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

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

Harshvardhan Uppaluru * Hossein Rastgoftar **

* Scalable Move and Resilient Traversability (SMART) Lab, Aerospace and Mechanical Engineering Department, University of Arizona, Tucson, AZ 85721, USA (e-mail: huppaluru@arizona.edu).

** Scalable Move and Resilient Traversability (SMART) Lab, Aerospace and Mechanical Engineering Department, University of Arizona, Tucson, AZ 85721, USA (e-mail: hrastgoftar@arizona.edu).

Abstract: COVID-19 is a global health crisis that has had unprecedented, widespread impact on households across the United States and has been declared a global pandemic on March 11, 2020 by World Health Organization (WHO). According to Centers for Disease Control and Prevention (CDC), the spread of COVID-19 occurs through person-to-person transmission i.e. close contact with infected people through contaminated surfaces and respiratory fluids carrying infectious virus. This paper presents a data-driven physics-based approach to analyze and predict the rapid growth and spread dynamics of the pandemic. Temporal and spatial conservation laws are used to model the evolution of the COVID-19 pandemic. We integrate quadratic programming and neural networks to learn the parameters and estimate the pandemic growth. The proposed prediction model is validated through finite-time estimation of the pandemic growth using the total number of cases, deaths and recoveries in the United States recorded from March 12, 2020 until October 1, 2021.

Copyright © 2022 The Authors. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)

Keywords: COVID-19, Discrete Event Dynamic Systems, Estimation, Modeling and Validation, Pandemic Growth Analysis

1. INTRODUCTION

Since the outbreak of COVID-19 at the end of 2019 due to a new virus (SARS-Cov-2) belonging to the family of coronavirus, the pandemic has significantly impacted daily life in the United States and rest of the world and continues to impact the global economy and human health. The first confirmed case of COVID-19 in the United States was identified on January 23, 2020 and the first death caused by COVID-19 in the United States is believed to have occurred on February 6, 2020 in Santa Clara County, California. The virus has grown and spread rapidly across the country since then. Solely in the United States, more than 80 million cases have been identified causing more than 1 million deaths. As of April 22, 2022, COVID-19 has infected over 500 million people and killed more than 6 million across the planet.

1.1 Related Work

Kermack and McKendrick (1927, 1932, 1933) developed mathematical models for infectious diseases which are considered an important tool for analyzing epidemiological features and transmission. A large number of studies on COVID-19 have been conducted by researchers and organizations around the world to estimate the dynamics of this infectious disease. Some of the most popular methods are SIR (Susceptible, Infected, and Recovered) and SEIR (Susceptible, Exposed, Infected, and Recovered). A time-window based SIR model was proposed by Liao et al. (2020), for dynamic data analysis measuring the basic infection number and predicting the growth rate of the pandemic. A generalised SIR model was presented by Keeling et al. (2004) for dynamic data analysis measuring the basic infection number and predicting the growth rate of the pandemic. Metapopulation Dynamics was used to study and evaluate the effects of domestic and international travel limitations on the spread of COVID-19 (See Chinazzi et al. (2020); Grenfell and Harwood (1997); Hanski (1998); Keeling et al. (2004)). Fanelli and Piazza (2020) applied mean-field theory to analyze and predict the spread of COVID-19 in countries such as China, Italy and France. van Doremalen et al. (2020) used a Bayesian regression model to analyze the aerosol stability of COVID-19. Recently, Deep Learning (DL) methods have also been at the forefront to analyze and predict the spread of COVID-19. Chimmula and Zhang (2020) was the first study that proposed using a Long Short-Term Memory (LSTM) network to model the spread of COVID-19 in Canada based on past data. One of the key features of this approach was considering the recovery rate during the development of the model. Krbas et al. (2020) presented a comparative study between Auto-Regressive Integrated Moving Average (ARIMA), Nonlinear Autoregression Neural Network (NARNN) and Long Short-Term Memory (LSTM) approaches to model and estimate the total number of COVID-19 infections for 8 different European countries. Four deep learning models: Long Short-term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN) and Multivariate Convolutional Neural Network (MCNN) were compared and analyzed on the basis of consistency & accuracy - with CNN outperforming other deep learning models (See Nabi et al. (2021)).
1.2 Objectives and Contributions

The fundamental objective of this article is to develop a data-driven physics-based approach which can be used for finite time estimation of the spread of the COVID-19 pandemic across the United States. This finite estimation model can accurately estimate the growth of the pandemic for a shorter future time horizon compared to the existing methods. An early work on this topic was previously studied in Rastgoftar and Atkins (2021) where the authors assumed that evolution of the pandemic in every US state is not impacted by the growth in other US states. In contrast, the approach described in this paper advances existing work (Rastgoftar and Atkins (2021)) by quantifying the inter-state influences and incorporating them into learning and prediction of the pandemic growth. To this end, our primary objective is to apply the temporal and spatial conservation laws to model and learn the evolution of the COVID-19 pandemic based on the total number of cases, deaths and recoveries across the United States. We then obtain the interstate influences and the parameters for the proposed dynamics by relying on the model, the pandemic data and leveraging quadratic programming & neural networks. The secondary objective of this article is to verify the stability of the pandemic. Using the open-source data from OWM (2020), we model the spread dynamics and use the stability criterion proposed in Rastgoftar and Atkins (2021) to analyze the stability of the growth of the COVID-19 pandemic across the United States.

1.3 Organization of the paper

We detail the problem statement in Section II. In Section III, our proposed finite-time estimation framework is described. A novel approach to learning growth of the pandemic is provided in Section IV. We analyze the stability of the growth of the pandemic in Section V. Our evaluation and results are presented in Section VI. Finally, we provide discussion along with thoughts for future directions and conclude our work in Section VII.

2. PROBLEM STATEMENT

We consider the growth of the COVID-19 pandemic in the United States where 50 US states and District of Columbia are sorted alphabetically and identified by set \( V = \{1, \ldots, 51\} \). We define \( t_i[k], d_i[k], \) and \( r_i[k] \) as the total number of cases, deaths, and recoveries, respectively, with \( i \in V \) at day \( k = 1, \ldots, n \). \( n \) is the total number of days from the establishment of pandemic growth. Estimations of \( t_i[k], d_i[k], \) and \( r_i[k] \) are represented as \( \hat{t}_i[k], \hat{d}_i[k], \) and \( \hat{r}_i[k], \) respectively. We apply temporal and spatial conservation laws to model and learn the evolution of the pandemic. The temporal conservation law is given by

\[
\dot{x}_i[k + 1] = x_i[k] + u_i[k], \quad k = 1, 2, \ldots, n \tag{1}
\]

where \( x_i[k] = [\hat{t}_i[k], \hat{d}_i[k], \hat{r}_i[k]]^T \) is the state vector and \( u_i[k] = [\hat{u}_{t,i}[k], \hat{u}_{d,i}[k], \hat{u}_{r,i}[k]]^T \) is the input vector specifying number of new cases, deaths, and recoveries at day \( k \). The spatial conservation law is given by

\[
\dot{\hat{x}}_i[k] = \hat{t}_i[k] - \hat{d}_i[k] - \hat{r}_i[k], \quad k = 1, 2, \ldots, n \tag{2}
\]

where \( \hat{a}_i[k] \) represents the number of active cases at day \( k \). This paper assumes that the number of new cases, deaths, and recoveries at day \( k \) are proportional to the number of active cases over the past \( N_r \) days from day \( k \). Therefore, we can define \( \hat{u}_{t,i}[k], \hat{u}_{d,i}, \hat{u}_{r,i} \) as follows:

\[
\begin{align*}
\hat{u}_{t,i}[k] &= \beta[k] \sum_{h=1}^{N_r} x_{i,i,h}[k] \hat{a}_i[k - h] \\
    &\quad + \left( 1 - \beta[k] \right) \sum_{h=1, j \in V}^{N_r} x_{i,j,h}[k] \hat{a}_j[k - h], \quad \forall i \in V, \tag{3a}
\end{align*}
\]

\[
\begin{align*}
\hat{u}_{d,i}[k] &= \beta[k] \sum_{h=1}^{N_r} x_{i,i,h}[k] \hat{a}_i[k - h] \\
    &\quad + \left( 1 - \beta[k] \right) \sum_{h=1, j \in V}^{N_r} x_{i,j,h}[k] \hat{a}_j[k - h], \quad \forall i \in V, \tag{3b}
\end{align*}
\]

\[
\begin{align*}
\hat{u}_{r,i}[k] &= \beta[k] \sum_{h=1}^{N_r} x_{i,i,h}[k] \hat{a}_i[k - h] \\
    &\quad + \left( 1 - \beta[k] \right) \sum_{h=1, j \in V}^{N_r} x_{i,j,h}[k] \hat{a}_j[k - h], \quad \forall i \in V. \tag{3c}
\end{align*}
\]

where \( x_{i,j,h}[k] : V \times V \times \{1, \ldots, N_r\} \rightarrow \mathbb{R}_{\geq 0}, \lambda_{i,j,h}[k] : V \times V \times \{1, \ldots, N_r\} \rightarrow \mathbb{R}_{\geq 0}, \) and \( \theta_{i,j,h}[k] : V \times V \times \{1, \ldots, N_r\} \rightarrow \mathbb{R}_{\geq 0} \) assign the influence of the estimated number of active cases in state \( j \in V \) from day \( k - h, \ldots, k - N_r \) on the number of new active cases, deaths, and recoveries in state \( i \in V \) at day \( k \). Scaling parameter \( \beta[k] \in [0, 1] \). If \( \beta[k] = 1 \), the second term on the right-hand side of Eqs. (3a), (3b), and (3c) vanishes and the pandemic independently grows in every individual state. However if \( \beta[k] = 0 \), the first term on the right-hand side of Eqs. (3a), (3b), and (3c) is 0 and the pandemic in a state grows dependent on other states. Otherwise (\( \beta \in [0, 1] \)), evolution of pandemic in every US state is influenced by the growth in other US states.

We first develop a finite time estimation model to predict the number of new cases, deaths, and recoveries in a finite number of future days and analyze the stability of growth of a pandemic disease in Section 3. Then, we offer an integration of quadratic programming and neural network to learn parameters \( \omega_{i,j,h}[k], \lambda_{i,j,h}[k], \theta_{i,j,h}[k] \) and \( \beta[k] \) at every day \( k \) and characterize the interstate influences on transmitting the current pandemic disease (see Section 4).

3. FINITE-TIME ESTIMATION OF THE PANDEMIC GROWTH

Per definition of the number of active cases, given in (2), the control input \( u_i[k] \) is obtained as follows:

\[
u_i[k] = \sum_{h=1, j \in V}^{N_r} G_{i,j,h}[k] x_i[k - h], \quad \forall i \in V, \tag{4}
\]

where

\[
G_{i,j,h} = [\omega_{i,j,h} - \lambda_{i,j,h} - \theta_{i,j,h}] \otimes 1_{3 \times 1} \in \mathbb{R}^{3 \times 3}. \tag{5}
\]

We define

\[
\hat{y}_i[k] = [\hat{x}_i^T[k] \cdots \hat{x}_i^T[k - N_r + 1]]^T \in \mathbb{R}^{3N_r \times 1}, \tag{6a}
\]
\[ \hat{Y} = [\hat{y}_1^T \cdots \hat{y}_{51}^T]^T \in \mathbb{R}^{153N_t \times 1}, \] (6b)

\[
L_{i,j}[k] = 
\begin{bmatrix}
0 & \mathbf{I}_3 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \mathbf{I}_3 \\
g_{i,j,N}[k] & g_{i,j,N-1}[k] & \cdots & \mathbf{I} + g_{i,j,1}[k]
\end{bmatrix} \in \mathbb{R}^{N_t \times 3N_t}
\] (6c)

\[ \Lambda_0 = 
\begin{bmatrix}
L_{1,1}[k] & L_{1,2}[k] & \cdots & L_{1,51}[k] \\
L_{2,1}[k] & L_{2,2}[k] & \cdots & L_{2,51}[k] \\
\vdots & \vdots & \ddots & \vdots \\
L_{51,1}[k] & L_{51,2}[k] & \cdots & L_{51,51}[k]
\end{bmatrix}
\] (6d)

\[ \Lambda_1 = 
\begin{bmatrix}
0 & 0 & \cdots & L_{51,51}[k] \\
L_{1,1}[k] & 0 & \cdots & 0 \\
L_{2,1}[k] & L_{2,2}[k] & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & L_{51,51}[k]
\end{bmatrix}
\] (6e)

where “⊗” is the Kronecker product symbol, \( \mathbf{I}_3 \in \mathbb{R}^{3 \times 3} \) is the identity matrix, \( \mathbf{1}_{3 \times 1} \in \mathbb{R}^{3 \times 1} \) is the one-entry vector. Then, the estimation of the pandemic growth is obtained by the following \( M \)-step prediction dynamics:

\[
\begin{align*}
\dot{\hat{Y}}[k + M] &= \mathbf{L}^M[k] \hat{Y}[k] \\
\dot{x}_i[k + M] &= [\mathbf{0}_{M_t \times 3} \mathbf{I}_3 \mathbf{0}_{M_t \times 3}] \hat{Y}[k + M], \quad i \in \mathcal{V}
\end{align*}
\] (7)

where

\[
M_1 = ((i - 1)N_t + 3(M - 1))
\]

\[
M_2 = (51 - i)N_t + 3(N_t - M)
\]

and \( \dot{x}_i[k + M] = [\dot{r}_i[k + M] \ d_i[k + M] \ r_i[k + M]]^T \) aggregates the prediction of day \( k \) for the total number of cases, deaths, and recoveries at day \( k + M \).

4. LEARNING OF THE PANDEMIC GROWTH

The parameters \( \omega_{i,j,h}[k] : \mathcal{V} \times \mathcal{V} \times \{1, \ldots, N_t\} \rightarrow \mathbb{R}_{\geq 0} \), \( \lambda_{i,j,h}[k] : \mathcal{V} \times \mathcal{V} \times \{1, \ldots, N_t\} \rightarrow \mathbb{R}_{\geq 0} \), and \( \theta_{i,j,h}[k] : \mathcal{V} \times \mathcal{V} \times \{1, \ldots, N_t\} \rightarrow \mathbb{R}_{\geq 0} \) must be all non-negative at every day \( k \) since the number of new cases, deaths, and recoveries are positively impacted by the number of active cases. Therefore, matrices \( \Lambda_0 \) and \( \Lambda_1 \) are non-negative. Matrices \( \Lambda_0 \) and \( \Lambda_1 \) are learned by solving quadratic programming problem defined in Sections 4.1 and 4.2, respectively. To characterize the interstate influences, parameter \( \beta \) is learned by a neural network presented in Section 4.3.

4.1 Matrix \( \Lambda_0 \)

In this section, we assume that the pandemic grows under quarantined condition in every US state \( i \in \mathcal{V} \). Therefore, the spread of pandemic in a particular state is independent of other US states. Therefore, \( \beta[k] = 1 \); the second term on the right-hand side of Eq. (6f) vanishes; and matrix \( \mathbf{L} = \Lambda_1 \) is obtained by minimizing cost function

\[
C_q[k] = \frac{1}{2} \left( \Lambda_1[k] \hat{Y}[k] - \mathbf{Y}[k] \right)^T \mathbf{W}_q \left( \Lambda_1 \hat{Y}[k] - \mathbf{Y}[k] \right)
\] (9)

subject to the \( \omega_{i,i,h} \geq 0 \), \( \lambda_{i,i,h} \geq 0 \), and \( \theta_{i,i,h} \geq 0 \) for every \( i \in \mathcal{V} \) and \( h \in \{1, \ldots, N_t\} \), where \( \mathbf{Y}[k] \) aggregates the actual number of infected cases, deaths, and recoveries recorded at days \( k - N_t + 1 \) through \( k \) and \( \mathbf{W}_q \in \mathbb{R}^{153N_t \times 153N_t} \) is a positive definite weight matrix. Note that cost function \( C_q \) is quadratic with respect to \( \omega_{i,i,h} \), \( \lambda_{i,i,h} \), and \( \theta_{i,i,h} \). Therefore, matrix \( \mathbf{L} = \Lambda_1 \) is assigned by solving a quadratic programming problem.

4.2 Matrix \( \Lambda_1 \)

In this section, we assume that pandemic grows under non-quarantined conditions in every US state \( i \in \mathcal{V} \). Therefore, the growth of pandemic disease in every US state \( i \in \mathcal{V} \) is influenced by other US states and \( \beta[k] = 0 \), the first term on the right-hand side of Eq. (6f) vanishes, and \( \mathbf{L}[k] = \Lambda_0[k] \). Therefore, matrix \( \mathbf{L} = \Lambda_0 \) is determined by minimizing quadratic cost function

\[
C_n[k] = \frac{1}{2} \left( \Lambda_0[k] \hat{Y}[k] - \mathbf{Y}[k] \right)^T \mathbf{W}_n \left( \Lambda_0 \hat{Y}[k] - \mathbf{Y}[k] \right)
\]

subject to the \( \omega_{i,j,h} \geq 0 \), \( \lambda_{i,j,h} \geq 0 \), and \( \theta_{i,j,h} \geq 0 \) for every \( i,j \in \mathcal{V} \) and \( h \in \{1, \ldots, 14\} \), where \( \mathbf{W}_n \in \mathbb{R}^{153N_t \times 153N_t} \) is a positive definite weight matrix. Similar to the quarantined condition, matrix \( \mathbf{L} = \Lambda_0 \) can be assigned by solving a quadratic programming problem.

4.3 Parameter \( \beta \)

![Fig. 1. Neural Network (NN) architecture: \( t_{i,j,k} \) corresponds to the total number of cases. \( d_{i,j,k} \) corresponds to the total number of deaths. \( r_{i,j,k} \) corresponds to the total number of recoveries. \( i = 1 \cdots 51 \) corresponds to the different states. \( j = 1 \cdots N_t \) corresponds to the past data from day \( k \). \( \beta \) is either 0 or 1. Therefore the fully connected neural network consists of 14 x 3 x 51 neurons in the input layer. The hidden layer contains 51 neurons connected to a single output \( \beta \). The activation function used for the neural network is ReLU.](image-url)
network. For example, in the case of Vermont, \( \beta = 1 \), from March 12, 2020 until October 1, 2020 because of the spread of COVID was controlled to an extent. From October 2, 2020 until April 23, 2021, the state of Vermont had an increase in cases and deaths over this period. Therefore, we have \( \beta = 0 \) from October 2, 2020 till April 23, 2021 and the data recorded over this period i.e., October 2, 2020 till April 23, 2021 is training data and is used to train the network. The data from April 24, 2021 until October 1, 2021 was used to test the neural network and obtain the estimated number of cases, deaths and recoveries (See Figure 1).

5. STABILITY OF THE PANDEMIC GROWTH

Following the stability analysis first provided in Rastgoftar and Atkins (2021), we define active cases, originally given by (2) as follows:

\[
a_i[k] = a \hat{x}_i[k], \quad k = 1, 2, \ldots, 569
\]

(10)

where

\[
a = [1 \quad -1 \quad -1] \in \mathbb{R}^{1 \times 3}
\]

From (4), we can see that

\[
G_{i,j,h}[k] = K_{i,j,h}[k]a
\]

(12)

where

\[
K_{i,j,h}[k] = [\omega_{i,j,h}[k] \lambda_{i,j,h}[k] \theta_{i,j,h}[k]]^T
\]

(13)

Using (12), we can rewrite (4) as

\[
\hat{u}_i[k] = \sum_{h=1}^{N_t} \sum_{j \in \mathcal{V}} K_{i,j,h}[k] a[k-h], \quad \forall i \in \mathcal{V},
\]

(14)

Substituting (14) in (1) and premultiplying both sides by \( a \), we get

\[
a \hat{x}[k] = a \hat{x}[k-1] + \sum_{h=1}^{N_t} \sum_{j \in \mathcal{V}} aK_{i,j,h}[k] a[k-h] \quad \forall i \in \mathcal{V}
\]

(15)

Let

\[
\gamma_{i,j,h}[k] := aK_{i,j,h}[k] = \omega_{i,j,h}[k] - \lambda_{i,j,h}[k] - \theta_{i,j,h}[k]
\]

(16)

Hence the terms in (15) can be replaced to obtain the dynamics of the spread of active cases given by the following equation

\[
a[k] = (1 + \gamma_{i,j,1}[k]) a[k-1] + \cdots + \gamma_{i,j,N_t}[k] a[k-N_t]
\]

(17)

Using the dynamics of the spread of active cases given by (17), the growth of the total number of active cases, defined by (2), will reach stability if there exists a day \( k_s \) such that eigenvalues of

\[
\Gamma[k] = \begin{bmatrix}
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 \\
\gamma_{i,j,N_t}[k] & \gamma_{i,j,N_t-1}[k] & \cdots & 1 + \gamma_{i,j,1}[k]
\end{bmatrix} \in \mathbb{R}^{N_t \times N_t}
\]

(18)

are all inside the unit disk centered at the origin at every day \( k > k_s \).

6. RESULTS

We define the following relative error formulations before evaluating the effectiveness of our proposed finite-time estimation method.

\[
e_{t,M_s}[k] = \frac{t_{p,M_s}[k] - t[k]}{t[k]}
\]

(19)

\[
e_{d,M_s}[k] = \frac{d_{p,M_s}[k] - d[k]}{d[k]}
\]

(20)

\[
e_{r,M_s}[k] = \frac{r_{p,M_s}[k] - r[k]}{r[k]}
\]

(21)

The data is collected from OWM (2020). The spread of COVID-19 is considered over a period of \( n = 569 \) days from March 12, 2020 to October 1, 2021. For \( k = 1, \ldots, 569 \), the total number of cases, deaths and recoveries are represented as \( t[k], d[k], r[k] \). We chose \( N_t = 14 \) for all of our experiments, to obtain the parameters of the model described in Section IV. Therefore, \( t[k], d[k], r[k] \) for \( k = 1, \ldots, 14 \) (March 12, 2020 to March 25, 2020) are used to learn the parameters of the proposed model for \( k \geq 15 \) by solving the optimization problem defined in Section IV subject to constraints mentioned. We obtain gains \( K_1[k], \ldots, K_{N_t}[k] \) for day \( k \). These gains are then used in predicting the growth in total number of cases, deaths and recoveries from March 26, 2020. For example, the data from December 18, 2020 to December 31, 2020 is used to learn the parameters of the proposed model and predict the number of cases, deaths, and recoveries for January 1, 2021.

![Graph](image1)

(a) Actual number of deaths, (b) Predicted number of deaths for \( M_s = 1 \), (c) Actual number of infected cases, (d) Predicted number of infected cases for \( M_s = 4 \), (e) Actual number of recoveries, (f) Predicted number of recoveries for \( M_s = 5 \).
6.2 Estimation and Prediction in the state of New York

We test the approach described in Section III particularly for the state of New York because it is one of the most affected states in US due to the pandemic by assigning $\beta = 0$. Therefore, for $k = 15, \cdots, 569$, we use $t[k], d[k], r[k]$ from all 50 states and District of Columbia to predict and estimate the total number of cases, deaths and recoveries in New York. Figure 5 shows the plots of actual and predicted number of infected cases, deaths and recoveries in New York for $k = 15, \cdots, 569$ for $\tau = 3$. Figure 6 plots the relative errors in predicting the number of cases, deaths and recoveries in the state of New York at day $k = 15, \cdots, 569$ for $\tau = 3$. It can be seen that the relative errors in our estimations are less than 1%. 

6.3 Estimation and Prediction of Non-Quarantined or Quarantined state using Neural Network

The objective of this experiment is to determine whether the US state of Vermont is in a quarantined ($\beta = 1$) or non-quarantined ($\beta = 0$) state using Neural Networks. As described in Section IV, the input to the Neural Network is $t[k], d[k], r[k]$ for the past $k = t - 14, \cdots, t$ for all 51 states making the input of shape $3 \times 14 \times 51$ for a particular day $k$. The hidden layer is a fully connected layer with ReLU activation function. The output from the Neural Network is the value $\beta$. Using (3a) (3b) (3c), we obtain the estimated number of new cases, deaths, and recoveries. Finally using (1) we get the predicted number of cases, deaths and recoveries. Using this approach, we can see
that the relative estimation errors are less than 0.01% as shown in Figure 6 (d).

7. CONCLUSION

We presented a new physics-based data-driven approach to model and predict the spread of COVID-19 pandemic in the United States. Using a time-sliding window and historical data, we described an algorithm to learn the parameters of the proposed approach and developed a prediction model using spatial and temporal conservation laws, to estimate in finite-time, in a receding-time horizon. The accuracy of the prediction model is established by estimating the number of cases, deaths and recoveries in United States and also in the state of New York during this time period.

We also analyzed the stability of the growth of COVID-19 pandemic. Our results are coinciding with

1. the growth of the pandemic considering the devastating effects and spread of the delta variant since May 2021,
2. the actual data reported (Figure 2(c)) in which we can see the another wave i.e. an increase in the number of cases.

In future, we plan on using Markov Decision Process (MDP) to model the spread of the COVID-19 pandemic under the finite-time estimation approach described in this article. Assuming full-state observability, this novel decision-making model could potentially be used to effectively make state-wide and nationwide recommendations by optimizing for non-pharmaceutical actions.

REFERENCES

Chinazzi, M., Davis, J.T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., y Piontti, A.P., Mu, K., Rossi, L., Sun, K., et al. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*. Fanelli, D. and Piazza, F. (2020). Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons & Fractals*, 134, 109761. Grenfell, B. and Harwood, J. (1997). (Meta) population dynamics of infectious diseases. *Trends in ecology & evolution*, 12(10), 395–399.

Hanski, I. (1998). *Metapopulation dynamics*. *Nature*, 396(6706), 41–49.

Keeling, M.J., Bjørnstad, O.N., and Grenfell, B.T. (2004). *Metapopulation dynamics of infectious diseases*. In *Ecology, genetics and evolution of metapopulations*, 415–445. Elsevier.

Kermack, W.O. and McKendrick, A.G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of London. Series A, Containing papers of a mathematical and physical character*, 115(772), 700–721.

Kermack, W.O. and McKendrick, A.G. (1932). Contributions to the mathematical theory of epidemics. ii.—the problem of endemicity. *Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character*, 138(834), 55–83.

Kermack, W.O. and McKendrick, A.G. (1933). Contributions to the mathematical theory of epidemics. iii.—further studies of the problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 141(843), 94–122.

Kurban, I., Sözen, A., Tuncer, A.D., and Kazancıoğlu, F.S. (2020). Comparative analysis and forecasting of COVID-19 cases in various european countries with ARIMA, NARNN and LSTM approaches. *Chaos, Solitons & Fractals*, 138, 110015.

Liao, Z., Lan, P., Liao, Z., Zhang, Y., and Liu, S. (2020). TW-SIR: time-window based sir for COVID-19 forecasts. *Scientific reports*, 10(1), 1–15.

Nabi, K.N., Tahmid, M.T., Rafi, A., Kader, M.E., and Haider, M.A. (2021). Forecasting COVID-19 cases: A comparative analysis between recurrent and convolutional neural networks. *Results in Physics*, 24, 104137.

Rastgoftar, H. and Atkins, E. (2021). A mass-conservation model for stability analysis and finite-time estimation of spread of COVID-19. *IEEE Transactions on Computational Social Systems*.

Singh, P. and Gupta, A. (2021). Generalized sir (GSIR) epidemic model: An improved framework for the predictive monitoring of covid-19 pandemic. *ISA transactions*.

van Doremalen, N., Bushmaker, T., Morris, D.H., Holbrook, M.G., Gamble, A., Williamson, B.N., Tamin, A., Harcourt, J.L., Thornburg, N.J., Gerber, S.I., et al. (2020). Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *New England Journal of Medicine*. Warner et al. (2020). *Worldometer*. https://www.worldometers.info/coronavirus/country/us/.

Chimmula, V.K.R. and Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals*, 135, 109864.