



# Toward the Impact of Non-pharmaceutical Interventions and Vaccination on the COVID-19 Pandemic With Time-Dependent SEIR Model

Yuexin Li<sup>1</sup>, Linqiang Ge<sup>2</sup>, Yang Zhou<sup>3</sup>, Xuan Cao<sup>4</sup> and Jingyi Zheng<sup>1\*</sup>

<sup>1</sup> Department of Mathematics and Statistics, Auburn University, Auburn, AL, United States, <sup>2</sup> TSYS School of Computer Science, Columbus State University, Columbus, GA, United States, <sup>3</sup> Department of Computer Science and Software Engineering, Auburn University, Auburn, AL, United States, <sup>4</sup> Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH, United States

## OPEN ACCESS

### Edited by:

Hong Qin,  
University of Tennessee at  
Chattanooga, United States

### Reviewed by:

Ramaraju Rudraraju,  
University of Alabama at Birmingham,  
United States  
Zongliang Yue,  
University of Alabama at Birmingham,  
United States

### \*Correspondence:

Jingyi Zheng  
jingyi.zheng@auburn.edu

### Specialty section:

This article was submitted to  
Medicine and Public Health,  
a section of the journal  
Frontiers in Artificial Intelligence

**Received:** 31 December 2020

**Accepted:** 22 February 2021

**Published:** 18 March 2021

### Citation:

Li Y, Ge L, Zhou Y, Cao X and Zheng J  
(2021) Toward the Impact of  
Non-pharmaceutical Interventions and  
Vaccination on the COVID-19  
Pandemic With Time-Dependent SEIR  
Model. *Front. Artif. Intell.* 4:648579.  
doi: 10.3389/frai.2021.648579

The outbreak of COVID-19, caused by the SARS-CoV-2 coronavirus, has been declared a pandemic by the World Health Organization (WHO) in March, 2020 and rapidly spread to over 210 countries and territories around the world. By December 24, there are over 77M cumulative confirmed cases with more than 1.72M deaths worldwide. To mathematically describe the dynamic of the COVID-19 pandemic, we propose a time-dependent SEIR model considering the incubation period. Furthermore, we take immunity, reinfection, and vaccination into account and propose the SEVIS model. Unlike the classic SIR based models with constant parameters, our dynamic models not only predicts the number of cases, but also monitors the trajectories of changing parameters, such as transmission rate, recovery rate, and the basic reproduction number. Tracking these parameters, we observe the significant decrease in the transmission rate in the U.S. after the authority announced a series of orders aiming to prevent the spread of the virus, such as closing non-essential businesses and lockdown restrictions. Months later, as restrictions being gradually lifted, we notice a new surge of infection emerges as the transmission rates show increasing trends in some states. Using our epidemiology models, people can track, timely monitor, and predict the COVID-19 pandemic with precision. To illustrate and validate our model, we use the national level data (the U.S.) and the state level data (New York and North Dakota), and the resulting relative prediction errors for the infected group and recovered group are mostly lower than 0.5%. We also simulate the long-term development of the pandemic based on our proposed models to explore when the crisis will end under certain conditions.

**Keywords:** COVID-19, epidemiology, dynamic modeling, reinfection, vaccination, time-dependent SEIR model

## 1. INTRODUCTION

On March 11, 2020, the World Health Organization (WHO) declared that the outbreak of the novel coronavirus (COVID-19) can be characterized as a pandemic. The COVID-19 outbreak started in Wuhan, China in December, 2019. By the end of January, 2020, the confirmed cases in China went up to 11,791. Only 1 month later, the number increased nearly seven-fold to 80,134 and the COVID-19 cases gradually showed up in other countries. Starting from March, 2020, the outbreak spread to more than 100 countries. By the end of 2020, the pandemic has led to 77.5M confirmed cases and more than 1.72M fatalities worldwide. **Figure 1** summarizes the percentage of global confirmed cases contributed by each country. As of December 24, the United States, India, and Brazil are the three countries most impacted by the COVID-19 pandemic. The trajectories of the confirmed cases in the three countries are also displayed.

The COVID-19 virus has caused a great disruption to the human health, social life, developments, and economics. To stop the spread of COVID-19 virus, governments have carried out numerous preventive measures such as stay-at-home orders, travel restrictions, school closure, mask-wearing mandate, and so forth. The impact on the society came later in all aspects, including rising unemployment, protests against restrictions, and psychological anxiety and stress brought to the public. However, a significant decrease in the transmission rate occurred, which proved that these mitigation measures were effective. Months later, many states in the U.S. have loosened their restrictions and lifted orders to allow businesses to reopen to the public. Consequently, the diagnoses of daily confirmed cases have displayed a consequential increasing trend after the reopen in some states such as Alabama. By looking at the numbers only, it is difficult to assess what stage we are at in the COVID-19 pandemic and when it is going to end. Hence, mathematical models considering the epidemiological characteristics of COVID-19 become crucial and significant to track and forecast the trend of the spread.

The classic epidemiology model exhibits compelling results, especially during the early period of the pandemic. The compartmental models, which are the simplified versions of mathematical models for infectious diseases, divide the population into different compartments between which people may progress. Different diseases are represented by different compartmental models (Schmidt, 1981; Sharomi and Gumel, 2011; Gao et al., 2016). The Susceptible-Infectious-Recovered (SIR) model, as one of the simplest and most classic compartmental models, characterizes the dynamic changes in each compartment using ordinary differential equations. There are three compartments in this model: susceptible ( $S$ ), infectious ( $I$ ), and recovered/deceased ( $R$ ). The number of individuals in each compartment varies over time. The deterministic SIR and its derivatives are widely used to predict infectious diseases like COVID-19 (Chen et al., 2020; Katul et al., 2020; Toda, 2020). Besides compartmental models, statistical learning techniques are also widely used in biomedical fields (Zheng et al., 2018,

2019; Hsieh and Zheng, 2019; Ganyani et al., 2020; Murray, 2020; You et al., 2020). For example, IHME team (Murray, 2020) employed a statistical model to predict the number of deaths, the demand of hospital beds, ICU beds and ventilators in a few months.

In this paper, we develop a time-dependent Susceptible-Exposed-Infectious-Recovered (SEIR) model with coefficients estimated by Least Absolute Shrinkage and Selection Operator (LASSO) regression. This model is inspired by the SIR model and takes the existence of incubation period (the time from exposure to development of symptoms) into consideration. The individuals who have been infected but are not yet infectious are labeled as exposed ( $E$ ). Instead of the constant parameters used in traditional SIR based models, we propose to model the dynamic with time-dependent parameters. Additionally, we extend our SEIR model to accommodate other crucial factors such as immunity, reinfection, and vaccination cases into account. With the epidemiology models, we aim at answering the following questions:

- What is the trajectory of transmission rate, incubation rate, and recovery rate?
- Has the inflection point been reached. If so, when?
- How does the reopen order affect the spread of the pandemic?
- How do reinfection and vaccination affect the pandemic?
- When will the mortality reach the peak?
- How many cases do we expect to have when the pandemic is over?

The remainder of the paper is organized as follows: we build the time-dependent SEIR model in section 2. Then we extend the model to include the vaccinated group as well as analyze the asymptotic stability of its disease-free equilibrium in section 3. To validate our model, we perform numerical analysis, prediction, and model simulation using national level data of the United States, and the state level data of two selected states, New York and North Dakota. The results are presented in section 4. Lastly, we conclude this paper in section 5.

## 2. THE TIME-DEPENDENT SEIR MODEL

Our proposed SEIR model with time-dependent parameters describes the transmission dynamic of an epidemic. It is assumed that there are totally four states in which an individual would experience: susceptible, exposed, infected, and recovered. In the susceptible state, the individual does not have the disease but can be infected by someone infectious through an effective contact. Once being infected, the individual moves to the exposed state. The exposed individual is not able to infect others until the incubation period is over. Eventually, the infected individual recovers from the disease. Altogether the four groups of individuals at different states compose the entire population and we denote the number of individuals in each group at time  $t$  by  $S(t)$ ,  $E(t)$ ,  $I(t)$ , and  $R(t)$ . In this model, a person is assumed to be immune to the virus after recovery and will not return to the susceptible state. Accordingly, the number of deaths caused by the disease is also counted in the recovered group  $R(t)$  since

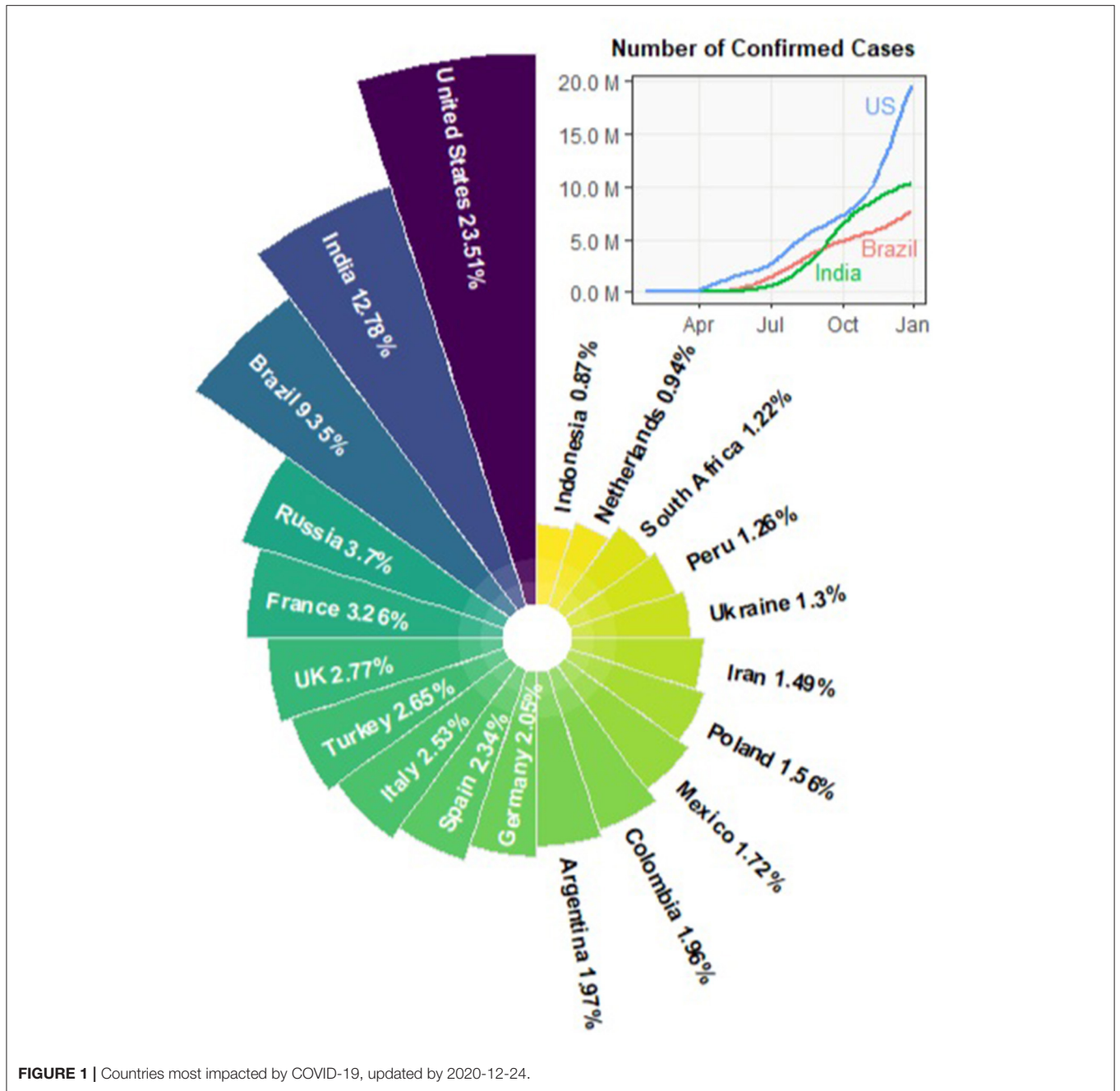


FIGURE 1 | Countries most impacted by COVID-19, updated by 2020-12-24.

neither of the recovered and dead has any more impact on the spread of the virus.

The differential equations that govern the trajectories of the four compartments are formulated as:

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

$$\frac{dE}{dt} = \frac{\beta_t S(t)I(t)}{N} - \sigma_t E(t), \tag{2}$$

$$\frac{dI}{dt} = \sigma_t E(t) - \gamma_t I(t), \tag{3}$$

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

with a constant total population  $N$ ,

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

and therefore, we have:

$$\frac{dS}{dt} + \frac{dE}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0. \tag{6}$$

Three time-dependent parameters, the transmission rate  $\beta_t$ , the transition rate  $\sigma_t$ , and the recovery rate  $\gamma_t$  are introduced in this model, which are all assumed to vary with respect to time. The descriptions and empirical ranges are listed in **Table 1**.

The proportion of susceptible and infected individuals in the population at time  $t$  are  $\frac{S(t)}{N}$  and  $\frac{I(t)}{N}$ , respectively. Given the transmission rate  $\beta_t$ , which describes the flow of susceptible becoming exposed to the virus, and the total population  $N$ , the number of newly exposed people is  $\frac{\beta_t S(t) I(t)}{N}$ . Later, the exposed individuals make the transition to the infected state at the transition rate  $\sigma_t$ , which is the inverse of the incubation period. The number of exposed individuals who complete the transition at time  $t$  is  $\sigma_t E(t)$ . Similarly, people recovered at time  $t$  is  $\gamma_t I(t)$ , given the recovery rate  $\gamma_t$ , which is the number of individuals recover from the infected state per person per time.

### 2.1. Discrete Time-Dependent SEIR Model

Since the COVID-19 case report is updated daily, we revise the differential Equations (1)–(4) into discrete time difference equations as follows:

$$S(t + 1) - S(t) = -\frac{\beta_t S(t) I(t)}{N}, \tag{7}$$

$$E(t + 1) - E(t) = \frac{\beta_t S(t) I(t)}{N} - \sigma_t E(t), \tag{8}$$

$$I(t + 1) - I(t) = \sigma_t E(t) - \gamma_t I(t), \tag{9}$$

$$R(t + 1) - R(t) = \gamma_t I(t), \tag{10}$$

with the four variables satisfying (5) and

$$\begin{aligned} S(t + 1) - S(t) + E(t + 1) - E(t) + \\ I(t + 1) - I(t) + R(t + 1) - R(t) = 0. \end{aligned} \tag{11}$$

Assuming historical data for a certain time period  $0 \leq t \leq T$  is available, i.e., we have  $\{S(t), E(t), I(t), R(t) | 0 \leq t \leq T\}$ . By deduction from (7) to (10), we can compute historical values of the parameter series  $\{\beta_t, \sigma_t, \gamma_t | 0 \leq t \leq T - 1\}$  using the following formulas:

$$\beta_t = \frac{N(E(t + 1) - E(t) + I(t + 1) - I(t) + R(t + 1) - R(t))}{S(t)I(t)}, \tag{12}$$

$$\sigma_t = \frac{I(t + 1) - I(t) + R(t + 1) - R(t)}{E(t)}, \tag{13}$$

$$\gamma_t = \frac{R(t + 1) - R(t)}{I(t)}. \tag{14}$$

Now predicting future values of the parameters  $\{\beta_t, \sigma_t, \gamma_t | t \geq T\}$  given historical values can be converted to a regression problem.

### 2.2. Tracking the Transmission Rate $\beta_t$ , Transition Rate $\sigma_t$ , and Recovery Rate $\gamma_t$

There are several approaches predicting future values of the time-dependent parameters. For instance, we can use linear models (e.g., linear regression), nonlinear methods (e.g., spline), or time series models (e.g., autoregressive model), etc. In this subsection, we fit the following LASSO regression models:

$$\hat{\beta}_{t+1} = a_0 + \sum_{i=1}^I a_i \beta_{t-i}, \tag{15}$$

$$\hat{\sigma}_{t+1} = b_0 + \sum_{j=1}^J a_j \sigma_{t-j}, \tag{16}$$

$$\hat{\gamma}_{t+1} = c_0 + \sum_{k=1}^K a_k \gamma_{t-k}, \tag{17}$$

where  $I, J$ , and  $K$  are the orders of the autoregressive process, and  $\{a_i | 0 \leq i \leq I\}$ ,  $\{b_j | 0 \leq j \leq J\}$  and  $\{c_k | 0 \leq k \leq K\}$  are the regression coefficients.

These coefficients are determined by minimizing the following loss functions, which are composed of the residual sums of squares (RSS) and regularization terms:

$$L(\beta) = \sum_{t=I+1}^{T-1} (\beta_t - a_0 - \sum_{i=1}^I a_i \beta_{t-i})^2 + \lambda_\beta \sum_{i=0}^I |a_i^2|, \tag{18}$$

$$L(\sigma) = \sum_{t=J+1}^{T-1} (\sigma_t - b_0 - \sum_{j=1}^J b_j \sigma_{t-j})^2 + \lambda_\sigma \sum_{j=0}^J |b_j^2|, \tag{19}$$

$$L(\gamma) = \sum_{t=K+1}^{T-1} (\gamma_t - c_0 - \sum_{k=1}^K c_k \gamma_{t-k})^2 + \lambda_\gamma \sum_{k=0}^K |c_k^2|, \tag{20}$$

$\lambda_\beta, \lambda_\sigma$ , and  $\lambda_\gamma$  are the regularization parameters deciding the penalty to the flexibility of model, and all regularization parameters can be optimized by cross-validation.

### 2.3. Estimating the Exposed $\hat{E}(t)$ , Infections $\hat{I}(t)$ , and Recovered $\hat{R}(t)$ Groups

Given the historical data  $\{S(t), E(t), I(t), R(t), 0 \leq t \leq T\}$ , we first compute the time-dependent parameter series  $\{\beta_t, \sigma_t, \gamma_t, 0 \leq t \leq T - 1\}$  introduced in section 2.1. Then we predict future values  $\{\hat{\beta}_t, \hat{\sigma}_t, \hat{\gamma}_t, t \geq T\}$  using the model built in section 2.2. According to (8), (9), (10), and (5), we can further predict the number of cases for the future as follows:

$$\hat{E}(t + 1) = \hat{E}(t) + \frac{\hat{\beta}_t \hat{S}(t) \hat{I}(t)}{N} - \hat{\sigma}_t \hat{E}(t), \quad t \geq T + 1, \tag{21}$$

**TABLE 1 |** Model parameters.

Parameter	Description	Empirical range	References
$\beta_t$	Transmission rate (effective contact rate) at a given time	0.5–1.5 day <sup>-1</sup>	Ngonghala et al., 2020; Read et al., 2020; Shen et al., 2020
$\sigma_t$	Transition rate from exposed to infections at a given time	$\frac{1}{5.1}$	Fairoza Amira et al., 2020; Ngonghala et al., 2020
$\gamma_t$	Recovery rate at a given time	$\frac{1}{10}$	Fairoza Amira et al., 2020; Ngonghala et al., 2020
$v_t$	Fraction of susceptible individuals vaccinated at a given time		
$w$	Fraction of infections gain immunity after recovery		

$$\hat{I}(t + 1) = \hat{I}(t) + \hat{\sigma}_t \hat{E}(t) - \hat{\gamma}_t \hat{I}(t), \quad t \geq T + 1, \quad (22)$$

$$\hat{R}(t + 1) = \hat{R}(t) + \hat{\gamma}_t \hat{I}(t), \quad t \geq T + 1, \quad (23)$$

$$\hat{S}(t + 1) = N - \hat{E}(t + 1) - \hat{I}(t + 1) - \hat{R}(t + 1), \quad t \geq T + 1, \quad (24)$$

Note that for the special case when estimating  $\{\hat{S}(t), \hat{E}(t), \hat{I}(t), \hat{R}(t) | t = T + 1\}$ , i.e., the numbers of cases at  $t = T + 1$ , we use the true values of  $\{S(t), E(t), I(t), R(t) | t = T\}$  instead of using the estimated values  $\{\hat{S}(t), \hat{E}(t), \hat{I}(t), \hat{R}(t) | t = T\}$  as in the formulas (21), (22), (23), and (24). The detailed steps of the entire procedure are summarized in Algorithm 1.

---

**Algorithm 1:** Tracking discrete time time-dependent SEIR model

---

**Input:**  $\{E(t), I(t), R(t) | 0 \leq t \leq T\}$ , regularization parameters  $\lambda_\beta, \lambda_\sigma$  and  $\lambda_\gamma$ , orders of autoregressive process  $I, J, K$ , prediction window  $t_w$ .

**Output:**  $\{\beta_t, \sigma_t, \gamma_t | 0 \leq t \leq T - 1\}$ ,  
 $\{\hat{\beta}_t, \hat{\sigma}_t, \hat{\gamma}_t | T \leq t \leq T + t_w - 1\}$ ,  
 $\{\hat{E}(t), \hat{I}(t), \hat{R}(t) | T + 1 \leq t \leq T + t_w\}$ .

Compute  $\{\beta_t, \sigma_t, \gamma_t | 0 \leq t \leq T - 1\}$  using (12), (13), and (14);

Train the LASSO regression models using  $\{\beta_t, \sigma_t, \gamma_t | 0 \leq t \leq T - 2\}$  as the predictors and  $\{\beta_{t_1}, \sigma_{t_2}, \gamma_{t_3} | I + 1 \leq t_1 \leq T - 1, J + 1 \leq t_2 \leq T - 1, K + 1 \leq t_3 \leq T - 1\}$  as the response;

**while**  $T \leq t \leq T + t_w - 1$  **do**

- Predict  $\hat{\beta}_t, \hat{\sigma}_t$  and  $\hat{\gamma}_t$  using (15), (16), and (17);
- Estimate  $\hat{E}(t + 1), \hat{I}(t + 1)$  and  $\hat{R}(t + 1)$  using (21), (22), and (23), respectively;

---

### 3. SEIR VARIATION CONSIDERING IMMUNITY, REINFECTION, AND VACCINATION

The human immune system protects the body against diseases with two parts. The first part, known as the innate

immune response, includes the release of chemicals that cause inflammation and white blood cells that can destroy infected cells. It is always ready to take actions as soon as any foreign invader is detected inside the body. However, this part is not specific to coronavirus. It will not learn and develop immunity to the virus. Instead, the second part: the adaptive immune response produces targeted antibodies that can stick to the virus and stop the spread to the body. The T cells<sup>1</sup> would attack the cells infected by the virus.

Existing research shows that most COVID-19 patients had an antibody response at 10 days or later after onset of symptoms (To et al., 2020). If the adaptive immune response is powerful enough, it could leave a lasting memory of the infection that will provide protection in the future. Other findings also suggest that strong responders (with higher antibody level) are significantly higher in severe patients, while it is unclear whether the asymptomatic or mildly symptomatic patients will develop sufficient adaptive immune response and gain immunity to the disease after recovery (Tan et al., 2020). In fact, there have been several reported cases of COVID reinfection in China, Hong kong, Belgium, the Netherlands, and the U.S. (Tan et al., 2020), and the reinfection case are indeed increasing. This implies the necessity of taking reinfection into consideration.

On the other hand, the worldwide endeavor to create a safe and effective COVID-19 vaccine is beginning to bear fruit. A wide variety of vaccines has already been authorized around the globe while many more remain in development. According to the U.S. CDC, as of December 13, 2020, the Pfizer-BioNTech COVID-19 vaccine has been authorized and large-scale (Phase 3) clinical trials are in progress or being planned for three other vaccines in the United States. Currently the supply of COVID-19 vaccine in the U.S. is limited, but it will increase in the upcoming weeks and months. Once large quantities are available, the increasingly large-scale vaccination will have a substantial impact on the pandemic.

#### 3.1. The Time-Dependent SEVIS Model

To take the factors of immunity, reinfection, and vaccination into account, we modify the proposed SEIR model by removing the recovered group  $R(t)$  and adding a vaccinated group  $V(t)$ , which represents the vaccinated individuals. In this susceptible, exposed, vaccinated, and infected modeling framework, the previous assumption for the SEIR model that an infected

<sup>1</sup>T cells are one of the important white blood cells of the immune system, and play a central role in the adaptive immune response.

individual will not become susceptible again after recovery is no longer employed. Instead, we assume that a fraction of the infected individuals gain immunity after recovery through producing antibodies while the rest return to the susceptible state. The former is counted in the  $V(t)$  group along with the vaccinated individuals since, epidemiologically speaking, both are immune to the virus and can no longer be infected. The new SEVIS model is governed by the following differential equations:

$$\frac{dS}{dt} = -\frac{\beta_t S(t)I(t)}{N} - \nu_t S + (1-w)\gamma_t I(t), \tag{25}$$

$$\frac{dE}{dt} = \frac{\beta_t S(t)I(t)}{N} - \sigma_t E(t), \tag{26}$$

$$\frac{dV}{dt} = \nu_t S + w\gamma_t I(t), \tag{27}$$

$$\frac{dI}{dt} = \sigma_t E(t) - \gamma_t I(t), \tag{28}$$

with a constant total population  $N$ ,

$$N = S(t) + E(t) + V(t) + I(t), \tag{29}$$

and therefore, we have:

$$\frac{dS}{dt} + \frac{dE}{dt} + \frac{dV}{dt} + \frac{dI}{dt} = 0. \tag{30}$$

The parameter settings of the transmission rate  $\beta_t$ , the transition rate  $\sigma_t$ , and the recovery rate  $\gamma_t$  remain the same as in the SEIR model. The vaccination rate  $\nu_t$  is low at the beginning of vaccine administration and gradually increasing as supply is growing.  $w \in [0, 1]$  is the fraction of infected cases that become immune after recovery. In addition, we assume it to be constant in this model. Hence, the number of infected individuals recover at time  $t$  is  $\gamma_t I(t)$ , and  $w\gamma_t I(t)$  join the  $V(t)$  group while  $(1-w)\gamma_t I(t)$  fail to gain immunity and return to the susceptible state  $S(t)$ .

### 3.2. Baseline Epidemiological Parameters

In previous studies, the transmission rate,  $\beta$  (as a constant), ranges from around 0.5 to 1.5 per person per day (Ngonghala et al., 2020; Read et al., 2020; Shen et al., 2020) and decreases as time goes. Based on existing literature, the incubation period (the time from exposure to development of symptoms) of COVID-19 and other coronaviruses ranges from 2 to 14 days. On average, symptoms show up in the newly exposed person about 5.1 days after contact (Fairoza Amira et al., 2020; Ngonghala et al., 2020). Thus, the transition rate, which is the inverse of the incubation period, is estimated to be  $\frac{1}{5.1}$ .

### 3.3. Basic Reproduction Number and Asymptotic Stability of Disease-Free Equilibrium

In this subsection we give the closed-form expression for the time-dependent basic reproduction number of the SEVIS model using the next generation operator method (Diekmann et al.,

1990; van den Driessche and Watmough, 2002). The basic reproduction number  $\mathcal{R}_0$  is defined as the average number of secondary infections caused by a single infectious individual who enters an entirely susceptible population. That actually is the special case where all parameters and compartments are at their initial state at time  $t = 0$ . Since we propose the parameters to be time-dependent in our model, we revise the basic reproduction number to a time-dependent version  $\mathcal{R}_t$  as well. When  $\mathcal{R}_t > 1$ , the infection will be able to start spreading in the population and develop into an epidemic. Generally speaking, it is more difficult to control the epidemic with the larger the value of the basic reproduction number.

Let  $X$  be the vector of infected classes and  $Y$  be the vector of uninfected classes. For the SEVIS model (25)–(28), we have:

$$X = \begin{bmatrix} E \\ I \end{bmatrix}, Y = \begin{bmatrix} S \\ V \end{bmatrix}.$$

Next we define the matrix of new infection terms  $\mathcal{F}$ , which only includes the flow from  $X$  to  $Y$ , and matrix of all other terms  $\mathcal{V}$ , which includes flows within  $X$  and flows leaving the system. For each compartment, in-flow in  $\mathcal{V}$  is negative and out-flow in  $\mathcal{V}$  is positive.

$$\mathcal{F} = \begin{bmatrix} \frac{\beta_t SI}{N} \\ 0 \end{bmatrix}, \mathcal{V} = \begin{bmatrix} \sigma_t E \\ -\sigma_t E + \gamma_t I \end{bmatrix}.$$

The next generation matrix is defined as  $FV^{-1}$  where:

$$F = \frac{\partial \mathcal{F}}{\partial X} \Big|_{DFE}, V = \frac{\partial \mathcal{V}}{\partial X} \Big|_{DFE}.$$

The disease-free equilibrium (DFE) of the SEVIS model is given by:  $(S^*, E^*, V^*, I^*) = (N, 0, 0, 0)$ , and we have

$$F = \begin{bmatrix} 0 & \beta_t \\ 0 & 0 \end{bmatrix}, V = \begin{bmatrix} \sigma_t & 0 \\ -\sigma_t & \gamma_t \end{bmatrix}.$$

Therefore, the next generation matrix is:

$$FV^{-1} = \begin{bmatrix} \beta_t/\gamma_t & \beta_t/\gamma_t \\ 0 & 0 \end{bmatrix}.$$

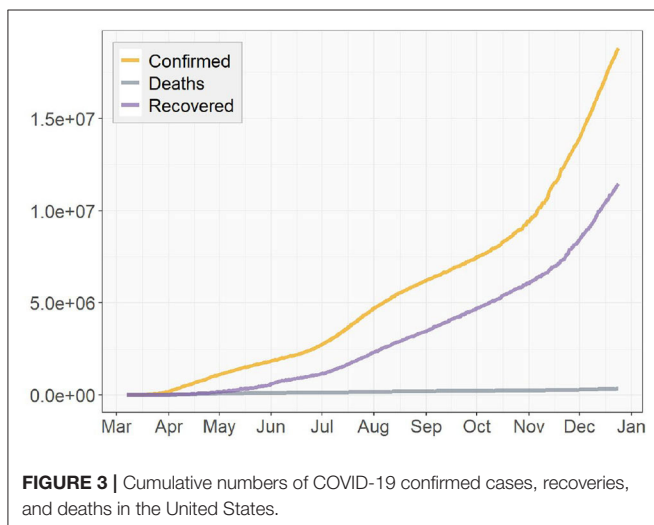
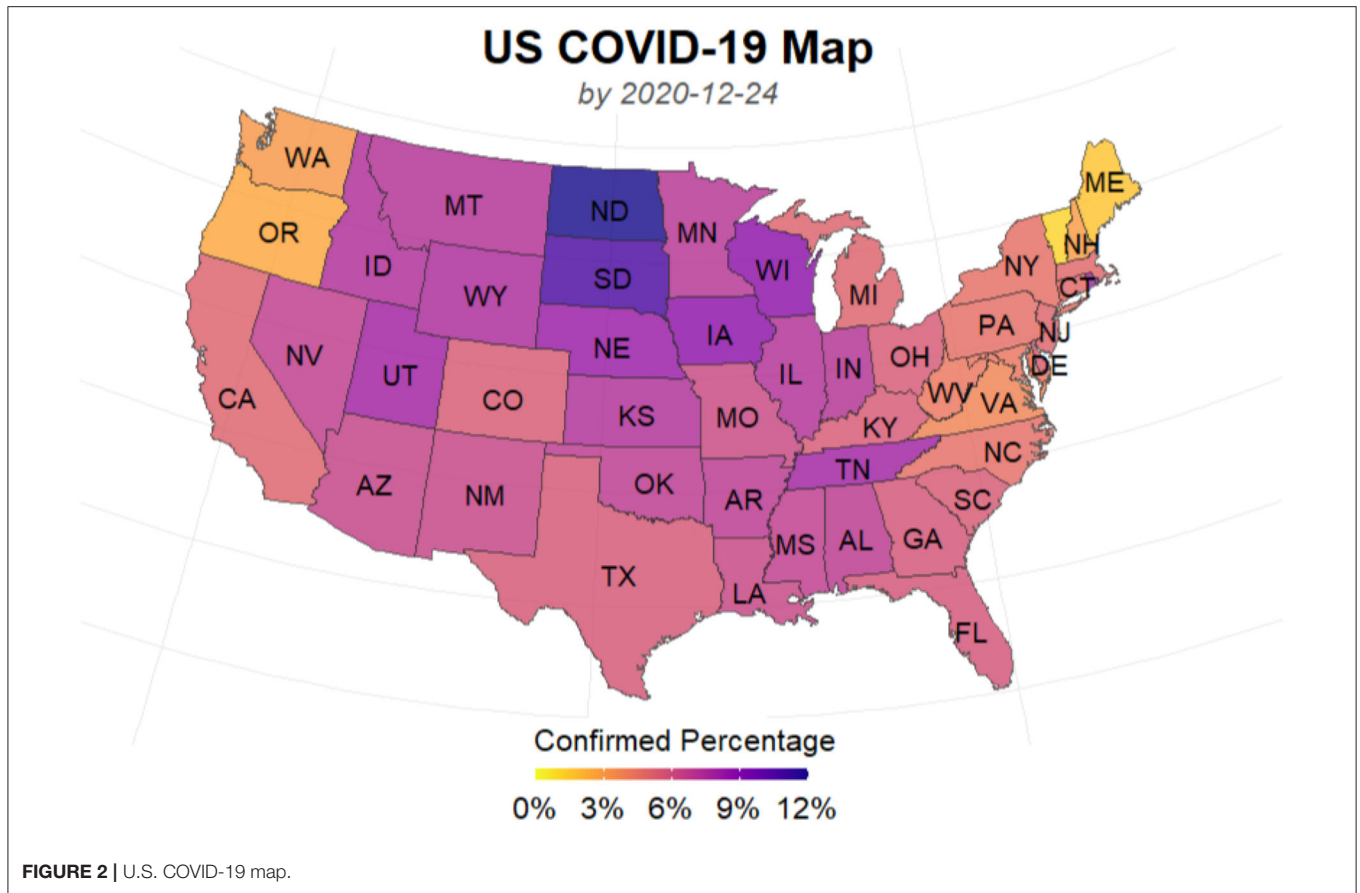
$\mathcal{R}_t$ , the basic reproduction number at time  $t$ , is given by the dominant eigenvalue of  $FV^{-1}$ :

$$\mathcal{R}_t = \frac{1}{2} \left( \frac{\beta_t}{\gamma_t} + \sqrt{\frac{\beta_t}{\gamma_t} \left( \frac{\beta_t}{\gamma_t} + 4 \right)} \right). \tag{31}$$

Similarly, we can obtain the same basic reproduction number for the time-dependent SEIR model. The DFE is locally asymptotically stable if  $\mathcal{R}_t < 1$ , and unstable if  $\mathcal{R}_t > 1$ .

## 4. NUMERICAL RESULTS, PREDICTIONS, AND SIMULATIONS

In this section, we will give the numeric results obtained by implementing Algorithm 1 on the national level data of the

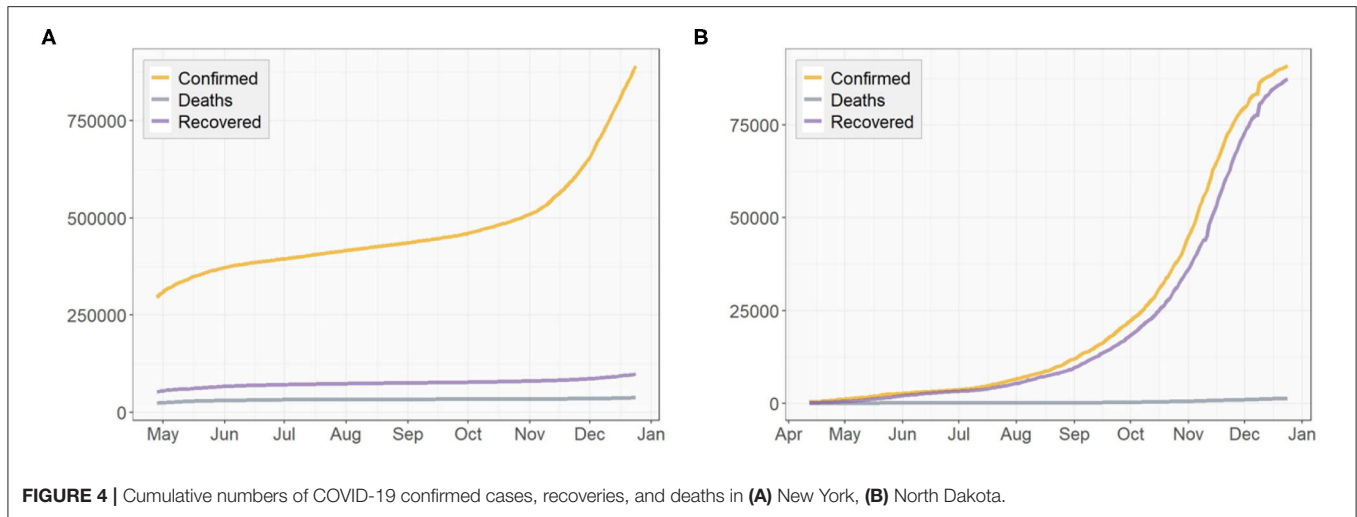


of the country. Though some new batches of hotspots have emerged across the country during the past months, the state of New York (NY) is still a region worth studying. On the other hand, as of December 24, a pack of northern states close to the Canada-US border have the highest percentages of cumulative confirmed cases in their populations as shown in **Figure 2**. The top one, North Dakota, has 11.94% of its population infected cumulatively, followed by South Dakota (10.69%), Wisconsin (8.61%), and some other nearby states. In this case, as a representative of this particular area, we take North Dakota (ND) as another example to illustrate our algorithm. We used the dataset that was collected from the COVID-19 data repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (Dong et al., 2020) and the **nCov2019** R package (Wu et al., 2020). The dataset contains time series of the numbers of confirmed cases, recovered cases and deaths up to December 24, 2020. The starting date of the training set used for model training varies according to the actual spread of the pandemic in each of the three regions: US, NY, and ND. For each region, a different start date of training set is chosen for model fitting according to the time when a relatively clear trend emerges.

**Figures 3, 4** presents the cumulative numbers of COVID-19 confirmed cases, recoveries and deaths reported in US, NY, and ND. The data starts at the beginning of the pandemic for US and ND, but it starts a while after the initial point for NY. The

United States (US) as well as the state level data of a few representative states.

In spring 2020, the New York Metropolitan Area experienced the largest COVID-19 outbreaks. As thousands of cases were being confirmed daily in New York, the state was the epicenter of the nation’s crisis and on a different scale than the rest



**FIGURE 4** | Cumulative numbers of COVID-19 confirmed cases, recoveries, and deaths in **(A)** New York, **(B)** North Dakota.

reason is that, back when the pandemic first started, a series of well-recorded numbers of recoveries were not available for many states, including NY. To obtain complete data on the three type of cases for computation, a cut-off is made. Therefore, the starting point of the data we collected for NY is about 2 months later than the actual date when the first case of COVID-19 was confirmed.

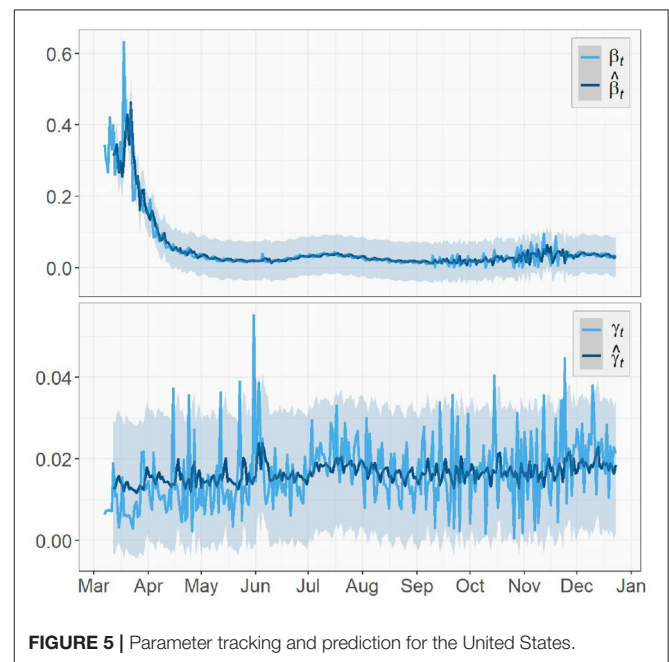
Due to the unavailability of the numbers of the exposed individuals  $E(t)$  in any of these regions, we substitute our model in section 2.1 with a simplified version as in Chen et al. (2020) that only includes the other three compartments  $S(t)$ ,  $I(t)$ , and  $R(t)$ . To validate our algorithm, we compare the prediction results with known data to see how well it performs, or how large the prediction errors are. Then we implement the algorithm again to predict how the COVID-19 pandemic will spread in the future.

At the end of this section, we simulate the long-term development of the pandemic based on the epidemiology models proposed in sections 2, 3 by constructing certain conditions and assigning assumed values to the parameters listed in **Table 1**. Based on the results, we discuss what they indicate as well as what differences we expect to see in reality compared to the simulation.

### 4.1. Parameter Tracking and Prediction

First we compute the true values of the transmission rate  $\beta_t$  and the recovery rate  $\gamma_t$  using (12), (13), and (14). Then starting from the sixth day in the parameter series, we take the value of a time-dependent parameter for each day as a subject for testing and a 5-day window before it as a corresponding observation used for training, i.e.,  $I, K = 5$  in section 2.2. By doing this, we construct the training and testing sets for model fitting. The R package **glmnet** is used to fit the LASSO regression models and choose the optimal values of  $\lambda_\beta$  and  $\lambda_\gamma$  that yield the minimum mean cross-validated errors.

**Figures 5, 6** depict the true values  $\{\beta_t, \gamma_t | 0 \leq t \leq T - 1\}$  and predicted values  $\{\hat{\beta}_{t_1}, \hat{\gamma}_{t_2} | I + 1 \leq t_1 \leq T - 1, K + 1 \leq t_2 \leq T - 1\}$

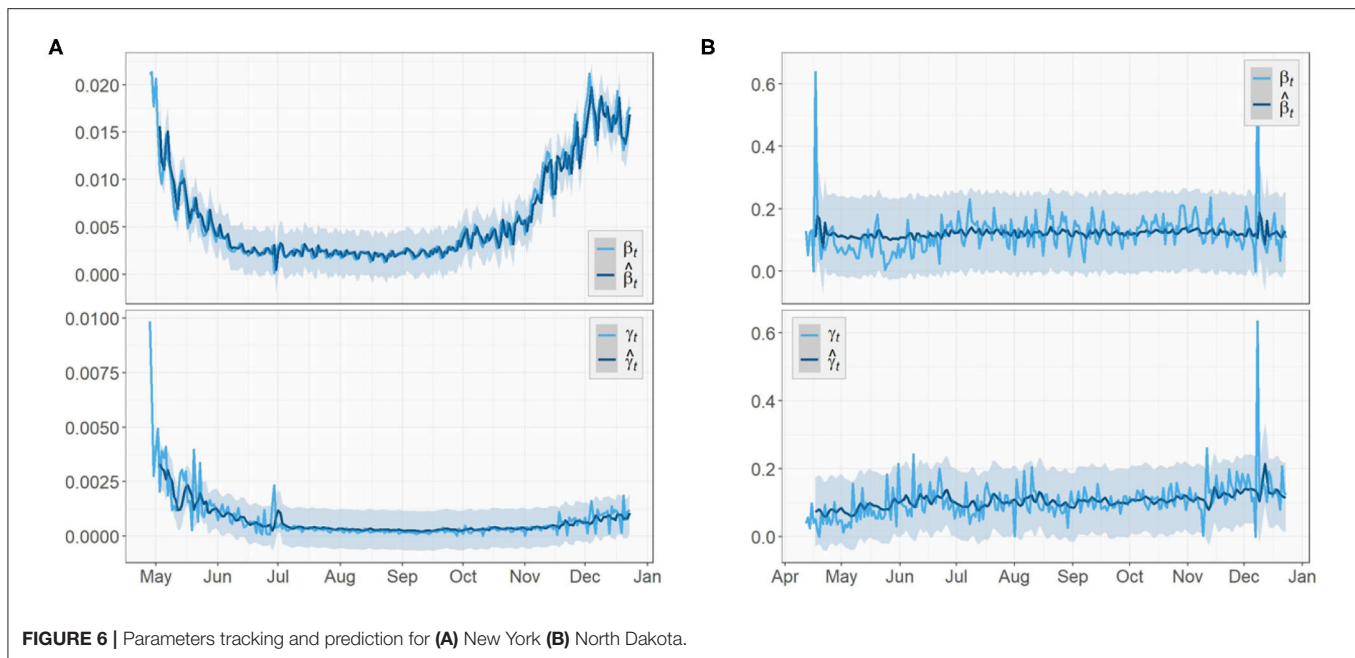


**FIGURE 5** | Parameter tracking and prediction for the United States.

of both the transmission rate and the recovery rate of US, NY, and ND, respectively. The 95% prediction intervals are shown as the gray bands above and below the curves.

For the U.S. case, there was a sharp decrease in the transmission rate from mid-March to May, just about 1 week after the spread of the virus started. This was an evidence that the social distancing measures and community lockdowns implemented across the country have effectively and significantly slowed down the spread of the pandemic. It kept decreasing for about a month before a surge appeared in July, which is possibly caused by the nationwide celebration of Independence Day. In the fall, starting from early September, the transmission rate slowly rose again with increasingly larger oscillations, which showed consistency with the surge in the fall that





pushed the total number of confirmed cases in US past 11M. This could be a result of a series of events prior to that (e.g., school opening, Halloween), and a prelude to the upcoming large gathering (e.g., Thanksgiving, Black Friday, Christmas). We expect this increase in the transmission rate to continue toward early 2021 and start to gradually decrease after the vaccination is administrated at a large scale in U.S. The recovery rate also had a slight increase around the same time in July but not as large as the one in the transmission rate. Overall, the recovery rate of U.S. is relatively steady and does not show any significant increasing or decreasing trend.

Similar to the US case, the transmission rate of NY started high and then reduced rapidly in the next few weeks. The trend maintained stationary for about 3 months until a rise appeared in late September and kept increasing toward the end. By December, the transmission rate is nearly as high as when it first started. The recovery rate of NY also had a large initial value followed by a 2-month-long decrease, but no clear trend was shown after a small spike at the beginning of July.

As for the ND case, the recovery rate started with a mild increase in the first 2 month. Later on, it remained steady just like the previous two regions. For the transmission rate, the overall trend is much more stationary compared to the results of US and NY and no significant change could be observed. However, the true values of the two parameters of ND have the greatest oscillations, i.e., the largest ranges of oscillations, among the three regions. Note the two unusually acute spikes in the transmission rate respectively in May and December and one in the recovery rate in December that deviate from the entire curves. In the absence of any pre or post trend, we consider these points as outliers in this paper and exclude them in model training.

## 4.2. Algorithm Validation and Relative Percentage Errors

In this section, we use the computed values of the parameters to estimate the three variables  $S(t)$ ,  $I(t)$ , and  $R(t)$  as in section 2.3. Instead of directly predicting future values for  $t > T$ , we use the historical data  $\{I(t), R(t) | T - t_w \leq t \leq T - 1\}$  and the predicted parameter series  $\{\hat{\beta}_t, \hat{\gamma}_t | T - t_w \leq t \leq T - 1\}$  to estimate the last  $t_w$  days of the entire period of time by which the data is covered, i.e., predict  $\{\hat{I}(t), \hat{R}(t) | T - t_w + 1 \leq t \leq T\}$ . Moreover, we also compare the proposed model with the classic SIR model with constant parameters by replacing the time-dependent parameter series with their means.

We evaluate the model performance using the relative percentage errors ( $RPE$ ) of the prediction for the infected group  $I(t)$  and the recovered group  $R(t)$  as follows:

$$RPE_I = \frac{|I(t) - \hat{I}(t)|}{I(t)}, \quad T - t_w + 1 \leq t \leq T, \quad (32)$$

$$RPE_R = \frac{|R(t) - \hat{R}(t)|}{R(t)}, \quad T - t_w + 1 \leq t \leq T. \quad (33)$$

To assess the predictions of the proposed method and compare with the classic SIR model, we compute the  $RPE$  series for the past week (i.e.,  $t_w = 7$ ) for the two models. The  $RPE$  series for US, NY, and ND are displayed in **Figures 7, 8** respectively, with their means summarized in the top-left corner of each figure. Using the proposed model with time-dependent parameters, the mean relative percentage errors for  $I(t)$  and  $R(t)$ , i.e.,  $RPE_I$  and  $RPE_R$ , are 2.35 and 0.39% for US, 0.2 and 0.2% for NY, and 4.67 and 0.09% for ND, respectively. Using the classic SIR model with

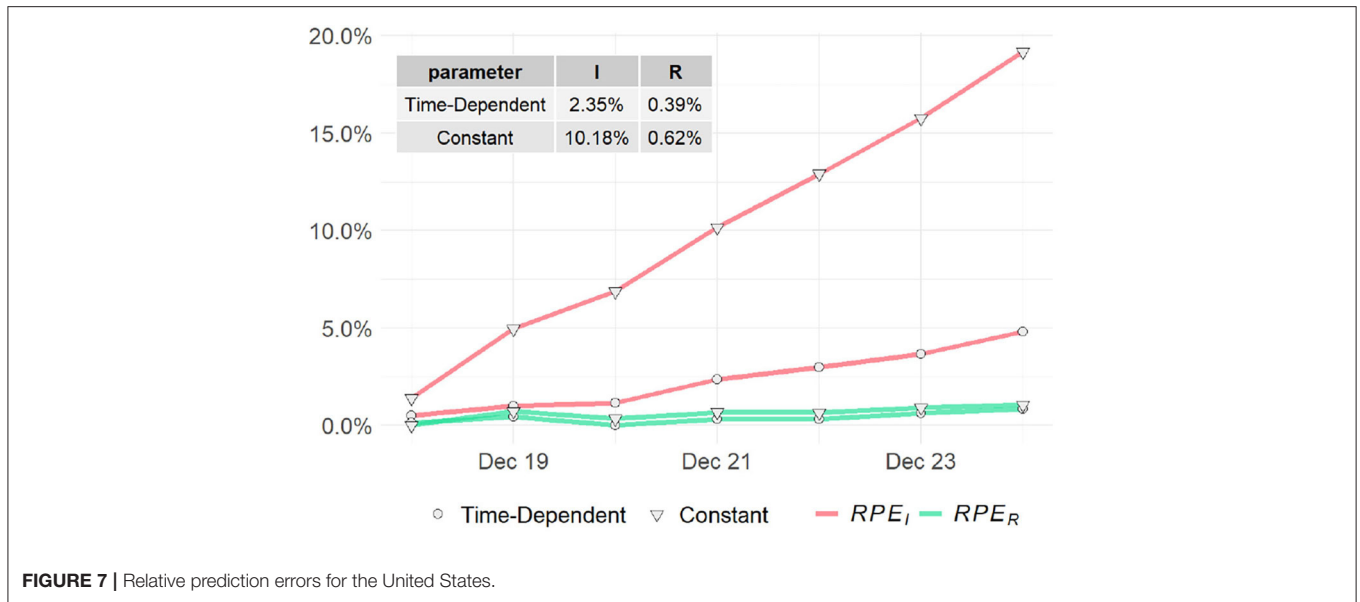


FIGURE 7 | Relative prediction errors for the United States.

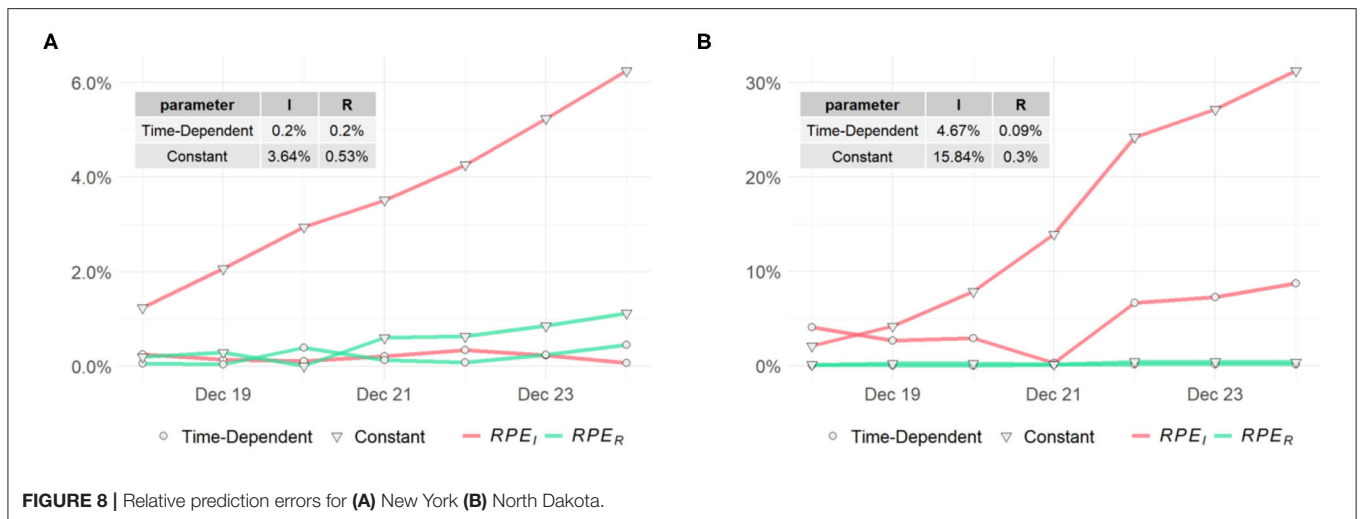


FIGURE 8 | Relative prediction errors for (A) New York (B) North Dakota.

constant parameters,  $RPE_I$  and  $RPE_R$  are 10.18 and 0.62% for US, 3.64 and 0.53% for NY, and 15.84 and 0.3% for ND, respectively. All errors are significantly larger than the former, which clearly shows the proposed time-dependent model yields better results in predicting the spread of the pandemic than the traditional SIR model with fixed parameters. Details of the model training and validation process are summarized in Table 2.

### 4.3. One-Day Prediction for $I(t)$ , $R(t)$ , and Basic Reproduction Numbers

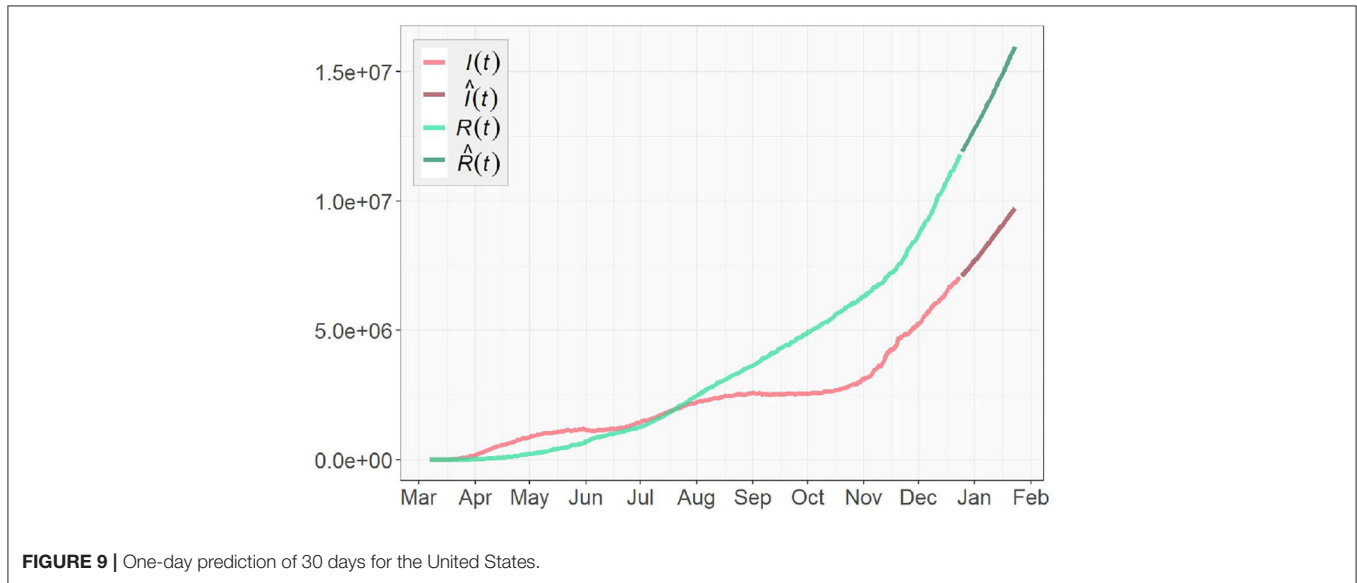
Next we implement Algorithm 1 to predict the number of infected  $I(t)$  and recovered individuals  $R(t)$  for the future  $\{\hat{I}(t), \hat{R}(t) | T + 1 \leq t \leq T + t_w\}$ . We reset the prediction window  $t_w$  to be 30, as we are to predict the spread of COVID-19 pandemic in the next 30 days after December 24, 2020. The

results of 1-day prediction for US, NY, and ND are shown in Figures 9, 10, respectively. For NY, the sharp increase in the infected group since November is predicted to continue toward the next year, due to the oscillatory rise in the transmission rate shown in Figure 6. On the other hand, the growth of the recovered group remains slow. For ND, the number of infected will stay low after the small surge was contained in November, while the rapid growth in the recovered group is expected to be continuous but might slow down. For US, the prediction shows that both curves will keep climbing at a high rate, which indicates that there will still be a long way to go before the pandemic finally ends. The prediction results are summarized in Table 3.

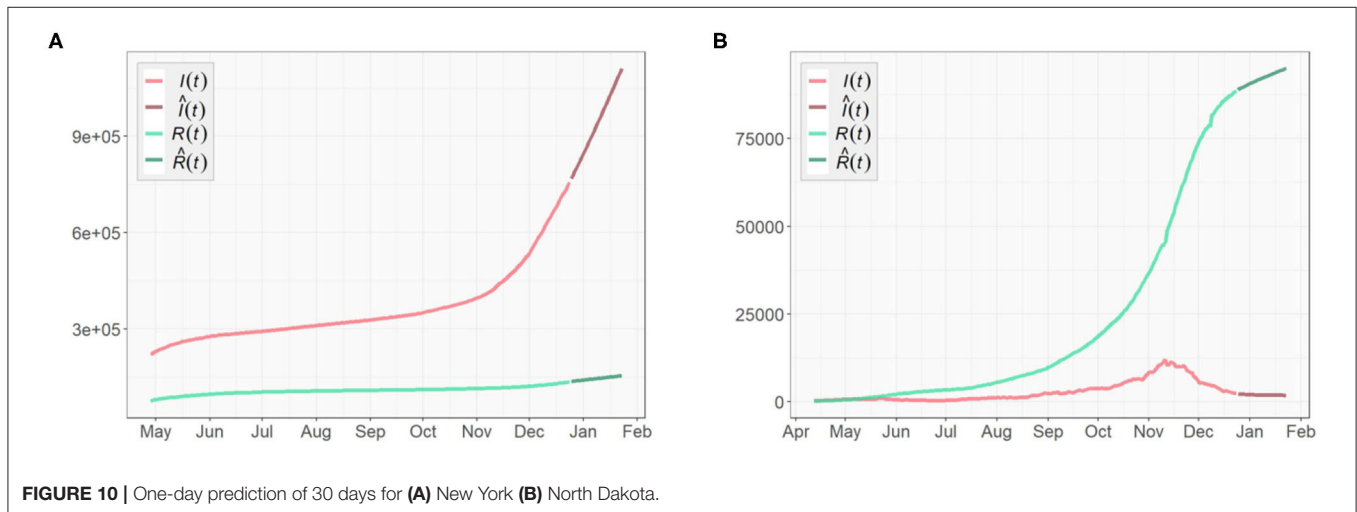
To assess the spread of COVID-19, we also obtain the 1-day prediction for the time-dependent basic reproduction number  $\mathcal{R}_t$  using (31). The results for the three regions

**TABLE 2** | Modeling training and validation.

Region	Start date of training data	Size of training set	Order	Prediction window $t_w$	Mean $RPE_I$ (%)	Mean $RPE_R$ (%)
United States	2020 – 03 – 07	287	5	7	2.35	0.39
New York	2020 – 04 – 28	235	5	7	0.2	0.2
North Dakota	2020 – 04 – 12	251	5	7	4.67	0.09



**FIGURE 9** | One-day prediction of 30 days for the United States.



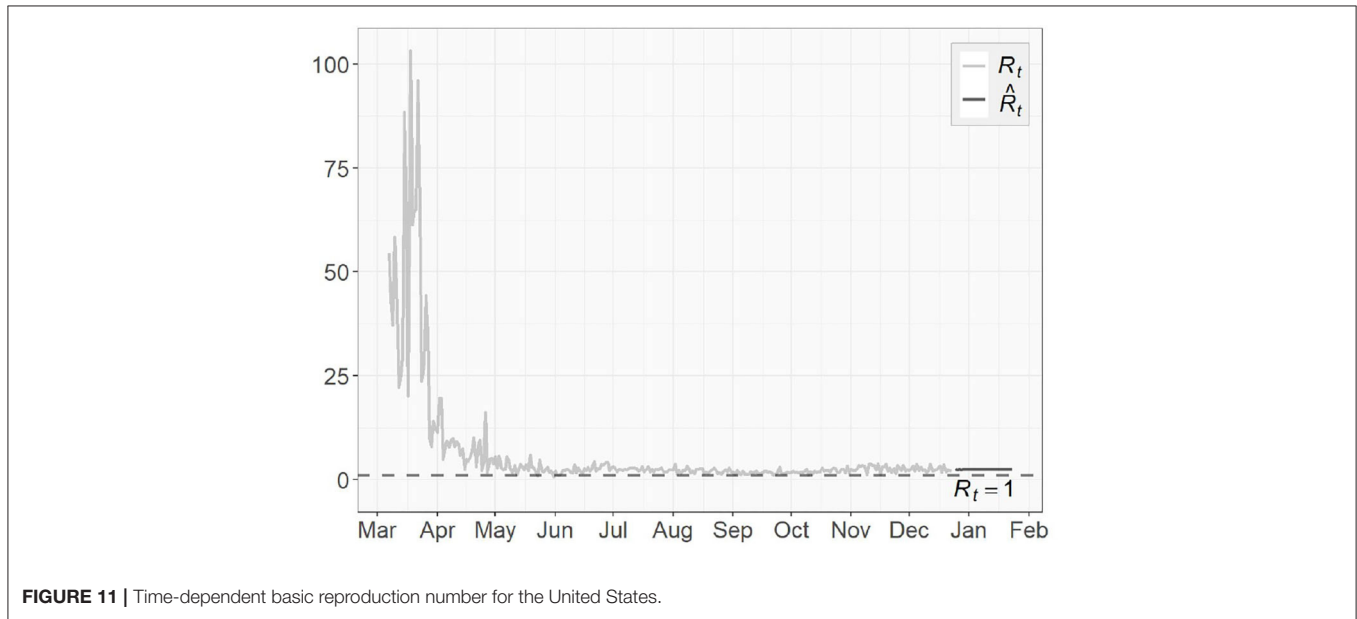
**FIGURE 10** | One-day prediction of 30 days for (A) New York (B) North Dakota.

are presented in **Figures 11, 12**, with horizontal lines representing  $\mathcal{R}_t = 1$ . As discussed in section 3.3, the virus will decline and gradually die out when  $\mathcal{R}_t < 1$ . Otherwise, it will continue to spread. According to the results shown in **Figures 11, 12**, only very few points fall below the horizontal line, while the majority lies above it. For NY, the surge in fall, 2020 and some scattered large values agree with the increasing trends in both the confirmed cases and the transmission rate we see in **Figures 4, 6**, respectively.

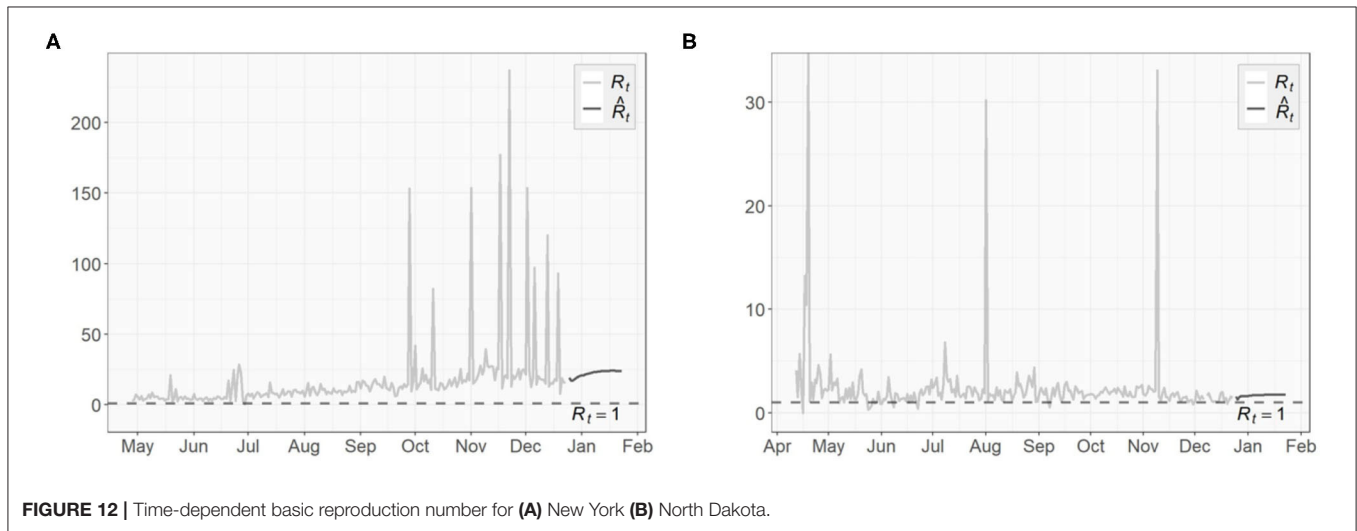
The basic reproduction numbers  $\mathcal{R}_t$  for each of the next 30 days are estimated to be  $>1$  for all three regions. The means of predicted values are found to be 2.48 for US, 22.28 for NY and 1.68 for ND, which suggests the inflection point, where  $\mathcal{R}_t$  stabilizes below 1 afterwards, has not been reached yet, especially for the NY case, where instead of having a decreasing trend, an increasing  $\mathcal{R}_t$  actually emerges over time. For US and ND, the curves gradually approaching the horizontal lines of  $\mathcal{R}_t < 1$  indicates that the measures taken to tackle the pandemic are taking effect, but at this point it is still too early to relax them.

**TABLE 3** | Prediction results.

Region	Total confirmed cases on last day	Prediction window $t_w$	$\hat{I}(t)$	$\hat{R}(t)$	Predicted total confirmed cases
United States	18,829,816	30	9,723,682	15,971,038	25,694,720
New York	891,270	30	1,111,117	153,277	1,264,394
North Dakota	90,947	30	1,760	95,016	96,776



**FIGURE 11** | Time-dependent basic reproduction number for the United States.



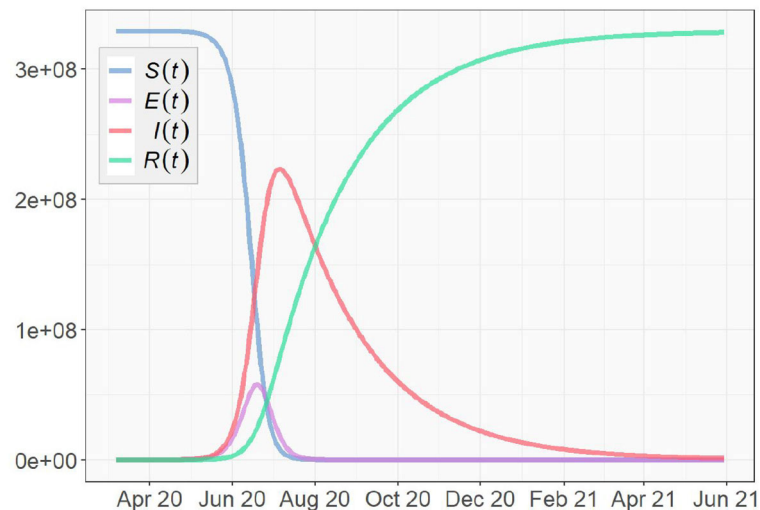
**FIGURE 12** | Time-dependent basic reproduction number for (A) New York (B) North Dakota.

### 4.4. Simulation Results for the SEIR and SEVIS Models

We also simulate the long-term development of the COVID-19 pandemic based on SEIR and SEVIS models. March 17, 2020, the first day in our US data, is chosen as the starting date of the pandemic in the simulations.

For the SEIR model, we set the transition rate to  $\sigma_t = \frac{1}{5.1}$  according to **Table 1**. To simulate as close to the reality

as possible, we set the transmission rate  $\beta_t$  and the recovery rate  $\gamma_t$  to the means of their true value series obtained in section 4.1. To construct the initial conditions of the system, we use the initial values  $I(0) = 311$  and  $R(0) = 27$  obtained from the data as well. In previous studies, the average Infected-Suspected ratio in China, one of the earliest hot spots of the global COVID-19 outbreak, was found to be 2.399 (e.g., Fairiza Amira et al., 2020). In this simulation, due to the lack of data



**FIGURE 13** | Simulation based on the SEIR Model for the U.S.

of the exposed group, we use the same ratio to initialize  $E(t)$ , i.e.,  $E(0) = \frac{1}{2.399}I(0) \approx 130$ . According to the U.S. and World Population Clock (United States Census Bureau, 2020), the U.S. population is  $N = 329,227,746$ . Using (5), we have:  $S(0) = N - E(0) - I(0) - R(0) \approx 329,227,278$ .

With the aforementioned parameter settings and initial conditions, we simulate the COVID-19 pandemic for the US. As shown in **Figure 13**, the number of infected people reaches a peak in early July, 2020, and the pandemic gradually dies out in summer 2021. It is important to note that the simulation is only theoretical and restricted by given conditions. These conditions can be dramatically different in reality. Moreover, no mitigation measure of any kind that can possibly prevent or limit the spread of the virus is considered in the simulation, such as wearing facial coverings, social distancing, community lockdowns, and work-from-home policies. Being free of the influences of such factors indicates that the pandemic might develop slower in the simulation than in reality. Since many states of the U.S. are following the strict guidelines set by CDC, the pandemic is highly likely to end earlier than the simulation result.

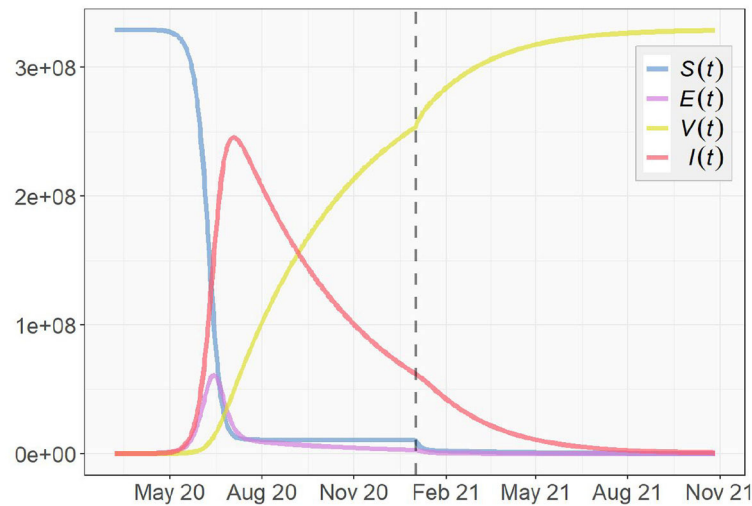
Next, we take immunity, reinfection and vaccination into account, and simulate the pandemic according to the SEVIS model proposed in section 3.1. The parameter settings of  $\beta_t$ ,  $\sigma_t$ , and  $\gamma_t$  remain the same as in the SEIR simulation. For the vaccination rate  $v_t$ , we clarify a starting date of vaccination  $t_v$ . Before the vaccination starts, i.e., for  $t < t_v$ ,  $v_t = 0$ . When  $t \geq t_v$ ,  $v_t$  becomes positive and based on the discussion in section 3, we assume  $v_t$  to start at a low value in reality and exponentially increase as time goes on. Here, we simplify this process by assuming the mean of  $\{v_t | t \geq t_v\}$  to be 1% per day and assigning it to  $v_t$ , and let the vaccination start on January 1, 2021. As for the last parameter  $w$  in **Table 1**, the fraction of infected cases that become immune after recovery is currently unknown. In this simulation, we assume  $w$  to be 0.5.

**Figure 14** shows the simulation result with the vertical dashed line representing  $t = t_v$ , (i.e., the first day of 2021). We notice that the trajectories obtained from the SEVIS model before the vaccination are nearly identical to the previous SEIR simulation. Once vaccination begins, the growth of the immunity group  $V(t)$  and the decrease of the infected group  $I(t)$  clearly accelerate. However, different from SEIR model which assumes no reinfection, the SEVIS model does allow reinfection, which leads to a longer time for the virus to die out. To speed up the process, we can employ a larger value for  $w$ , i.e., increased flows from  $I(t)$  to  $V(t)$  and reduced flows from  $I(t)$  to  $S(t)$ .

## 5. CONCLUSION

Considering the incubation period of COVID-19, we first proposed a time-dependent SEIR model with the time-dependent parameters estimated by LASSO regression. The proposed model is validated using the national level data (the United States) and state level data (New York and North Dakota). Overall, our proposed model outperforms the SIR model with smaller prediction errors. Furthermore, by taking immunity, reinfection, and vaccination into account, we proposed a time-dependent SEVIS model without assuming guaranteed immunity after recovery as in the SEIR model. Simulations are performed using the proposed two models to predict the spread of COVID-19 pandemic for the United States.

With the daily recorded data in the U.S., our algorithm predicts that the numbers of the infected and recovered individuals will keep increasing at a high rate in the short future. The total number of confirmed cases in the U.S. is estimated to reach close to 25.7M by late January, 2021, while North Dakota and New York will face 1.26 and 0.96M total confirmed cases, respectively. Given the historical transmission and recovery rate of the COVID-19, the simulation of SEVIS model predicts that



**FIGURE 14** | Simulation based on the SEVIS Model for the U.S.

the pandemic will die down in fall 2021, assuming the mean vaccination rate to be 1% per day and the probability of gaining immunity after recovery to be 50%. Note that this prediction is subject to change with more accurate parameters chosen according to the real data once vaccination starts.

In addition, it is crucial to understand that neither of the prediction and simulation takes any mitigation measures that can prevent or limit the growth of the pandemic into consideration, such as social distancing, facial covering, lockdown restrictions, and closing non-essential businesses. As a result, the end of the pandemic in reality is highly likely to come earlier than the numeric outcome. However, at this point the spread of the pandemic is still ongoing and has not been contained yet, as the time-dependent basic reproduction number for US is still steadily positive. Also, in some particular parts of US (e.g., New York), a new surge in the transmission rate was detected as the end of the year 2020 approaches. These all could serve as an alert that it is too early to relax the measures already implemented to tackle the pandemic. Fortunately, these measures have been proven effective by evidences. We expect them to continue taking effect over time and suggest the necessity of bring in more. Hopefully, with effort made by people around the world and the upcoming release of vaccine, we will be able to conquer this global crisis in no time.

## REFERENCES

- Chen, Y.-C., Lu, P.-E., Chang, C.-S., and Liu, T.-H. (2020). A time-dependent sir model for covid-19 with undetectable infected persons. *IEEE Trans. Netw. Sci. Eng.* 7, 3279–3294. doi: 10.1109/TNSE.2020.3024723
- Diekmann, O., Heesterbeek, J. A. P., and Metz, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio  $r_0$  in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* 28, 365–382. doi: 10.1007/BF00178324

Another limitation of the proposed time-dependent SEVIS model is that, it assume absolute immunity to the virus after vaccination, while in reality, the effectiveness of the vaccine is not 100% guaranteed. For example, as reported by the BBC news, a single dose of the Moderna vaccine can provide 80.2% protection. When a second dose is injected after a period of time, the effectiveness rise to 95.6%. In the future, we would like to extend the model by factoring in changing effectiveness at different stage of the vaccination.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

## AUTHOR CONTRIBUTIONS

JZ designed the study. YL collected data for analysis and interpreted the results and drafted the manuscript. JZ and YL analyzed the data and developed the models. LG, YZ, XC, and JZ revised the manuscript. All authors gave final approval for publication.

- Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track covid-19 in real time. *Lancet Infect. Dis.* 20, 533–534. doi: 10.1016/S1473-3099(20)30120-1
- Fairoza Amira, B. H., Cher, H., Hafeez, N., Dominic, L., Guanhua, L., Mohammad, S., et al. (2020). *Coronatracker: World-Wide Covid-19 Outbreak Data Analysis and Prediction*. Bulletin of the World Health Organization.
- Ganyani, T., Kremer, C., Chen, D., Torneri, A., Faes, C., Wallinga, J., et al. (2020). Estimating the generation interval for covid-19 based on symptom onset data, March 2020. *Eurosurveillance* 25:2000257. doi: 10.2807/1560-7917.ES.2020.25.17.2000257

- Gao, D., Porco, T. C., and Ruan, S. (2016). Coinfection dynamics of two diseases in a single host population. *J. Math. Anal. Appl.* 442, 171–188. doi: 10.1016/j.jmaa.2016.04.039
- Hsieh, F., and Zheng, J. (2019). Unraveling pattern-based mechanics defining self-organized recurrent behaviors in a complex system: a zebrafish's calcium brain-wide imaging example. *Front. Appl. Math. Stat.* 5:13. doi: 10.3389/fams.2019.00013
- Katul, G. G., Mrad, A., Bonetti, S., Manoli, G., and Parolari, A. J. (2020). Global convergence of COVID-19 basic reproduction number and estimation from early-time sir dynamics. *PLoS ONE* 15:e239800. doi: 10.1371/journal.pone.0239800
- Murray, C. J. (2020). Forecasting COVID-19 impact on hospital bed-days, icu-days, ventilator-days and deaths by us state in the next 4 months. *medRxiv [Preprint]*. doi: 10.1101/2020.03.27.20043752
- Ngonghala, C. N., Iboi, E., Eikenberry, S., Scotch, M., MacIntyre, C. R., Bonds, M. H., et al. (2020). Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel coronavirus. *Bellman Prize Math. Biosci.* 325:108364. doi: 10.1016/j.mbs.2020.108364
- Read, J. M., Bridgen, J. R., Cummings, D. A., Ho, A., and Jewell, C. P. (2020). Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. *medRxiv [Preprint]*. doi: 10.1101/2020.01.23.20018549
- Schmidt, W. (1981). Eisen, M.: Mathematical models in cell biology and cancer chemotherapy. Lecture notes in biomathematics, vol. 30. Springer-Verlag, Berlin-Heidelberg-New York 1979. IX, 431 s., 70 abb., 17 tab., DM 39,-. *Biometr. J.* 23, 519–520. doi: 10.1002/bimj.4710230517
- Sharomi, O., and Gumel, A. (2011). Dynamical analysis of a sex-structured chlamydia trachomatis transmission model with time delay. *Nonlin. Anal. Real World Appl.* 12, 837–866. doi: 10.1016/j.nonrwa.2010.08.010
- Shen, M., Peng, Z., Xiao, Y., and Zhang, L. (2020). Modelling the epidemic trend of the 2019 novel coronavirus outbreak in china. *Innovation* 1:100048. doi: 10.1016/j.xinn.2020.100048
- Tan, W., Lu, Y., Zhang, J., Wang, J., Dan, Y., Tan, Z., et al. (2020). Viral kinetics and antibody responses in patients with covid-19. *medRxiv [Preprint]*. doi: 10.1101/2020.03.24.20042382
- To, K. K. W., Tsang, O. T. Y., Leung, W. S., Tam, A. R., Wu, T. C., Lung, D. C., et al. (2020). Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by sars-CoV-2: an observational cohort study. *Lancet Infect. Dis.* 20, 565–574. doi: 10.1016/S1473-3099(20)30196-1
- Toda, A. A. (2020). Susceptible-infected-recovered (sir) dynamics of COVID-19 and economic impact. *arXiv [Preprint] arXiv:2003.11221*.
- United States Census Bureau (2020). *U.S. and World Population Clock*. United States Census Bureau. Available online at: <https://www.census.gov/> (accessed December 24, 2020).
- van den Driessche, P., and Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Bellman Prize Math. Biosci.* 180, 29–48. doi: 10.1016/S0025-5564(02)00108-6
- Wu, T., Ge, X., Yu, G., and Hu, E. (2020). Open-source analytics tools for studying the COVID-19 coronavirus outbreak. *medRxiv [Preprint]*. doi: 10.1101/2020.02.25.20027433
- You, C., Deng, Y., Hu, W., Sun, J., Lin, Q., Zhou, F., et al. (2020). Estimation of the time-varying reproduction number of COVID-19 outbreak in china. *Int. J. Hyg. Environ. Health* 228:113555. doi: 10.1016/j.ijheh.2020.113555
- Zheng, J., Fushing, H., and Ge, L. (2019). A data-driven approach to predict and classify epileptic seizures from brain-wide calcium imaging video data. *IEEE/ACM Trans. Comput. Biol. Bioinformatics* 17, 1858–1870. doi: 10.1109/TCBB.2019.2895077
- Zheng, J., Liang, M., Ekstrom, A. D., Ge, L., Yu, W., and Hsieh, F. (2018). “On association study of scalp EEG data channels under different circumstances,” in *International Conference on Wireless Algorithms, Systems, and Applications* (New York, NY), 683–695. doi: 10.1007/978-3-319-94268-1\_56

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Li, Ge, Zhou, Cao and Zheng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.