



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Original articles

# SEIR model with unreported infected population and dynamic parameters for the spread of COVID-19

Ziren Chen<sup>a</sup>, Lin Feng<sup>a</sup>, Harold A. Lay Jr.<sup>b,\*</sup>, Khaled Furati<sup>c,\*</sup>, Abdul Khaliq<sup>a,\*</sup><sup>a</sup> Department of Mathematics, Middle Tennessee State University, Murfreesboro, TN 37132, USA<sup>b</sup> Thompson Machinery Commerce Corporation, 1245 Bridgestone Blvd LaVergne, TN 37086, USA<sup>c</sup> Department of Mathematics & Statistics, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia

Received 20 June 2021; received in revised form 9 December 2021; accepted 18 February 2022

Available online 25 February 2022

## Abstract

Coronavirus disease 2019 (COVID-19) is a contagious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that can be transmitted through human interaction. In this paper, we present a Piecewise Susceptible–Exposed–Infectious–Unreported–Removed model for infectious diseases and discuss qualitatively and quantitatively. The parameters are explored by mathematical and statistical methods. Numerical simulations of these models are performed on COVID-19 US data and Python is used in the visualization of results. Outbreak factor is generated by piecewise model to explore the future trend of the US pandemic. Several error metrics are given to discuss the accuracy of the models. The main achievement of this paper is to propose the piecewise model and find the relationship between spread of pandemic and mitigation measures to control it by observing the results of numerical simulations. Performance analysis of piecewise model is presented based on COVID-19 data obtained by ‘worldmeter’.

© 2022 International Association for Mathematics and Computers in Simulation (IMACS). Published by Elsevier B.V. All rights reserved.

**Keywords:** Dynamic parameters; SEIUR; Outbreak factor; COVID-19; Unreported population

## 1. Introduction

Since the end of 2019, the COVID-19 pandemic [1,2], caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is raging around the world [16]. In December 2019, the first cases of COVID-19 were detected in Wuhan, Hubei Province, China. As of March 11, 2021, more than 118 million confirmed cases have been reported in 219 countries and regions, of which more than 2.621 million died and 66.926 million recovered. At present, the number of confirmed cases is still rising rapidly [3].

Hamer, Ross, and others have done a lot of work in the establishment of mathematical models of infectious diseases until 1927 [7,10]. Kermack and McKendrick studied the SIR compartmental model during the Black Death in London [7,14], and the SIS model was established in 1932 [15,19]. Based on the study of these models, the threshold theory in the dynamics of infectious diseases was proposed [17]. The SIR model of Kermack and McKendrick is the most classic and basic model among infectious disease models and has made a foundational

\* Corresponding authors.

E-mail address: [jj.lay@tmcat.com](mailto:jj.lay@tmcat.com) (H.A. Lay Jr.).

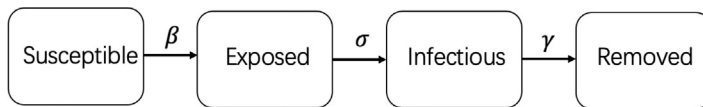


Fig. 1. SEIR model.

contribution to the study of infectious disease dynamics. The SIR model divides the total population into the following three populations: susceptible  $S$ , which represents the people who are not infected but are likely to be infected by this type of disease; infectious  $I$ , which represents the people who have been infected as patients and has the capability of transmitting the disease; and removed  $R$ , which stands for the people who have been removed from the infected, see also [12,13] and references therein. Nowadays, the SEIR model [8] is widely used to analyze infectious diseases, see also [9,20] and references therein. In this model, individuals experience a long incubation duration (the “exposed” category), such that the individual is infected but not yet symptomatic. For example, chicken pox and even vector-borne diseases such as dengue hemorrhagic fever have a long incubation duration where the individual cannot yet transmit the pathogen to others.

The SEIR model is as follows;

$$\frac{dS(t)}{dt} = -\frac{\beta S(t)I(t)}{N} \tag{1}$$

$$\frac{dE(t)}{dt} = \frac{\beta S(t)I(t)}{N} - \sigma E(t) \tag{2}$$

$$\frac{dI(t)}{dt} = \sigma E(t) - \gamma I(t) \tag{3}$$

$$\frac{dR(t)}{dt} = \gamma I(t) \tag{4}$$

The meaning of each parameter and variable is shown below:

$N$ : Total population.

$S(t)$ : The number of susceptible individuals at time  $t$ . They are not infected at this point.

$E(t)$ : The number of asymptomatic individuals exposed to the virus at time  $t$ . They are in the incubation period and capable of transmitting the disease but have not exhibited any symptoms.

$I(t)$ : The number of symptomatic infected individuals at time  $t$ . This group has tested positive and corresponds to the Worldmeter data for active cases. [5]

$R(t)$ : The number of individuals who have been removed from the infected at time  $t$  due to either death or recovery.

$\beta$ : Transmission rate.

$\gamma$ : Removed rate.

$\sigma$ : Incubation rate.

The SEIR diagram below shows how individuals move through each compartment in the model (see Fig. 1):

The SEIR model can be applied to most infectious diseases, but there are some limitations when applied to the COVID-19 data. Due to the characteristics of COVID-19, exposed individuals also have the capability of transmitting the disease. The traditional SEIR model assumes that the members of the  $E$  compartment are infected but not infectious during the incubation period[8]. The SEIR model also ignores the group of unreported cases. Due to the lack of medical resources and variability in the testing policy, there is no guarantee that all infected people are tested and reported during a wide-spread pandemic. This group of people is likely to become hidden transmitters in the population. The SEIR model with constant parameters cannot be used for long-term simulation. This is because the transmission rate  $\beta$  and removed rate  $\gamma$  must be changed over a long period of time.

In this paper, we propose the piecewise SEIUR model and the use of a type of least-squares method to estimate the parameters, which overcomes the main difficulty in the SEIR model.

## 2. SEIUR model

We expect to address the two weaknesses of the SEIR model. First, to overcome the second weakness, the SEIUR model divides the infection equation (3) from SEIR model into two new parts: an unreported symptomatic

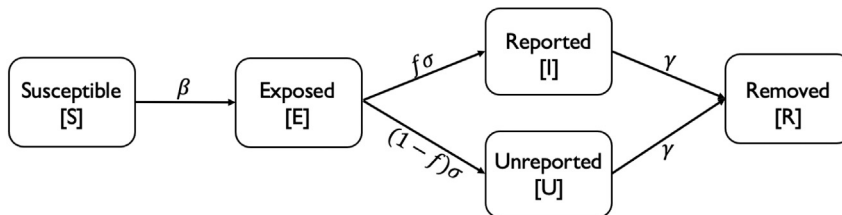


Fig. 2. SEIUR model.

infection equation (13) and a reported symptomatic infection equation (12). Compartment  $I$  represents individuals who are infectious and have been tested (and thus reported). The new compartment  $U$  represents people who are symptomatic and infectious but have not been tested for various reasons. Let  $f$  be the proportion from  $E$  to  $I$ , and  $1 - f$  be the proportion from  $E$  to  $U$ . To handle the first weakness, the infectious population  $I$  in Eqs. (10) and (11) is replaced by the sum of exposed population  $E$  and unreported infectious cases  $U$ , because the individuals in groups  $U$  and  $E$  have the ability to spread the virus to others during the infection. On the contrary, people in group  $I$  were required to be quarantined at home or in hospital and could not infect others.

The establishment of the SEIUR model is based on the following assumptions:

- Keep first three assumptions same as SEIR model
- Exposed and Unreported infected individuals are capable of transmitting the disease
- Reported infectious individuals do not have capability of transmitting the disease
- Unreported infected individuals cannot become the reported infected individuals

The SEIUR model is given by;

$$\frac{dS(t)}{dt} = -\frac{\beta S(t)(E(t) + U(t))}{N} \tag{5}$$

$$\frac{dE(t)}{dt} = \frac{\beta S(t)(E(t) + U(t))}{N} - \sigma E(t) \tag{6}$$

$$\frac{dI(t)}{dt} = \sigma f E(t) - \gamma I(t) \tag{7}$$

$$\frac{dU(t)}{dt} = \sigma(1 - f)E(t) - \gamma U(t) \tag{8}$$

$$\frac{dR(t)}{dt} = \gamma(I(t) + U(t)) \tag{9}$$

Keep  $S(t)$ ,  $E(t)$ ,  $R(t)$ ,  $\beta$ ,  $\gamma$ , and  $\sigma$  same as SEIR Model.

$I(t)$ : The number of reported infected individuals at time  $t$  who are symptomatic and have been tested.

$U(t)$ : The number of individuals infected with the virus but have not been tested at time  $t$  and thus not reported.

$f$ : The proportion of Exposed individuals that become reported infected individuals (Reported fraction) (see Fig. 2).

The transition diagram is shown as:

No matter how the SEIR model changes, the core parameters are always  $\beta$  and  $\gamma$ . Therefore, the estimation of  $\beta$  and  $\gamma$  is particularly vital for both SEIR model and SEIUR model.

### 2.1. Piecewise SEIUR model

The SEIR and SEIUR models have a fixed value for  $\beta$  and  $\gamma$  for the entire time period. Since the values of  $\beta$  and  $\gamma$  are affected by many factors, such as government epidemic prevention measures, quarantine, and vaccination, it is unreasonable to set  $\beta$  and  $\gamma$  as constants. In addition, another popular method is to use a function to fit  $\beta$  and  $\gamma$  respectively. This approach makes the values of  $\beta$  and  $\gamma$  continuous, and the changes are too frequent. Therefore, this paper combines these two ideas, assuming that the values of  $\beta$  and  $\gamma$  are constant in a period of time, and when entering the next period, the values of  $\beta$  and  $\gamma$  are updated. The selection of update points is based on the

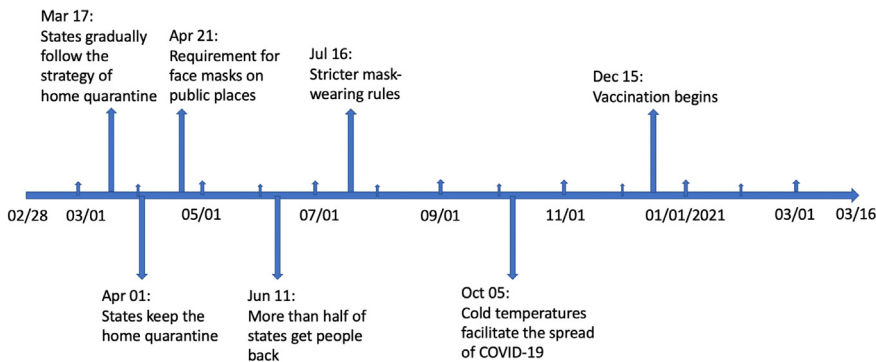


Fig. 3. Strategies [18] release timeline.

time when the government implemented the relevant strategies or policies, and the phase between two consecutive update points is defined as the “period”. The strategies release timeline is shown in Fig. 3.

Using Fig. 3, we separate the whole time period from 2/28/2020 to 3/16/2021 into eight different parts:

1. From 2/28/2020 to 3/16/2020
2. From 3/17/2020 to 3/31/2020
3. From 4/1/2020 to 4/20/2020
4. From 4/21/2020 to 6/10/2020
5. From 6/11/2020 to 7/15/2020
6. From 7/16/2020 to 10/4/2020
7. From 10/5/2020 to 1/15/2021
8. From 1/15/2021 to 3/16/2021

For each period  $i$ , we have an independent SEIUR model, and corresponding specific parameters  $\beta_i$  and  $\gamma_i$ . The piecewise SEIUR model is shown below:

$$\frac{dS(t)}{dt} = -\frac{\beta_i S(t)(E(t) + U(t))}{N} \tag{10}$$

$$\frac{dE(t)}{dt} = \frac{\beta_i S(t)(E(t) + U(t))}{N} - \sigma E(t) \tag{11}$$

$$\frac{dI(t)}{dt} = \sigma f E(t) - \gamma_i I(t) \tag{12}$$

$$\frac{dU(t)}{dt} = \sigma(1 - f)E(t) - \gamma_i U(t) \tag{13}$$

$$\frac{dR(t)}{dt} = \gamma_i(I(t) + U(t)) \tag{14}$$

$\beta_i$ : Transmission rate during the period  $i$ .

$\gamma_i$ : Removed rate during the period  $i$ .

Other notations have the same meaning as before and the transition diagram for period  $i$  is shown in Fig. 4.

In an epidemic, we can obviously judge whether the epidemic is breaking out or disappearing through the change rate of the number of infected people. In the piecewise SEIUR model,  $E$ ,  $I$  and  $U$  represent the number of infections. The difference is that those in  $E$  are in the incubation period, while those in  $I$  and  $U$  have been infected.

Thus, we can conclude that the epidemic is still in outbreak when the change rate of the summation of  $E$ ,  $I$  and  $U$  is greater than 0, that is,

$$\frac{d(E + I + U)}{dt} > 0$$

which implies

$$\frac{dE}{dt} + \frac{dI}{dt} + \frac{dU}{dt} > 0$$

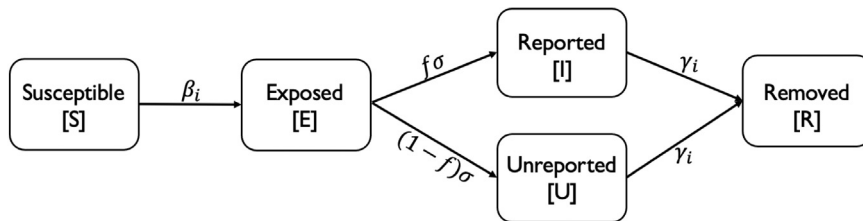


Fig. 4. SEIUR model for period  $i$ .

By the SEIUR model

$$\frac{\beta_i S(t)(E(t) + U(t))}{N} - \sigma E(t) + \sigma E(t) - \gamma_i(I(t) + U(t)) > 0$$

It follows that  $\sigma E(t)$

$$\frac{\beta_i S(t)(E(t) + U(t))}{N} - \gamma_i(I(t) + U(t)) > 0$$

Since  $\gamma_i, I(t), U(t)$  are all positive, we have

$$\frac{\beta_i(E(t) + U(t))S(t)}{\gamma_i(I(t) + U(t))N} > 1$$

When the rate of change of  $E + I + U < 0$ , we have that the virus is disappearing.

Using the previous approach, we have that

$$\frac{\beta_i(E(t) + U(t))S(t)}{\gamma_i(I(t) + U(t))N} < 1$$

When the rate of change of  $E + I + U = 0$ , we have that the virus is under control and may coexist with humans for a long time.

Therefore, we have

$$\frac{\beta_i(E(t) + U(t))S(t)}{\gamma_i(I(t) + U(t))N} = 1$$

By the mathematical analysis above, in each period  $i$ , the formula of the outbreak factor at time  $t$  generated by the piecewise SEIUR model is defined by

$$(O(t))_i = \frac{\beta_i(E(t) + U(t))S(t)}{\gamma_i(I(t) + U(t))N}$$

And we define  $O_i = \overline{O(t)}_i$  to represent the average value of outbreak factor in period  $i$ .

### 3. Learning of the piecewise SEIUR model

#### 3.1. Data

The data used in this paper is from the reference website ‘Worldmeter’ [5]. This website provides a variety of real-time statistical data. The website belongs to Dadax, an independent digital media company in the United States. This website records all pandemic data in the United States from February 15, 2020 to the present, and the data is updated daily.

The data is comprised of three different contents. The first is active cases starting from 2/15/2020 to 3/16/2021. The second and third contents are in same period and include the total cases and total deaths respectively. The detail of the contents is shown in Table 1:

We shift the initial data point from 2/15/2020 to 2/28/2020. Although the United States announced the first case of the coronavirus on 1/21/2020, the lack of effective detection methods and insufficient understanding of the coronavirus have led to the low accuracy of the data in the early stage of the pandemic. Thus, we discarded the data from the early period of the pandemic and set 2/28/2020 as the new initial data point.

**Table 1**  
Data content of total cases, active cases and total deaths.

Data content	Data description	Data period	Data size
Total cases	Total number of infections (including deaths and recoveries)	2/15/2020 to 3/16/2021	(1, 396)
Active cases	Current number of infections (excluding deaths and recoveries)	2/15/2020 to 3/16/2021	(1, 396)
Total deaths	Total number of deaths	2/15/2020 to 3/16/2021	(1, 396)

**Table 2**  
Split data according to strategies [18].

Period #	Time period	Data size	Related policy
Period 1	2/28/2020 to 3/16/2020	(1, 18)	No control measures implemented
Period 2	3/17/2020 to 3/31/2020	(1, 15)	States gradually follow the strategy of home quarantine
Period 3	4/1/2020 to 4/20/2020	(1, 20)	States keep the home quarantine
Period 4	4/21/2020 to 6/10/2020	(1, 51)	Requirement for face masks on public places
Period 5	6/11/2020 to 7/15/2020	(1, 35)	More than half of states get people back to work
Period 6	7/16/2020 to 10/4/2020	(1, 81)	Stricter mask-wearing rules
Period 7	10/5/2020 to 12/14/2020	(1, 71)	Cold temperatures facilitate the spread of COVID-19
Period 8	12/15/2020 to 3/16/2021	(1, 92)	Vaccination begins

In addition to the above three sets of data, this paper also needs the total removed cases and total recovered cases. These two sets of data can be obtained by simple calculations between total cases, active cases, and total deaths. The formulas is followed:

$$\text{Total Removed Cases} = \text{Total Cases} - \text{Active Cases}$$

$$\text{Total Recovered Cases} = \text{Total Removed Cases} - \text{Total Deaths}$$

In previous sections, we clearly explained the reasons for constructing the piecewise model and how we segmented the data into distinct periods.

Table 2 shows the details about each time period.

According to the piecewise SEIUR model, in each period  $i$ , there is a corresponding SEIUR model. In addition to showing the accuracy of SEIUR model in each period  $i$ , this paper also illustrates the performance of this model in the short future term. In other words, we obtain the SEIUR model in period  $i$ , and use it to estimate the data at the beginning of period  $i + 1$ . For that reason, the training data for SEIUR model in period  $i$  is all data from itself, and the test data consists of the first seven days from period  $i + 1$ . For the last period, since the next period data is not available, we divide the period itself into training data and test data with test data being the last seven days.

### 3.2. Reported rate estimation and algorithms

Our first goal is to estimate an appropriate value of the proportion  $f$ . Recalling the definition explained in the previous section,  $f$  stands for the proportion of  $E$  that become  $I$ . Let  $E(t)$  represent the exposed individuals in time  $t$ ,  $\gamma$  represent the removed rate,  $\sigma$  is the incubation rate,  $n(t)$  represent the daily COVID-19 tests per thousand people at time  $t$ ,  $p(t)$  stand for the daily positive rate (the share of COVID-19 tests that are positive) at time  $t$ , and  $N$  is the total population. And we assume that all individuals in  $E(t)$  are newly infected and the effect of new exposed cases from  $t$  to  $t + \frac{1}{\sigma}$  on the results is ignored, then the proportion  $f(t)$  at time  $t$  is given by

$$f(t) = \frac{\sum_{i=t}^{t+\frac{1}{\sigma}} \left( \frac{n(i)p(i)N}{1000} \right)}{E(t)}. \tag{15}$$

In fact, the reported rate at time  $t$  can be approximately equal to the proportion of  $E(t)$  that would be converted into reported cases. Consequently, to estimate  $f(t)$ ,  $E(t)$  is served as the denominator, and the numerator should be the number of newly reported cases from time  $t$  to the time that all  $E(t)$  is converted into reported cases or unreported cases. By the definition, we have that  $\frac{n(t)p(t)N}{1000}$  represents the daily positive tests at time  $t$ .  $\frac{1}{\sigma}$  stands for the average number of days it takes for a newly infected individual (in  $E$ ) to move to the next compartment ( $I$  or  $U$ ). Hence,  $\sum_{i=t}^{t+\frac{1}{\sigma}} (\frac{n(i)p(i)N}{1000})$  can be used to estimate the number of newly reported cases from time  $t$  to the time that all  $E(t)$  is converted into reported cases or unreported cases. The result is the formula (15).

Since the number of exposed people cannot be directly counted, it is almost impossible to find relevant statistical data on  $E$ . Here we use the SEIR model to generate historical data of  $E(t)$ .  $\sigma$  is  $\frac{1}{6}$  assumed in previous section.  $\gamma$  is obtained by the least square method.  $N$  is the total population of United States. For  $n(t)$  and  $p(t)$ , we use data from ‘Our World in Data’ [11], which provide the daily COVID-19 tests per thousand people and the daily positive rate. Now, we can generate the proportion  $f(t)$  at any day  $t$ .

For simplicity, we define  $f = \overline{f(t)}$  as the constant proportion rate for all data and use it in the numerical simulation.

By the formula (15) and the provided data, we estimate  $f = \overline{f(t)} = 0.5937$

We can easily obtain the numerical solution of the SEIUR model and piecewise SEIUR model. The algorithms are shown as Algorithms 1 and 2.

---

**Algorithm 1** Determining the numerical solution of SEIUR model at all time  $t \in \{1, 2, \dots, n\}$

---

**Input:**

- The initial value of variables:  $S[0], E[0], I[0], U[0]$  and  $R[0]$ ;
- The optimal parameters:  $\beta, \gamma, f$  and  $\sigma$ ;
- Total population of the United States:  $N$ ;
- The number of iterations or days:  $n$ ;

**Output**

$S[t], E[t], I[t], U[t]$  and  $R[t]$  at all time  $t \in \{1, 2, \dots, n\}$

---

**Procedure**

**For**  $i$  in 1 to  $n$

$$S[i] = S[i - 1] - \frac{\beta S[i-1]I[i-1]}{N}$$

$$E[i] = E[i - 1] + \frac{\beta S[i-1]I[i-1]}{N} - \sigma E[i - 1]$$

$$I[i] = I[i - 1] + \sigma f E[i - 1] - \gamma I[i - 1]$$

$$U[i] = U[i - 1] + \sigma(1 - f)E[i - 1] - \gamma U[i - 1]$$

$$R[i] = R[i - 1] + \gamma(I[i - 1] + U[i - 1])$$

**Return**  $S[i], E[i], I[i], U[i]$  and  $R[i]$

---



---

**Algorithm 2** Determining the numerical solution of piecewise SEIUR model at time  $t \in \{p_i, p_i + 1, \dots, p_{i+1} - 1\}$  for  $i = 1 \dots n$  where  $p_i$  is the start point of period  $i$  and  $p_1 = 0$

---

**Input:**

- The initial value of variables:  $S[0], E[0], I[0], U[0]$  and  $R[0]$ ;
- The optimal parameters for each period  $i$ :  $\beta_i, \gamma_i, f$  and  $\sigma$ ;
- Total population of the United States:  $N$ ;
- The number of iterations or days:  $n$ ;

**Output**

$S[t], E[t], I[t]$  and  $R[t]$  at all time  $t \in \{1, 2, \dots, n\}$

---



---

**Procedure**

**For**  $i$  in 1 to  $n$

**For**  $j$  in  $p_i$  to  $p_{i+1} - 1$

$$S[j] = S[j - 1] - \frac{\beta_i S[j-1](I[j-1])}{N}$$

$$E[j] = E[j - 1] + \frac{\beta_i S[j-1](I[j-1])}{N} - \sigma E[j - 1]$$

$$I[j] = I[j - 1] + \sigma f E[j - 1] - \gamma_j I[j - 1]$$

$$U[j] = U[j - 1] + \sigma(1 - f)E[j - 1] - \gamma_j U[j - 1]$$

$$R[j] = R[j - 1] + \gamma_j(I[j - 1] + U[j - 1])$$

**Return**  $S[j], E[j], I[j], U[j]$  and  $R[j]$

---

In the SEIUR model, there are four unknown parameters to be estimated:  $\beta, \sigma, f,$  and  $\gamma$ . We have estimated the proportion of  $E$  to  $I$  as 0.6, that is,  $f = 0.6$ . The incubation rate of the COVID-19 has been estimated between 2 to 11 days (2.5th to 97.5th percentile) and the mean incubation period as 6.4 days (95% CI: 5.6–7.7) [6]. Thus, we assume  $\sigma = \frac{1}{6}$ . Least square method is employed to find the optimal parameters of  $\beta$  and  $\gamma$ . The algorithm is shown as Algorithm 3.

---

**Algorithm 3** Parameter Estimation for SEIUR model

---

**Input**

Sequence  $\{\beta_j\}_{j=0}^{99} = \{0, 0.01, 0.02, \dots, 0.99\}$ ;

Sequence  $\{\gamma_j\}_{j=0}^{99} = \{0, 0.01, 0.02, \dots, 0.99\}$ ;

Parameters:  $f = 0.6, \sigma = \frac{1}{6}$ ;

Numerical Solution (Algorithm 1) for  $I$  at day  $t$ , denoted as  $I(t, \beta, \gamma)$ ;

Real Data [5] for  $I$  at day  $t$ , denoted as  $AI(t)$

The number of iterations or days:  $n$ ;

**Output**

Optimal Parameters:  $\beta, \gamma$

---

**Procedure**

**For**  $i$  in 0 to 99

**For**  $j$  in 0 to 99

$$m_{ij} = \frac{1}{n} \sum_{t=0}^n (AI(t) - I(t, \beta_i, \gamma_j))^2$$

**Find**  $\min\{m_{ij}, \text{ for } i, j \text{ in } 0 \text{ to } 99\}$  and corresponding indices  $a$  and  $b$

**Return**  $\beta_a$  and  $\gamma_b$

---

In the piecewise SEIUR model, we also have  $\sigma = \frac{1}{6}$  and  $f = 0.6$ . Parameter estimation of  $\beta_i$  and  $\gamma_i$  for each period  $i$  is shown in Algorithm 4.

---

**Algorithm 4** Parameter Estimation for piecewise SEIUR model

---

**Input**

The number of periods:  $k$ ;

Sequence  $\{\beta_j\}_{j=0}^{99} = \{0, 0.01, 0.02, \dots, 0.99\}$ ;

Sequence  $\{\gamma_j\}_{j=0}^{99} = \{0, 0.01, 0.02, \dots, 0.99\}$ ;

Parameters:  $f = 0.6, \sigma = \frac{1}{6}$ ;

Numerical Solution (Algorithm 2) for  $I$  at day  $t$ , denoted as  $I(t, \beta, \gamma)$ ;

Real Data [5] for  $I$  at day  $t$ , denoted as  $AI(t)$ ;

The number of iterations or days:  $n$ ;

Note that:  $\{0, 1, 2, \dots, n\} = \cup_{i=1}^k \{p_i, p_i + 1, \dots, p_{i+1} - 1\}$ ;

where  $p_i$  is the start point of period  $i$  and  $p_1 = 0$ ;

**Output**

Optimal Parameters:  $(\beta_i, \gamma_i)$  for all  $i$  in 0 to  $k$

---

---

```

Procedure
For  $i$  in 1 to  $k$ 
  For  $j$  in 0 to 99
    For  $g$  in 0 to 99
       $m_{ijg} = \frac{1}{p_{i+1}-p_i} \sum_{t=p_i}^{p_{i+1}} (AI(t) - I(t, \beta_j, \gamma_g))^2$  (Mean Square Error)
    Find  $\min\{m_{ijg}, \text{ for } j, g \text{ in } 0 \text{ to } 99\}$  and corresponding indices  $a_i$  and  $b_i$ 
  Return  $\beta_{a_i}$  and  $\gamma_{b_i}$ 

```

---

#### 4. Simulation results

##### 4.1. Parameter estimation results

Based on the parameter estimation Algorithm 3, we estimate the parameters  $\beta$  and  $\gamma$  for SEIUR model in this section. Fig. 5 is the 3D error graph for different values of  $\beta$  and  $\gamma$ .

We know that the optimal parameters can be derived from the lowest error point. For the SEIUR model, the optimal parameters are  $\beta = 0.13$  and  $\gamma = 0.08$

Based on the parameter estimation Algorithm 4, we estimate the parameters  $\beta_i$  and  $\gamma_i$  for piecewise SEIUR model in each period  $i$ . Figs. A.18 to A.21 are the 3D error graphs showing the optimal values of parameters for each period.

Applying the same method from the SEIUR model to the piecewise SEIUR model, we can find the  $\beta$  and  $\gamma$  related to the lowest error for each period  $i$ . This allows us to generate the optimal parameters for all periods in Table 3.

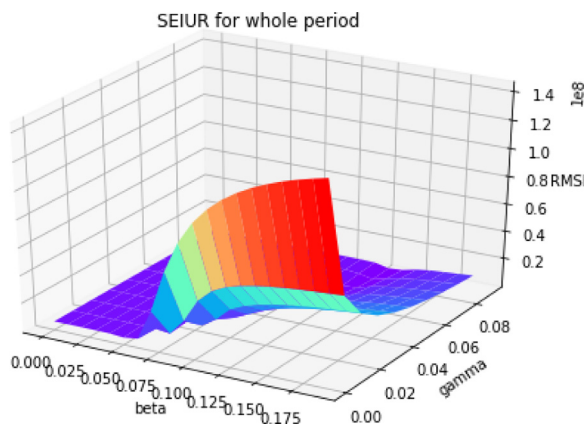


Fig. 5. 3D error graph with different  $\beta$  and  $\gamma$  for SEIUR model.

**Table 3**  
Optimal model parameters by period.

Period	$\beta$	$\gamma$
1	0.44	0.36
2	0.26	0.10
3	0.10	0.04
4	0.03	0.01
5	0.08	0.03
6	0.06	0.03
7	0.11	0.06
8	0.10	0.05

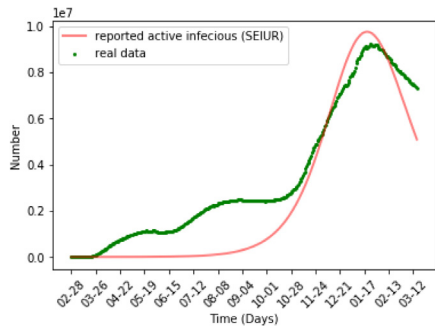


Fig. 6. Numerical simulation for SEIUR.

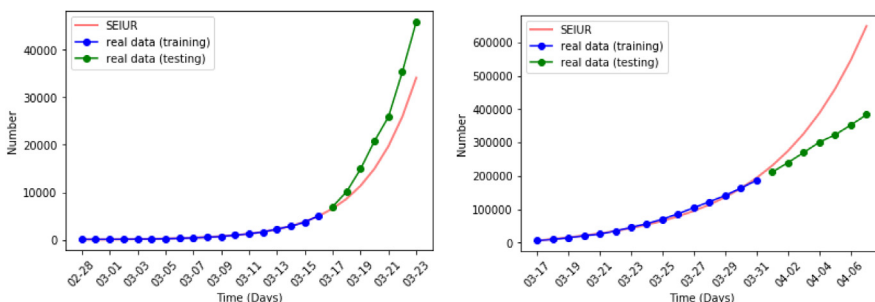


Fig. 7. Numerical simulation in periods 1 and 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

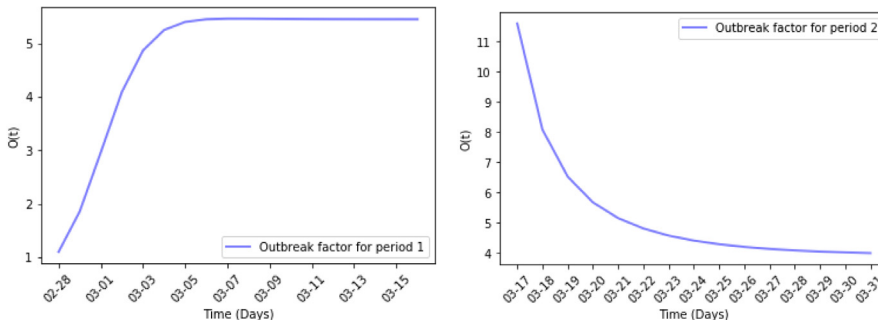


Fig. 8. Outbreak factor in periods 1 and 2.

4.2. Performance

Fig. 6 shows the numerical simulation results of the SEIUR model on all data. We find that when simulating all of the data at once, the SEIUR model can successfully simulate the trend of data, but it cannot accurately reflect the true value of the data at each point. We have a more intuitive reflection in the error analysis in the next section.

Fig. 7 shows the estimated reported active cases and the real data in periods 1 and 2 where the blue dots are training data and the green dots are test data. Fig. 8 shows the daily outbreak factor in periods 1 and 2. In period 1, no control measures were implemented which led to an exponential outbreak, and we calculate that  $O_1 = 4.9461$  in period 1. This is much greater than 1. This is consistent with what we proved in previous section: when  $O(t)$  is much greater than 1, the pandemic is in a major outbreak stage. Observing the simulation results on the test data, the SEIUR model in the first period can still simulate the data at the beginning of the second period very well. This also means that at the beginning of the second period the home quarantine strategy has not yet contributed to a significant impact, and the pandemic was still breaking out. The  $O(t)$  of the entire second period also confirmed

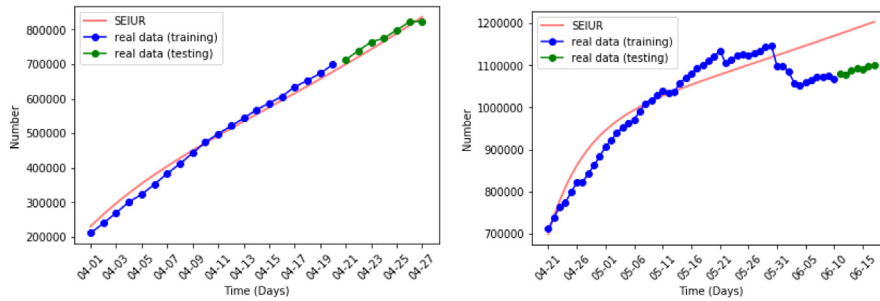


Fig. 9. Numerical simulation in periods 3 and 4.

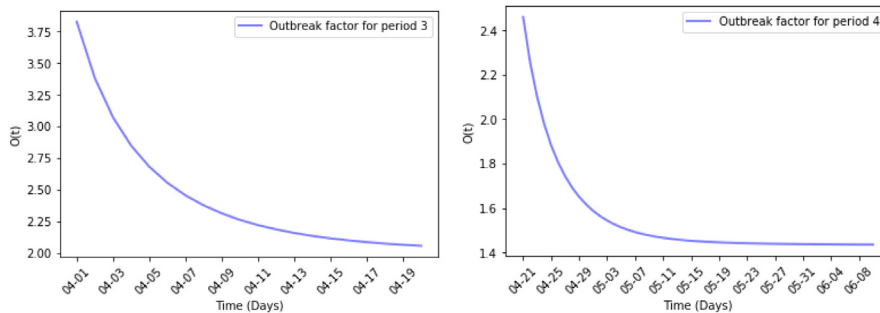


Fig. 10. Outbreak factor in periods 3 and 4.

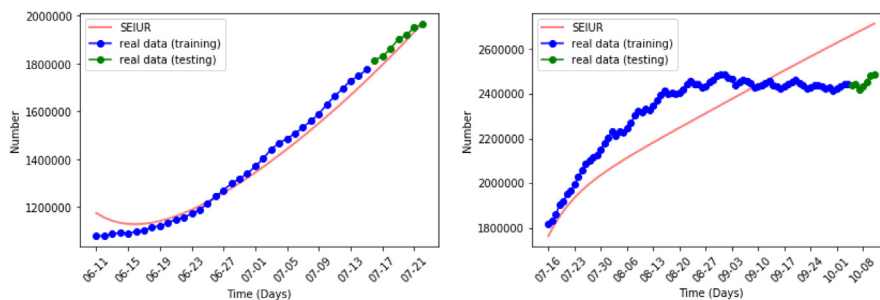


Fig. 11. Numerical simulation in periods 5 and 6.

this point. In this period,  $O_2 = 4.8743$ , which is slightly lower than that of the period 1 but still far greater than 1. This indicates that in the middle and late stages of period 2, the home quarantine strategy has achieved initial results, but the epidemic was still in outbreak. The test data of period 2 showed that at the beginning of the period 3, the true value of the test data is significantly smaller than the results obtained by SEIUR model generated from parameters estimated in period 2. This implies that the home quarantine strategy implemented in the period 2 achieved a significant effect in the period 3.

Figs. 9 and 10 show the results of the numerical simulations and the outbreak factor in periods 3 and 4.  $O_3 = 2.3412$  in period 3, which has a 52% decrease compared to period 2. This confirms the analysis made in the discussion of period 2. The home quarantine strategy has slowed the spread of the epidemic. In period 4, mask wearing further reduced the spread of pandemic. The figure about period 4 illustrates the growth rate of the active cases has decreased compared to period 3 and reached the first turning point at the end of May. Quantitatively,  $O_4 = 1.5362$  in period 4, which is a decrease of 34% compared to period 3.

Figs. 11 and 12 show the results of numerical simulations and outbreak factors in periods 5 and 6. In period 5, many states are relaxing their restrictions, with more than half set to be partially reopened. This led to the pandemic rebounding rapidly after reaching the first turning point and active cases starting to rise again. The average outbreak

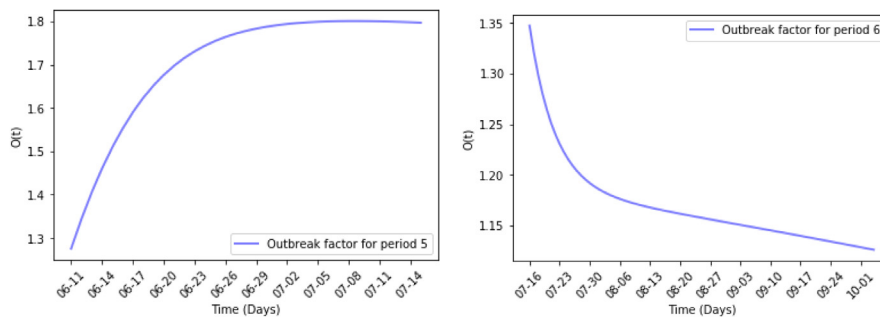


Fig. 12. Outbreak factor in periods 5 and 6.

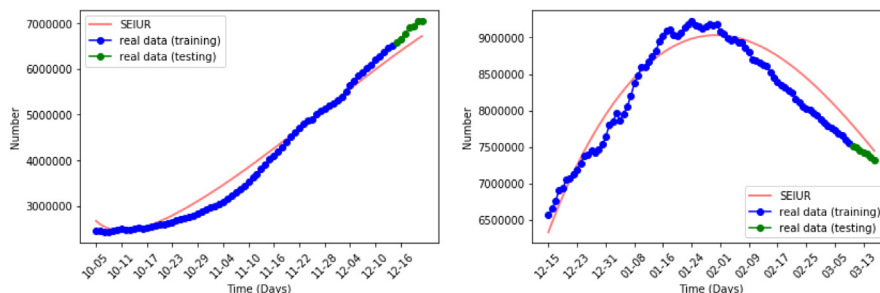


Fig. 13. Numerical simulation in periods 7 and 8.

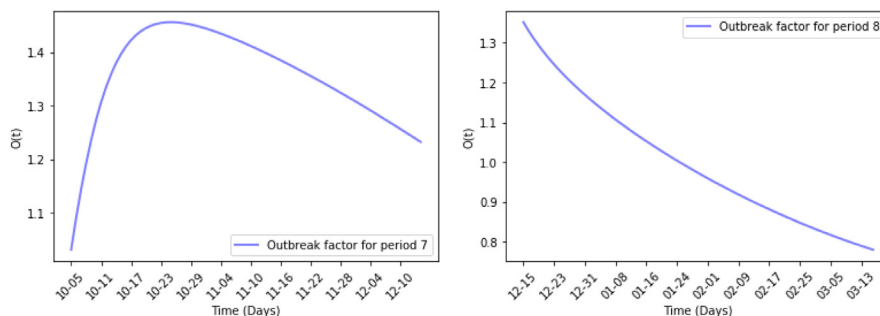


Fig. 14. Outbreak factor in periods 7 and 8.

factor also rose from 1.5362 to 1.7176. In period 6, stricter mask-wearing rules slowed the spread of pandemic, and the second turning point was reached at the end of August with  $O_6 = 1.1660$ .

Figs. 13 and 14 show the results of numerical simulations and outbreak factors in periods 7 and 8. Winter brings shorter days and lower temperatures to the United States which facilitate the spread of COVID-19. Period 7 shows this process. With the advent of winter, active cases began to increase again after reaching the second turning point, and  $O_7$  rose back to 1.3414. Period 8 is the latest period so far, and vaccinations began on the first day of this period. The graph in period 8 shows that the vaccination quickly controlled the spread of the pandemic, and the third turning point appeared at the end of January 2021. So far, the reported active cases have continued to decline. In period 8,  $O_8 = 0.9991$  which is less than 1. If this continues, the epidemic is likely to be completely controlled in the next few months.

Fig. 15 shows the numerical simulation results of the piecewise SEIUR model and SEIUR model, where the result of piecewise SEIUR is obtained by combining the graphs of eight periods. It can be seen from the figure that the piecewise SEIUR model performs better than the SEIUR model. This also means that it is reasonable to divide a long period of time into different sub-period according to policy dates, and then use the SEIUR model to simulate each period respectively.

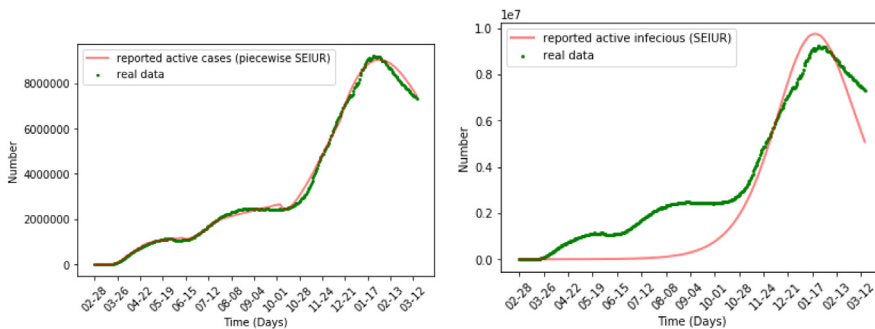


Fig. 15. Numerical simulation for piecewise SEIUR and SEIUR.

### 4.3. Error metrics

After performing the numerical simulation of the model, error analysis is an effective way to assess the accuracy of the model. There are many error metrics, such as mean absolute error (MAE), mean absolute percentage error (MAPE), symmetric mean absolute percentage error (SMAPE), mean squared error (MSE), root mean squared error (RMSE) and  $R^2$  score. In this section we discuss several representative error metrics for error analysis.

The Mean Absolute Percentage Error (MAPE) [4] is one of the most commonly used to measure model accuracy. It is the average of the relative error. It is given by the following formula:

$$MAPE = \frac{1}{n} \cdot \sum_{i=1}^n \left| \frac{real_i - estimate_i}{real_i} \right| \times 100\%$$

The Mean Absolute Error (MAE) [4] is a popular error metric to measure model accuracy. As the name implies, it is the average of the absolute error. The formula is shown below:

$$MAE = \frac{1}{n} \cdot \sum_{i=1}^n |real_i - estimate_i|$$

The Mean Squared Error (MSE) [4] is an error metric to measure model accuracy. Since MSE is a continuous function, it is often used together with the least square method and gradient descent method. It is defined as the average squared error. The formula is shown below:

$$MSE = \frac{1}{n} \cdot \sum_{i=1}^n (real_i - estimate_i)^2$$

The  $R^2$  score [4] function computes the coefficient of determination, usually denoted as  $R^2$ . It represents the proportion of variance (of  $y$ ) that has been explained by the independent variables in the model. It provides an indication of goodness of fit and therefore a measure of how well unseen samples are likely to be predicted by the model through the proportion of explained variance.

If  $\hat{y}_i$  is the predicted value of the  $i$ th sample and  $y_i$  is the corresponding true value for total  $n$  samples, the estimated  $R^2$  is defined as:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$$MAPE_{PW-SEIUR} = 4.083\%$$

$$MAPE_{SEIUR} = 59.526\%$$

From the results shown in Fig. 16, it can be seen that the piecewise SEIUR model has a small MAPE in each period and over the entire data. Specifically, the MAPE from period 3 to period 8 is lower than 5%, the MAPE in

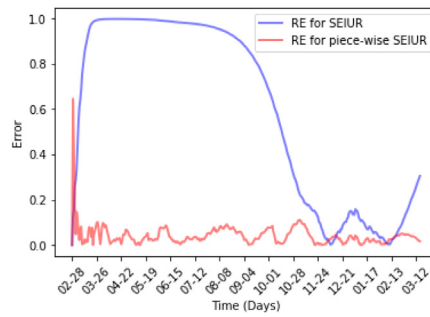


Fig. 16. Relative error for Piecewise SEIUR and SEIUR.

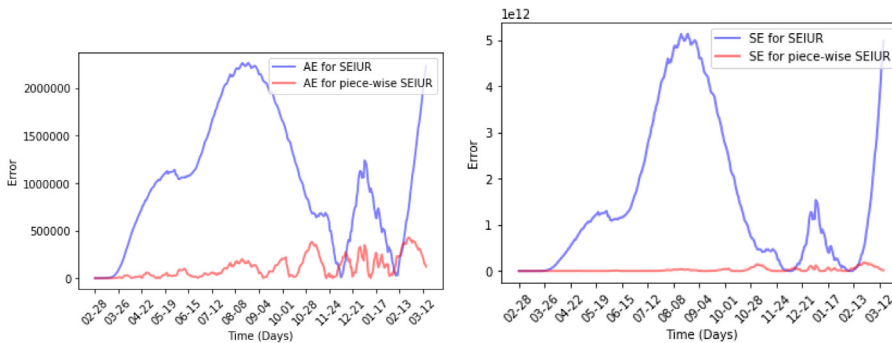


Fig. 17. Absolute error (AE) and Square error (SE) for Piecewise SEIUR and SEIUR.

Table 4  
Error table for SEIUR and Piecewise SEIUR models.

Models name	Error metrics			
	MAPE	$R^2$	MAE	MSE
Piecewise SEIUR model	4.083%	0.9972	$1.12 \times 10^5$	$2.50 \times 10^{10}$
SEIUR model	59.526%	0.8239	$1.041 \times 10^6$	$1.55 \times 10^{12}$

all periods is lower than 10%, and the MAPE in the entire period is only 4%. In contrast, the MAPE of SIEIUR model exceeds 50%.

$$MAE_{PW-SEIUR} = 1.12 \times 10^5, MSE_{PW-SEIUR} = 2.50 \times 10^{10}$$

$$MAE_{SEIUR} = 1.041 \times 10^6, MSE_{SEIUR} = 1.55 \times 10^{12}$$

MAE and MSE are shown in Fig. 17 and produce the same results. No matter which error metric is used, the piecewise SEIUR model has a better performance than the SEIUR model. Quantitatively, the error of piecewise SEIUR model under the MAE metric is only one-tenth of that of the SEIUR, and under the MSE metric it is only one-sixth.

$$R^2_{PW-SEIUR} = 0.9972$$

$$R^2_{SEIUR} = 0.8239$$

The results from  $R^2$  score also illustrate that piecewise SEIUR model has an almost perfect fitting. The  $R^2$  score of piecewise SEIUR model is approximately equal to 1 (see Table 4).



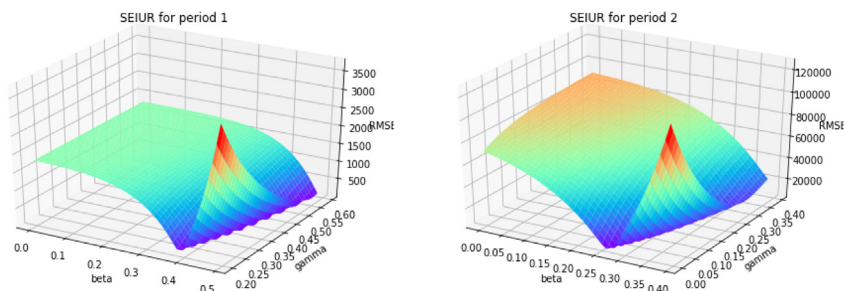


Fig. A.18. 3D error graph for different  $\beta$  and  $\gamma$  for periods 1 and 2.

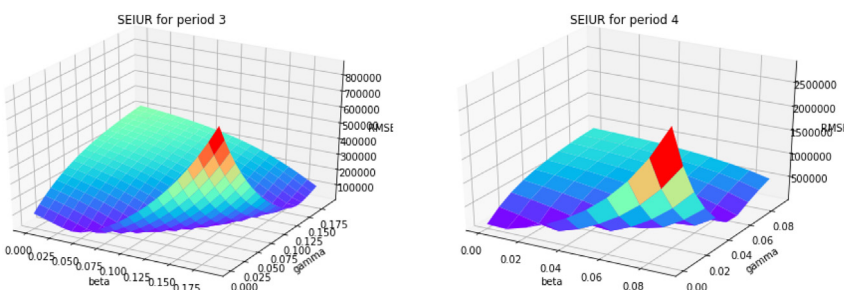


Fig. A.19. 3D error graph for different  $\beta$  and  $\gamma$  for periods 3 and 4.

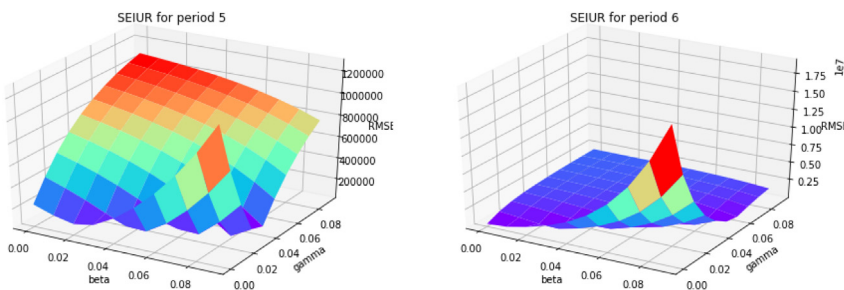


Fig. A.20. 3D error graph for different  $\beta$  and  $\gamma$  for periods 5 and 6.

### 5. Conclusion

In this paper, we have presented the SEIUR model and the piecewise SEIUR model. These two models are tested on COVID-19 data in the United States to demonstrate their performance. We estimated piecewise parameters  $\beta$  and  $\gamma$  for each period of the model. These two models were applied to COVID-19 data in the United States to demonstrate their performance. The piecewise SEIUR model is seen to produce higher simulation accuracy than the SEIUR model. The MAPE of the piecewise SEIUR is only 4%. The error of piecewise SEIUR model under MAE and MSE metrics is far less than that of the SEIUR model. The  $R^2$  score of the piecewise SEIUR model is 0.9972 which is close to 1. The outbreak factor generated by the piecewise SEIUR model can be applied to highlight the impact of the epidemic prevention strategies in each period. This also provides a mathematical tool for future research on the impact of different strategies on epidemics.

### Appendix. Error graphs for the seieur model

See Figs. A.18–A.21.



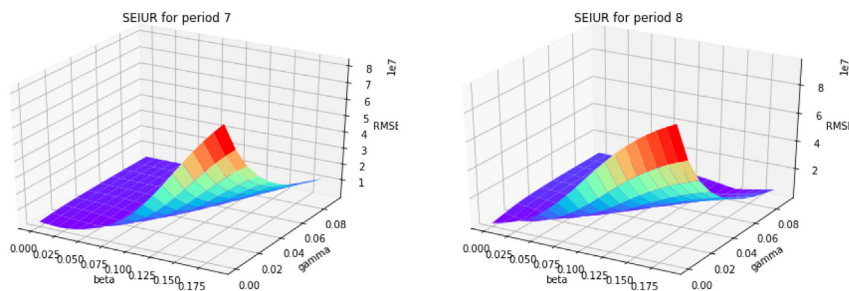


Fig. A.21. 3D error graph for different  $\beta$  and  $\gamma$  for periods 7 and 8.

## References

- [1] Covid-19 Pandemic, Wikipedia, Wikimedia Foundation, [https://en.wikipedia.org/wiki/COVID-19\\_pandemic](https://en.wikipedia.org/wiki/COVID-19_pandemic).
- [2] Coronavirus Disease (Covid-19) - Events as They Happen, World Health Organization. World Health Organization, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>.
- [3] Coronavirus Cases: Worldometer. <https://www.worldometers.info/coronavirus/>.
- [4] 3.3. Metrics And scoring: Quantifying the quality of predictions, Scikit [https://scikit-learn.org/stable/modules/model\\_evaluation.html#regression-metrics](https://scikit-learn.org/stable/modules/model_evaluation.html#regression-metrics).
- [5] United states covid cases, Worldometer <https://www.worldometers.info/coronavirus/country/us/>.
- [6] Jantien A. Backer, Don Klinkenberg, Jacco Wallinga, Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from wuhan, China, 20–28 2020, *Eurosurveillance* 25 (5) (2020) 2000062.
- [7] Fred Brauer, *Mathematical epidemiology: Past, present, and future*, *Infect. Dis. Model.* 2 (2) (2017) 113–127.
- [8] José M. Carcione, Juan E. Santos, Claudio Bagaini, Jing Ba, A simulation of a COVID-19 epidemic based on a deterministic SEIR model, *Front. Public Health* 8 (2020) 230.
- [9] Youssef Hamdy, Najat Alghamdi, Magdy A. Ezzat, Alaa A. El-Bary, Ahmed M. Shawky, Study on the SEIQR model and applying the epidemiological rates of COVID-19 epidemic spread in Saudi Arabia, *Infect. Dis. Model.* 6 (2021) 678–692.
- [10] W.N. Hammer, Epidemic disease in england—the evidence of variability and the persistence of type, *Lancet* II (1906) 733–739.
- [11] Ritchie Hannah, Edouard Mathieu, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Esteban Ortiz-Ospina, Joe Hasell, Bobbie Macdonald, Diana Beltekian, Max Roser, Coronavirus (COVID-19) testing - statistics and research. Our world in data, 2020, <https://ourworldindata.org/coronavirus-testing>.
- [12] Sergey O. Ilyin, A recursive model of the spread of COVID-19: Modelling study, *JMIR Public Health Surveill.* 7 (4) (2021) e21468.
- [13] Long Jie, A.Q.M Khaliq, K.M. Furati, Identification and prediction of time-varying parameters of COVID-19 model: a data-driven deep learning approach, *Int. J. Comput. Math.* 98 (2021) 1617–1632.
- [14] . Kermack, William Ogilvy, Anderson G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. R. Soc. Lond. Ser. A* 115 (772) (1927) 700–721.
- [15] . Kermack, William Ogilvy, Anderson G. McKendrick, Contributions to the mathematical theory of epidemics, ii.—The problem of endemicity, *Proc. R. Soc. Lond. Ser. A* 138 (834) (1932) 55–83.
- [16] B.S Mohan, V. Nambiar, Covid-19: an insight into SARS-CoV-2 pandemic originated at wuhan city in hubei province of China, *J. Infect. Dis. Epidemiol.* 6 (4) (2020) 146.
- [17] Constantinos I. Siettos, Lucia Russo, Mathematical modeling of infectious disease dynamics, *Virulence* 4 (4) (2013) 295–306.
- [18] The New York Times, See Reopening Plans and Mask Mandates for All 50 States, The New York Times. The New York Times, 2020, <https://www.nytimes.com/interactive/2020/us/states-reopen-map-coronavirus.html>.
- [19] Junyuan Yang, Yuming Chen, Fei Xu, Effect of infection age on an SIS epidemic model on complex networks, *J. Math. Biol.* 73 (5) (2016) 1227–1249.
- [20] Hamdy M. Youssef, Najat A. Alghamdi, Magdy A. Ezzat, Alaa A. El-Bary, Ahmed M. Shawky, A new dynamical modeling SEIR with global analysis applied to the real data of spreading COVID-19 in Saudi Arabia, *Math. Biosci. Eng.* 17 (2020) 7018–7044.