




## Exploring the potential of learning methods and recurrent dynamic model with vaccination: A comparative case study of COVID-19 in Austria, Brazil, and China

Seyed Ali Rakhshan <sup>\*</sup>, Marzie Zaj, and Fatemeh Helen Ghane <sup>†</sup>  
Department of Mathematics, Ferdowsi University of Mashhad, Mashhad, Iran

Mahdi Soltani Nejad  
Department of Railway Engineering, Iran University of Science and Technology, Tehran, Iran

 (Received 14 June 2023; accepted 11 December 2023; published 9 January 2024)

In order to effectively manage infectious diseases, it is crucial to understand the interplay between disease dynamics and human conduct. Various factors can impact the control of an epidemic, including social interventions, adherence to health protocols, mask-wearing, and vaccination. This article presents the development of an innovative hybrid model, known as the Combined Dynamic-Learning Model, that integrates classical recurrent dynamic models with four different learning methods. The model is composed of two approaches: The first approach introduces a traditional dynamic model that focuses on analyzing the impact of vaccination on the occurrence of an epidemic, and the second approach employs various learning methods to forecast the potential outcomes of an epidemic. Furthermore, our numerical results offer an interesting comparison between the traditional approach and modern learning techniques. Our classic dynamic model is a compartmental model that aims to analyze and forecast the diffusion of epidemics. The model we propose has a recurrent structure with piecewise constant parameters and includes compartments for susceptible, exposed, vaccinated, infected, and recovered individuals. This model can accurately mirror the dynamics of infectious diseases, which enables us to evaluate the impact of restrictive measures on the spread of diseases. We conduct a comprehensive dynamic analysis of our model. Additionally, we suggest an optimal numerical design to determine the parameters of the system. Also, we use regression tree learning, bidirectional long short-term memory, gated recurrent unit, and a combined deep learning method for training and evaluation of an epidemic. In the final section of our paper, we apply these methods to recently published data on COVID-19 in Austria, Brazil, and China from 26 February 2021 to 4 August 2021, which is when vaccination efforts began. To evaluate the numerical results, we utilized various metrics such as RMSE and R-squared. Our findings suggest that the dynamic model is ideal for long-term analysis, data fitting, and identifying parameters that impact epidemics. However, it is not as effective as the supervised learning method for making long-term forecasts. On the other hand, supervised learning techniques, compared to dynamic models, are more effective for predicting the spread of diseases, but not for analyzing the behavior of epidemics.

DOI: [10.1103/PhysRevE.109.014212](https://doi.org/10.1103/PhysRevE.109.014212)

### I. INTRODUCTION

The global community is increasingly focusing on research and control measures to curb the spread of epidemics caused by various viruses leading to infectious diseases [1–4]. Predicting the outbreak of such diseases is crucial in epidemiology, as it enables health officials to prepare and respond to potential outbreaks effectively.

Compartmental dynamic models serve as flexible tools for predicting different epidemiological parameters [5]. They can also evaluate the impact of lockdowns and other disease control measures on disease spread [6–11], the efficacy of vaccines, and their distribution strategies [12–16]. For instance, in a study by Faranda *et al.* [17], a susceptible-exposed-infected-recovered (SEIR) model was used to forecast the long-term dynamics of an epidemic in two countries. The model was

adjusted for uncertainties in estimating COVID-19 prevalence and the presence of superspreaders using stochastic perturbations and a lognormal distribution. Another study by Alberti *et al.* [18] focused on statistical predictions of COVID-19 infections by fitting asymptotic distributions to real-world data. The results showed significant uncertainties in predictions during the initial stages of the epidemic's growth, indicating the challenges in accurately forecasting the spread of COVID-19 due to limited data and inherent complexities.

It is also worth noting that many infectious diseases, including COVID-19, exhibit recurrent behavior [19–21]. Therefore, incorporating recurrence in epidemic models is crucial [22]. A recurrent compartmental model, a more complex version of the common model, suggests that the number of susceptible individuals can increase over time, leading to potential recurrent outbreaks. This model can help researchers better understand disease dynamics and the impact of different interventions on disease transmission [23,24].

The dynamics of disease immunity and transmission in a society in the setting of vaccination can be modeled by

<sup>\*</sup>seyedalirakhshan@yahoo.com

<sup>†</sup>ghane@math.um.ac.ir

epidemiological models. There are different types of works on the coronavirus infection with dynamical analysis approaches to controls the infection of disease [25,26], especially, models including the epidemiological impact of vaccination has a long and rich history.

Epidemic models with vital dynamics are a category of mathematical models that take into account the interplay between disease dynamics and the demographic attributes of a population. These models serve as valuable tools for gaining insights into the long-term impact of diseases on populations and the influence of factors like birth and death rates on the progression of an epidemic [27–29]. These rates have a significant influence on population size, and vulnerability, thereby playing a pivotal role in shaping the transmission and consequences of epidemics. Vital dynamics models have been employed to mathematically model COVID-19 [30,31]. Birth and death rates have also a significant impact on the stability of disease equilibria. These rates can lead to the emergence of bistable states, where both disease-free and endemic states coexist [32–34].

While dynamic models are beneficial for analyzing epidemic diseases, deep learning and machine learning models have emerged as promising tools for forecasting disease outbreaks. They can analyze large data sets and identify patterns that traditional methods might miss. Consequently, these learning methods are often used to predict the likely number of COVID-19 cases and deaths [35–39]. Advanced techniques like intelligent algorithms and machine learning methods are increasingly being used to solve mathematical biology problems, including modeling and analyzing infectious diseases like COVID-19. These methods can help researchers identify patterns and trends in epidemiological data, develop more accurate predictive models, and optimize disease control strategies [40–43].

For example, a study by Muhammad *et al.* [44] used various machine learning algorithms to develop models for predicting COVID-19 patient recovery. The decision tree model was found to be the most efficient, with an overall accuracy of 99.85%. In another study [45], Convolutional neural network were used to automatically predict the prevalence of COVID-19 cases. Rashed *et al.* [40] used machine learning based on long short-term memory (LSTM) to predict the spread of COVID-19, demonstrating the effectiveness of vaccination in controlling the outbreak. Syeda *et al.* [46] evaluated six different models, including artificial neural networks, convolutional neural networks, LSTM, and recurrent neural networks, to accurately identify COVID-19 patients. Their predictive models achieved an accuracy of 86.66%, F1 score of 91.89%, precision of 86.75%, recall of 99.42%, and Area under the curve of 62.50%. They also discussed the role of artificial intelligence in combating COVID-19, specifically in epidemiology, diagnosis, and disease progression. A study by Rasmy *et al.* [47] developed a recurrent neural network-based model that uses online data from electronic health records to predict outcomes of COVID-19 patients without feature selection or missing data imputation. These studies highlight the potential of machine learning and deep learning techniques in accurately diagnosing and managing COVID-19, enabling healthcare providers to make informed decisions and control the disease spread.

This article presents the combined dynamic-learning model, an approach that merges classical dynamic models and contemporary learning methods to analyze and predict infectious diseases, with a particular focus on COVID-19 over a 160-day period. The study emphasizes the integration of dynamic systems theory for understanding complex nonlinear processes related to disease recurrence, along with advanced learning techniques for processing large data sets and identifying intricate patterns. The training and testing data undergo processing using decision tree learning, Bidirectional long short-term memory (BiLSTM), gated recurrent unit (GRU), and a combined deep learning method. The model also incorporates considerations for vaccinations and acknowledges partial immunity in vaccinated individuals or those with active antibodies. It utilizes officially reported data for parameter estimation and accommodates uncertainties through a stochastic approach. This innovative model aims to enhance our understanding of disease dynamics and improve our forecasting capabilities, ultimately contributing to more effective prevention and treatment strategies.

## II. METHODS

### A. The proposed Combined Dynamic-Learning Model

The proposed dynamic-learning model in this study integrates two approaches: a recurrent compartmental model to analyze disease outbreaks and fit the available data, and learning algorithms to enhance prediction accuracy for future outbreak scenarios. The model utilizes dynamic systems theory in conjunction with various learning methods, including decision tree learning, bidirectional long short-term memory (BiLSTM), GRU, and a combined deep learning approach.

The recurrent compartmental epidemic model uses piecewise constant parameters to fit released data and provide various analyses of epidemic behavior. This model is particularly effective in understanding the dynamics of disease immunity and transmission in a society, especially in the context of vaccination.

The model also accounts for temporal variability in disease outbreaks, which can exhibit gradual or abrupt changes over time. By calibrating the model’s parameters based on real epidemic data and assuming piecewise constant parameters, the model aims to capture the fluctuations in the values of parameters and the number of population groups resulting from control measures implemented by authorities.

In our model, we have divided the total population into five categories based on their disease status: Susceptible ( $S$ ), Exposed ( $E$ ), Infectious ( $I$ ), Vaccinated ( $V$ ), and Recovered ( $R$ ). To simplify our calculations, we use fractions instead. Thus,  $S, E, V, I,$  and  $R,$  respectively, represent the fraction of the population that is susceptible, exposed, vaccinated, infectious, and recovered. The interactions among these categories are illustrated in Fig. 1.

The model presented in this article consists of a set of differential equations that represent the progression of the population in each compartment. The model is

$$\frac{dS(t)}{dt} = \sigma - \beta(t)S(t)E(t) + \zeta(t)V(t) + \rho(t)R(t) - (v(t) + \mu)S(t),$$

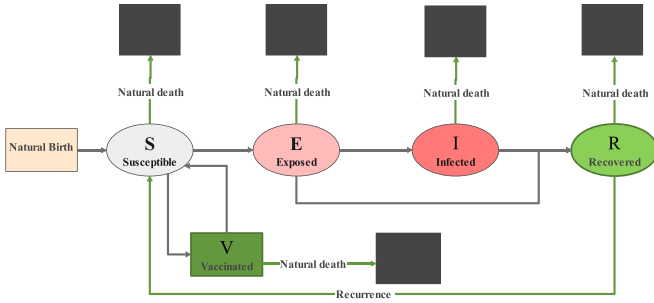


FIG. 1. Scheme of the SEIRS-V model transmission.

$$\begin{aligned}
 \frac{dE(t)}{dt} &= \beta(t)S(t)E(t) - \alpha(t)E(t) - \eta(t)E(t) - \mu E(t), \\
 \frac{dV(t)}{dt} &= v(t)S(t) - \zeta(t)V(t) - \mu V(t), \\
 \frac{dI(t)}{dt} &= \alpha(t)E(t) - \gamma(t)I(t) - \mu I(t), \\
 \frac{dR(t)}{dt} &= \gamma(t)I(t) + \eta(t)E(t) - \rho(t)R(t) - \mu R(t). \quad (1)
 \end{aligned}$$

The parameters of the model are piecewise constant and positive, each representing different aspects of disease transmission and progression. These include the transmission rate,  $\beta(t)$ , diagnosis rate,  $\alpha(t)$ , recovery rate,  $\gamma(t)$ , rate of recovered individuals becoming susceptible again,  $\rho(t)$ , vaccination rate,  $v(t)$ , waning rate of the vaccine,  $\zeta(t)$ , symptomatic recovery rate,  $\eta(t)$ , constant natural birth rate,  $\sigma$ , and constant natural death rate,  $\mu$ . The model assumes that a specific period of interest can be divided into multiple stages, each characterized by relatively stable parameters within the virus-human system. However, these parameters may vary across different stages due to changes in population behavior, interventions, or the natural evolution of the pathogen.

The model simplifies several aspects as assumptions, including limited immunity leading to susceptibility after recovery, the possibility of vaccinated individuals getting infected, and the consideration of birth and natural death rates. Incorporating birth and death processes in our model offers a more comprehensive representation of population dynamics, aligning with real-world scenarios and allowing us to capture observed demographic changes [48]. Given the sustained presence of the coronavirus, we include both birth and death rates in our model. Infection-related deaths are often negligible compared to the natural death rate.

In the second part, this research offers a description and comparison of some learning models. ‘‘Learning’’ refers to the use of machine learning or statistical techniques to extract patterns and insights from data. In the context of epidemic analysis and forecasting, learning methods can be applied to historical data on epidemic outbreaks, population demographics, intervention measures, and other relevant factors [40–43]. These methods can identify relationships, trends, and hidden patterns in the data, enabling the development of predictive models. We use some supervised learning methods that can be used to predict outbreaks and control the spread of infectious diseases. We also propose an efficient deep learning method that combines two

well-known learning methods to improve the accuracy of predictions. The architecture of the learning methods used in this article is shown in Fig. 2, which includes both a feedforward neural network and a recurrent neural network. The feedforward network can capture complex relationships between input features and output predictions, while the recurrent network can model dynamic temporal patterns in the data. Based on Fig. 2, the learning methods employed in the article can be classified into the following categories:

(1) Regression tree learning (RT): Decision tree learning is a widely used supervised learning method that constructs a hierarchical structure, called a decision tree, to model data. Decision trees are composed of nodes representing features or attributes and edges representing relationships between them [49]. They are effective for both classification and regression tasks, especially when dealing with complex relationships and numerous features. The accuracy of a decision tree relies on strategic splits, where the most informative feature is chosen at each node. Information gain and other metrics are used to maximize separation between different classes or categories. Classification trees are used for discrete outputs, while regression trees are used for continuous outputs. Algorithms like the Gini index, entropy, and information gain determine optimal splits and subnode creation. The cost function evaluates model performance and aids in split point selection, with the goal of minimizing prediction error. The sum of squared errors is commonly used for regression trees. By minimizing the cost function, the decision tree identifies optimal splits and creates subnodes for improved accuracy. Homogeneous branches, exhibiting similar responses, are sought to enhance predictability. The decision tree evaluates each feature in the training data, calculates costs for potential splits, and selects the split with the lowest cost [50].

(2) Gated recurrent unit (GRU): Gated recurrent units (GRUs) were introduced in 2014 by Cho *et al.* as a gating mechanism for recurrent neural networks (RNNs) [51]. They are similar to LSTM networks but have fewer parameters and lack output gates, making them more computationally efficient. GRUs use gating mechanisms, including reset and update gates, to control information flow and address the vanishing gradient problem in traditional RNNs. The reset gate determines what information to discard from the previous time step, while the update gate controls the amount of information to keep from the current time step. GRUs perform similarly to LSTMs in tasks like polyphonic music modeling, speech signal modeling, and natural language processing. However, the choice between GRUs and LSTMs depends on the specific task and available data. Full-gated units have variations based on previous hidden states and bias, and there is a simplified version called the minimal-gated unit. The fully gated unit is defined as follows [52]:

$$z_t = \sigma_g(W_z x_t + U_z h_{t-1} + b_z), \quad (2)$$

$$r_t = \sigma_g(W_r x_t + U_r h_{t-1} + b_r), \quad (3)$$

$$\hat{h}_t = \phi_h(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h), \quad (4)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \hat{h}_t, \quad (5)$$

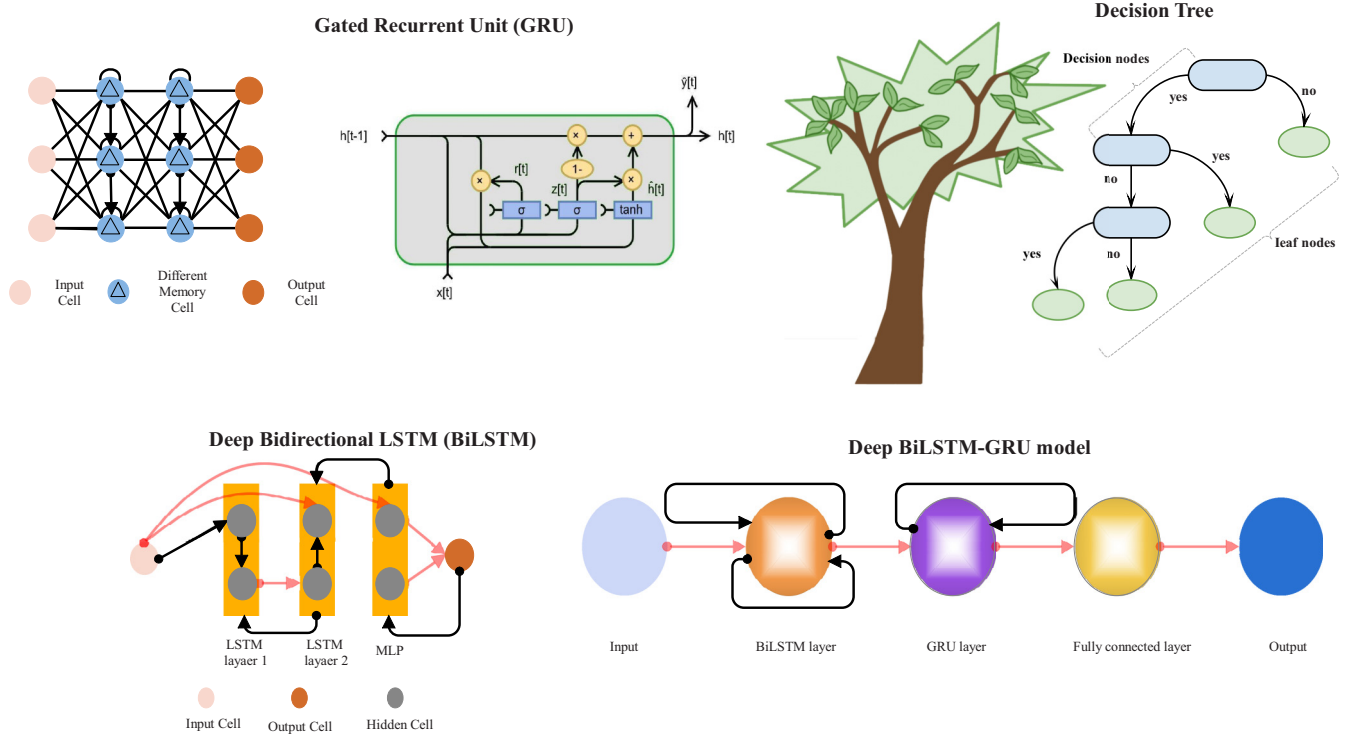


FIG. 2. Architecture of learning methods.

where the operator  $\odot$  denotes the Hadamard product, the subscript  $t$  indexes the time step,  $x_t$  denotes the input vector,  $h_t$  denotes output vector,  $\hat{h}_t$  denotes activation vector,  $z_t$  denotes update gate vector,  $r_t$  denotes reset gate vector, and  $W$ ,  $U$ , and  $b$  denote parameter matrices and vector. Active functions are as follows:

- (a)  $\sigma_g$ : sigmoid function;
- (b)  $\sigma_h$ : hyperbolic tangent function.

Additionally, the minimal gated unit is similar to the fully gated unit, except that the reset and update gates are merged into a single forget gate. In this case, we must also change the equation for the output vector:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f), \quad (6)$$

$$\hat{h}_t = \phi_h(W_h x_t + U_h (f_t \odot h_{t-1}) + b_h), \quad (7)$$

$$h_t = (1 - f_t) \odot h_{t-1} + f_t \odot \hat{h}_t. \quad (8)$$

(3) Bidirectional long short-term memory (BiLSTM): A BiLSTM is a RNN that combines two LSTMs, one moving in the forward direction and the other in the backward direction, to process sequential data. The forward LSTM reads the input sequence in a forward direction, while the backward LSTM reads it in reverse. This enables the BiLSTM to capture information from both past and future inputs, enhancing contextual understanding and increasing available information. During training, input data is first fed into the forward LSTM, followed by training with a reversed input sequence using another LSTM model [53]. This allows the model to capture both forward and backward contextual information, making it highly effective for tasks involving sequence element relationships. BiLSTMs have proven effective in various applications, including natural language processing, speech

recognition, and music modeling. They are particularly valuable for predicting and managing infectious disease outbreaks. By capturing both forward and backward contextual information, BiLSTMs offer insights into temporal dynamics and transmission patterns, aiding in the formulation of public health policies and interventions. Studies have demonstrated that BiLSTM models outperform regular LSTMs [54].

(4) New combined method BiLSTM-GRU: The BiLSTM-GRU is a combined method that leverages the strengths of both BiLSTM and GRU neural network architectures [55]. To train our data, we adopt a combined approach that leverages the strengths of BiLSTM and GRUs to create a powerful deep-learning model. Our network architecture, as illustrated in Fig. 2, comprises three layers. The first layer employs a BiLSTM network, which processes input data in both forward and backward directions, capturing information from past and future inputs. This layer enables the network to learn intricate temporal patterns and dependences in the data. The output of the BiLSTM layer serves as input to the second layer, which features a GRU network. The GRU network utilizes its gating mechanism to regulate information flow and mitigate the vanishing gradient problem often encountered in conventional RNNs. The final layer of our network is a fully connected layer responsible for generating the ultimate prediction for the output variable. By integrating the strengths of BiLSTM and GRU, our network can effectively capture complex and nuanced relationships in the data, resulting in enhanced predictions and more efficient control strategies for managing infectious disease outbreaks.

By integrating modeling and learning, this method aims to provide a more comprehensive understanding of epidemic dynamics. It leverages the strengths of both approaches to capture the complexity of disease spread, consider the impact

of various factors (such as interventions and human behavior), and make accurate forecasts about future epidemic trends. The integrated approach allows for a more nuanced analysis of the dynamics of infectious diseases. It can help in evaluating the effectiveness of intervention strategies, assessing the potential impact of different control measures, and providing insights for public health decision making. Ultimately, the goal is to enhance our ability to manage and mitigate the spread of epidemics by combining the strengths of traditional modeling and modern learning techniques.

**B. Mathematical analysis of dynamic model**

The qualitative analysis aims to establish the consistency between the numerical simulation and analytical results.

$$X_0 = \left( \frac{\zeta + \mu}{\zeta + v + \mu}, 0, \frac{v}{\zeta + v + \mu}, 0, 0 \right), \tag{9}$$

and

$$X_1 = (S^*, E^*, V^*, I^*, R^*), \quad S^* = \frac{\alpha + \eta + \mu}{\beta}, \quad V^* = \frac{v(\alpha + \eta + \mu)}{\beta(\zeta + \mu)}, \quad I^* = \frac{\alpha}{\mu + \gamma} E^*, \quad R^* = \frac{\gamma\alpha + \eta(\gamma + \mu)}{(\gamma + \mu)(\rho + \mu)} E^*,$$

$$E^* = \frac{(\zeta + \mu)(v + \mu)(\alpha + \eta + \mu) - \mu\beta(\zeta + \mu) - \zeta v(\alpha + \eta + \mu)}{\rho\gamma\alpha + \rho\eta(\gamma + \mu) - (\gamma + \mu)(\rho + \mu)(\alpha + \eta + \mu)} \frac{(\gamma + \mu)(\rho + \mu)}{(\zeta + \mu)\beta}. \tag{10}$$

The next theorem gives the basic reproduction number  $R_0$ . The basic reproduction number is known as the outbreak threshold of epidemiology models.

*Theorem B.2.* The basic reproduction number of model (1) is defined by

$$R_0 = \frac{\beta(\zeta + \mu)}{(\alpha + \eta + \mu)(\zeta + v + \mu)}. \tag{11}$$

*Proof.* See Supplemental Material [56]. ■

The following result presents the local stability condition for the disease-free equilibrium through the analysis of the eigenvalues of the Jacobian matrix of system (1) at this equilibrium.

*Theorem B.3.* The disease-free equilibrium  $X_0 = (\frac{\zeta + \mu}{\zeta + v + \mu}, 0, \frac{v}{\zeta + v + \mu}, 0, 0)$  of system (1) is locally asymptotically stable in  $\Xi$  if  $R_0 < 1$  and it is unstable if  $R_0 > 1$ .

*Proof.* See Supplemental Material [56]. ■

**C. Sensitivity analysis of  $R_0$**

We proceed with a sensitivity analysis of the model to examine its response to variations in each parameter. Understanding the sensitivity of the model is crucial in identifying the most influential parameters affecting disease transmission, particularly regarding the endemic threshold (10). Sensitivity analysis is a popular technique used to evaluate the robustness of model predictions to parameter values, especially when there is incomplete or inaccurate data. It is important to note that the sensitivity index can be affected by various parameters in the system, but it can also remain constant regardless of any parameter. This suggests that some parameters play a more significant role in determining the sensitivity index than others.

Moreover, qualitative analysis can present graphical representations of the behavior of the model's solutions. In this regard, we investigate some theorems and lemmas about the proposed model (1) as outlined below:

*Theorem B.1.* The simplex  $\Xi = \{(S, E, V, I, R) \in \mathbb{R}_+^5 : S(t), E(t), V(t), I(t), R(t) \geq 0, S(t) + E(t) + V(t) + I(t) + R(t) \leq 1\}$  is positively flow invariant for model (1).

*Proof.* See Supplemental Material [56]. ■

Our next objective is to estimate the reproduction number and show that the disease-free equilibrium point is locally asymptotically stable if the reproduction number is less than unity. The equilibrium points of the model can be easily obtained. The two disease-free equilibrium points,  $(X_0)$ , and the endemic equilibrium point,  $(X_1)$ , are as follows:

When estimating sensitive parameters, caution must be exercised as even slight perturbations in such parameters can lead to substantial quantitative changes in the model's predictions. Conversely, parameters with a low sensitivity index require less attention during estimation, suggesting that the model's predictions are relatively insensitive to changes in these parameters. Our objective is to conduct a sensitivity analysis on  $R_0 = \frac{\beta(\zeta + \mu)}{(\alpha + \eta + \mu)(\zeta + v + \mu)}$  with respect to the parameters  $\alpha, \beta, \zeta, \eta,$  and  $v$ .

*Definition C.1.* The normalized forward sensitivity index of  $R_0$  which is differentiable concerning a given parameter  $\theta$ , is defined by

$$\Upsilon_{\theta}^{R_0} = \frac{\partial R_0}{\partial \theta} \frac{\theta}{R_0}.$$

The sensitivity indices' values for the parameter values can be experimentally determined in a manner that accurately reflects the mathematical model's ability to describe real-world data. The next theorem examines the sensitivity of our system to changes in modeling parameters.

*Theorem C.1.* The normalized sensitivity index of  $R_0$  concerning model parameters in  $\Xi$  can be estimated as follows:

$$\Upsilon_{\beta}^{R_0} = 1, \quad -1 < \Upsilon_{\alpha}^{R_0} < 0, \quad -1 < \Upsilon_{\eta}^{R_0} < 0, \\ -1 < \Upsilon_v^{R_0} < 0, \quad 0 < \Upsilon_{\zeta}^{R_0} < 1. \tag{12}$$

*Proof.* See Supplemental Material [56]. ■

Based on these calculations and parameter conditions, we can conclude that

(1) The transmission rate  $\beta$  has the highest sensitivity index of 1, indicating that it has the most substantial impact on the model's predictions. For example, decreases (increases) in transmission rate by 10% would decrease (increase)  $R_0$  by

10%. Therefore, controlling this parameter may be one of the most effective ways to control the spread of the disease.

(2) Since  $0 < \frac{\alpha}{\alpha+\eta+\mu} < 1$ , thus it is trivial that  $-1 < \frac{-\alpha}{\alpha+\eta+\mu} < 0$ . So,  $\Upsilon_{\alpha}^{R_0}$  being negative which means that  $R_0$  decreases when  $\alpha$  increases. The same holds for  $\Upsilon_{\eta}^{R_0}$  and  $\Upsilon_{\mu}^{R_0}$ . In other words, increasing the values of these three parameters is beneficial to controlling the epidemic.

(3) It is obvious that  $\frac{\xi v}{(\xi+\mu)(\xi+v+\mu)}$  is always positive and smaller than 1. Note that,  $\Upsilon_{\xi}^{R_0}$  being positive means that increasing  $\xi$  implies growth of  $R_0$  and it has a negative impact on the spread of COVID-19.

### III. NUMERICAL ANALYSIS

#### A. Datasets

In our study, we utilized a data set consisting of COVID-19 reports from the World Health Organization (WHO) [57]. The data set included information on infected, recovered, and vaccinated individuals. Confirmed cases were classified based on the WHO's definition, which considered individuals positive for the nucleic acid amplification test or meeting clinical and/or epidemiological criteria and testing positive using SARS-CoV-2 Antigen-RDT. Additionally, we included individuals who had completed the recommended vaccination protocol, which typically involves receiving two doses, or in some cases, three doses for certain types of vaccines.

To analyze the COVID-19 outbreak, we selected three countries from different continents: Austria (Europe), Brazil (South America), and China (Asia). Daily data from February 26, 2021, to August 4, 2021, was divided into training and testing sets. Each country had 160 daily data points used for training, testing, and evaluating the proposed methods for predicting COVID-19 cases. This timeframe was chosen because it represented a critical period in the pandemic, marked by widespread vaccination efforts and the emergence of the Gamma variant in Brazil. Various vaccines, such as Johnson & Johnson's, AstraZeneca's, Pfizer-BioNTech, and Moderna, were being administered during this time. The Delta variant also became dominant, leading to a third wave of infections, particularly in the United States. Analyzing this period allowed us to assess the impact of vaccination on virus spread and develop accurate deep-learning models for predicting COVID-19 cases. We employed a dynamic model to calculate disease parameters and trained our learning algorithms and dynamic model using 80% of the data. The remaining 20% was used to evaluate the prediction process and assess model accuracy. By comparing the results of the dynamic model with the predictions generated by the deep learning model, we validated the effectiveness of our approach in predicting and controlling infectious disease outbreaks.

#### B. Estimating parameters and fitting process of the dynamic model and uncertainty quantification

In this study, we used MATLAB software to develop an optimization algorithm based on the least square method. The algorithm was employed to determine the parameters for the dynamic model proposed in the article. Key parameters, such as transmission rate, diagnosis rate, recovery rate, vaccination rate, waning rate of vaccine, recovery rate of exposed

individuals, and recurrence rate, were calculated for 160 days in the targeted countries. These parameters provided valuable insights into the dynamics of the COVID-19 outbreak in each country and could inform public health policies and interventions. The least-squares method optimization algorithm allowed for accurate parameter estimation and a better understanding of the impact of various factors on the virus's spread.

All parameters in the study were assumed to be piecewise constant, meaning they remained constant within specific time intervals. The 160 days were divided into 20 equal intervals, and the optimization algorithm was applied in each interval to estimate the parameters of the dynamic model. The parameter estimation results for Austria are presented in Fig. 3 (See Supplemental Material [56] for Brazil, and China), offering insights into the outbreak dynamics and aiding in policy and intervention decisions.

The optimization algorithm minimized the square of the absolute error in the objective function using data on infected, vaccinated, and recovered individuals. This allowed for parameter estimation and predictions of susceptible and exposed individuals in each country based on the estimated parameters and the assumption of a constant total population. Estimating the number of susceptible and exposed individuals provided a deeper understanding of the outbreak dynamics and identified areas for targeted interventions.

Figure 3 depicted valuable insights into the outbreak dynamics, revealing trends such as high transmission rates at the beginning of the period, and emphasizing the importance of early intervention measures. The figure also highlighted the high recurrence rate of the virus despite vaccination efforts, emphasizing the need for ongoing monitoring and control. Additionally, the low diagnosis and recovery rates indicated potential challenges in information, equipment, medical staff, and emerging variants. After estimating the parameters, the fourth-order Runge-Kutta method was used to solve the dynamic model and obtain numerical solutions for variables such as susceptible, exposed, vaccinated, infected, and recovered individuals. The model's output was compared with actual data, showing good curve fitting and demonstrating the accuracy and effectiveness of the proposed method in predicting outbreak dynamics. This comparison provided insights into the model's accuracy and areas for improvement.

In this research paper, we present an analysis that utilizes various sets of parameters and multiple stochastic realizations. Our methodology consists of several steps. Firstly, we identify the parameter set that best fits the data, exhibiting the lowest error in daily cases across different realizations. Secondly, we select all parameter sets that are within a 20% distance from the estimated parameters, using a uniform distribution. Finally, we combine the results from these selected parameter sets and their respective stochastic realizations, providing the mean values, as well as the 2.5th and 97.5th percentiles (representing the 95% percentile range), and the 25th and 75th percentiles (representing the 50% percentile range).

To capture a wide range of potential outcomes, we employ the Monte Carlo method to sample 10 000 parameter sets and run the model using each of them. Each parameter set generates results for 160 days, divided into 20 subintervals. By simulating with parameter sets that are within a 20% distance from the estimated parameters of the best fit,

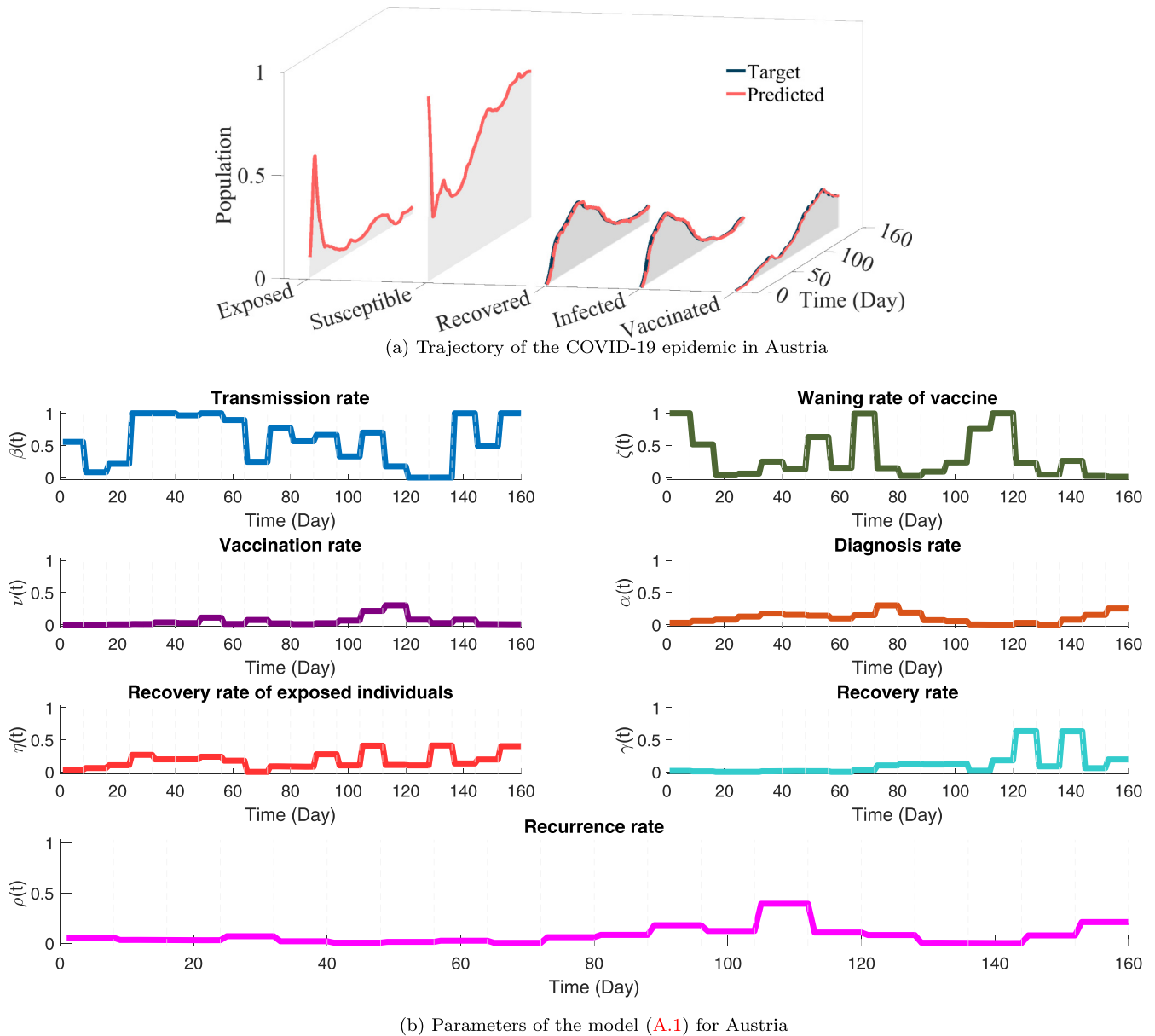


FIG. 3. The figure shows the trajectory of the COVID-19 epidemic in Austria (see Supplemental Material [56] for Brazil, and China). The daily results are displayed for all susceptible, exposed, vaccinated, infected, and recovered individuals. In plot (a), the trajectory of the COVID-19 outbreak is shown. The model output is shown in solid lines, while the actual data is shown in dotted lines. Also, panel plots of transmission rate, waning rate of vaccine, diagnosis rate, recovery rates, and recurrence rate are shown for Austria in the plot (b).

we can demonstrate the uncertainty in the model parameters (see Fig. 4). The findings highlight the stability of the model, even when there are 20% variations in the parameters. This suggests that 95% of the results closely align with real-world data and the approximations presented in our article. This observation is further supported by the average outcomes obtained through 10 000 repetitions, reinforcing the reliability of our analysis. For more detailed results regarding Brazil and China, please refer to the Supplemental Material [56].

Finally, we investigate the relationship between the length of subintervals within this 160-day period and the fitting of the data. By dividing the 160-day duration into subintervals of varying lengths (the number of subintervals is 5, 10, 15,

20, 25, 30, 35, and 40), we examine how well the data aligns with the predictions within each interval. This analysis allows us to assess the accuracy and reliability of the model used for forecasting COVID-19 cases. By comparing the actual counts with the predicted values for each subinterval, the article provides insights into the performance of the model over different time spans. Based on the insights gained from Fig. 5 and Table I, it is evident that the accuracy of the method does not necessarily improve with an increasing number of subintervals. This finding highlights the importance of carefully selecting the appropriate number of subintervals to achieve satisfactory fitting results when utilizing the method presented in this article. It emphasizes the need to find the optimal balance in determining the number of subintervals in order

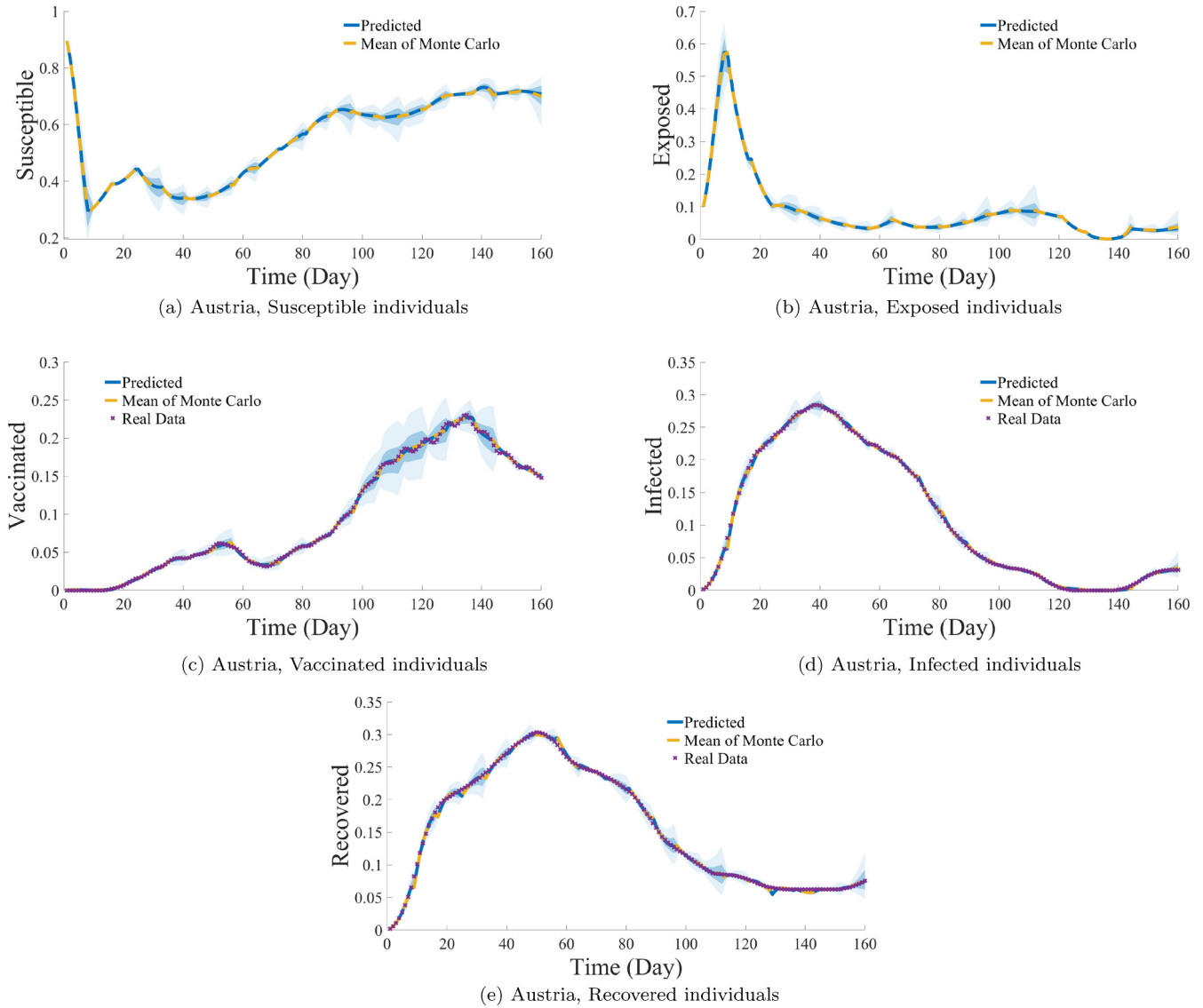


FIG. 4. Daily confirmed cases of Austria (purple) overlaid on the median modeled detected cases (vivid orange line), and predicted cases (strong blue lines) with shaded areas representing the 25th-75th centile (mid blue), and 2.5th-97.5th centile (light blue) of estimated detected cases.

to obtain accurate and reliable outcomes using the described approach. Based on the information provided in Table I, it can be inferred that the optimal number of subintervals for a 160-day duration in Austria falls within the range of 15 to 25. This range is identified as the one that yields the most accurate and reliable predictions based on the evaluation criteria utilized in the study.

**C. Quantitative analysis of model (1)**

Proving the stability of the endemic equilibrium posed a challenge due to the complexity of our explicit solution. To overcome this, we turned to MATLAB simulations to examine the long-term behavior. We applied our methods to recently released COVID-19 data for Brazil, specifically from February 26, 2021, to August 4, 2021, which coincided with the commencement of vaccination efforts. To facilitate our analysis, we divided the total 160-day time frame into

20 subintervals. To estimate the system’s parameters accurately using real data from Brazil, we developed an optimal numerical design. For detailed results, please refer to the Supplemental Material [56].

**D. Numerical results of reproduction number and its sensitivity analysis**

In this section, we focus on calculating the reproduction number and its sensitivity index to various parameters of the model for particular countries. The sensitivity index enables us to understand how different factors impact  $R_0$ , providing valuable insights into how we can better manage and control the spread of diseases. In Fig. 6, plot (a) depicts the reproduction number for Austria (See Supplemental Material [56] for Brazil, and China). This plot demonstrates that the reproduction number is a function that is piecewise constant and is obtained in 20 distinct subintervals. By breaking down



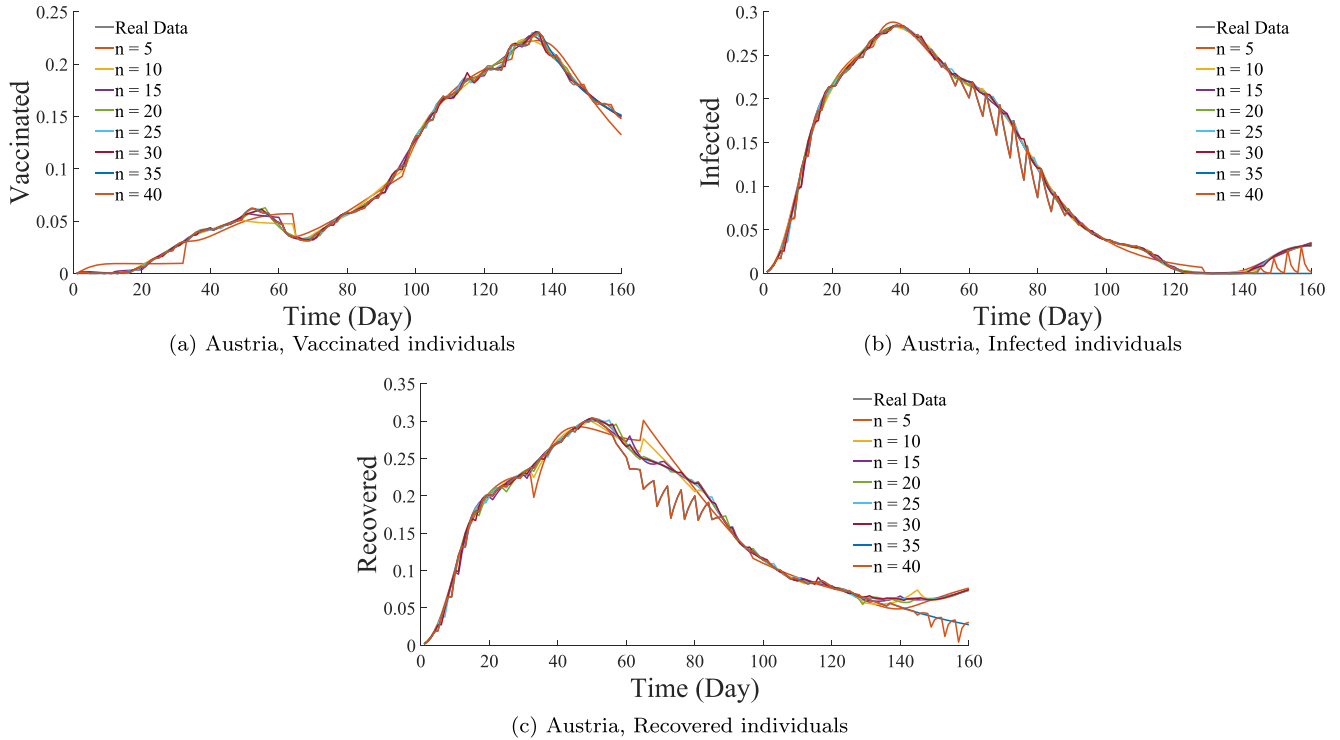


FIG. 5. The figure displays the daily counts of vaccinated, infected, and recovered cases in Austria over a period of 160 days. The fitting of the data is presented for different subintervals, ranging from 5 to 40, within the 160-day duration.

the reproduction number into multiple subintervals, we can better understand how it changes over time, which is critical when it comes to predicting and controlling the spread of diseases. Plot (b) in Fig. 6 shows the sensitivity indices of  $R_0$  corresponding to different parameters for Austria (See Supplemental Material [56] for Brazil, and China). It is worth noting that the figure displays indices for only those parameters that appear in the  $R_0$  formula, as  $R_0$  has no direct dependence on any other parameter. By examining the signs of the indices, we can determine whether  $R_0$  will increase or decrease in response to changes in specific parameters. Additionally, the magnitude of the indices provides insight into the extent of this change. In Fig. 6, the bar graph represents the sensitivity indices of  $R_0$  to different parameters. The bars facing upwards indicate that  $R_0$  increases as the parameter value increases. On the other hand, the bars facing

downwards indicate that  $R_0$  decreases as the parameter value increases.

Figure 6 provides valuable information on the sensitivity indices of  $R_0$  to different parameters. For instance,  $\Upsilon_{\zeta}^{R_0} = +0.5$  implies that  $R_0$  will increase by 0.5% if  $\zeta$  increases by 1%, and  $\Upsilon_{\eta}^{R_0} = -0.6$  means that  $R_0$  will decrease by 0.6% if  $\eta$  increases by 1%. Moreover, we can observe that  $R_0$  has the strongest negative correlation with  $\alpha$ ,  $\eta$ , and  $\nu$ , while it has the strongest positive correlation with  $\zeta$ . This means that changes in  $\alpha$ ,  $\eta$ , and  $\nu$  would decrease  $R_0$ , while changes in  $\zeta$  would increase it. Figure 7 also demonstrates the sensitivity analysis of the model for Austria and check how it responds to the variation of parameter  $\beta$  (For additional results on other parameters, please refer to the Supplemental Material [56]).

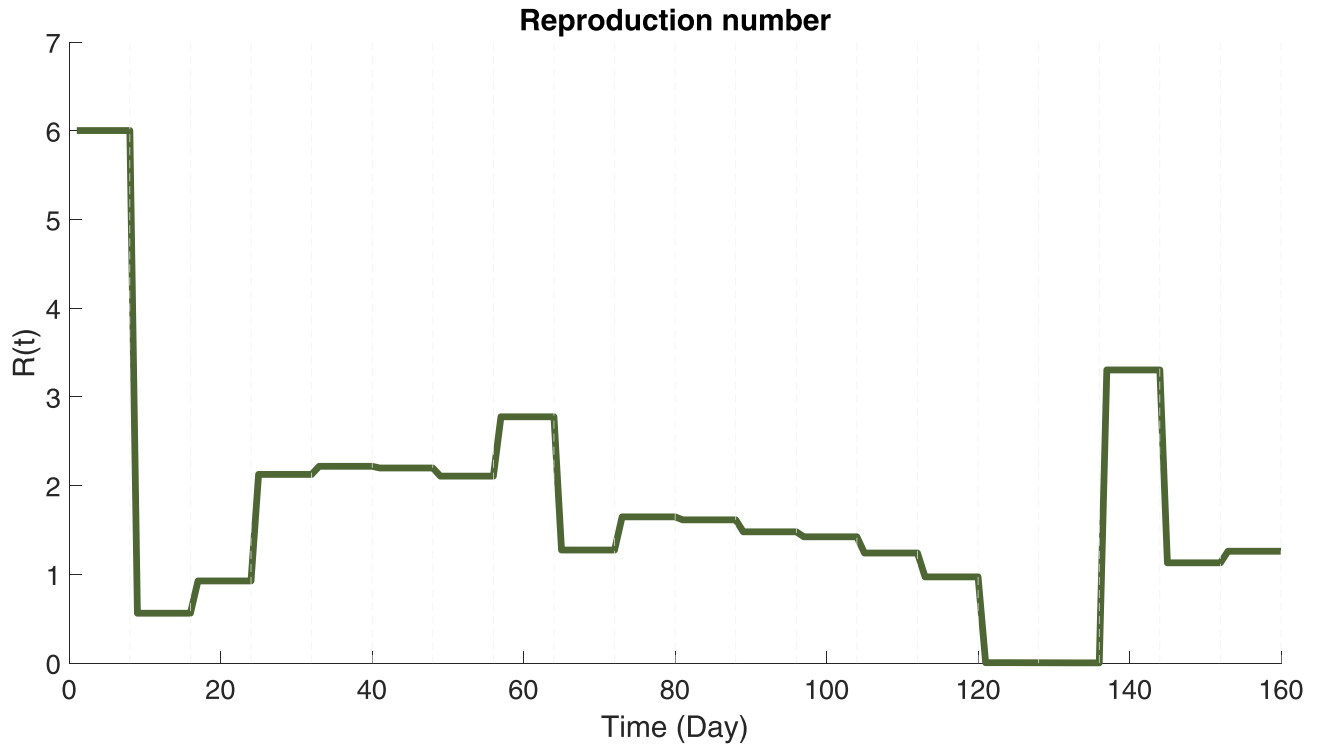
**E. Prediction process**

We conducted a comparative analysis of prediction methods, focusing on two main categories: classical and modern approaches. Classical methods relied on mathematical models like the recurrent dynamic model, using historical data to forecast disease transmission. Modern methods, on the other hand, employ machine learning techniques such as GRU, BiLSTM, regression tree, and the combined deep learning BiLSTM-GRU. These modern methods could learn from extensive data sets and make predictions about future disease spread. By thoroughly evaluating these methods, we aimed to identify the most accurate and reliable approach.

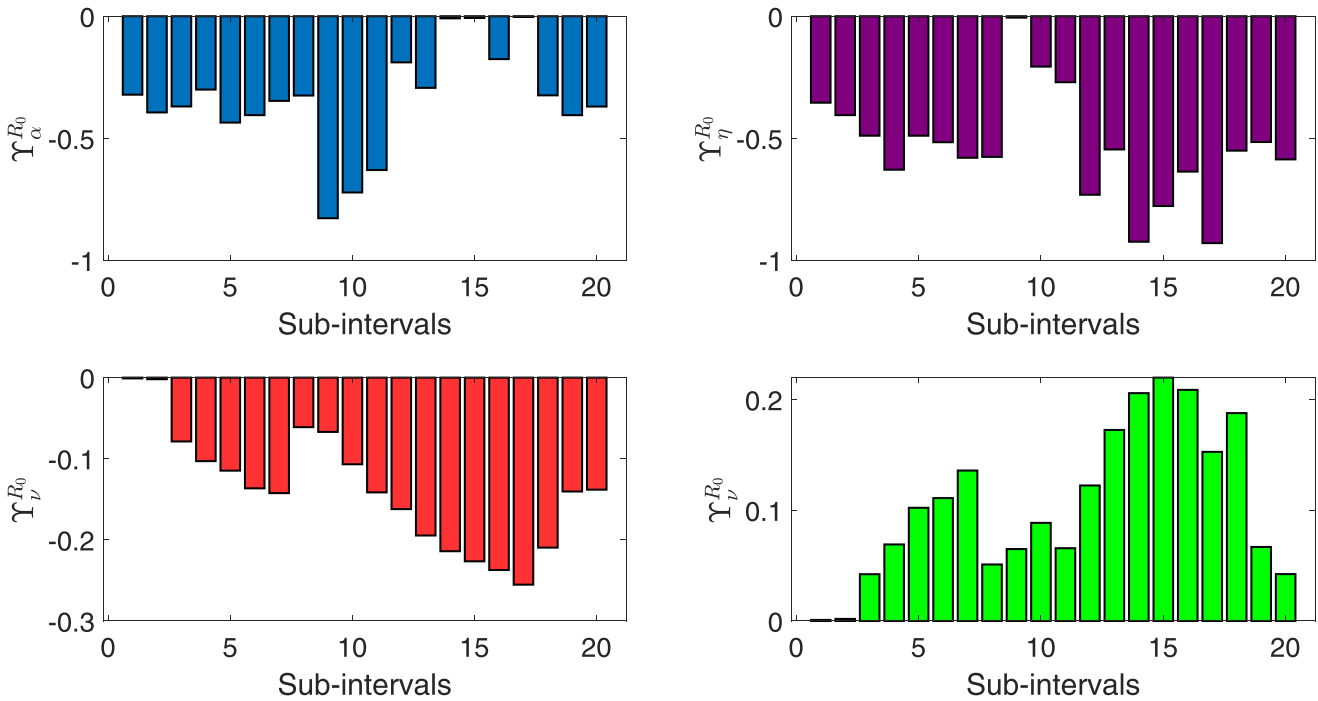
In our study, we utilized 128 days of data to train the selected prediction methods. All methods underwent

TABLE I. RMSE for different predictions of a 160-day duration in Austria, with varying subintervals.

Subintervals	Vaccinated	Infected	recovered	Mean
n = 5	0.0075	0.0038	0.0114	0.0076
n = 10	0.0041	0.0026	0.0056	0.0041
n = 15	0.0026	0.0024	0.0037	0.0029
n = 20	0.0026	0.0026	0.0035	0.0029
n = 25	0.0018	0.0023	0.0027	0.0023
n = 30	0.0019	0.0026	0.0028	0.0024
n = 35	0.0022	0.0120	0.0178	0.0107
n = 40	0.0020	0.0108	0.0188	0.0105



(a) Reproduction number in Austria



(b) Sensitivity index for Austria

FIG. 6. The figure displays the reproduction number in (a) and the sensitivity indices of the reproduction number concerning each of the system parameters linked to  $R_0$  for model system (1) in (b) for Austria.

calibration and training under identical conditions. Subsequently, we reserved 32 days of data for testing purposes to evaluate the performance of these methods in predicting disease spread. By comparing the prediction results for the 32-day testing period, we could assess the effectiveness of the

prediction methods. Figure 8 illustrates the prediction results for Austria (please refer to the Supplemental Material [56] for results on Brazil and China).

From Fig. 8, we observe that most of the prediction methods were well-trained, but their performances varied in the

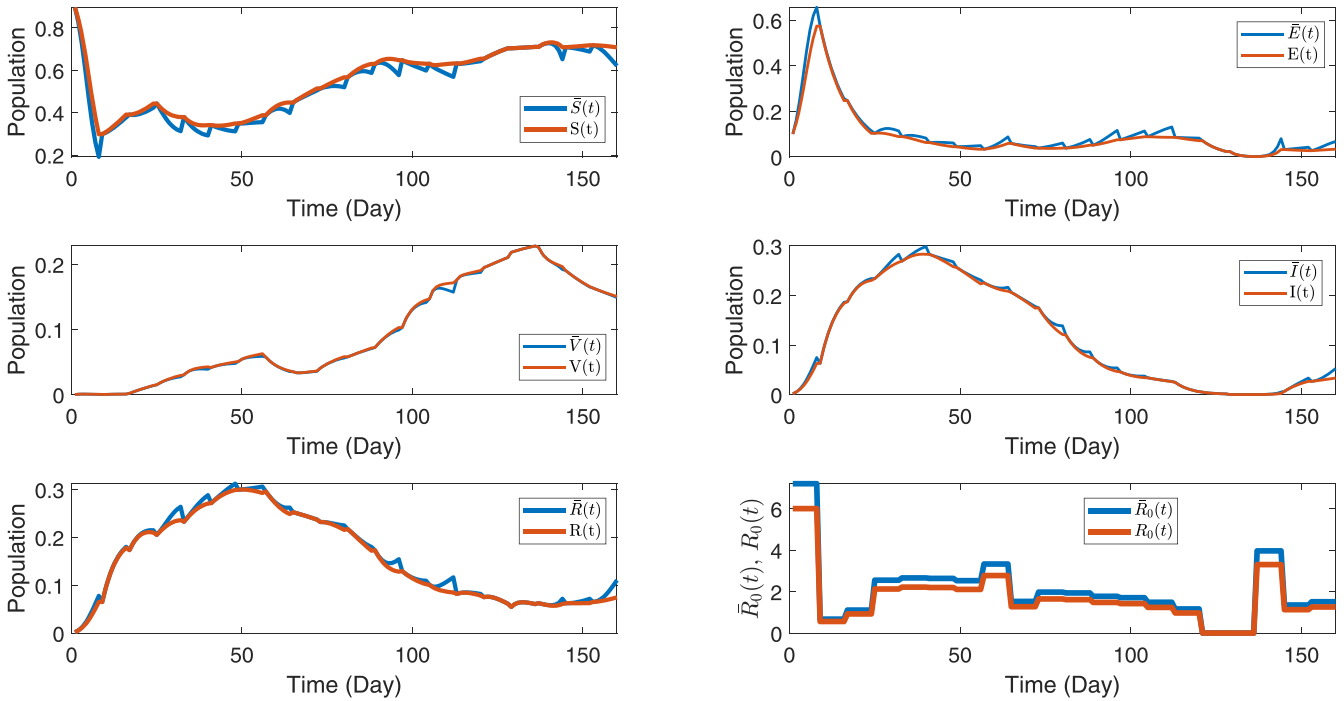


FIG. 7. Sensitivity analysis showing the effect of varying the  $\beta$  during the period of investigation for Austria. We multiply the experienced value of  $\beta$  by 1.2 and show the corresponding evolution of model (1) for its variables. The plot shows the variation in the susceptible, vaccinated, recovered, exposed, and infected individuals, and their reproduction numbers. The brown curves show the model fit up, and the blue curves show the variation in the required individuals. Increasing  $\beta$  significantly increases the reproduction number and the model is extremely sensitive to variations in the value of  $\beta$ .

prediction region. Graphically, it is apparent that the regression tree method and the dynamic system method did not perform as well in predicting the data compared to other learning methods. However, other learning methods such as BiLSTM, GRU, and the combined deep learning BiLSTM-GRU method exhibited good performance in predicting disease spread.

**F. Performance evaluation outcome**

To comprehensively assess the performance of a model, it is essential to utilize a variety of performance measures or metrics. Different metrics offer distinct perspectives on the model’s performance, aiding in identifying areas that may require improvement. In this study, we employed two commonly used metrics, root mean squared error (RMSE) and R-squared, defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\text{Predicted}_i - \text{Actual}_i)^2}{N}}, \tag{13}$$

$$R - \text{Squared} = 1 - \frac{\sum_{i=1}^N (\text{Predicted}_i - \text{Actual}_i)^2}{\sum_{i=1}^N (\text{Actual}_i)^2}. \tag{14}$$

Table II presents the performance evaluation of five prediction methods across three categories of individuals (vaccinated, infected, and recovered) and two data regions (training and testing). We used RMSE as the metric to assess the methods’ performance. The methods were ranked based on their

RMSE scores for each category and data region, with the best-performing method highlighted in a distinct color to differentiate it from the others.

As indicated in Table II, the dynamic system method exhibited superior performance in predicting the spread of COVID-19 among vaccinated individuals in Austria’s training data region. Similarly, the GRU method demonstrated the best performance in predicting the number of vaccinated individuals in Austria’s testing data region. These findings underscore the significance of assessing the performance of different prediction methods across various regions and categories of individuals.

Furthermore, Table II presents an overall measure, calculated to facilitate a comprehensive comparison of the five selected prediction methods using the RMSE metric. We computed the mean performance of each method for the training and testing data regions and ranked them accordingly. According to the RMSE metric, the dynamic system method (DS) exhibited the best performance in the training data region, followed by BiLSTM-GRU, regression tree (RT), GRU, and BiLSTM in descending order of performance. These findings can assist healthcare professionals and policymakers in selecting the most suitable method for predicting the spread of infectious diseases in the training data region. Similarly, in the testing data region, the BiLSTM-GRU method demonstrated the best performance, followed by GRU, BiLSTM, DS, and RT in descending order of performance. For results on the R-squared metric, please refer to the Supplemental Material [56].

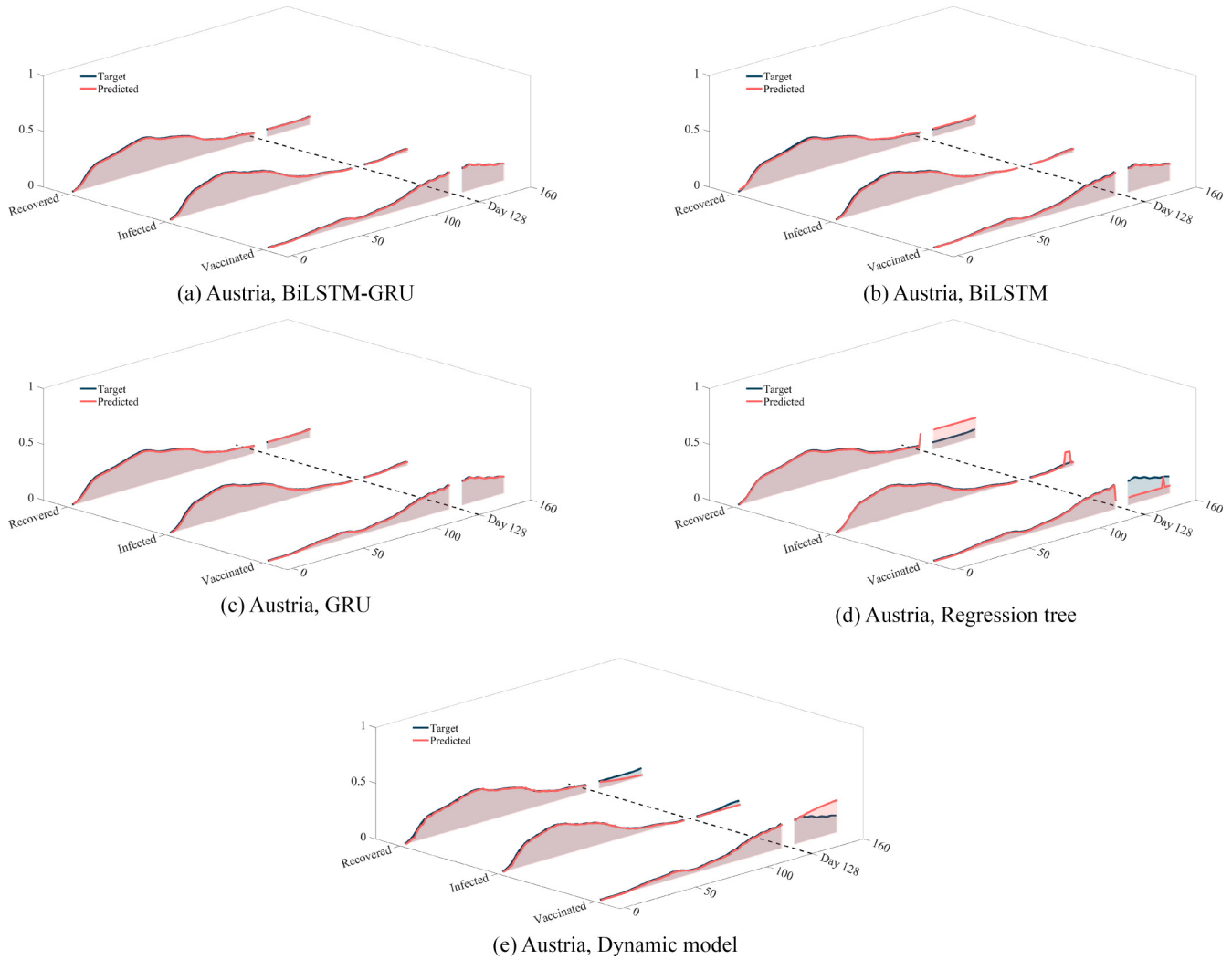


FIG. 8. Trajectory of the COVID-19 prediction in Austria. Prediction scenario for vaccinated, infected, and recovered individuals are done by five methods, including recurrent dynamical model (1) and learning methods GRU, regression tree, BiLSTM, and BiLSTM-GRU.

IV. DISCUSSION

The article presents an innovative hybrid model called the integrated dynamic-learning model.” This model merges a traditional recurrent dynamic model with modern learning techniques to evaluate and predict the transmission of infectious diseases, with a specific focus on COVID-19. The research delves into the utilization of dynamic systems theory in conjunction with a variety of learning methodologies. By combining dynamic and learning approaches, it becomes feasible to achieve more accurate and comprehensive predictive results. Hybrid models that integrate dynamic models and learning methods find applications in diverse domains. In epidemiology, for instance, these models serve as invaluable tools for predicting the transmission of infectious diseases, adapting to evolving transmission dynamics, and guiding public health decision making [58,59].

Our dynamic model is an epidemiological compartmental model governed by ordinary differential equations to elucidate the epidemic dynamics. It utilizes piecewise constant parameters, allowing it to effectively account for variations

in population groups resulting from control measures implemented by authorities, shifts in population behavior, and alterations in epidemic characteristics. Within our model, parameters are derived from officially reported data, providing a reflection of the genuine dynamics of infectious diseases.

Infectious diseases featuring recurrent patterns, like seasonal influenza, cyclic outbreaks of diseases such as measles or COVID-19, and other diseases displaying periodicity, demand models capable of encapsulating these patterns and shedding light on their dynamics. Additionally, the traditional assumption of permanent immunity following recovery may not be applicable to diseases like COVID-19, where immunity can diminish over time, and reinfection remains a possibility. Our proposed model takes into account recurrence, waning immunity, and the potential for reinfection, providing a more realistic portrayal of disease dynamics. Additionally, our proposed model includes the impact of vaccination, which is crucial in controlling diseases like COVID-19 and making it more applicable to modeling of epidemic diseases.

However, within the existing literature, there is a scarcity of studies that simultaneously address certain critical aspects:

TABLE II. RMSE evaluation. The best method for any type of data in any compartmental is indicated by bolding the corresponding number based on the RMSE.

Country	Data type	Compartments	RT	GRU	BiLSTM	BiLSTM-GRU	DS
Austria	Trained	Vaccinated	0.001690	0.004488	0.004755	0.004294	<b>0.000218</b>
		Infected	<b>0.000035</b>	0.002371	0.001128	0.001971	0.000104
		Recovered	0.004188	0.000910	0.003434	0.000140	<b>0.000135</b>
	Tested	Vaccinated	0.350605	<b>0.001993</b>	0.006039	0.004517	0.069424
		Infected	0.039854	0.005739	0.013515	<b>0.004970</b>	0.011888
		Recovered	0.350149	0.014903	0.032928	<b>0.011788</b>	0.028809
Brazil	Trained	Vaccinated	0.001442	0.001153	0.001398	0.001465	<b>0.000054</b>
		Infected	0.003311	0.003909	0.007633	0.003183	<b>0.000184</b>
		Recovered	0.002222	0.002659	0.003611	0.002276	<b>0.000193</b>
	Tested	Vaccinated	0.239629	<b>0.010747</b>	0.021142	0.011377	0.046999
		Infected	0.143138	0.049816	0.082708	<b>0.031578</b>	0.020276
		Recovered	0.010494	0.006078	0.010150	<b>0.005133</b>	0.019832
China	Trained	Vaccinated	0.003082	0.001581	0.003378	0.001180	<b>0.000079</b>
		Infected	0.001156	0.000112	0.001294	0.000146	<b>0.000052</b>
		Recovered	0.001767	0.001992	0.003065	0.001025	<b>0.000310</b>
	Tested	Vaccinated	0.130682	<b>0.001904</b>	0.003507	0.002649	0.013394
		Infected	0.049393	<b>0.000891</b>	0.002119	0.001017	0.010821
		Recovered	0.089787	0.022960	0.071510	<b>0.014165</b>	0.031023
	Mean of trained data		0.002099	0.002130	0.003299	0.001742	<b>0.000014</b>
	Ranking in training zone		3	4	5	2	1
	Mean of tested data		0.155970	0.012781	0.027068	<b>0.009688</b>	0.028051
	Ranking in testing zone		5	2	3	1	4

time-varying infectivity, which allows for the capture of epidemic characteristic changes, waning immunity, and recurrent patterns resulting from reinfection.

Integrating birth and death processes into the model provides a more holistic depiction of population dynamics, aligning with real-world scenarios and enabling us to account for observed demographic shifts and make long-term predictions. Usually, traditional models overlook births and deaths in the population. These rates have a significant influence on population size, and vulnerability, thereby playing a pivotal role in shaping the transmission and consequences of epidemics. Our proposed model takes these factors into account, providing a more accurate representation of population dynamics and their impact on disease spread.

While dynamic models are valuable for analyzing epidemics, machine learning models serve as auxiliary tools for studying real data, including COVID-19. Various studies demonstrate the potential of machine learning models in creating roadmaps for disease outbreak control and prediction [35–37,39]. In the realm of classification, machine learning shows promise by seeking out hidden relationships between inputs and outputs. These methods are also applied to forecast the number of confirmed cases and mortality in upcoming seasons. In the second approach, this research offered a description and comparison of some machine learning models including decision tree learning, BiLSTM, GRU, and a combined deep learning approach. These methods were applied to analyze and forecast time series data related to the COVID-19 outbreak. They were chosen for their capacity to conduct highly accurate simulations, their training capabilities, and their adaptability as function approximators. We assessed the performance of these models in terms of training, learning, and prediction. Our learning models exhibit several

appealing features, such as the ability to address temporal dependencies within time series data and their flexibility in modeling nonlinear features. To the best of our knowledge, this kind of comparative analysis has been scarcely documented in existing works. Furthermore, this article stands out for its accurate long-term analysis and forecasting of the COVID-19 disease.

In the final section of the article, we applied these various methods to study the spread of COVID-19 in specific countries such as Austria, Brazil, and China. We first used an optimization method to fit the models to historical data and then employed the fitted models for prediction, aiming to forecast the future spread of the disease. The parameters were estimated based on officially reported data, ensuring accuracy and reliability. By utilizing high-quality data, we were able to accurately estimate key parameters related to transmission rates, recovery rates, diagnosis rates, vaccination rates, recurrent rates, waning rates, and symptomatic recovery rates. The analysis involved multiple sets of parameters and stochastic realizations, providing mean values, percentiles, and ranges. The stability of the model and its alignment with real-world data were demonstrated through careful validation processes. The findings offered reliable insights and specific results for Austria, Brazil, and China.

To evaluate the performance of the proposed method, we used metrics such as RMSE and R-squared. By comparing the performance of each method in both the fitting and prediction stages, we identified the most effective method for each country. According to the results, the DS method performed exceptionally well in the data training phase, while the BiLSTM-GRU method exhibited superior performance in the prediction phase. These findings emphasized the importance of selecting appropriate methods for specific stages

of disease analysis to ensure accurate predictions and effective decision making.

The results also highlighted the high recurrence rate of the virus despite vaccination efforts, indicating the need for ongoing monitoring and control. Additionally, the low diagnosis and recovery rates suggested potential challenges in information, equipment, medical staff, and emerging variants. Calculating the reproduction number and its sensitivity to various parameters of the model for specific countries provided valuable insights into how different factors impact disease spread, aiding in better management and control.

Our results emphasized the need for caution when extrapolating epidemic counts in the long term, as they depended not only on data quality but also on the stage of the epidemic due to the non-linear nature of the underlying dynamics. The study investigated how uncertainty changes during different stages of the epidemic, providing valuable insights into the dynamics of infectious diseases.

In conclusion, this article effectively examined the transmission of infectious diseases by integrating dynamics and contemporary approaches. The recurrent SEIR-V model and deep learning techniques demonstrated their effectiveness in predicting disease spread. The outcomes derived from evaluating diverse methods at different analysis stages offered valuable perspectives for healthcare experts and policymakers to make well-informed choices concerning disease management and prevention strategies.

*Model limitations and strengths.* Firstly, like traditional SIR and SEIR models, our proposed model assumes that individuals in the population have equal chances of contact, which is not realistic. Factors such as social structures, geographical barriers, and individual behaviors significantly influence disease transmission.

Another limitation of traditional models is that they assume constant disease transmission and recovery rates over time [48]. However, in reality, these rates can change due to public health interventions, changes in human behavior, and the emergence of new virus variants. The dynamic model we propose addresses this by incorporating time-varying parameters, allowing for more flexibility in real-world scenarios. Additionally, traditional models overlook births and deaths in

the population, which are important for long-term predictions or diseases with high mortality rates. Our proposed model takes these factors into account, providing a more accurate representation of population dynamics and their impact on disease spread. Moreover, the assumption of permanent immunity after recovery in traditional models may not hold true for diseases like COVID-19, where immunity can wane over time and reinfection is possible. Our proposed model considers waning immunity and possible reinfection, offering a more realistic depiction of disease dynamics. Furthermore, our proposed model, like traditional SIR and SEIR models, does not consider individual characteristics such as age, sex, or health status, which can significantly influence disease susceptibility and outcomes [60].

While traditional models ignore spatial structure and geographical factors that can impact disease spread [61], our proposed model also does not explicitly include them. However, it has been successfully implemented in different countries with varying geographies, indicating its adaptability and relevance to different regions. Similarly, changes in human behavior, like social distancing or mask-wearing, are not incorporated into traditional models or the proposed model. Moreover, traditional models do not consider the impact of vaccination on disease dynamics, which is crucial in controlling diseases like COVID-19 [5]. Our proposed model includes the impact of vaccination, making it more applicable to diseases where vaccination plays a key role. Lastly, neither traditional models nor the proposed model explicitly accounts for the emergence of new virus variants, which can have different transmissibility and virulence characteristics. Finally, we propose a hybrid approach to overcome the limitations of traditional models in disease prediction. By combining a traditional model with modern supervised methods, our aim is to significantly improve the accuracy of disease prediction.

#### ACKNOWLEDGMENT

M.Z. was supported by a grant from Basic Sciences Research Fund (No. BSRF-math-399-01).

We declare no competing interests.

- 
- [1] L. J. Abu-Raddad, H. Chemaitelly, P. Coyle, J. A. Malek, A. A. Ahmed, Y. A. Mohamoud, S. Younuskuju, H. H. Ayoub, Z. Al Kanaani, E. Al Kuwari *et al.*, SARS-CoV-2 antibody-positivity protects against reinfection for at least seven months with 95% efficacy, *EClinicalMedicine* **35**, 100861 (2021).
- [2] B. Choi, M. C. Choudhary, J. Regan, J. A. Sparks, R. F. Padera, X. Qiu, I. H. Solomon, H.-H. Kuo, J. Boucau, K. Bowman *et al.*, Persistence and evolution of SARS-CoV-2 in an immunocompromised host, *N. Engl. J. Med.* **383**, 2291 (2020).
- [3] V. J. Hall, S. Foulkes, A. Charlett, A. Atti, E. J. Monk, R. Simmons, E. Wellington, M. J. Cole, A. Saei, B. Oguti *et al.*, SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative healthcare workers in England: A large, multicentre, prospective cohort study (SIREN), *Lancet* **397**, 1459 (2021).
- [4] E. J. Rubin, L. R. Baden, and S. Morrissey, Audio interview: What's gone right in our battle against covid-19, *New England J. Medicine* **385**, e95 (2021).
- [5] S. A. Rakhshan, M. S. Nejad, M. Zaj, and F. H. Ghane, Global analysis and prediction scenario of infectious outbreaks by recurrent dynamic model and machine learning models: A case study on covid-19, *Comput. Biol. Med.*, **158**, 106817 (2023).
- [6] M. Khairulbahri, The Seir model incorporating asymptomatic cases, behavioral measures, and lockdowns: Lesson learned from the covid-19 flow in Sweden, *Biomed. Signal Process. Control* **81**, 104416 (2023).
- [7] S. Ansumali, S. Kaushal, A. Kumar, M. K. Prakash, and M. Vidyasagar, Modelling the covid-19 pandemic: asymptomatic patients, lockdown and herd immunity, *IFAC-PapersOnLine* **53**, 823 (2020).

- [8] M. Fošnarič, T. Kamenšek, J. Žganec Gros, and J. Žibert, Extended compartmental model for modeling covid-19 epidemic in Slovenia, *Sci. Rep.* **12**, 16916 (2022).
- [9] M. A. Khan, A. Atangana, E. Alzahrani *et al.*, The dynamics of covid-19 with quarantined and isolation, *Adv. Differ. Equ.* **2020**, 425 (2020).
- [10] S. İ. Araz, Analysis of a covid-19 model: Optimal control, stability and simulations, *Alexandria Eng. J.* **60**, 647 (2021).
- [11] A. Atangana, Modelling the spread of covid-19 with new fractal-fractional operators: Can the lockdown save mankind before vaccination? *Chaos, Solitons Fractals* **136**, 109860 (2020).
- [12] C. E. Wagner, C. M. Saad-Roy, and B. T. Grenfell, Modelling vaccination strategies for covid-19, *Nat. Rev. Immunol.* **22**, 139 (2022).
- [13] J. Arino, F. Brauer, P. Van Den Driessche, J. Watmough, and J. Wu, A model for influenza with vaccination and antiviral treatment, *J. Theor. Biol.* **253**, 118 (2008).
- [14] Z. Yu, J. Liu, X. Wang, X. Zhu, D. Wang, and G. Han, Efficient vaccine distribution based on a hybrid compartmental model, *PLoS ONE* **11**, e0155416 (2016).
- [15] M. A. Acuña-Zegarra, S. Díaz-Infante, D. Baca-Carrasco, and D. Olmos-Liceaga, Covid-19 optimal vaccination policies: A modeling study on efficacy, natural and vaccine-induced immunity responses, *Math. Biosci.* **337**, 108614 (2021).
- [16] J. K. K. Asamoah, C. Bornaa, B. Seidu, and Z. Jin, Mathematical analysis of the effects of controls on transmission dynamics of SARS-CoV-2, *Alexandria Eng. J.* **59**, 5069 (2020).
- [17] D. Faranda and T. Alberti, Modeling the second wave of COVID-19 infections in France and Italy via a stochastic SEIR model, *Chaos* **30**, 111101 (2020).
- [18] T. Alberti and D. Faranda, On the uncertainty of real-time predictions of epidemic growths: A covid-19 case study for China and Italy, *Commun. Nonlinear Sci. Numer. Simul.* **90**, 105372 (2020).
- [19] T. Kuniya, Recurrent epidemic waves in a delayed epidemic model with quarantine, *J. Biol. Dyn.* **16**, 619 (2022).
- [20] X.-C. Duan, J.-F. Yin, and X.-Z. Li, Global hopf bifurcation of an sirs epidemic model with age-dependent recovery, *Chaos, Solitons Fractals* **104**, 613 (2017).
- [21] H. Cao, D. X. Yan, and A. Li, Dynamic analysis of the recurrent epidemic model, *Math. Biosci. Eng.* **16**, 5972 (2019).
- [22] Y. Nakata and T. Kuniya, Global dynamics of a class of Seirs epidemic models in a periodic environment, *J. Math. Anal. Appl.* **363**, 230 (2010).
- [23] O. N. Bjørnstad, K. Shea, M. Krzywinski, and N. Altman, The Seirs model for infectious disease dynamics., *Nat. Methods* **17**, 557 (2020).
- [24] B. K. Mishra and D. K. Saini, Seirs epidemic model with delay for transmission of malicious objects in computer network, *Appl. Math. Comput.* **188**, 1476 (2007).
- [25] L. R. Baden, H. M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S. A. Spector, N. Rouphael, C. B. Creech *et al.*, Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine, *N. Engl. J. Med.* **384**, 403 (2021).
- [26] M. Dashtbali and M. Mirzaie, A compartmental model that predicts the effect of social distancing and vaccination on controlling covid-19, *Sci. Rep.* **11**, 8191 (2021).
- [27] M. Teti, E. Schatz, and L. Liebenberg, Methods in the time of COVID-19: The vital role of qualitative inquiries, *Internatl. J. Qualitative Methods* **19**, 1609406920920962 (2020).
- [28] A. Page, S. Y. Diallo, W. J. Wildman, G. Hodulik, E. W. Weisel, N. Gondal, and D. Voas, Computational simulation is a vital resource for navigating the covid-19 pandemic, *Simul. Healthcare* **17**, e141 (2022).
- [29] E. Rechtman, P. Curtin, E. Navarro, S. Nirenberg, and M. K. Horton, Vital signs assessed in initial clinical encounters predict covid-19 mortality in an NYC hospital system, *Sci. Rep.* **10**, 21545 (2020).
- [30] G. Vivekanandhan, M. N. Zavareh, H. Natiq, F. Nazarimehr, K. Rajagopal, and M. Svetec, Investigation of vaccination game approach in spreading covid-19 epidemic model with considering the birth and death rates, *Chaos, Solitons Fractals* **163**, 112565 (2022).
- [31] L. Taghizadeh, A. Karimi, and C. Heitzinger, Uncertainty quantification in epidemiological models for the covid-19 pandemic, *Comput. Biol. Med.* **125**, 104011 (2020).
- [32] H. Leng, Y. Zhao, J. Luo, and Y. Ye, Simplicial epidemic model with birth and death, *Chaos* **32**, 093144 (2022).
- [33] Q. Wu and K. A. Kabir, Compact pairwise methods for susceptible–infected–susceptible epidemics on weighted heterogeneous networks, *Physica A* **621**, 128805 (2023).
- [34] D. Han, J. Wang, and Q. Shao, On epidemic spreading in metapopulation networks with time-varying contact patterns, *Chaos* **33**, 093142 (2023).
- [35] D. Painuli, D. Mishra, S. Bhardwaj, and M. Aggarwal, Forecast and prediction of COVID-19 using machine learning, in *Data Science for COVID-19*, edited by U. Kose, D. Gupta, V. H. C. de Albuquerque, and A. Khanna (Academic Press, Cambridge, MA, USA, 2021), pp. 381–397.
- [36] L. Ali, S. E. Alnawayseh, M. Salahat, T. M. Ghazal, M. A. Tomh, and B. Mago, AI-based intelligent model to predict epidemics using machine learning technique., *Intell. Autom. Soft Comput.* **36**, 1095 (2023).
- [37] S. Palaniappan, V. Ragavi, B. David, and S. P. Nisha, Prediction of epidemic disease dynamics on the infection risk using machine learning algorithms, *SN Comput. Sci.* **3**, 47 (2022).
- [38] A. N. Desai, M. U. Kraemer, S. Bhatia, A. Cori, P. Nouvellet, M. Herringer, E. L. Cohn, M. Carrion, J. S. Brownstein, L. C. Madoff *et al.*, Real-time epidemic forecasting: challenges and opportunities, *Health Secur.* **17**, 268 (2019).
- [39] H. Verma, S. Mandal, and A. Gupta, Temporal deep learning architecture for prediction of covid-19 cases in india, *Expert Syst. Appl.* **195**, 116611 (2022).
- [40] E. A. Rashed, S. Kodera, and A. Hirata, Covid-19 forecasting using new viral variants and vaccination effectiveness models, *Comput. Biol. Med.* **149**, 105986 (2022).
- [41] Z. Liao, P. Lan, Z. Liao, Y. Zhang, and S. Liu, Tw-sir: time-window based sir for covid-19 forecasts, *Sci. Rep.* **10**, 22454 (2020).
- [42] J. Sun, X. Chen, Z. Zhang, S. Lai, B. Zhao, H. Liu, S. Wang, W. Huan, R. Zhao, M. T. A. Ng *et al.*, Forecasting the long-term trend of covid-19 epidemic using a dynamic model, *Sci. Rep.* **10**, 21122 (2020).
- [43] Z. Liao, P. Lan, X. Fan, B. Kelly, A. Innes, and Z. Liao, Sirvd-dl: A covid-19 deep learning prediction model based on time-dependent SIRVD, *Comput. Biol. Med.* **138**, 104868 (2021).

- [44] L. Muhammad, M. M. Islam, S. S. Usman, and S. I. Ayon, Predictive data mining models for novel coronavirus (covid-19) infected patients' recovery, *SN Comput. Sci.* **1**, 206 (2020).
- [45] A. Narin, C. Kaya, and Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, *Pattern Anal. Appl.* **24**, 1207 (2021).
- [46] H. B. Syeda, M. Syed, K. W. Sexton, S. Syed, S. Begum, F. Syed, F. Prior, and F. Yu Jr, Role of machine learning techniques to tackle the covid-19 crisis: systematic review, *JMIR Med. Inf.* **9**, e23811 (2021).
- [47] L. Rasmy, M. Nigo, B. S. Kannadath, Z. Xie, B. Mao, K. Patel, Y. Zhou, W. Zhang, A. Ross, H. Xu *et al.*, Recurrent neural network models (CovRNN) for predicting outcomes of patients with COVID-19 on admission to hospital: model development and validation using electronic health record data, *Lancet Digital Health* **4**, e415 (2022).
- [48] G. Massonis, J. R. Banga, and A. F. Villaverde, Structural identifiability and observability of compartmental models of the COVID-19 pandemic, *Ann. Rev. control* **51**, 441 (2021).
- [49] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu *et al.*, Top 10 algorithms in data mining, *Knowl. Inf. Syst.* **14**, 1 (2008).
- [50] M. Studer, G. Ritschard, A. Gabadinho, and N. S. Müller, Discrepancy analysis of state sequences, *Sociol. Methods Res.* **40**, 471 (2011).
- [51] K. Cho, B. Van Merriënboer, D. Bahdanau, and Y. Bengio, On the properties of neural machine translation: Encoder-decoder approaches, *Syntax, Semantics Struct. Stat. Trans.*, 103 (2014).
- [52] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, in *NIPS 2014 Workshop on Deep Learning, December 2014* (2014).
- [53] J. Gao, H. Liu, and E. T. Kool, Expanded-size bases in naturally sized dna: Evaluation of steric effects in Watson–Crick pairing, *J. Am. Chem. Soc.* **126**, 11826 (2004).
- [54] P. Baldi, S. Brunak, P. Frasconi, G. Soda, and G. Pollastri, Exploiting the past and the future in protein secondary structure prediction, *Bioinformatics* **15**, 937 (1999).
- [55] T. Young, D. Hazarika, S. Poria, and E. Cambria, Recent trends in deep learning based natural language processing, *IEEE Comput. Intell. Mag.* **13**, 55 (2018).
- [56] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.109.014212> for further details on exploring the potential of learning methods and recurrent dynamic model with vaccination: A comparative case study of covid-19 in Austria, Brazil, and China.
- [57] W. H. Organization, *Coronavirus Disease (Covid-19) Pandemic* (W.H.O., Geneva, Switzerland, 2022).
- [58] S. A. Nawaz, J. Li, U. A. Bhatti, S. U. Bazai, A. Zafar, M. A. Bhatti, A. Mehmood, Q. u. Ain, and M. U. Shoukat, A hybrid approach to forecast the covid-19 epidemic trend, *PLoS ONE* **16**, e0256971 (2021).
- [59] K. K. Hwang, C. J. Edholm, O. Saucedo, L. J. Allen, and N. Shakiba, A hybrid epidemic model to explore stochasticity in covid-19 dynamics, *Bull. Math. Biol.* **84**, 91 (2022).
- [60] K. Prem, Y. Liu, T. W. Russell, A. J. Kucharski, R. M. Eggo, N. Davies, S. Flasche, S. Clifford, C. A. Pearson, J. D. Munday *et al.*, The effect of control strategies to reduce social mixing on outcomes of the covid-19 epidemic in Wuhan, China: A modelling study, *Lancet Public Health* **5**, e261 (2020).
- [61] Y. Tsori and R. Granek, Epidemiological model for the inhomogeneous spatial spreading of covid-19 and other diseases, *PLoS ONE* **16**, e0246056 (2021).