LSTM-based Estimation of Time-Varying Parameters in a Spatiotemporal PDE Model for Prediction of Epidemic Spread

Deepak Antony David * **Logan Street** *
Subramanian Ramakrishnan ** **Manish Kumar** *

* *Department of Mechanical and Materials Engineering, University of Cincinnati, Cincinnati, OH USA. email: daviddy@mail.uc.edu, streetln@mail.uc.edu, kumarmu@ucmail.uc.edu*
** *Department of Mechanical and Aerospace Engineering, University of Dayton, Dayton OH, USA. email: sramakrishnan1@udayton.edu*

Abstract: In this paper we apply a Long-Short-Term-Memory (LSTM) deep learning method for forecasting time-varying parameters of a Partial Differential Equation (PDE)-based compartmental dynamic model of epidemic spread. The predictive efficacy of such models depends on accurately estimating and updating time-varying model parameters, based on empirical infection data updated daily by public health systems during an epidemic. Investigating the role of deep learning methods is important in this context and motivates our work. We first note that numerical data used in this work correspond to empirical COVID-19 infection data for the state of Ohio, USA. The LSTM is subject to an iterative training process for a total period of 30 days such that model parameters are generated for each day. Each iteration generates parameter values corresponding to a specific day and therefore yields a comprehensive representation of the temporal dynamics of the infection progression. Moreover, the training process is designed to ensure both the model's accuracy and predictive reliability. Using the day-to-day infection parameters yielded by the LSTM as input to the PDE model, a forecast of infection spread is obtained from the latter and validated against empirical COVID-19 spread data. In summary, by combining advanced deep learning techniques with epidemiological modeling, this study advances both our understanding of the complex, time-varying dynamics of epidemics and our ability to accurately forecast the dynamics from available empirical data.

Keywords: LSTM, Online Learning, Pattern Recognition, Partial Differential Equations, Epidemic Models, Time-Varying Parameters, Numerical Optimization.

1. INTRODUCTION

The COVID-19 pandemic caused by the SARS-COV-2 virus had an unprecedented and tremendous global impact, the aftershocks of which continue to affect our lives, including aspects of public health and socioeconomic activity. The virus was first reported in Wuhan, China, and then spread rapidly as recorded in the Center for Disease Control (CDC) Museum COVID-19 timeline. The CDC (2023) timeline provides a detailed account of the global evolution of the pandemic. Dynamic epidemic modeling - critical to developing effective countermeasures to combat epidemics - has naturally found resurgent interest following the emergence of COVID-19. In particular, this has led to an accelerated expansion of the corresponding literature, also for instance reported by Park et al. (2020) and Rahimi et al. (2023), which presents a systematic review of articles addressing early-stage epidemic models developed in response to COVID-19.

A preeminent class of models in mathematical epidemiology since the seminal work of Kermack and McKendrick (1927) is *compartmental models*. In this approach, a population is divided into compartments - typically, *Susceptible* (S), *Infected* (I), and *Recovered* (R) compartments. Continuous-time epidemic dynamic models are characterized by a system of coupled differential equations for the S, I, and R population densities, with individuals successively transitioning between the S, I, and R compartments according to prescribed rules. Notably, *spatiotemporal* epidemic dynamics can only be understood by treating the densities as functions of both time and space - correspondingly the model equations will be *partial differential equations* (PDE).

While compartmental models have been successfully employed to describe a variety of epidemics, their accuracy in capturing the essential dynamics is dependent on the accurate estimation of the model parameters (i.e. coefficients of the terms in the model equations). In particular, such parameter estimation is crucial when a model is used - along with empirical infection spread data - to *forecast* epidemic spread dynamics.

⋆ This work was supported by the National Science Foundation [Grant Numbers- 2140405 (PI- Ramakrishnan), 2140420 (PI- Kumar)].

The broad success of Machine Learning (ML) methods, particularly when large data sets are available, naturally suggests employing them for parameter estimation in epidemic modeling and fundamentally motivates the present effort. Indeed data-driven approaches based on ground truth data as input have been gaining traction in epidemic spread forecasting (for instance, Olumoyin et al. (2021)). However, combining data-driven approaches with mechanistic models (such as the PDE compartmental models discussed above) offers fertile ground for significant advances and is yet to be fully explored (Feng et al. (2022)). In recent times, physics-informed and bayesian machine learning (Niraula et al. (2022)) approaches have been employed to predict the flow of infections over certain periods, but not without their shortcomings.

We reiterate that parameter estimation is a crucial step towards robust predictions of epidemic spread based on available data at any given time. In turn, the prediction can serve public health policy development aimed at spread mitigation or prevention. In this context, we note our previous research (Majid et al. (2021, 2022)) that developed a PDE-based $SLIR$ model for epidemic spread forecasting. However, in contrast to the present work, Machine Learning methods were not considered in our previous efforts. Specifically, model parameter estimation was accomplished using a gradient descent toolbox which employed a root mean square error(RMSE) method to fit the parameters across a designated timeframe. The PDE model parameters estimated using this method are considered optimal for the observed trend and are assumed constant throughout the simulation period. However, the latter is a tenuous assumption, particularly during volatile phases of an epidemic. More specifically, parameters (defined in Section 2) such as the infection rate (ϕ) and diffusion coefficients (η_S & η_I) demand treatment as functions of time for improved accuracy in capturing the epidemic dynamics(Setianto and Hidayat (2023), Giordano et al. (2020)).

Motivated thus, in this paper we present a framework that leverages both the benefits of a compartmental epidemic model and the potential of an LSTM network. The empirical COVID-19 dataset for the state of Ohio is utilized to optimize the parameters of the PDE model. We propose an approach to extract the initial pattern of day-to-day infection parameters from the PDE model, which is subsequently utilized to train an LSTM to predict the upcoming day's infection parameter values in an online learning manner. We also propose two distinct perspectives for optimizing and extracting the initial pattern of the day-to-day infection parameters. One approach involves looking at data from 5 days prior and the other by looking at 10 days prior data.

The rest of the paper is set as follows. We present the foundational PDE model in Section 2.1, followed by the structure of LSTM in Section 2.2. In Section 3 we illustrate the various optimization methods that have been employed, followed by an experimental case study of the proposed methodology on the real-time COVID-19 data for the state of Ohio in Section 4. The paper concludes with a discussion of the results in Section 5 and an outlook for further research in Section 6.

2. MATHEMATICAL MODELING

2.1 Epidemic Model

We consider a compartmental epidemic model(Kermack and McKendrick (1927)). Here, the entire population in the domain is divided into three compartments: S, I, and R. However, an infection typically has an incubation period - the time delay between exposure and the emergence of the first infection symptoms. We use a time-delay (τ) to represent this latency in infection spread Li and Zou (2009) and include a Latent compartment (L) in our model. We treat the transfer of population in the system to be unidirectional, that is, the transitions are restricted from S to L to I, and to R exclusively in that order. The dynamics of the compartmental densities, including the compartmental transitions are described by a set of differential equations. Spatiotemporal heterogeneities significantly influence epidemic dynamics (Brauer et al. (2008)), and therefore motivate the use of a spatiotemporal model described by PDE as the mathematical basis for this research. Additionally, appropriately chosen PDE models can also account for complex dynamic characteristics that drive epidemic spread such as the random movement of individuals within the compartments.

The epidemic model used in this paper builds upon our previous work reported in Majid et al. (2021) & Majid et al. (2022) that focused on developing and validating a PDE-based $SLIR$ epidemic modeling framework. As such, the epidemic model includes the S, L, I, and R compartments. Here S represents the susceptible population, L the Latent population, I the infected population, and R the recovered population. The following equations represent the dynamics (Majid et al. (2021) & Majid et al. (2022)):

$$\frac{\partial S}{\partial t} = \lambda + \eta_S \nabla^a S - \theta S - u(x,y,t)\phi IS \tag{1}$$

$$\frac{\partial L}{\partial t} = u(x,y,t)\phi IS - u(x,y,t)\epsilon\phi \tag{2}$$
$$\times \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y,t-\tau) S(x,y,t-\tau) f_a(x,y) dxdy \right)$$

$$\frac{\partial I}{\partial t} = \eta_I \nabla^a I - \delta I + u(x,y,t)\epsilon\phi \tag{3}$$
$$\times \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y,t-\tau) S(x,y,t-\tau) f_a(x,y) dxdy \right)$$

$$\frac{\partial R}{\partial t} = \eta_R \nabla^a R - \theta R + \omega I \tag{4}$$

Where,

- λ - Birth rate in the domain.
- η_S - Diffusion coefficient representative of intensity of the random motion of the susceptible population.
- η_I - Diffusion coefficient representative of intensity of the random motion of the infected population.
- η_R - Diffusion coefficient representative of intensity of the random motion of the recovered population.
- θ - Mortality rate due to natural deaths.
- ϕ - Infection rate.
- δ - Removal of individuals from the infected compartment due to death or recovery.
- ω - Recovery rate from COVID-19.
- ϵ - Fraction of the infected population that survive the latency period and enter the infected category.
- $u(x,y,t)$ - The control parameter which is a function

of space and time.
$f_a(x, y)$ - Gaussian kernel defining the extent of the mobility of the latent population.

To facilitate the numerical computation, Eqs. 1 - 4 are discretized by the application of Euler's forward method (Majid et al. (2021, 2022)):

$$S_{x,y}^{T+1} = N\lambda + \eta_S(S_{x,y+1}^T + S_{x,y-1}^T + S_{x+1,y}^T + S_{x-1,y}^T) + (1 - 4\eta_S - \theta - u\frac{\phi}{N}I_{x,y}^T)S_{x,y}^T \quad (5)$$

$$L_{x,y}^{T+1} = u\frac{\phi}{N}I_{x,y}^T S_{x,y}^T - u\epsilon\frac{\phi}{N}\sigma_{x,y}^T \quad (6)$$

$$I_{x,y}^{T+1} = \eta_I(I_{x,y+1}^T + I_{x,y-1}^T + I_{x+1,y}^T + I_{x-1,y}^T) + (1 - 4\eta_I - \delta)I_{x,y}^T + \epsilon u\frac{\phi}{N}\sigma_{x,y}^T \quad (7)$$

$$R_{x,y}^{T+1} = \eta_R(R_{x,y+1}^T + R^T x, y - 1 + R_{x+1,y}^T + R_{x-1,y}^T) + (1 - 4\eta_R - \theta)R_{x,y}^T + \omega I_{x,y}^T \quad (8)$$

$$\sigma_{x,y}^{T+1} = \sum_{i=1}^{n}\sum_{j=1}^{n} I_{x,y}^{T-\tau} S_{x,y}^{T-\tau} f_a l^2 \quad (9)$$

$$f_a(x, y) = \frac{1}{\sqrt{4\pi\alpha}} e^{-\frac{(i-x)^2+(j-y)^2}{4\alpha}} \quad (10)$$

Here $X_{x,y}^{T+1}$ denotes the value of each of the compartment variables, i.e., $(X \in \{S, L, I, R\})$, on the day $(T+1)$ corresponding to the cell that is represented by the discretized spatial coordinates (x, y) in the two-dimensional domain under consideration. Here, n is the number of cells in one dimension, N is the unit population and σ^T is the latent population corresponding to a cell on the given day (T).

2.2 Deep Learning and LSTM Models

Machine learning models are a class of Artificial Intelligence (AI) algorithms that are designed to automatically learn the specific patterns of hierarchial representations from large amount of training data. An Artificial Neural Network (ANN) is an algorithmic framework that mimics the working principle of the human brain. Since the advent of ANN in 1980s, research in this field has become both extensive and diverse. The article Basheer and Hajmeer (2000) presents a good overview of ANN and its applications. A Deep Neural Network (DNN) can be viewed as a more complex version of an ANN comprising of multiple hidden layers that enable it to learn the pattern of features from the input data. A Recurrent Neural Network (RNN) is a type of DNN that can learn from the past, by using feedback loops, and consider this information while making a future decision. The structure of an RNN is provided in the Fig. 1. where,

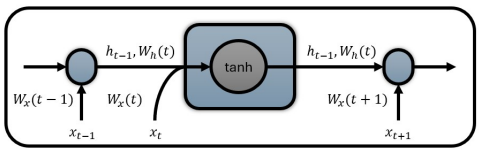

Fig. 1. RNN Structure

x_t - Input at time t
h_t - Hidden memory of the cell at time t
$W_{x(t)}$ - Weight matrix of x at time t
$W_{h(t)}$ - Weight matrix of h_{t-1} at time t

At each time step, the input to the current cell, which is the hidden memory, and the stored memory from the previous cell are combined into a new vector. The new vector contains information from both the previous cell and the new hidden memory which is then processed using a non-linear activation function - typically either the sigmoid or the the hyperbolic tangent $tanh()$, the latter given in Eqn. 11.

$$h_t = tanh(W_{h(t)} * h_{t-1} + W_{x(t)} * x_t + b) \quad (11)$$

For the goals of this research, RNNs can be implemented to generate the parameters of interest. However, we focus on the Long-Short-Term-Memory (LSTM) model first reported in Hochreiter and Schmidhuber (1997). The LSTM model is a type of RNN capable of accounting for long-term factors and sequentially recording updated information. The structure of the LSTM model used in this work is presented in Figure 2.

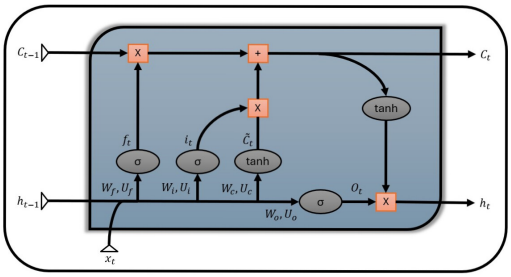

Fig. 2. LSTM Structure

The cell state combined with the three gates forms the backbone of the LSTM model. We now discuss the details.

Cell State: The cell state acts as a memory reservoir active through the entire processing sequence, also capable of capturing necessary details throughout the process and transmitting them forward. It also manages the retention and transmission of information across the course of the pattern, effectively acting as the neural network's "memory box". As the sequential processing of data unfolds, the cell state - influenced by the various gates - dynamically includes or excludes information. These gates learn information that needs to be retained or discarded from the memory during the training phase. The cell state is obtained using the following equation:

$$C_t = (f_t)(C_{t-1}) + (i_t)(\tilde{C}_t) \quad (12)$$

The 3 different gates interacting with the recurrent module of an LSTM network are given below (Feng et al. (2022)).

Forget Gate: The forget gate in an LSTM determines the information from the previous cell state to be discarded and aids in memory management and preventing irrelevant information from being propagated further. The forget gate performs the following computation:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (13)$$

Input Gate: The input gate regulates the flow of new information into the cell state, allowing the LSTM to selectively add relevant new information to the old information while filtering out irrelevant data.

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (14)$$

$$\tilde{C}_t = tanh(W_C x_t + U_C h_{t-1} + b_C) \quad (15)$$

Output Gate: The output gate controls the flow of information from the cell state to the output, determining the

information to be passed on to the next recurrent module of the LSTM model:

$$O_t = \sigma(W_O x_t + U_O h_{t-1} + b_O) \quad (16)$$

$$h_t = (O_t)(tanh(C_t)) \quad (17)$$

3. MODEL PARAMETER OPTIMIZATION

The parameter optimization of the S, L, I, R compartments represents the learning phase where the optimal model parameters are obtained through comparison with the ground truth dataset. To perform the optimization of the model parameters, we make use of the MATLAB Optimization Toolbox and the *fmincon* (a gradient descent-based method) to minimize the error of the following cost function:

$$L_m = \sqrt{\frac{1}{3T} \sum_{X \in \{S,I,R\}} \sum_{n=1}^{N_c} \sum_{t=1}^{T} (X(n,t) - \hat{X}(n,t))^2} \quad (18)$$

Where,

X - Ground truth $S - I - R$ data of the n^{th} county
$\hat{X}$ - Model output of the n^{th} county
L_m - Model Error (Computed using RMSE Approach)
N_c - Total Number of Counties in Ohio (88 in Total)
T - Simulation Duration (Total Number of Days)
t - Time Index (counted in days)
n - County Index

We carry out the parameter optimization process in three different phases. We optimize the model parameters for T number of days, which we call the training period. Following this, depending on the use case, we employ those model parameters to extrapolate the results from the PDE model, which is called the testing period.

3.1 Phase-1 (Static Initial Parameters)

The parameters of the PDE model in Eqn. 5 to Eqn. 8 govern various aspects of the dynamics. Optimizing these parameters by fitting real-time data ensures that the model is appropriately calibrated and reflects the ground truth. As mentioned earlier, we use a gradient-based approach to search for an optimal set of parameters. The approach minimizes the objective function given by Eqn. 18 that represents the error between the model prediction and the ground truth. During this phase, we optimize all 10 parameters in Eqs. 1-4 of the model using the the previous 30 day's ground truth data ($T = 30$). These parameters are then used as a baseline in $Phase-2$ of optimizations and also in the final model evaluation.

3.2 Phase-2 (Infection Parameters Evolving in Time)

While the parameters of the model from $Phase-1$ are obtained based on the past 30-day data, it is to be noted that there is a continuous temporal evolution of all the parameters - especially the ones that maximally influence the infection spread. This is due to the fundamentally dynamic nature of the infection spread induced by temporal variations in the transmissibility characteristics of the virus (or pathogen, in general). Additionally, factors such as human mobility and immunity towards the virus also evolve. Indeed, an accurate modeling framework must account for these dynamically changing features by allowing the model parameters to correspondingly change. For this reason, we propose to adapt the daily values of those parameters based on a prediction algorithm described below. In this paper, we choose three of the above ten parameters which directly influence the dynamics of infection spread. These *parameters* are: infection rate (ϕ), the diffusion coefficient of Susceptible population (η_S), and the diffusion coefficient of Infected population (η_I). These parameters are considered to adaptively change daily in our prediction framework. To capture this temporal evolution of infection parameters, we run the optimization process iteratively for each day spanning the next 30 days by looking at the previous T_r days of ground truth data. In this we have considered two values: ($T_r = 5$) and ($T_r = 10$). In this process, out of the ten total number of parameters, seven are assumed to be static and the remaining three (ϕ, η_S, and η_I) are optimized daily. The rationale for considering a shorter time frame for $Phase-2$ of parameter optimization is based on two reasons: 1. To capture the recent trend in the progression of the temporal parameters, and 2. To reduce the computational burden. After completing the optimization in $Phase-2$, the day-to-day values of the infection parameters along with the static parameters evaluated in $Phase-1$ are passed on to the next step - Training LSTM to forecast future parameters.

3.3 Phase-3 (Training an LSTM for infection parameters)

The novelty in the proposed work is the use of an LSTM network to predict the daily values of infection parameters. The LSTM is trained on the initial pattern of the parameters obtained from $Phase-2$. Among the multiple deep neural network architectures available, an LSTM is used in this study as it is particularly effective in modeling temporal patterns and sequences. The LSTM's potential to learn and replicate the intricate temporal structures along with its ability to selectively remember or forget information using the gate mechanism makes it well suited for our task. To train the LSTM network, we input the parameter values from the preceding T_r days as the training data. The output of the LSTM corresponds to the values of the parameters for the current day. It may be noted that since the training data generation (as mentioned in Phase 2) requires previous T_r days, we will not be having training data for the initial T_r days. To address this, we utilized a constant padding scheme wherein we use the static parameter values available from $Phase-1$ as a way of data imputation for the missing values. A simple pseudo algorithm for the LSTM training operation is given below in Alg.1.

4. CASE STUDY: OHIO, USA

The approach discussed earlier is utilized to optimize the day-to-day infection parameters of the PDE model applied to the State of Ohio using the ground truth COVID-19 data available for the period April 20 - June 05, 2020. This process involves dividing the shapefile for the state of Ohio into a mesh grid of size 50×50 where each grid is assigned the following values: 1. The name of the county the cell belongs to, 2. A logical value to state if the cell belongs to the state of Ohio, 3. A logical value to state if the cell belongs to the boundary, and 4. The initial values

Algorithm 1 Training an LSTM for time-varying infection parameters

Input:
data (day-to-day infection parameters from Phase-2):
Parameters = $[[\eta_{S,1}, \phi_1, \eta_{I,1}], \ldots, [\eta_{S,n}, \phi_n, \eta_{I,n}]]$
epochs = 500; timesteps = 5 *and* 10;
batch size = 30; testing period = 15;
features = $[\eta_S, \phi, \eta]$
Layers = [sequenceInputLayer(length(features)),
lstmLayer(128, 'OutputMode','last'),
fullyConnectedLayer(length(features)), regressionLayer]
Options = trainingOptions(optimizer = "adam", ...
loss = "mse")

Output:
Fitting and prediction of η_S, ϕ *and* η_I

Params = Parameters;
for day = 1:testing period **do**
  # Creating a data structure with time steps
  x_train = []; y_train = []
  for i = 1:batch size **do**
    $x\_train(i)$ = Params(i − timesteps : i − 1);
    $y\_train(i)$ = Params(i)
  end for
  # Train the LSTM on Input Data
  procedure LSTM TRAINING(Layers)
    net = trainNetwork(x_train, y_train, Layers, ...
      Options);
    return model
  end procedure
  # Forecast
  procedure FORECAST_LSTM(net)
    y_predict = predict(net, x_train)
    $y\_predict$ = model.predict(x_train)
    return $y\_predict$
  end procedure
  # Consider the adjacent 30 day's data
  Parameters(length(Parameters)+1) = y_predict;
  Params = Parameters(1+day:n+day);
end for

---

of S, I and R. The methodology employed in our previous study by Majid et al. (2021, 2022) is being deployed to predict the trajectories of S, I, and R for Ohio throughout the simulation. The empirical COVID-19 dataset for the state of Ohio, obtained from The New York Times (2021) is considered the ground truth dataset. We note that the dataset contains the daily count of infections and deaths across each county in Ohio and we assume that 75% of people who got infected 10 days before the current day are recovered. The specific timeframe from April 20, 2020 to June 05, 2020 is being considered for this analysis. The infection counts for a specific county in the ground truth dataset is evenly distributed among all cells corresponding to that county within the grid structure. The initial value of S is the total population of the county, and for each following day, we subtract the value of I from the previous day's S. After the computational environment is set up in this manner, we carry out the process of parameter optimization, training, and prediction. The flowchart of the process is presented in Fig. 3

5. RESULTS AND DISCUSSIONS

The results are now presented. In each of the figures (Figs. 3-5) presented below, the model training happens using data from Day 0-30. The trained model is used to predict

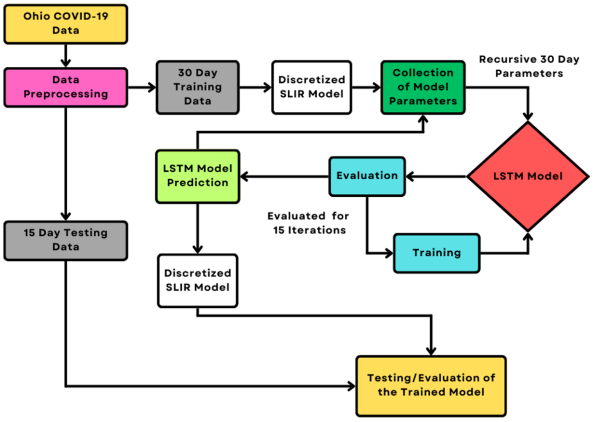

Fig. 3. Parameter Optimization Flowchart

the infection spread for the testing period from Day 31-45. Hence, in each figure the blue lines represent the validation phase (where the trained model is used to obtain results for the training period of Day 0-30) and the dotted orange line shows the testing phase where the trained model is used for prediction in the testing period of Day 31-45. Fig. 4 shows the results for Test Case -1 in which the PDE model outputs using the static parameters obtained from $Phase-1$ of optimization. As noted in Section 3.2, day-to-day optimization of the infection parameters was conducted for two distinct scenarios of T_r (number of previous day's data for LSTM training): a 5-day case (Test Case -2) and a 10-day case (Test Case-3), and their results are presented in Fig. 5 and 6 respectively. For Test Case-1, the outputs from the model that uses static parameters deviate from the ground truth significantly as the change in spread dynamics is not captured using static parameters ($Phase-1$). Conversely, while using the day-to-day infection parameters using the adaptive LSTM approach proposed in $Phase-2$ & 3, the model prediction more closely aligns with the progression of the ground truth data. Root mean square error (RMSE) plots comparing the predictions from the three test cases with the ground truth are shown in Fig. 7. This evaluation shows that while the 10-day optimization case performs slightly better than the 5-day case during the training phase, the model outputs using the LSTM-predicted parameters in the 10-day case exhibit a slight divergence from the ground truth compared to the 5-day case for the prediction/testing phase. This could be because to the larger variance in input data for the 10-day case and also due to the rapidly evolving attribute of the pathogen is well captured by 5-day case.

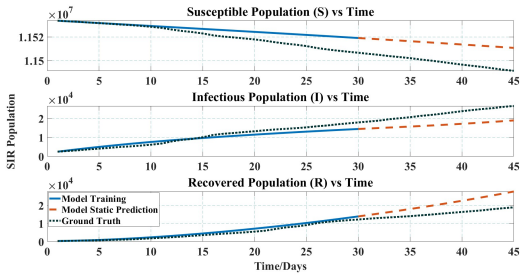

Fig. 4. Static Parameters (Test Case-1)

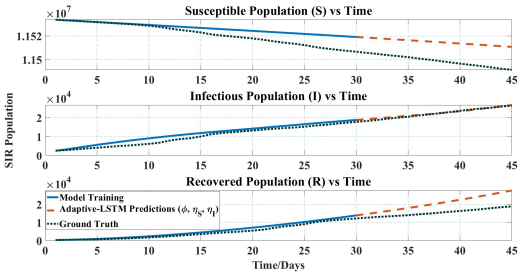

Fig. 5. LSTM Predictions for parameters optimized using 5 day data (Test Case-2)

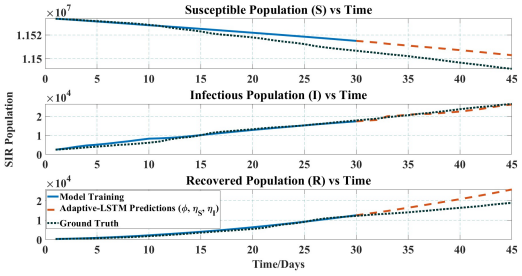

Fig. 6. LSTM Predictions for parameters optimized using 10 day data (Test Case-3)

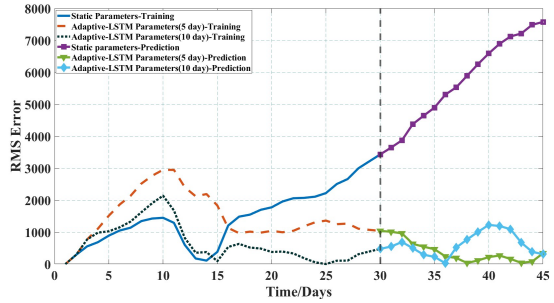

Fig. 7. RMSE of the three test cases with the ground truth

6. CONCLUSIONS

We present a novel framework that integrates a spatiotemporal $SLIR$ epidemic model with an LSTM network to improve the accuracy of epidemic spread forecast by the model. This is achieved by adaptively generating significant model parameters using a pre-trained LSTM. With previous parameter values as input, the LSTM yields parameter values for the current day. Using empirical COVID-19 data for the state of Ohio, the PDE model parameters are optimized to extract the initial pattern of daily infection parameters. This pattern is then used to train an LSTM network for predicting future parameter values on a daily basis. We then apply the parameters learned by the LSTM network to simulate the PDE-based $SLIR$ model and compare the results with those obtained using static parameters. The framework not only enhances the predictive capabilities of the epidemic model but also yields deeper insights into the spatiotemporal spread dynamics. Moreover, forecasts obtained from applying the framework can help public health systems respond more effectively to future epidemics by supporting the development of effective interventions in real time. Potential future research directions include the implementation of advanced optimization techniques to improve computational efficiency and the exploration of data from longer time periods as well as larger spatial domains.

REFERENCES

Basheer, I. and Hajmeer, M. (2000). Artificial Neural Networks: Fundamentals, Computing, Design, & Application. *Journal of Microbiological Methods*, 43(1), 3–31.

Brauer, F., van den Driessche, P., and Wu, J. (2008). *Mathematical Epidemiology*. Lecture Notes in Mathematics. Springer, Berlin/Heidelberg.

CDC (2023). CDC Museum COVID-19 Timeline. Publication Title: Centers for Disease Control and Prevention.

Feng, L., Chen, Z., Lay Jr., H.A., Furati, K., and Khaliq, A. (2022). Data driven time-varying SEIR-LSTM/GRU algorithms to track the spread of COVID-19. *MBE*, 19(9), 8935–8962.

Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., and Colaneri, M. (2020). Modelling the COVID-19 Epidemic and Implementation of Population-Wide Interventions in Italy. *Nature Medicine*, 26, 855–860.

Hochreiter, S. and Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780.

Kermack, W.O. and McKendrick, A.G. (1927). A Contribution to the Mathematical Theory of Epidemics. *Proc. R. Soc. Lond. A*, 115(772), 700–721.

Li, J. and Zou, X. (2009). Modeling Spatial Spread of Infectious Diseases with a Fixed Latent Period in a Spatially Continuous Domain. *Bull. Math. Biol.*, 71(8), 2048–2079.

Majid, F., Deshpande, A.M., Ramakrishnan, S., Ehrlich, S., and Kumar, M. (2021). Analysis of Epidemic Spread Dynamics Using a PDE Model and COVID-19 Data from Hamilton County OH USA. *IFAC-PapersOnLine*, 54(20), 322–327.

Majid, F., Gray, M., Deshpande, A.M., Ramakrishnan, S., Kumar, M., and Ehrlich, S. (2022). Non-Pharmaceutical Interventions as Controls to mitigate the spread of epidemics: An analysis using a spatiotemporal PDE model and COVID–19 data. *ISA Transactions*, 124, 215–224.

Niraula, P., Mateu, J., and Chaudhuri, S. (2022). A Bayesian machine learning approach for spatio-temporal prediction of COVID-19 cases. *Stochastic Environmental Research and Risk Assessment*, 36(8), 2265–2283.

Olumoyin, K.D., Khaliq, A.Q.M., and Furati, K.M. (2021). Data-Driven Deep-Learning Algorithm for Asymptomatic COVID-19 Model with Varying Mitigation Measures and Transmission Rate. *Epidemiologia*, 2(4), 471–489.

Park, M., Cook, A.R., Lim, J.T., Sun, Y., and Dickens, B.L. (2020). A Systematic Review of COVID-19 Epidemiology Based on Current Evidence. *JCM*, 9(4), 967.

Rahimi, I., Chen, F., and Gandomi, A.H. (2023). A review on COVID-19 forecasting models. *Neural Comput & Applic*, 35(33), 23671–23681.

Setianto, S. and Hidayat, D. (2023). Modeling the Time-Dependent Transmission Rate Using Gaussian Pulses for Analyzing the COVID-19 Outbreaks in the World. *Scientific Reports*, 13(1), 4466.

The New York Times (2021). U.S. COVID-19 Cases and Deaths.