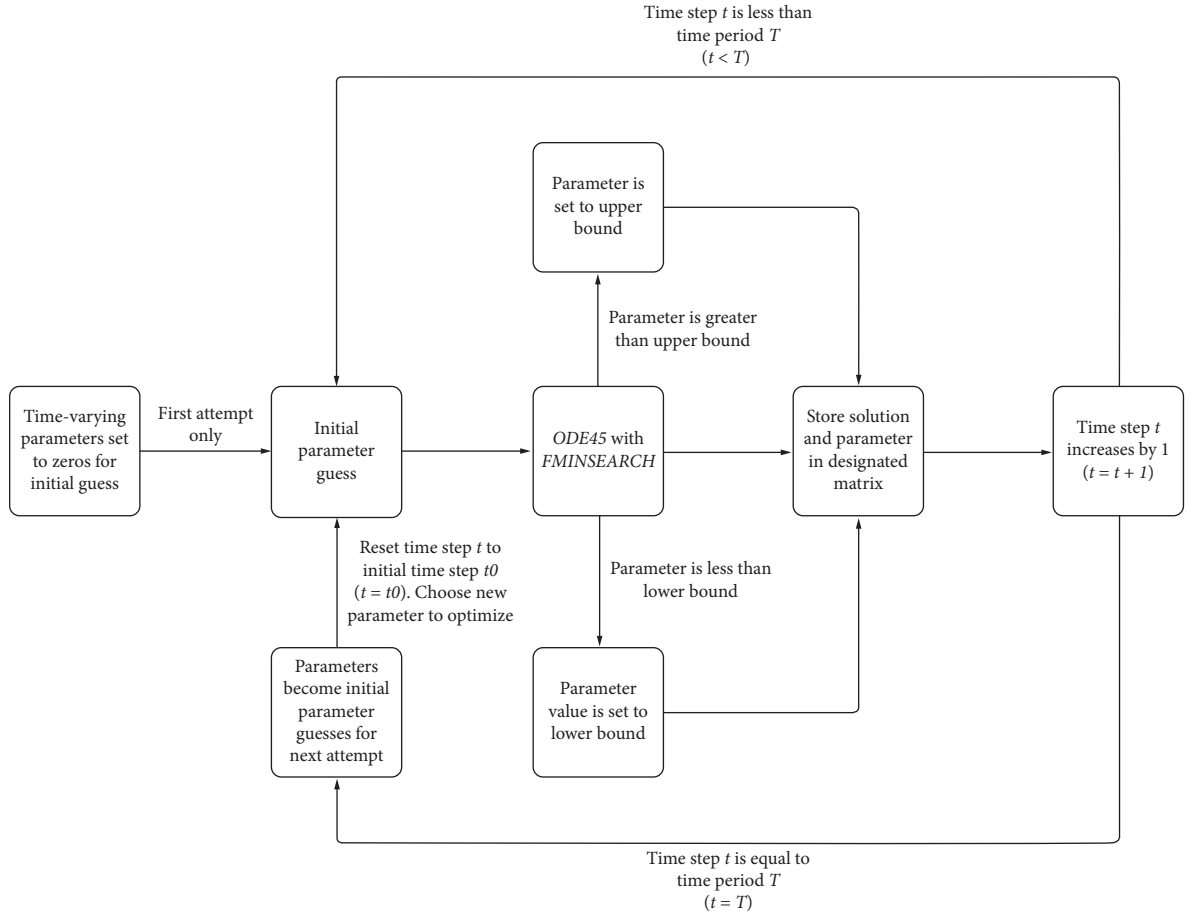

Figure 2: Parameter estimation flowchart.

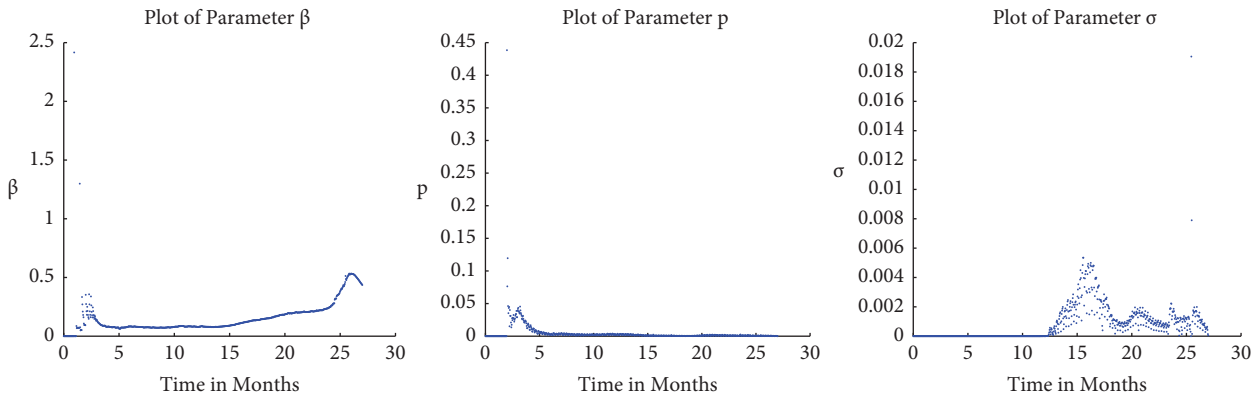

Figure 3: Plots of parameters used in order to attain optimized solutions.

which is around late March to late May of 2020 (dates of lockdown vary by state). The parameter β steadily increases between March 2021 and February 2022, then rapidly increases after, which corresponds well to the real behavior of people practicing strict safety precautions during lockdown, then gradually increasing transmission, and currently having little to no safety precautions after month 25. The death probability p was initially high in the first few days of COVID-19, then quickly dropped to less than 0.05 and eventually converged to 0 around May 2020. The vaccination rate increases rapidly from around January until May 2021, which was when vaccines were first issued to the public in the United States.

The first peak in the vaccination rate likely corresponds to when the COVID-19 vaccines were first released to the public, where large numbers of people were receiving the vaccines at around the same time. Following that first peak, there are smaller peaks that follow, which suggests that there are people who became vaccinated in the months following the period in which the vaccines were first rolled out. The vaccination rate in later months was significantly lower than the vaccines that were first available, suggesting that everyone who wanted the vaccine had received it. The parameter values yielded from the parameter estimation method correlate well with people's behavior over the course of the pandemic.

4. Basic Reproduction Number

The basic reproduction number denoted by $\mathscr{R}_0$ is the expected number of cases from one infected case. Since the parameters are functions of time, the basic reproduction number is also computed as a function of time. Here, $\mathscr{R}_0$ is computed using two different methods, one based on the transmission rate and mean infectious time, while the other uses Van den Driessche's next-generation approach [23]. It is able to be observed that the next-generation approach yields better results when compared to the reproduction rate data [21].

4.1. Computing Basic Reproduction Number with Transmission Rate and Infectious Period. The basic reproduction number $\mathscr{R}_0$ is often approximated using the simple definition of the transmission rate multiplied by the exit rate from the infectious compartment.

$$\mathscr{R}_0 = \beta\tau, \tag{8}$$

where

$$\tau = (\gamma + \mu)^{-1}. \tag{9}$$

The $\mathscr{R}_0$ values obtained from the transmission rate and mean infectious time agree with the reproduction rate data up until around month 15 in Figure 4. Its inability to match the trend of the data throughout the entire time span suggests another method may be able to yield results closer to the data.

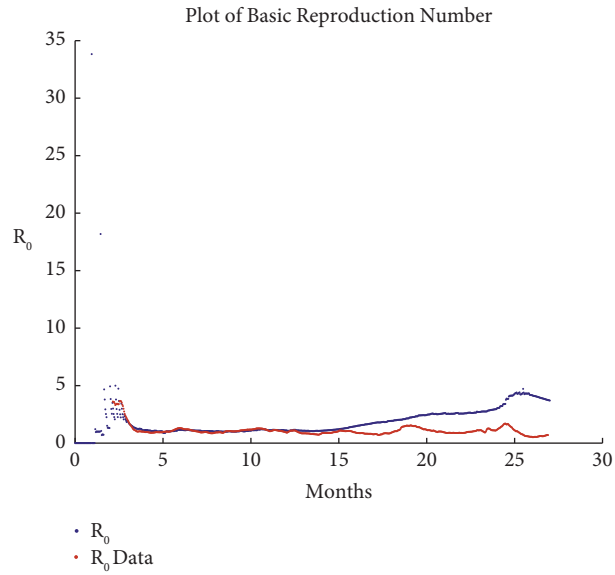

Figure 4: Plot of basic reproduction number obtained by the transmission rate and the mean infectious time.

4.2. Basic Reproduction Number (Van den Driessche and Watmough Next Generation). In the attempt to obtain better results for $\mathscr{R}_0$, Van den Driessche's next-generation approach is used [23]. Let $\mathscr{F}$ denote the terms entering the infected compartment and V be the exiting terms from the infected compartment.

$$\begin{aligned} \mathscr{F} &= \frac{\beta \mathrm{SI}}{N} + r\frac{\beta \mathrm{VI}}{N}, \\ \mathscr{V} &= (\gamma + \mu)I. \end{aligned} \tag{10}$$

Let F and V be partial derivatives with respect to the active infected cases I at initial value x_0.

$$\begin{aligned} F &= \frac{\partial \mathscr{F}(x_0)}{\partial I} = \frac{\beta S_0}{N} + r\frac{\beta V_0}{N}, \\ V &= \frac{\partial \mathscr{V}(x_0)}{\partial I} = (\gamma + \mu), \end{aligned} \tag{11}$$

where $\mathscr{R}_0$ is defined as $\mathrm{FV} - 1$. Therefore, substituting F and V yields

$$\mathscr{R}_0 = \mathrm{FV}^{-1} = \frac{\beta(S_0 + rV_0)}{N(\gamma + \mu)}. \tag{12}$$

The $\mathscr{R}_0$ values obtained from the transmission rate and mean infectious time in Figure 5 are able to agree with the reproduction rate data for only around the first year. The next-generation approach, however, agrees well with the reproduction rate data throughout the entire two years. The dynamical system is solved repeatedly with different parameters at each time step; the initial values S_0 and V_0 are considered to be the endpoints of the previous solution values. Since the $\mathscr{R}_0$ values from the next-generation

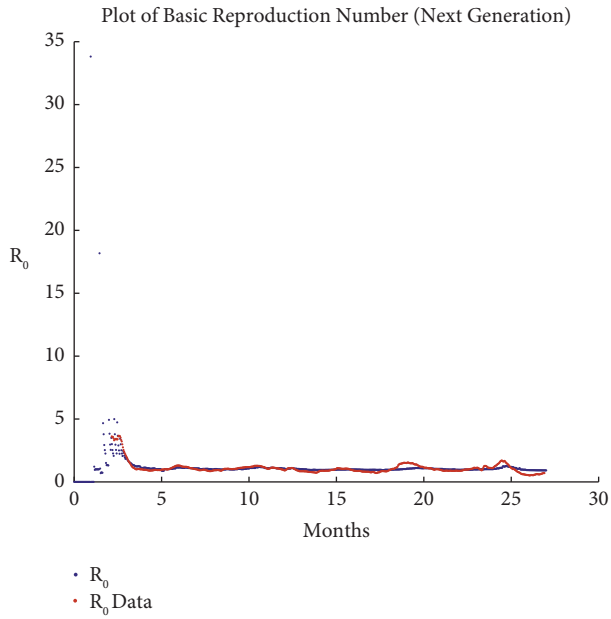

Figure 5: Plot of the basic reproduction number obtained by Van den Driessche's next-generation approach.

approach agree well with the data, it can be stated that the estimated values of the time-varying parameter β and the chosen constant parameters r, γ, and μ are suitable.

5. Results and Discussion

5.1. Numerical Solutions and Data Comparison. The data for active infected cases, deaths from COVID-19, and vaccinations are available by day; therefore, numerical results of compartments I, D, and V are able to be compared to real data.

After obtaining the solutions for active infected cases, deaths, and vaccinations, each class is plotted against their respective real data for visual comparison (Figure 6). The solutions for all three compartments seem to agree well with the real data across the entire time span. The major advantage of varying parameter values by day is the ability to capture every detail such as small and large peaks in active infected cases and to account for sudden increases or decreases. It can also be seen that the plots for vaccinations and deaths are able to capture changes in rates despite being monotonically increasing solutions. The plot of vaccinations here is the count for the number of vaccinations, which neglects the population that has exited the compartment and has become infected or died (from reasons other than COVID-19) after vaccinations.

5.2. Numerical Solutions for All Compartments. The solutions for all compartments are plotted in the same figure in order to compare the populations in each compartment and the magnitude of each trajectory relative to one another.

The numerical solutions shown in the plot of all solutions (Figure 7) are representative of the population currently in each class. It should be noted that the solution curve for the vaccinated class V here represents the number of vaccinated people who have not become infected after vaccination, not to be confused with the vaccinated plot previously shown (Figure 6), which is the count of the number of people who are fully vaccinated. Similarly, the solution curve for the recovered class R represents the number of people who have recently recovered and are not yet susceptible to reinfection.

The rapid decline of the solutions to the susceptible class S suggests that the number of susceptible people will soon approach 0, which suggests the number of COVID-19 infections would likely also reach 0 at around the same time. Although the vaccinated class V is similar to the susceptible class S in the way that people in both compartments are able to become infected, the plots show that significantly more people are protected by the vaccine, since at the moment the solution for V is significantly higher than the solution for S.

Despite the high number of deaths, the magnitude of the solution curve for deaths is relatively low throughout the entire time span when compared to solutions of the other compartments. The curve of recovered class R follows the same trends as the curve of infected class I, confirming the fact that most people who are infected with COVID-19 do indeed recover, keeping in mind that R is continuously losing its population to S at the rate of α.

6. Forecasting Parameters and Solutions

6.1. Forecasting Parameters. An autoregressive (AR) model based on MATLAB's *ar* function is used on the parameters β, p, and σ in order to predict the future parameter values for the next 365 days based on all past respective parameter values. Previous research utilizing the AR model used the method directly on the solutions of the dynamical system [24]. Meanwhile, in this research, it is applied to the parameters, which are then used when solving the dynamical system. White noise is then added to the forecasted parameters in order to simulate the noise present in the parameter values obtained through the parameter estimation method.

6.2. Forecasting Solutions. Past parameter values and forecasted parameter values are shown together on the same plots in Figure 8. The forecasted parameter values are then used in the model in order to forecast the solutions for I, D, and V. The prediction shows that all parameter values will eventually converge rapidly to 0.

Additional data corresponding to these compartments are added in order to evaluate the performance of the forecasting method in Figure 9. From the plot of active infected cases, it is able to be observed that the solution using the forecasted parameters follows the same trend as the real

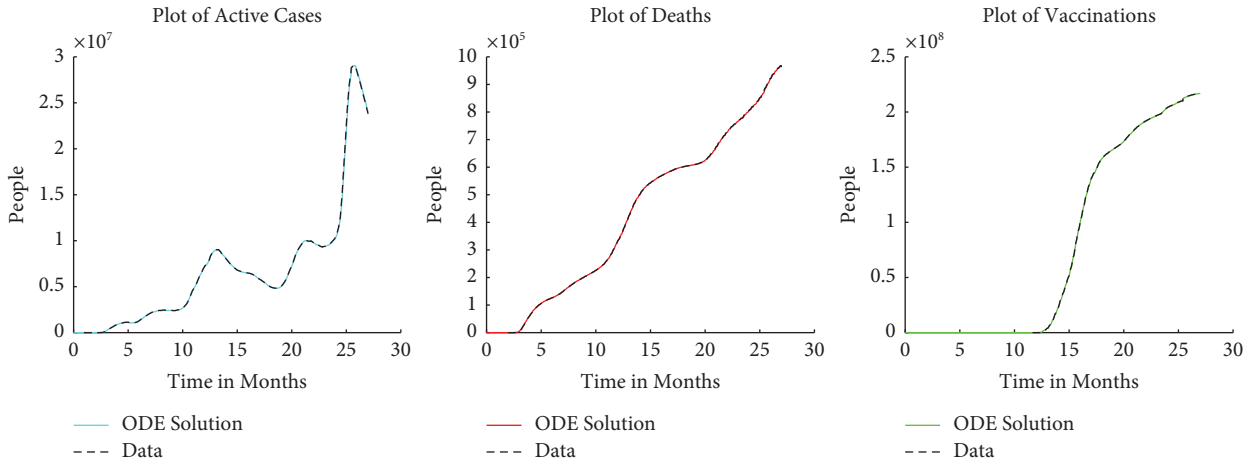

Figure 6: Plots of solutions compared to real data.

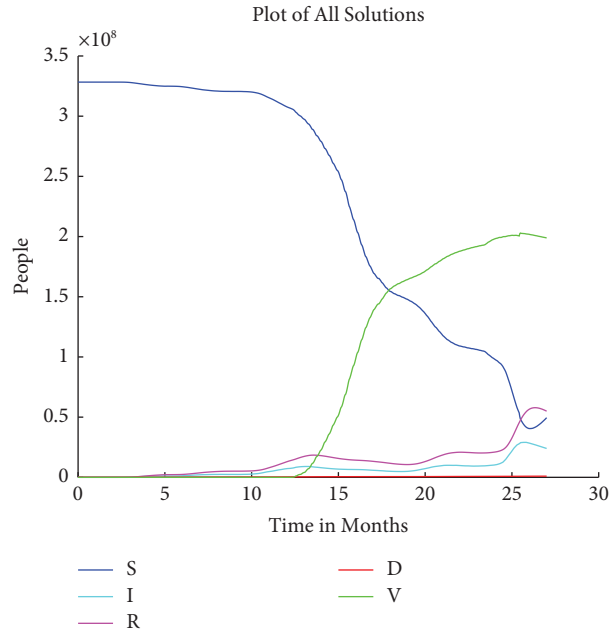

Figure 7: Plot of numerical solutions for all compartments.

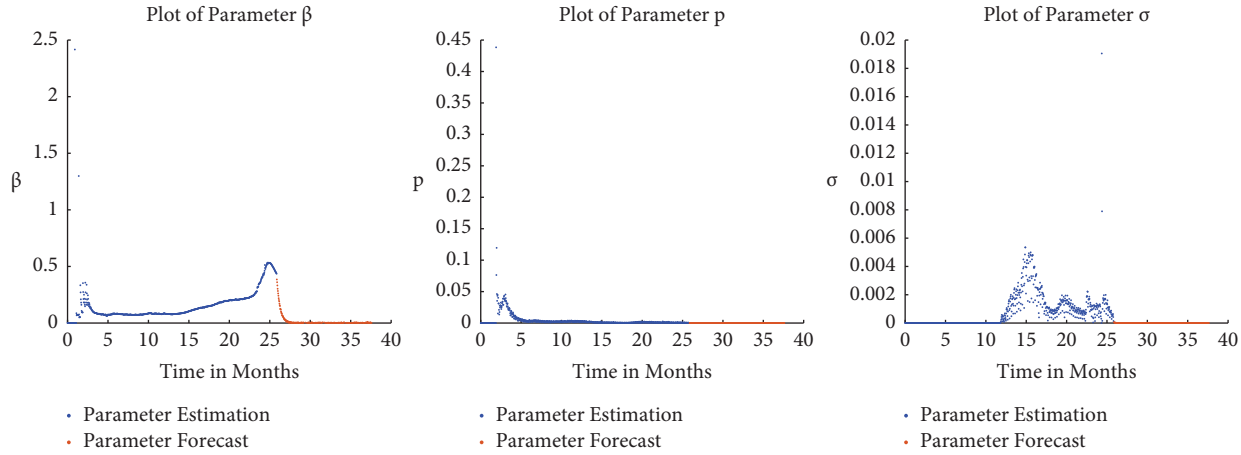

Figure 8: Forecast of parameter values.

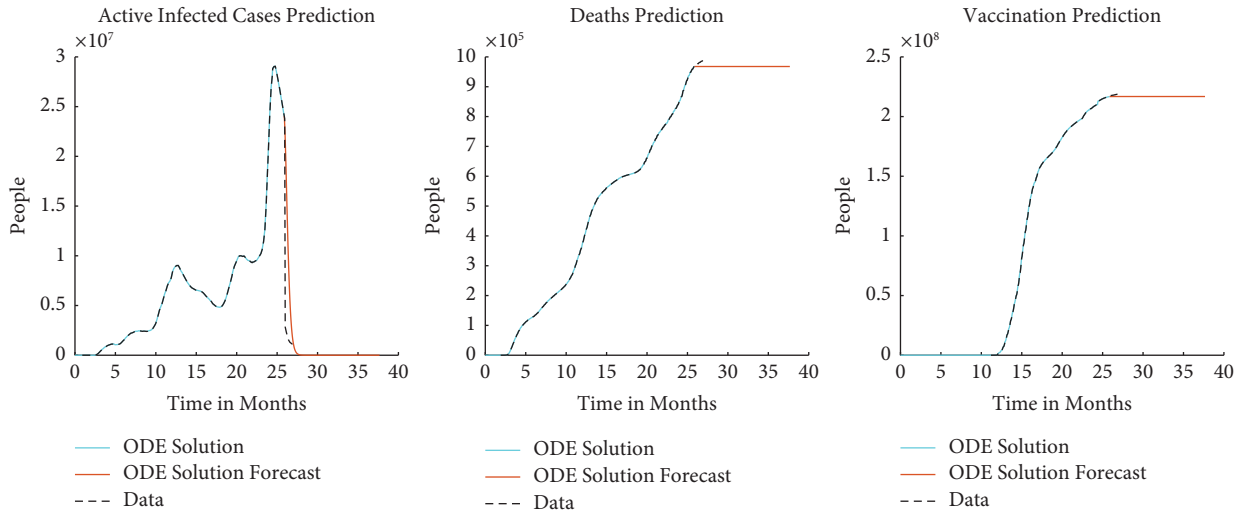

Figure 9: Forecast of solutions.

data, except the real data decrease at a slightly faster rate. The predictions for deaths and vaccinations suggest that both classes will no longer increase due to the parameter values for p and σ being 0 in the forecast.

7. Sensitivity and Elasticity Analysis

7.1. Sensitivity. The objective of the sensitivity analysis is to determine which parameters are most influential on the solution of the model. Sensitivity is defined as follows:

$$\mathcal{S} = \frac{\partial Q}{\partial P}. \tag{13}$$

In which the sensitivity of quantity Q is being analyzed with respect to the parameter P. The reproduction number $\mathcal{R}_0$ obtained by the next-generation method is the quantity whose sensitivity is chosen to be analyzed. Let the disease-free equilibrium be $\mathcal{E} = (S^*, I^*, R^*, D^*, V^*) = (N, 0, 0, 0, 0)$. The reproduction number at disease-free equilibrium is

$$\mathcal{R}_0 = \frac{\beta}{\gamma + \mu}. \tag{14}$$

Taking the partial derivative of $\mathcal{R}_0$ with respect to each parameter yields

$$\mathcal{S}^{\gamma}_{\mathcal{R}_0} = \frac{\partial \mathcal{R}_0}{\partial \gamma} = -\frac{\beta}{(\gamma + \mu)^2}, \tag{15}$$

$$\mathcal{S}^{\mu}_{\mathcal{R}_0} = \frac{\partial \mathcal{R}_0}{\partial \mu} = -\frac{\beta}{(\gamma + \mu)^2}, \tag{16}$$

$$\mathcal{S}^{\beta}_{\mathcal{R}_0} = \frac{\partial \mathcal{R}_0}{\partial \beta} = \frac{1}{\gamma + \mu}. \tag{17}$$

Since $\mathcal{S}_{\mathcal{R}_0}\beta$ only depends on constant parameters, it may be easily computed as a constant scalar value.

$$\mathcal{S}^{\beta}_{\mathcal{R}_0} = 0.9998. \tag{18}$$

However, $\mathcal{S}_{\mathcal{R}_0}\gamma$ and $\mathcal{S}_{\mathcal{R}_0}\mu$ depend on β, which is a variable parameter which can be a wide range of values. Many different values of sensitivity will arise from this simple formula. Therefore, a range of sensitivity values may be observed.

$$\begin{aligned} \max_{\beta} \mathcal{S}^{\gamma}_{\mathcal{R}_0} \le \mathcal{S}^{\gamma}_{\mathcal{R}_0} \le \min_{\beta} \mathcal{S}^{\gamma}_{\mathcal{R}_0}, \\ \max_{\beta} \mathcal{S}^{\mu}_{\mathcal{R}_0} \le \mathcal{S}^{\mu}_{\mathcal{R}_0} \le \min_{\beta} \mathcal{S}^{\mu}_{\mathcal{R}_0}. \end{aligned} \tag{19}$$

Since $\mathcal{S}_{\mathcal{R}_0}\gamma = \mathcal{S}_{\mathcal{R}_0}\mu$, upon substituting with appropriate values for sensitivity formulae (15) or (16), yields the sensitivity range

$$-473.2930 \le \mathcal{S}^{\gamma}_{\mathcal{R}_0}, \mathcal{S}^{\mu}_{\mathcal{R}_0} \le 0. \tag{20}$$

The average values of $\mathcal{S}_{\mathcal{R}_0}\gamma$ and $\mathcal{S}_{\mathcal{R}_0}\mu$ are obtained by letting β be the mean value of all the β values.

$$\overline{\mathcal{S}^{\gamma}_{\mathcal{R}_0}}, \overline{\mathcal{S}^{\mu}_{\mathcal{R}_0}} = -18.6214. \tag{21}$$

In order for the model to match the data towards the beginning of the pandemic, the transmission rate β had to be significantly higher due to the small initial infected population I_0. Therefore, when computing $\mathcal{S}_{\mathcal{R}_0}\gamma$ and $\mathcal{S}_{\mathcal{R}_0}\mu$ which depend on β, the sensitivity spikes all the way to -473.2930. To demonstrate the sensitivity of each parameter, a perturbation of $+0.001$ is applied to one parameter (β, γ, or μ), while the other parameters, both variables and constants, are kept at their original values after optimization. The solutions using the perturbed parameter values are then plotted in Figure 10 against the real data for all three classes in order to observe the effects of the perturbation.

The active cases and deaths are greatly affected by the small changes in parameter values. Vaccinations, however, experience slight changes throughout perturbations of all three parameters, although they are still visually observable. As expected, the solutions exhibit the same characteristics as the original parameter values, although they increase and decrease at different rates.

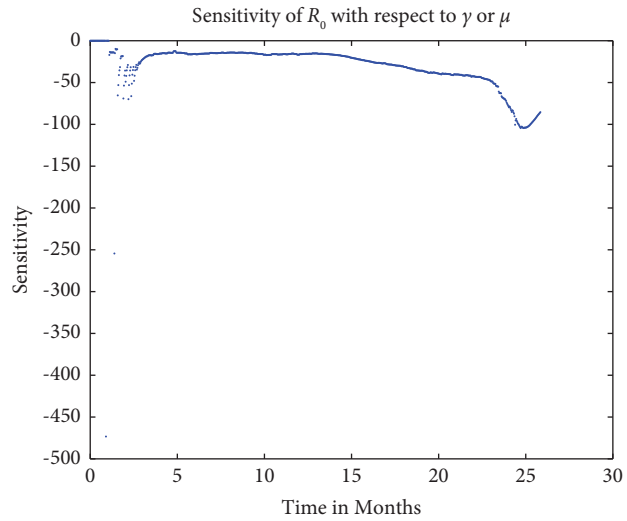

FIGURE 10: Plot of the sensitivity of $\mathscr{R}_0$ with respect to γ or μ.

7.2. Elasticity. Since the sensitivity analysis is local, it does not take into account the range of values that can be input as parameters [2]. Therefore, the more useful concept elasticity is computed, where elasticity is defined as follows:

$$E_Q^P = \frac{\partial Q}{\partial P}\frac{P}{Q}. \tag{22}$$

Using the results previously obtained from the section on sensitivity, elasticity may be computed. Since formulae obtained for elasticity are dependent only on constant parameters, singular scalar values for elasticities may be obtained.

$$\begin{aligned} E_{\mathscr{R}_0}^{\gamma} &= \frac{\partial \mathscr{R}_0}{\partial \gamma}\frac{\gamma}{\mathscr{R}_0} = -\frac{\beta}{(\gamma+\mu)^2}\frac{\gamma}{\mathscr{R}_0} = -\frac{\gamma}{\gamma+\mu} = -0.9998, \\ E_{\mathscr{R}_0}^{\mu} &= \frac{\partial \mathscr{R}_0}{\partial \mu}\frac{\mu}{\mathscr{R}_0} = -\frac{\beta}{(\gamma+\mu)^2}\frac{\mu}{\mathscr{R}_0} = -\frac{\mu}{\gamma+\mu} = -2.4794\times 10^{-4}, \\ E_{\mathscr{R}_0}^{\beta} &= \frac{\partial \mathscr{R}_0}{\partial \beta}\frac{\beta}{\mathscr{R}_0} = \frac{1}{(\gamma+\mu)}\frac{\beta}{\mathscr{R}_0} = 1. \end{aligned} \tag{23}$$

The value obtained for $E_{\mathscr{R}_0}\beta$ is 1, which is expected. The $\mathscr{R}_0$ observes a 0.9998% decrease after a 1% increase in γ. Similarly, $\mathscr{R}_0$ decreases by 2.4794×10^{-4}% after a 1% increase in μ. The impacts of β and γ are very similar, meanwhile μ has a significantly less significance (Figure 11).

8. Model Validation

Model validation is used in order to check accuracy and performance of a certain model. Since the nature of this research is able to be interpreted as a model solved separately over small time steps or as one large model covering a long period of time, the validation process is performed using both interpretations.

8.1. Error Analysis. In order to measure the error between the solution and the data, some basic regression analysis techniques are used [25]. The sum of squared errors (SSE) and the sum of squared total (SST) are computed. Points where there are no data present are excluded from the computation of SSE. Let x_i denote the value of the class at time step i and $\hat{x}_i$ denote corresponding data at the same time step. Furthermore, let $\overline{x}$ be the mean of the corresponding data. The time step i is a single day.

$$\begin{aligned} \text{SSE} &= \sum_{i=1}^{N}(x_i-\hat{x}_i)^2, \\ \text{SST} &= \sum_{i=1}^{N}(x_i-\overline{x})^2, \\ \text{FVU} &= \frac{\text{SSE}}{\text{SST}}, \\ R^2 &= 1-\text{FVU}. \end{aligned} \tag{24}$$

Table 2 shows the SSE, SST, R^2, and FVU of infected, deaths, and vaccinated compartments. The SSE values computed for each class are notably high for infected and vaccinated classes, which is contributed by several different factors. The model handles very large populations; therefore, relatively small errors between the solutions and the real data are in reality quite large when compared to models that work with smaller populations. The model covers a long time span and therefore contains more data, which are prone to having higher SSE values. The parameter estimation also has limitations, as the optimized solution is the best possible solution obtained by *fminsearch*, which is inferior to more sophisticated optimization techniques in machine learning. The R^2 values, however, are approximately 1 for all classes, with the fraction of variance unexplained (FVU) being insignificantly small.

8.2. Single Time Step AIC. The Akaike Information Criterion (AIC) is applied to the model and data in order to determine the validity of the model. Since the model is solved multiple times in small time steps, the SSE and AIC values are calculated for every time step j. Let m denote the number of data points and k denote the number of parameters. For a single time step, $m = 2$ and $k = 7$.

$$\text{AIC}_j = m\ \log\left(\frac{\text{SSE}_j}{m}\right) + 2k. \tag{25}$$

From the plot of AIC values (Figure 12), the range of AIC values from where the spread of COVID-19 started in the United States appears to be in the typical range of AIC values for SIR-type models [2].

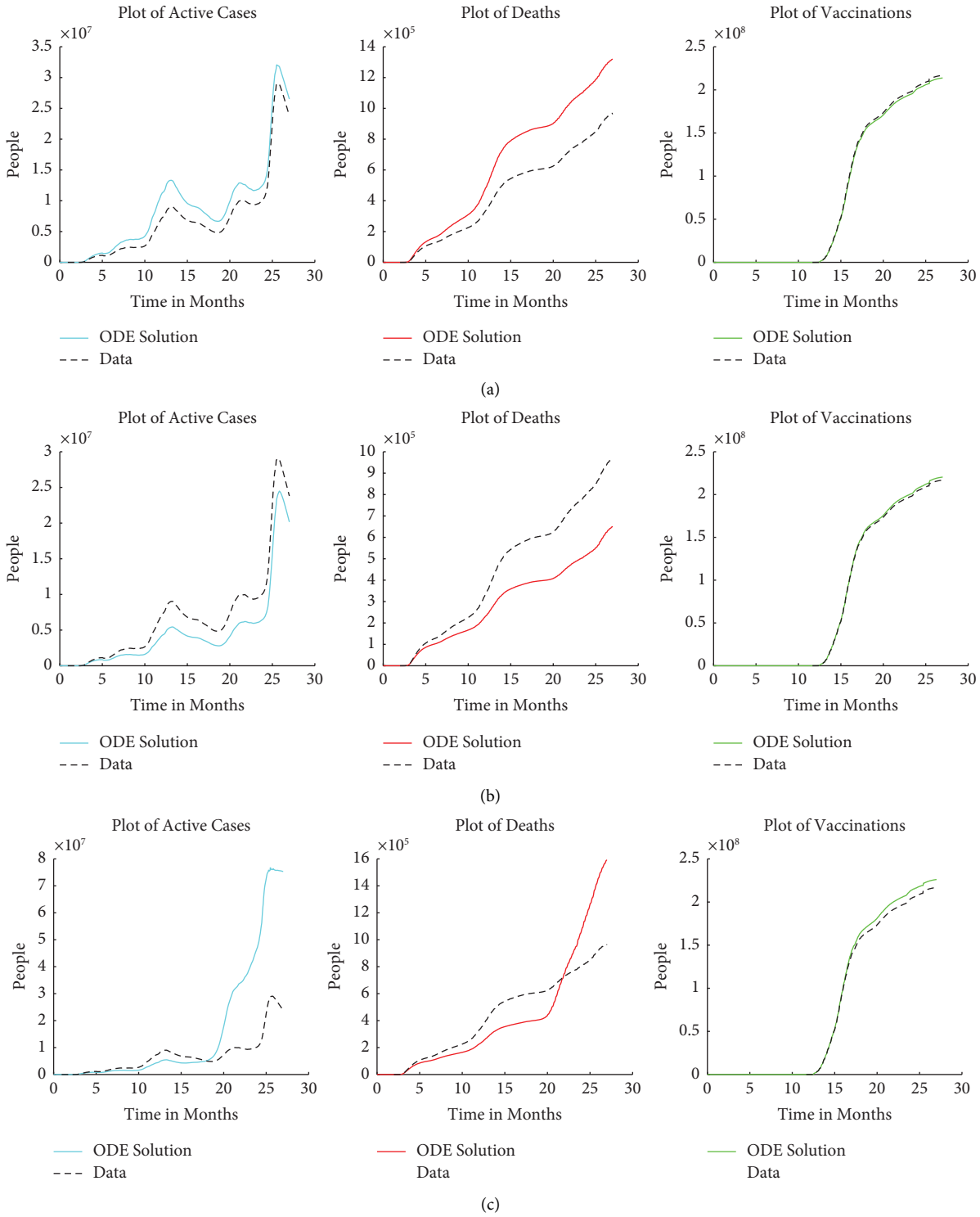

Figure 11: Solutions after perturbing each parameter value by +0.001: (a) perturbation of β, (b) perturbation of γ, and (c) perturbation of μ.

Table 2: Error.

Class	SSE	SST	R^2	FVU
Infected	3.3845×10^{11}	3.7574×10^{16}	1.0000	9.2507×10^{-6}
Deaths	6.3502×10^{7}	7.1267×10^{13}	1.0000	8.9103×10^{-7}
Vaccinated	6.5550×10^{12}	9.1546×10^{18}	1.0000	7.1663×10^{-7}

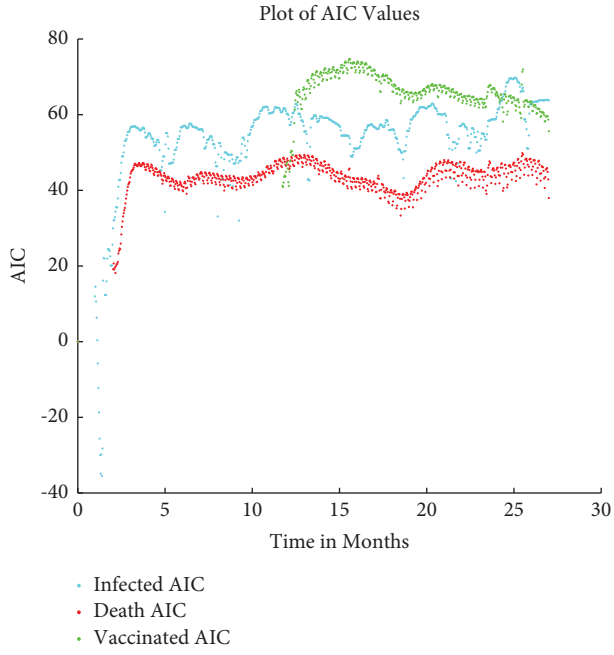

Figure 12: Plot of AIC values.

Table 3: Total AIC.

Class	Total AIC
Infected	20826
Deaths	13897
Vaccinated	23191

8.3. Total AIC. The total AIC is calculated using the SSE between the solutions and respective data over the entire time span previously calculated in Section 7.1, along with the total number of parameters used, which account for every single value for each varying parameter. For total AIC, $m = 805$ and $k = 3(805) + 4 = 2410$ (3 sets of varying parameters and 4 constant parameters).

The total AIC values are similar to those calculated based on the study of dynamic epidemic models via iterated filtering [26]. The AIC values for all three in Table 3 are significantly high due to the large sample size and number of parameters used in total.

9. Conclusion

By allowing the parameters to vary, the numbers of infections, deaths, and vaccinations are able to be matched relatively well to the real data on a daily basis for over two years by a very simple model and optimization technique. A major advantage of using the simplex algorithm for optimization of many thousand parameter values is its ability to yield satisfactory results within merely a few minutes on a personal machine. The basic reproduction number calculated based on the parameter values obtained by the parameter estimation method agrees well with the reproduction rate data. However, the model may be vulnerable to data overfitting, as the solutions depend heavily on the parameter values, where future parameter values may not be precisely determined, possibly leading to the future solutions being inaccurate.

Performing the sensitivity analysis shows that there is a wide range of sensitivity due to the many values of the transmission rate β. Elasticity may be computed as singular scalar values due to not depending on any variable parameters. The SSE along with the AIC values falls within the typical range for models of similar nature. From the results, it can be concluded that the model along with the parameter estimation algorithm is able to handle the United States' COVID-19 data well and therefore should be able to handle simulating the spread of other infectious diseases with more intricate behavior, in which models with constant parameters are unable to. A limitation worth mentioning is that the model is designed around COVID-19 data in the United States, which makes the use of this model applicable only to similar types of data. The model may be modified and adjusted accordingly in order to be applied to other infectious diseases that are similar in nature.

Data Availability

The data used to support the findings of this study are included within the website, https://www.worldometers.info/coronavirus/country/us/ [14].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

We would like to thank Dr. Casey Diekman and Soheil Sagafhi of New Jersey Institute of Technology for assisting in the groundwork and initial structure of the project. This research is supported by Ph.D Degree Program in Mathematics, Faculty of Science, Chiang Mai University, under the CMU Presidential Scholarship and Chiang Mai University.

References

[1] F. Brauer, *Mathematical Models in Epidemiology*, Springer, New York, NY, USA, 2019.

[2] M. Martcheva, *An Introduction to Mathematical Epidemiology*, Springer, New York, NY, USA, 2015.

[3] S. A. Alanazi, M. M. Kamruzzaman, M. Alruwaili, N. Alshammari, S. A. Alqahtani, and A. Karime, "Measuring and preventing COVID-19 using the SIR model and machine learning in smart health care," *Journal of Healthcare Engineering*, vol. 2020, Article ID 8857346, 12 pages, 2020.

[4] C. J. Wang, C. Y. Ng, and R. H. Brook, "Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing," *JAMA*, vol. 323, no. 14, pp. 1341-1342, 2020.

[5] B. Tang, X. Wang, Q. Li et al., "Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions," *Journal of Clinical Medicine*, vol. 9, no. 2, p. 462, 2020.

[6] M. Alenezi, F. Al-Anzi, and H. Alabdulrazzaq, "Building a sensible SIR estimation model for COVID-19 outspread in

Kuwait," *Alexandria Engineering Journal*, vol. 60, no. 3, pp. 3161–3175, 2021.

[7] I. Cooper, A. Mondal, and C. G. Antonopoulos, "A SIR model assumption for the spread of COVID-19 in different communities," *Chaos, Solitons and Fractals*, vol. 139, Article ID 110057, 2020.

[8] L. Lopez and X. Rodo, "A modified SEIR model to predict the COVID-19 outbreak in Spain and Italy: simulating control scenarios and multi-scale epidemics," *Results in Physics*, vol. 21, Article ID 103746, 2021.

[9] L. Peng, "Epidemic analysis of COVID-19 in China by dynamical modeling," p. 6563, 2020, https://arxiv.org/abs/2002.06563.

[10] K. Prem, Y. Liu, T. W. Russell et al., "The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study," *The Lancet Public Health*, vol. 5, no. 5, pp. e261–e270, 2020.

[11] N. M. Ferguson, "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand," *Imperial College London*, vol. 16, no. 16, pp. 1–20, 2020.

[12] G. Giordano, F. Blanchini, R. Bruno et al., "Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy," *Nature Medicine*, vol. 26, no. 6, pp. 855–860, 2020.

[13] J. M. Read, J. R. E. Bridgen, D. A. T. Cummings, A. Ho, and C. P. Jewell, "Novel coronavirus 2019-nCoV (COVID-19): early estimation of epidemiological parameters and epidemic size estimates," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 376, pp. 1–4, 2021.

[14] M. Harris, "WHO issues its first emergency use validation for a COVID-19 vaccine and emphasizes need for equitable global access," 2020, https://www.who.int/news/item/31-12-2020-who-issues-its-first-emergency-use-validation-for-a-covid-19-vaccine-and-emphasizes-need-for-equitable-global-access.

[15] R. Ghostine, M. Gharamti, S. Hassrouny, and I. Hoteit, "An extended SEIR model with vaccination for forecasting the COVID-19 pandemic in Saudi arabia using an ensemble kalman filter," *Mathematics*, vol. 9, no. 6, p. 636, 2021.

[16] G. Nastasi, C. Perrone, S. Taffara, and G. Vitanza, "A time-delayed deterministic model for the spread of COVID-19 with calibration on a real dataset," *Mathematics*, vol. 10, no. 4, p. 661, 2022.

[17] P. Girardi and C. Gaetan, "An SEIR model with time-varying coefficients for analyzing the SARS-CoV-2 epidemic," *Risk Analysis*, vol. 43, no. 1, pp. 144–155, 2021.

[18] A. Bousquet, W. H. Conrad, S. O. Sadat, N. Vardanyan, and Y. Hong, "Deep learning forecasting using time-varying parameters of the SIRD model for Covid-19," *Scientific Reports*, vol. 12, no. 1, p. 3030, 2022.

[19] C. You, Y. Deng, W. Hu et al., "Estimation of the time-varying reproduction number of COVID-19 outbreak in China," *International Journal of Hygiene and Environmental Health*, vol. 228, Article ID 113555, 2020.

[20] W. Worldometer's, "COVID-19 data," 2022, https://www.worldometers.info/coronavirus/country/us/.

[21] H. Ritchie, "Our world in data COVID-19 dataset," 2022, https://ourworldindata.org/coronavirus/country/united-states.

[22] W. C. Roda, M. B. Varughese, D. Han, and M. Y. Li, "Why is it difficult to accurately predict the COVID-19 epidemic?" *Infectious Disease Modelling*, vol. 5, pp. 271–281, 2020.

[23] P. Van den Driessche, "Reproduction numbers of infectious disease models," *Infectious Disease Modelling*, vol. 2, no. 3, pp. 288–303, 2017.

[24] S. K. Yadav and Y. Akhter, "Statistical modeling for the prediction of infectious disease dissemination with special reference to COVID-19 spread," *Frontiers in Public Health*, vol. 9, 2021.

[25] J. O. Rawlings, S. G. Pantula, and D. A. Dickey, *Applied Regression Analysis: A Research Tool*, Springer, New York, NY, USA, Second edition, 1998.

[26] T. Stocks, T. Britton, and M. Höhle, "Model selection and parameter estimation for dynamic epidemic models via iterated filtering: application to rotavirus in Germany," *Biostatistics*, vol. 21, no. 3, pp. 400–416, 2020.