# A Data-driven Approach for Assessing the Impact of Re-opening Order and Non-pharmaceutical Interventions on the Spread of COVID-19 using Time-dependent SEIR Model


Fauziya Ado Yakasai*
Auburn University
Auburn, Alabama
fay0002@auburn.edu

Ismail Abdulrashid*
Sacred Heart University
Fairfield, Connecticut
abdulrashidi@sacredheart.edu

Yuexin Li
Auburn University
Auburn, Alabama
yzl0233@auburn.edu

Nedret Billor
Auburn University
Auburn, Alabama
billone@auburn.edu

Yang Zhou
Auburn University
Auburn, Alabama
yangzhou@auburn.edu

Jingyi Zheng†
Auburn University
Auburn, Alabama
jingyi.zheng@auburn.edu



## ABSTRACT

The outbreak of COVID-19, caused by a coronavirus called SARS-CoV-2 virus, has been declared a pandemic by the World Health Organization (WHO) in March and rapidly spread to over 100 countries. As of May 20, there are over $4.9M$ cumulative confirmed cases with $324K$ deaths worldwide. To mathematically model the spread of the COVID-19 pandemic, we propose a time-dependent SEIR model considering the incubation period and both symptomatic and asymptomatic infectious cases. Unlike the classic SIR based models with constant parameters, our dynamic model not only predicts the number of cases in each state, but also monitors the trajectory of the changing rate associated with each state, such as transmission rate $\beta(t)$, recover rate $\gamma(t)$, the basic reproduction number $R_0(t)$ and so on. Tracking these changing rates, we observe the obvious decreasing in the transmission rate after the authority announced a series of orders, such as school closure and stay-at-home, to prevent the spread of the pandemic. As of May 20, most states lifted the restrictions but it may come with a surge of new infections as we see some states already showed the increasing trend in the confirmed cases. Using our time-dependent model, people can track and predict $\beta(t)$, $R_0(t)$ and other critical rates in order to timely monitor the pandemic. To validate our model, we use both the national level data, e.g. China, USA and Italy, and state level data, e.g. Iowa and Texas. Our one-day relative prediction error of the number of confirmed cases for the U.S. is around 0.95%. We also perform simulation analysis of the model considering both symptomatic and asymptomatic infections.



## KEYWORDS

COVID-19, epidemiology, dynamic modeling, time-dependent SEIR model, asymptomatic infections

**ACM Reference Format:**
Fauziya Ado Yakasai, Ismail Abdulrashid, Yuexin Li, Nedret Billor, Yang Zhou, and Jingyi Zheng. 2020. A Data-driven Approach for Assessing the Impact of Re-opening Order and Non-pharmaceutical Interventions on the Spread of COVID-19 using Time-dependent SEIR Model. In *Proceedings of 19th International Workshop on Data Mining in Bioinformatics (BIOKDD 2020).* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn


## 1 INTRODUCTION

On March 11, 2020, the World Health Organization (WHO) made the assessment that the outbreak of the novel coronavirus (COVID-19) can be characterized as a pandemic. The outbreak was first declared a Public Health Emergency of International Concern on 30 January 2020 after 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 one month later, the number increased almost 8-fold to 80, 134 and the COVID-19 cases gradually showed up in other countries. Starting March 2020, the outbreak spread to more than 100 countries. Within a few months, the pandemic led to almost 5 million confirmed cases and over $300K$ deaths worldwide.

As of May 20, there are $4.9M$ confirmed cases with $324K$ deaths and $1.71M$ recovery. Out of the $4.9M$, we compare the cumulative confirmed cases of each country, and summarize what percentage each country contributes to the confirmed cases worldwide. Figure 1 shows the top 20 countries most impacted by the COVID-19 pandemic. The percentage of cases in each country is also labeled. As of May 16, there are four countries, the United States, Spain, Russia, and United Kingdom, share more than 5% of the global cases. The trajectories of the confirmed cases of the four countries are also displayed in the top right corner of Figure 1.

The COVID-19 pandemic not only causes mortality and heavy stress on health systems, but also impacts people's daily life and the society. To fight the COVID-19 pandemic, governments across the globe have schools closed, travel and business restricted. The vast majority of states in the United States had stay-at-home order

*Both authors contributed equally to this research.
†Correspondence


Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).


*BIOKDD 2020, San Diego, CA,*


© 2020 Copyright held by the owner/author(s).


ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

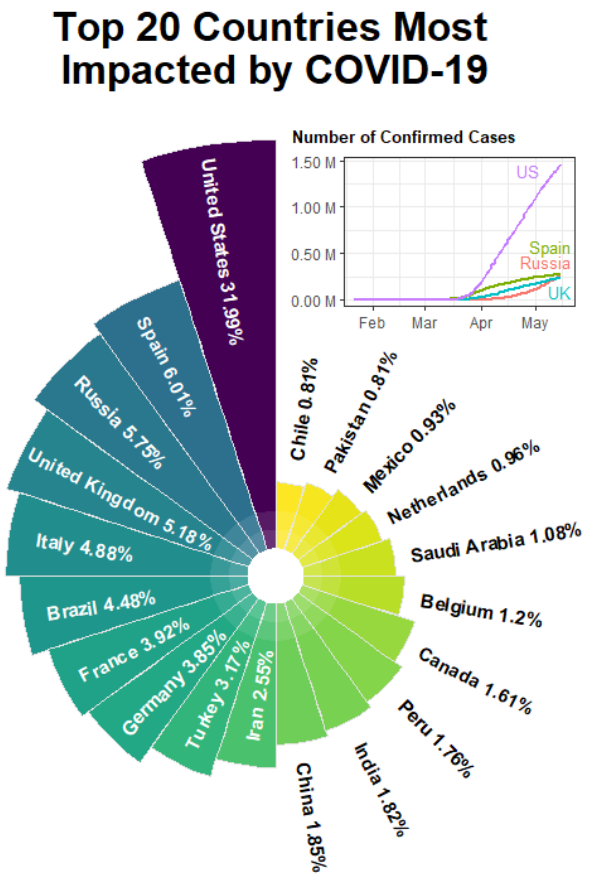


**Figure 1: Countries impacted by COVID-19**

officially placed in order to reduce the spread of the pandemic by minimizing close contact between people. Residents stayed at home, all non-essential businesses were closed and their employees worked from home only. The impact on the society comes later with the rising unemployment, psychological well-being of individuals and significant behavioral changes causing social disturbances.

As of May 20, almost all states have lifted the stay-at-home order and loosened their restrictions to allow some businesses to reopen. However, we notice that the number of daily confirmed cases displays a significant 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. Therefore, it is crucial to build a model to characterize and further predict the spread of the COVID-19 pandemic.

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 (see e.g., [2], [3], [9], [21], [22]). The SIR model is one of the classic and simple compartment models which is used to model a disease without an incubation period. There are three compartments in the model: susceptible (S), infectious (I), and recovered/deceased (R). The number in each compartment may vary over time. And SIR model characterizes the dynamic changes in each compartment using ordinary differential equations with details described in section 2. The deterministic SIR and derivatives are widely used to predict the infectious deceases like COVID-19 [4, 11, 26]. Besides compartmental models, statistical learning techniques are also widely used in biomedical field [8, 10, 16, 29–31]. IHME team [16] employs 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 propose to develop a time-dependent SEIR model with coefficients estimated by ridge regression. This model is obtained from SIR model and considers the existence of incubation period. Individuals who have been infected but are not yet infectious are labeled as exposed (E). Instead of using constant parameters in SIR based models, we propose to model the parameters as time-dependent. Additionally, we extend our model to include both symptomatic and asymptomatic infectious cases. With the proposed time-dependent SEIR model, we aim at answering the following questions:

- What is the trajectory of transmission rate, incubation rate and recovery rate?
- When is the inflection point?
- How does the reopen order affect the spread of the pandemic?
- How do the asymptomatic infections 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 and extend the model to include symptomatic and asymptomatic infections in Section 3. We analyze the asymptotic stability of the disease-free equilibrium in Section 4. To valid our model, we perform numerical analysis using national level data, state level data and simulations. The numerical results are summarized in Section 5. We conclude this paper in Section 6.

## 2 THE TIME-DEPENDENT SEIR MODEL

The traditional susceptible-infected-recovered (SIR) model, includes three states: the susceptible state,the infected state, and the recovered state. An individual in the susceptible state is the one who does not have the disease at time t yet, but may be infected if one is in contact with a person infected with the disease. The infected state refers to an individual who has a disease at time t and may infect a susceptible individual potentially (if they come into contact with each other). The recovered state refers to an individual who is either recovered or dead from the disease and is no longer contagious at time t. Also, a recovered individual is assumed not to be back to the susceptible state anymore. The reason why deaths is counted in the recovered state is that, from an epidemiological point of view, recovered and deaths both mean no more impact on the spread of the disease. As such, they can be effectively eliminated from the potential host of the disease. We use $S(t), I(t)$ and $R(t)$ to denote

the numbers of susceptible, infections, and recovered/death at time $t$. The resulting ODE model describing the transmission dynamics of COVID-19 is given by the following deterministic system of nonlinear differential equations:

$$\frac{dS}{dt} = -\frac{\beta(t)S(t)I(t)}{n}, \quad (1)$$

$$\frac{dI}{dt} = \frac{\beta(t)S(t)I(t)}{n} - \gamma I(t), \quad (2)$$

$$\frac{dR}{dt} = \gamma I(t), \quad (3)$$

where

$$S(t) + I(t) + R(t) = N$$

and $N$ is the population. Health care professionals have declared that individuals who have been exposed to COVID-19 show symptoms about 5 days on average. Considering the existence of latent period prior to the onset of infectiousness, we add one more compartment, exposed (E), in the dynamic model. We also assume the parameters to be time dependent. In particular, let $\sigma(t)$ be the rate at which individuals leave the E state (incubation rate), $\beta(t)$ be the transmission rate and $\gamma(t)$ be the recovering rate.

The SEIR model with time dependent parameters are formulated as follows:

$$\frac{dS}{dt} = \frac{-\beta(t)S(t)I(t)}{N}, \quad (4)$$

$$\frac{dE}{dt} = \frac{\beta(t)S(t)I(t)}{N} - \sigma(t)E(t), \quad (5)$$

$$\frac{dI}{dt} = \sigma(t)E(t) - \gamma(t)I(t), \quad (6)$$

$$\frac{dR}{dt} = \gamma(t)I(t), \quad (7)$$

given the total population $N$,

$$S(t) + E(t) + I(t) + R(t) = N \quad (8)$$

For a randomly chosen individual in the population, the probability of being a susceptible is $\frac{S(t)}{N}$. $\beta(t)$ is the average number of contacts per person per time, then the average number of contacts between susceptible and infection is $\frac{\beta(t)S(t)}{N}$. There are totally $I(t)$ infections, therefore the number of people newly exposed to the disease at time $t$ is $\frac{\beta(t)S(t)I(t)}{N}$. $\sigma(t)$ is the rate at which individual transition from the exposed state to the infectious state (i.e. incubation rate). As every individual in the infected state recover with rate $\gamma(t)$, there are (on average) $\gamma(t)I(t)$ people recovered at time $t$.

## 2.1 Discrete Time-Dependent SEIR Model

In this subsection, we rewrite the differential equations in (4)–(7) into discrete time difference equations since the COVID-19 data is updated in days.

$$S(t+1) - S(t) = \frac{-\beta(t)S(t)I(t)}{N}, \quad (9)$$

$$E(t+1) - E(t) = \frac{\beta(t)S(t)I(t)}{N} - \sigma(t)E(t), \quad (10)$$

$$I(t+1) - I(t) = \sigma(t)E(t) - \gamma(t)I(t), \quad (11)$$

$$R(t+1) - R(t) = \gamma(t)I(t), \quad (12)$$

with the variables $S(t), E(t), I(t)$ and $R(t)$ satisfying (8). Assuming most of the population are in the susceptible state i.e $\{S(t) \approx N, t \geq 0\}$. By simplifying (9)–(12), we get $\beta(t)$, $\sigma(t)$, $\gamma(t)$ of each day as follows:

$$\beta(t) = \frac{E(t+1) - E(t) + I(t+1) - I(t) + R(t+1) - R(t)}{I(t)} \quad (13)$$

$$\sigma(t) = \frac{I(t+1) - I(t) + R(t+1) - R(t)}{E(t)}. \quad (14)$$

$$\gamma(t) = \frac{R(t+1) - R(t)}{I(t)}. \quad (15)$$

Assuming historic data is available for each state during a time period. For instance, we have $\{S(t), E(t), I(t), R(t), 0 \leq t \leq T\}$ before time $T$. Then we can compute the rates series $\{\beta(t), \sigma(t), \gamma(t)\}$ using (13), (14), and (15) for the time period $0 \leq t \leq T-1$. Now the problem can be formulated as a regression problem, i.e. given $\{\beta(0), \beta(1), \ldots, \beta(T-1)\}$, how to predict the future $\beta(t)$, and the same for $\sigma(t)$ and $\gamma(t)$.

## 2.2 Tracking the Transmission Rate $\beta(t)$, Incubation Rate $\sigma(t)$ and Recovery Rate $\gamma(t)$

Given $\{\beta(0), \beta(1), \ldots, \beta(T-1)\}$, there are actually many ways to predict $\beta(t)$ for $t \geq T$. For instance, we can use moving average filter, autoregressive model and other time series models, nonlinear methods like spline, etc. Here we choose to build linear model with regularization as follows.

$$\min_{a_j} \sum_{t=J}^{T-2} (\beta(t) - a_0 - \sum_{j=1}^{J} a_j \beta(t-j))^2 + \lambda_1 \sum_{j=0}^{J} a_j^2 \quad (16)$$

$$\min_{b_k} \sum_{t=K}^{T-2} (\sigma(t) - b_0 + \sum_{k=1}^{K} b_k \sigma(t-k))^2 + \lambda_2 \sum_{k=0}^{K} b_k^2, \quad (17)$$

$$\min_{c_l} \sum_{t=L}^{T-2} (\gamma(t) - c_0 + \sum_{l=1}^{L} c_l \gamma(t-l))^2 + \lambda_3 \sum_{l=0}^{L} c_l^2, \quad (18)$$

where $\lambda_1, \lambda_2$ and $\lambda_3$ are the regularization parameters and $J, K$ and $L$ are the orders of the autoregressive models ($0 < J, K, L < T-2$), $a_j, j = 0, 1, ..., J$, $b_k, k = 0, 1, ..., K$ and $c_l, l = 0, 1, ..., L$ are the coefficients. The coefficients and tuning parameters can be estimated using cross validation.

**Table 1: Model parameters with brief description, likely ranges based on modeling and clinical studies, and default values**

| Parameter | Description | Likely range | References |
|---|---|---|---|
| $\beta$ | Transmission rate (infectious contact rate) | 0.5–1.5 day$^{-1}$ | [17], [20], [23] |
| $\sigma$ | Transition from exposed to infectious | $\frac{1}{5.1}$ | [5] [17] |
| $\eta$ | Infectiousness factor for asymptomatic carriers | 0.4 - 0.6 | [7], [13] |
| $\alpha$ | Fraction of infections that become symptomatic | 0.15 - 0.7 | [7], [13], [15], [27] |
| $\gamma$ | Recovery rate | $\frac{1}{10}$ | [17] [5] |
| $\gamma_A$ | Recovery rate for Asymptomatic | 1/14 - 1/3 day$^{-1}$ | [25], [32] |
| $\gamma_I$ | Recovery rate for symptomatic | 1/30 - 1/3 day$^{-1}$ | [25], [32] |

### 2.3 Predicting the Exposed Persons $\hat{E}(t)$, Infections $\hat{I}(t)$, and Recovered $\hat{R}(t)$

Given the historic data $\{S(t), E(t), I(t), R(t), 0 \le t \le T\}$, we first compute three rates series $\{\beta(t), \sigma(t), \gamma(t), 0 \le t \le T-1\}$ as in Section 2.1. For each rate, we predict the value for $t \ge T$ using the model in Section 2.2 and denote them as $\{\hat{\beta}(t), \hat{\sigma}(t), \hat{\gamma}(t), t \ge T\}$. Inspired by (10), (11), and (12), all we need to predict tomorrow's cases number are today's cases number and rates, i.e.

$$\hat{E}(t+1) = \hat{E}(t) + \hat{\beta}(t)\hat{I}(t) - \hat{\sigma}(t)\hat{E}(t), \quad t \ge T, \tag{19}$$

$$\hat{I}(t+1) = \hat{I}(t) + \hat{\sigma}(t)\hat{E}(t) - \hat{\gamma}(t)\hat{I}(t), \quad t \ge T, \tag{20}$$

$$\hat{R}(t+1) = \hat{R}(t) + \hat{\gamma}(t)\hat{I}(t), \quad t \ge T \tag{21}$$

Note that when predicting $S(T+1), E(T+1), R(T+1)$, i.e. $t = T$, the available data $S(T), E(T), R(T)$ are used. The detailed steps are summarized in Algorithm 1

**Algorithm 1:** Tracking Discrete Time Time-dependent SEIR Model

**input** : $\{E(t), I(t), R(t), 0 \le t \le T-1\}$, regularization parameters $\lambda_1$, $\lambda_2$ and $\lambda_3$, order of autoregressive model $J, K, L$, prediction window $W$.

**output** : $\{\beta(t), \sigma(t), \gamma(t).\ 0 \le t \le T-2\}$, $\{\hat{\beta}(t), \hat{\sigma}(t), \hat{\gamma}(t).\ 0 \le t \le T-2\}$, $\{\hat{E(t)}, \hat{I(t}, \hat{R(t)}, t \ge T\}$.

1 Measure $\{\beta(t), \sigma(t), \gamma(t).\ 0 \le t \le T-2\}$ using (15), (14), and (13) respectively.
2 Train the ridge regression and estimate $\hat{\beta}(T-1)$, $\hat{\sigma}(T-1)$ and $\hat{\gamma}(T-1)$ using (16), (17), and (18).
3 Estimate the number of exposed persons $\hat{E(t)}$, infected persons $\hat{I(t)}$ and recovered persons $\hat{R(t)}$ on the next day T using (19), (20) and (21) respectively.
4 **While** $T \le t \le T+W$ **do**
5 Estimate $\hat{\beta}(t)$, $\hat{\sigma}(t)$ and $\hat{\gamma}(t)$ by (13), (14) and (15) respectively.
6 Predict $\hat{E}(t+1)$, $\hat{I}(t+1)$ and $\hat{R}(t+1)$ using (19), (20), and (21) respectively.
7 **end while**

## 3 THE SEIR MODEL WITH SYMPTOMATIC AND ASYMPTOMATIC INFECTIOUS PERSONS

People with COVID-19 have had a wide range of symptoms reported ranging from mild symptoms to severe illness. In fact, it was reported the United States center for disease and control that symptoms of COVID-19 may appear 2-14 days after exposure to the virus, which include but not limited to coughing, shortness of breath or difficulty breathing, fever, chills, muscle, pain, sore throat, and loss of taste or smell. However, what make COVID-19 disease novel that the scientific community is still battling to understand is that there are people who do not show any of the symptoms. In a recent study published by the World Health Organisation in Jan 2020 [28], about 87.9% of the COVID-19 patients had fever and **67.7%** of them had cough. Moreover, the recent work by [14] confirmed the existence of asymptomatic COVID-19 patients that are unaware of their ability to transmit the disease. For this reason, we propose a susceptible, exposed, symptomatic infectious, asymptomatic infectious, and recovered modeling framework denoted by $\boldsymbol{S(t)}$, $\boldsymbol{E(t)}$, $\boldsymbol{I(t)}$, $\boldsymbol{A(t)}$, and $\boldsymbol{R(t)}$. We assumed that there are infectious individuals that can transmit the disease but without showing symptoms (i.e. asymptomatic persons). The resulting model for the transmission dynamics of COVID-19 is given by the following deterministic system of nonlinear differential equations:

$$\frac{dS}{dt} = -\beta(t)\,(I + \eta A)\,\frac{S}{N}, \tag{22}$$

$$\frac{dE}{dt} = \beta(t)\,(I + \eta A)\,\frac{S}{N} - \sigma E, \tag{23}$$

$$\frac{dI}{dt} = \alpha\sigma E - \gamma_I I, \tag{24}$$

$$\frac{dA}{dt} = (1-\alpha)\sigma E - \gamma_A A, \tag{25}$$

$$\frac{dR}{dt} = \gamma_I(t) I + \gamma_A(t) A, \tag{26}$$

where

$$S + E + I + A + R = N \tag{27}$$

is the total population in the community, and $\boldsymbol{\beta(t)}$ is the transmission rate which is assumed to be time dependent. Moreover, $\boldsymbol{\gamma_I(t)}$ and $\boldsymbol{\gamma_A(t)}$ are the recovery rates for the symptomatic and asymptomatic persons which is assumed to be varying with respect to time, $\boldsymbol{\eta}$ represent the relative infectiousness of asymptomatic persons (in comparison to symptomatic persons), $\boldsymbol{\sigma}$ is the transition

rate from the exposed to infectious class (so $1/\sigma$ is the disease incubation period), and $\alpha$ is the fraction of cases that are symptomatic for the population under consideration.

The above SEIR model is based on the modeling framework proposed in [24], and the variety of SEIR-style models recently employed in [6, 12].

# 4 ANALYTICAL RESULTS

## 4.1 Basic Reproduction Number for the SEIR Model with Asymptomatic Persons

In this section we give the closed-form expression for the basic reproduction number, $\mathcal{R}_0$, for the model with asymptomatic persons (22)–(26) for the special case when $\beta(t) \equiv \beta_0$, $\gamma_I(t) \equiv \gamma_I$, and $\gamma_A(t) \equiv \gamma_A$. The local stability of the disease free equilibrium (DFE), given by

$$\text{DFE: } (S^*, E^*, I^*, A^*, R^*) = (N(0), 0, 0, 0),$$

is explored using the next generation operator method [18], [19]. Using the notation in [18], it follows that the matrices $\mathcal{F}$ of new infection terms and $\mathcal{V}$ of the remaining transfer terms associated with the version of the model (22)–(26) are given, respectively, by

$$\mathcal{F} = \begin{bmatrix} 0 & \beta_0 & \beta_0\eta \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathcal{V} = \begin{bmatrix} \sigma & 0 & 0 \\ -\alpha\sigma & \gamma_I & 0 \\ -(1-\alpha)\sigma & 0 & \gamma_A \end{bmatrix}.$$

The basic reproduction number of the model (22)–(26), denoted by $\mathcal{R}_0$, is given by

$$\mathcal{R}_0 = \frac{\beta_0\alpha}{\gamma_I} + \frac{\beta_0\eta(1-\alpha)}{\gamma_A} \tag{28}$$

The result below follows from Theorem 2 of [19].

Theorem 4.1. *The disease-free equilibrium (DFE) of the model (22)–(26) is locally-asymptotically stable if $\mathcal{R}_0 < 1$, and unstable if $\mathcal{R}_0 > 1$.*

The basic reproduction number, $\mathcal{R}_0$, in this case measures the average number of new COVID-19 infections generated by an average infected individual introduced into a population. The newly symptomatic infectious person can infects on averaged $\frac{\beta_0}{\gamma_I}$ with probability $\alpha$. Simailarly, the asymptomatic individual can infects on average $\frac{\beta_0}{\gamma_A}$ with probability $(1-\alpha)$. The epidemiological implication of Theorem 4.1 is that a small influx of COVID-19 cases will not generate a COVID-19 outbreak if the basic reproduction number $\mathcal{R}_0$ is less than unity, and there will be an outbreak if $\mathcal{R}_0$ is greater than unity.

# 5 NUMERICAL RESULTS

## 5.1 Result For Time-dependent SIR Model

In this section, we implement a simplified version of our model as in [4] which includes only $S(t), I(t), R(t)$ since we do not have the recorded Exposed people in any of the countries or states. To valid our model, we apply our model to both national level data: Russia, China, Spain, Italy and the USA, as well as state level data: Iowa, Maryland and Texas.

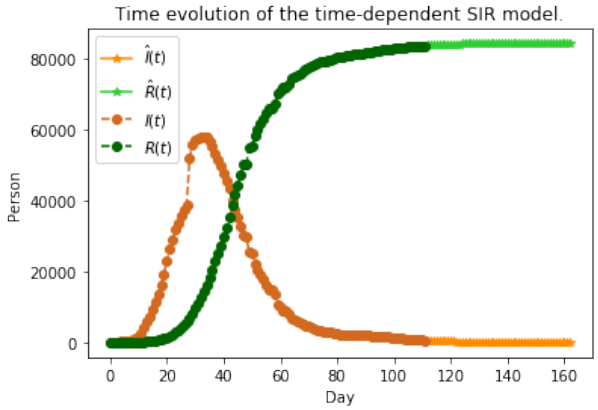


**Figure 2: China. The circle-marked solid curve with dark orange (resp. green) color is the real number of infected persons $I(t)$ (resp. recovered persons $R(t)$), the star-marked dashed curve with light orange (resp. green) color is the predicted number of infected persons $\hat{I}(t)$ (resp. recovered $\hat{R}(t)$).**

We collect the data from JHU open source and [1]. The data contains the number of confirmed cases, the number of recovered cases and the number of death. For each country, we choose a different start date for training according to the time when a relative clear trend is shown. We set up a range of 3 to 20 as the order of the autoregressive models for predicting $\beta(t)$ and $\gamma(t)$ in Algorithm 1 and choose the final model with the best performance. The scikit-learn library and GridSearchCV are used to carry out the algorithm. The prediction result is summarized in Table 2.

The result of China shows that there will be approximately 84,400 total confirmed cases in the next two months from $\mathbf{2020-05-05}$ and figure 2 confirms that the disease has ended even before the prediction window. In Russia, the results show that the disease will end in approximately two months (70 days) as shown in figure 3 and there will be 387,860 confirmed cases if the Russian government remains their control policy.

The result of the U.S. implies the possibility of having approximately $\mathbf{1.7M}$ confirmed cases by August 2020. Since these are a only few states we can get the whole records of the recovered cases, we pick three states (Maryland, Iowa and Texas) to see how reopen will impact the pandemic. Our prediction results indicate that there will be 33,804 confirmed cases in Maryland, 10,297 confirmed cases in Iowa and 47,083 in Texas by $\mathbf{2020-06-23}$. The infected number of people in Texas will start decreasing by $\mathbf{2020-05-25}$ as in Figure 4. Hence the reopen order issued may not increase the infection.

In Figure 5 we show the one day relative prediction error of the model using the USA data. The mean error for infected cases is $\mathbf{0.95\%}$ and the mean error for recovered cases is $\mathbf{9.04\%}$. The high prediction error for recovered might be caused by the incomplete record of the recovered cases since the recovered prediction error for China is actually less than $\mathbf{1\%}$.

**Table 2: Prediction Results for Countries**

| Country | Startdate of training dataset | Enddate of training dataset | Prediction Window | Predicted Total Confirmed Cases |
|---|---|---|---|---|
| China | **2020 − 01 − 15** | **2020 − 05 − 05** | 100 | 84,413 |
| Italy | **2020 − 02 − 24** | **2020 − 05 − 03** | 100 | 230,088 |
| Russia | **2020 − 04 − 06** | **2020 − 05 − 10** | 100 | 387,860 |
| Spain | **2020 − 03 − 03** | **2020 − 05 − 06** | 100 | 295,974 |
| the USA | **2020 − 03 − 04** | **2020 − 05 − 05** | 100 | 1,742,736 |
| UK | **2020 − 03 − 17** | **2020 − 05 − 10** | 100 | 258,071 |

**Table 3: Prediction Results for States in the U.S.**

| State | Start date of training dataset | End date of training dataset | Prediction Window | Predicted Total Confirmed Cases |
|---|---|---|---|---|
| Maryland | **2020 − 04 − 12** | **2020 − 05 − 04** | 50 | 33,804 |
| Iowa | **2020 − 04 − 12** | **2020 − 05 − 04** | 50 | 10,297 |
| Texas | **2020 − 04 − 12** | **2020 − 05 − 04** | 50 | 47,083 |

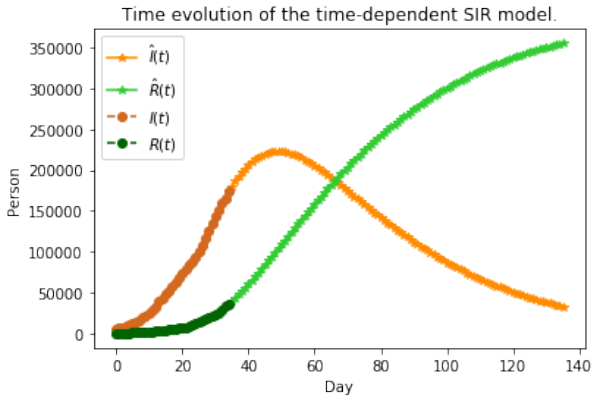


**Figure 3: Russia. The circle-marked solid curve with dark orange (resp. green) color is the real number of infected persons $I(t)$ (resp. recovered persons $R(t)$), the star-marked dashed curve with light orange (resp. green) color is the predicted number of infected persons $\hat{I}(t)$ (resp. recovered $\hat{R}(t)$).**

**Figure 4: Texas. The circle-marked solid curve with dark orange (resp. green) color is the real number of infected persons $I(t)$ (resp. recovered persons $R(t)$), the star-marked dashed curve with light orange (resp. green) color is the predicted number of infected persons $\hat{I}(t)$ (resp. recovered $\hat{R}(t)$).**

## 5.2 Result For Time-dependent SEIR Model

In this section, we use our time-dependent SEIR model in Section 2 to analyze and predict the trend of COVID-19 in China. Using the data source from [1], we lack the number of Exposed individuals. To get this information, we make the estimation that the number of infected individuals is the number of confirmed cases minus the number of recovered persons and the number of death, and the recovered person (R in our model) as the number of recovered cases plus death cases. Using two weeks of data of china from **2020 − 01 − 22**, the average Infected-Suspected ratio was found to be **2.399** which we used to define the number of Exposed individuals as a product of the number of infected individuals [5].

The result shows that there will be **84, 418** total confirmed cases, **1, 156** Exposed persons, **520** infected persons and **83, 886** recovered plus death persons by around end of June 2020. The infection reached the peak around mid February, while the disease ended before our prediction date **2020 − 05 − 06** as seen in Figure 6. This indicate that the reopen order do not increase the infection in China.

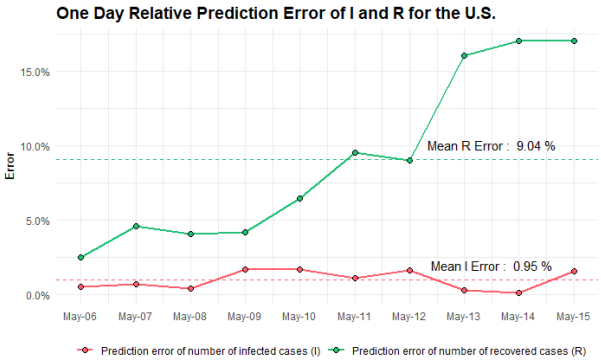


**Figure 5: One-day relative prediction error for the U.S.**

In figures 7, 8, and 9, we show the decreasing estimated transmission rate $\beta(t)$ and estimated incubation rate $\sigma(t)$ and the increasing recovery rate $\gamma(t)$. It agrees with what China is experiencing now

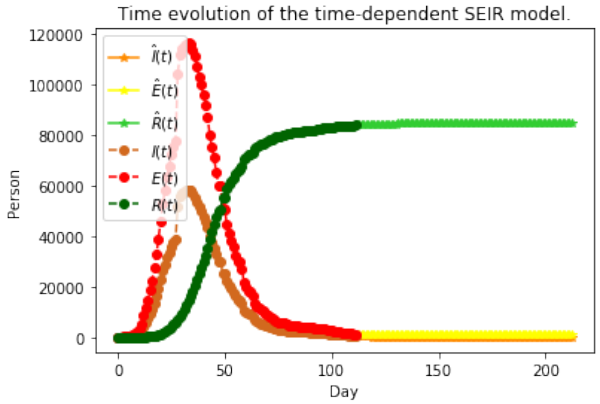


**Figure 6: China. The circle-marked solid curve with dark orange (resp. red, green) color is the real number of infected persons $I(t)$ (resp. exposed person $E(t)$, recovered $R(t)$), the star-marked dashed curve with light orange (resp. yellow, green) color is the predicted number of infected persons $\hat{I}(t)$ (resp. exposed $E\hat{(t)}$, recovered $\hat{R}(t)$).**

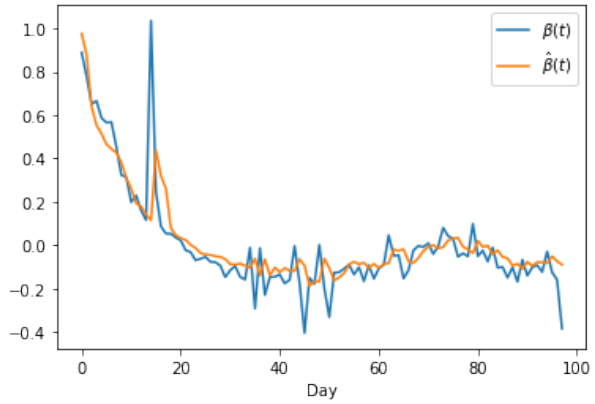


**Figure 7: Trajectory of the transmission rate $\beta(t)$ and predicted transmission rate $\hat{\beta}(t)$.**

with the pandemic due to the policy implemented by the Chinese government that tries to suppress the transmission rate by city-wide lock down and traffic halt [4]. In figure 10 we show that the reproduction number is decreasing and inflection point (turning point) that is when $R_0(t) = \frac{\beta(t)}{\gamma(t)} \leq 1$ occurred on $2020-02-10$ from which it goes below one suggesting that the disease will gradually dies down.

## 5.3 Simulation Result for the SEIR Model

In this section, we will model China and the USA trajectory of the infection cases, recovered cases and exposed cases using the SEIR model, 100 days from the start date of $2020-01-22$ for china and $2020-03-09$ for the USA. We set the parameters of the SEIR model as in Table 1. The population of China is assumed to be approximately 1.4 billion and the population of the U.S. is assumed

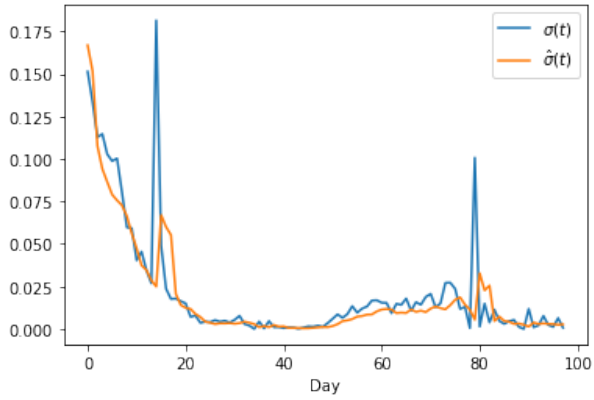


**Figure 8: Trajectory of the incubation rate $\sigma(t)$ and predicted incubation rate $\hat{\sigma}(t)$.**

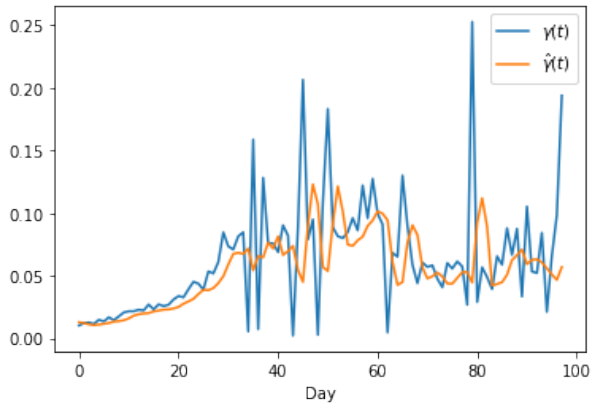


**Figure 9: Trajectory of the recovery rate $\gamma(t)$ and predicted recovery rate $\hat{\gamma}(t)$.**

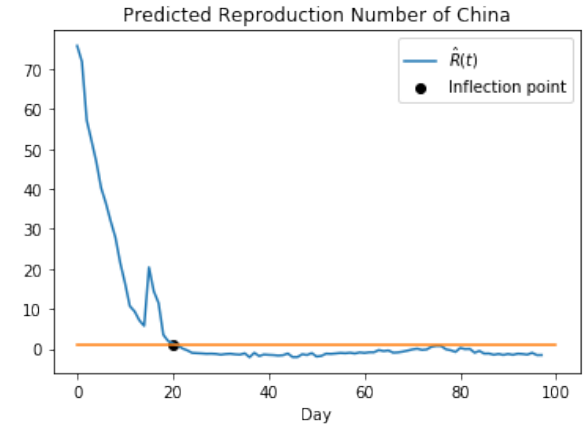


**Figure 10: Trajectory of the basic reproduction number $R_0(t)$ in China.**

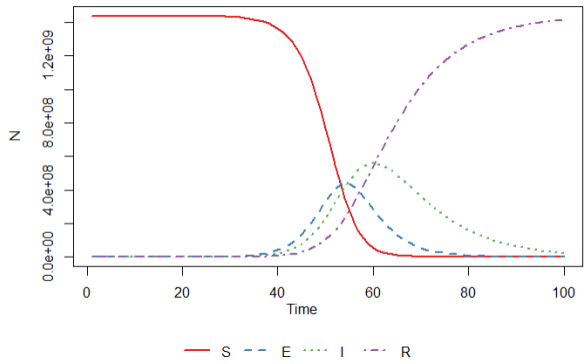


**Figure 11: Simulation Of the SEIR Model For China**

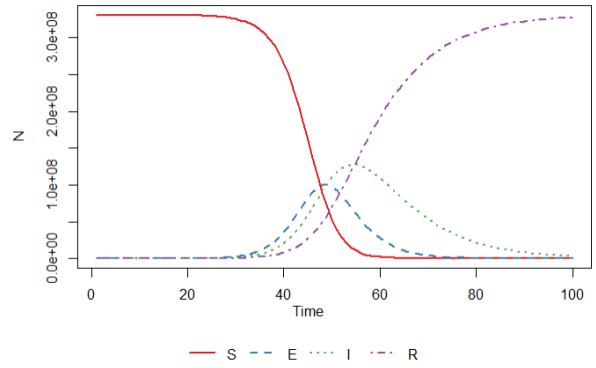


**Figure 12: Simulation Of the SEIR Model For the USA**

to be approximately 328 million. Similarly as in section 5.2, we make the estimation that the number of infected individuals today is the number of confirmed cases minus the number of recovered persons and the number of death persons. The average Infected-Suspected ratio was found to be **2.399** in China, which we use to initialize the number of exposed individuals as a product of the number of Infected individuals on **2020 − 01 − 22** [5]. In figure 11, we show that the infection in China will reach the peak around **2020−03−10** and the disease will end around **2020−03−24**. As for the USA, Figure 12 shows that the infection will be at peak around **2020 − 05 − 01** and the disease will end around **2020 − 05 − 16**. This clearly indicate that the pandemic is over for those US states that followed the guidelines set by center for disease control (CDC) for the re-opening. This result agreed with current situation in most of the US states except about twelves states that are currently experiencing a rise in the number of new cases, which is associated with their decision made on re-opening their states before reaching the guidelines. Thus, our findings showed that the decision as to when to re-open greatly affect the spread of the pandemic.

## 6 CONCLUSION

In this paper, we propose a time-dependent SEIR model considering the incubation period and both symptomatic and asymptomatic infectious cases, with coefficients estimated by ridge regression and parameterized using data relevant to the COVID-19 transmission dynamics in China, Italy, Russia, Spain, the USA, and UK, and three states in the U.S., Maryland, Iowa, and Texas. Using two weeks of data of China from 2020-01-22, our time-dependent SEIR model suggested the total confirmed cases to be **84, 418**, exposed persons to be **1, 156**, and recovery about **83, 886** with infection reaching the peak around mid of February. There is considerable ongoing debate on whether to recommend reopening. Some countries already had successfully curved the spread of the disease such as China, Germany, South Korea, Japan, and Italy. In the case of China, our result indicates that the transmission rate $\boldsymbol{\beta(t)}$ and incubation rate $\boldsymbol{\sigma(t)}$ decreases with time while the recovery rate $\boldsymbol{\gamma(t)}$ increases. Moreover, we find out that the inflection point for China occurred on February 10, 2020 from which the basic reproduction number is below one indicating that the disease is gradually dying down. This result agrees with what China is currently experiencing which clearly shows the effect of mitigation after non-pharmaceutical interventions were being implemented by the Chinese government. Furthermore, we simulated our SEIR model for China and the U.S. to model the trajectory of the infection, recovered cases, and exposed cases. In the further, we will consider both deterministic SEIR model and stochastic compartmental models to have a complementary insights of the COVID-19 pandemic.

## REFERENCES

[1] 2020. https://github.com/GuangchuangYu/nCov2019. (2020).
[2] Ismail Abdulrashid, Abdallah AM Alsammani, and Xiaoying Han. 2019. Stability analysis of a chemotherapy model with delays. *Discrete & Continuous Dynamical Systems-B* 24, 3 (2019), 989.
[3] Ismail Abdulrashid and Xiaoying Han. 2020. A mathematical model of chemotherapy with variable infusion. *Communications on Pure & Applied Analysis* 19, 4 (2020), 1875.
[4] Yi-Cheng Chen, Ping-En Lu, Cheng-Shang Chang, and Tzu-Hsuan Liu. 2020. A Time-dependent SIR model for COVID-19 with Undetectable Infected Persons. *arXiv preprint arXiv:2003.00122* (2020).
[5] Fairoza Amira Binti FA, Lau C, Nazri H, Ligot DV, Lee Raina G, and Tan H. CL. 2020. CoronaTracker: World-wide COVID-19 Outbreak Data Analysis and Prediction CoronaTracker Community Research Group. *Bull World Health Organ* (2020).
[6] Laydon D. Nedjati Gilani G. Imai N. Ainslie K. Baguelin M. ... Dighe A. Ferguson, N. 1979. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. (1979).
[7] Neil Ferguson, Daniel Laydon, Gemma Nedjati Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, ZULMA Cucunuba Perez, Gina Cuomo-Dannenburg, et al. 2020. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. (2020).
[8] Tapiwa Ganyani, Cecile Kremer, Dongxuan Chen, Andrea Torneri, Christel Faes, Jacco Wallinga, and Niel Hens. 2020. Estimating the generation interval for COVID-19 based on symptom onset data. *medRxiv* (2020).
[9] Daozhou Gao, Travis C Porco, and Shigui Ruan. 2016. Coinfection dynamics of two diseases in a single host population. *Journal of mathematical analysis and applications* 442, 1 (2016), 171–188.
[10] Fushing Hsieh and Jingyi Zheng. 2019. Unraveling Pattern-Based Mechanics Defining Self-Organized Recurrent Behaviors in a Complex System: A Zebrafish's Calcium Brain-Wide Imaging Example. *Frontiers in Applied Mathematics and Statistics* 5 (2019).
[11] Gabriel G. Katul, Assaad Mrad, Sara Bonetti, Gabriele Manoli, and Anthony J. Parolari. 2020. Global convergence of COVID-19 basic reproduction number and estimation from early-time SIR dynamics. *medRxiv* (2020).
[12] Pei S. Chen B. Song Y. Zhang T. Yang W. Shaman J. Li, R. 2020. *Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2).*
[13] Ruiyun Li, Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, and Jeffrey Shaman. 2020. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* 368, 6490 (2020), 489–493.
[14] B. F. Maier and D. Brockmann. 2020. Effective containment explains sub- exponential growth in confirmed cases of recent covid-19 outbreak in mainland china.

*arXiv preprint arXiv:2002.07572, 2020* (2020).

[15] Leah F Moriarty. 2020. Public health responses to COVID-19 outbreaks on cruise ships—worldwide, February–March 2020. *MMWR. Morbidity and mortality weekly report* 69 (2020).

[16] Christopher Jl Murray. 2020. Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months. *medRxiv* (2020).

[17] Calistus N. Ngonghala, Enahoro Iboi, Steffen Eikenberry, Matthew Scotch, Chandini Raina MacIntyre, Matthew H. Bonds, and Abba B. Gumel. 2020. Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. *medRxiv* (2020).

[18] J.A. Metz O. Diekmann, J.A.P. Heesterbeek. 1990. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology* 28 (1990), 365–382.

[19] J. Watmough P. van den Driessche. 2002. Reproduction numbers and subthreshold endemic equilibria for com- partmental models of disease transmission. *Mathematical Biosciences* 180 (2002), 29–48.

[20] Jonathan M Read, Jessica RE Bridgen, Derek AT Cummings, Antonia Ho, and Chris P Jewell. 2020. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. *MedRxiv* (2020).

[21] W Schmidt. 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,—. *Biometrical Journal* 23, 5 (1981), 519–520.

[22] O Sharomi and AB Gumel. 2011. Dynamical analysis of a sex-structured Chlamydia trachomatis transmission model with time delay. *Nonlinear Analysis: Real World Applications* 12, 2 (2011), 837–866.

[23] Mingwang Shen, Zhihang Peng, Yanni Xiao, and Lei Zhang. 2020. Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China. *BioRxiv* (2020).

[24] Enahoro Iboi Tin Phan Keenan Eikenberry Yang Kuang Eric Kostelich Steffen E. Eikenberry, Marina Mancuso and Abba B. Gumel. 2020. *To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic.* https://doi.org/abs/2004.03251

[25] Biao Tang, Nicola Luigi Bragazzi, Qian Li, Sanyi Tang, Yanni Xiao, and Jianhong Wu. 2020. An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov). *Infectious disease modelling* 5 (2020), 248–255.

[26] Alexis Akira Toda. 2020. Susceptible-Infected-Recovered (SIR) Dynamics of COVID-19 and Economic Impact. *arXiv preprint arXiv:2003.11221* (2020).

[27] Robert Verity, Lucy C Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko Imai, Gina Cuomo-Dannenburg, Hayley Thompson, Patrick GT Walker, Han Fu, et al. 2020. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases* (2020).

[28] WHO. 2016. Epilepsy Fact Sheet. *Biological Cybernetics* 42, 1 (2016), 9–15.

[29] Chong You, Yuhao Deng, Wenjie Hu, Jiarui Sun, Qiushi Lin, Feng Zhou, Cheng Heng Pang, Yuan Zhang, Zhengchao Chen, and Xiao-Hua Zhou. 2020. Estimation of the Time-Varying Reproduction Number of COVID-19 Outbreak in China. *medRxiv* (2020).

[30] Jingyi Zheng, Hsieh Fushing, and Linqiang Ge. 2019. A Data-driven Approach to Predict and Classify Epileptic Seizures from Brain-wide Calcium Imaging Video Data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (2019), 1–1.

[31] Jingyi Zheng, Mingli Liang, Arne D. Ekstrom, Linqiang Ge, Wei Yu, and Fushing Hsieh. 2018. On Association Study of Scalp EEG Data Channels Under Different Circumstances. In *International Conference on Wireless Algorithms, Systems, and Applications*. 683–695.

[32] Fei Zhou, Ting Yu, Ronghui Du, Guohui Fan, Ying Liu, Zhibo Liu, Jie Xiang, Yeming Wang, Bin Song, Xiaoying Gu, et al. 2020. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The lancet* (2020).