



# Global analysis and prediction scenario of infectious outbreaks by recurrent dynamic model and machine learning models: A case study on COVID-19

Seyed Ali Rakhshan <sup>a</sup>, Mahdi Soltani Nejad <sup>b</sup>, Marzie Zaj <sup>a</sup>, Fatemeh Helen Ghane <sup>a,\*</sup>

<sup>a</sup> Department of Mathematics, Ferdowsi University of Mashhad, Iran

<sup>b</sup> Department of Railway Engineering, Iran University of Science and Technology, Iran



## ARTICLE INFO

Dataset link: <https://www.who.int/countries/irn/>

**Keywords:**

Dynamic epidemic model  
Machine learning methods  
Bifurcation  
COVID-19  
Reproduction number  
Tipping phenomena

## ABSTRACT

It is essential to evaluate patient outcomes at an early stage when dealing with a pandemic to provide optimal clinical care and resource management. Many methods have been proposed to provide a roadmap against different pandemics, including the recent pandemic disease COVID-19. Due to recurrent epidemic waves of COVID-19, which have been observed in many countries, mathematical modeling and forecasting of COVID-19 are still necessary as long as the world continues to battle against the pandemic. Modeling may aid in determining which interventions to try or predict future growth patterns. In this article, we design a combined approach for analyzing any pandemic in two separate parts. In the first part of the paper, we develop a recurrent SEIRS compartmental model to predict recurrent outbreak patterns of diseases. Due to its time-varying parameters, our model is able to reflect the dynamics of infectious diseases, and to measure the effectiveness of the restrictive measures. We discuss the stable solutions of the corresponding autonomous system with frozen parameters. We focus on the regime shifts and tipping points; then we investigate tipping phenomena due to parameter drifts in our time-varying parameters model that exhibits a bifurcation in the frozen-in case. Furthermore, we propose an optimal numerical design for estimating the system's parameters. In the second part, we introduce machine learning models to strengthen the methodology of our paper in data analysis, particularly for prediction scenarios. We use MLP, RBF, LSTM, ANFIS, and GRNN for training and evaluation of COVID-19. Then, we compare the results with the recurrent dynamical system in the fitting process and prediction scenario. We also confirm results by implementing our methods on the released data on COVID-19 by WHO for Italy, Germany, Iran, and South Africa between 1/22/2020 and 7/24/2021, when people were engaged with different variants including Alpha, Beta, Gamma, and Delta. The results of this article show that the dynamic model is adequate for long-term analysis and data fitting, as well as obtaining parameters affecting the epidemic. However, it is ineffective in providing a long-term forecast. In contrast machine learning methods effectively provide disease prediction, although they do not provide analysis such as dynamic models. Finally, some metrics, including RMSE, R-Squared, and accuracy, are used to evaluate the machine learning models. These metrics confirm that ANFIS and RBF perform better than other methods in training and testing zones.

## 1. Introduction

### 1.1. Background

Mathematical epidemic models are applied to monitor the spread dynamics of the outbreak. We can quantitatively predict disease outbreaks and capture significant phenomena by using these models, leading to an improved understanding and control of diseases [1–4]. It is known that many diseases can cause secondary infections after recovery, such as Influenza, Hand, Foot, Mouth, Rotavirus, and the new disease COVID-19 [5–7]. These types of diseases can be classified

as recurrent epidemics. As one of the hottest topics in epidemiology, studies of recurrent epidemics have been attracting considerable attention for decades. Significant progress has been accomplished in the past few years, including the persistence of recurrent infectious diseases and the derivation of thresholds for outbreaks as well as the dynamic analysis of seasonal trends and non-seasonal trends, and so on [8–11]. Unfortunately, recurrent dynamic models have been less considered for COVID-19 disease due to inherent and computational complexities. Therefore, this is a fundamental weakness in epidemiological models

\* Corresponding author.

E-mail address: [ghane@math.um.ac.ir](mailto:ghane@math.um.ac.ir) (F.H. Ghane).

for COVID-19 that have not surveyed this challenge, and most models can be criticized.

Disease outbreak patterns vary widely across places, times, and diseases. Time series vary in frequency, with some showing outbreaks every year or every two years, others showing long interepidemic intervals, and others looking irregular, like COVID-19, which makes it more difficult to model [12,13]. The pattern of outbreaks for a given disease may also change over time, either gradually or abruptly. Mathematics epidemiology has been devoted to predicting the interepidemic interval, or the interval between successive outbreaks of diseases, since the beginning of the 20th century [14]. Therefore, the concept of mathematical modeling has become more and more important. Indeed, mathematical models of epidemics enable us to predict future outcomes and help inform public health interventions [15–17]. Indeed, using simple assumptions and a combination of statistics and mathematics, models for various infectious diseases are constructed and used to calculate the effects of interventions like mass vaccinations and population-wide strategies, including social distancing and lockdown.

There have been many mathematical modeling approaches proposed over the years which solve the problem from a different perspective. We can categorize them into (I) stochastic and (II) deterministic or (a) statistical methods for epidemic surveillance, (b) state-space models of epidemic spread, and (c) machine learning methods [18]. One of the most widely used methods of the category (b) is compartmental models that can be autonomous or nonautonomous (time-dependent). However, time-dependent epidemic models enable to capture of fluctuations in the number of susceptible, infectious, existing confirmed cases with controlled infectivity, hence, they perform better than autonomous models with constant parameters. In general, these models capture the trajectories of each population group well because the effective rates in the spread of epidemics, such as the transmission, recovery, and immunity rates, are extracted for each given time rather than as a constant over time [19–22].

Although dynamic models are beneficial for analyzing any epidemic diseases, machine learning models are employed as auxiliary tools to study actual data like COVID-19. There are several works that show machine learning models can create a road map for controlling and predicting disease outbreaks. In the field of classification, machine learning is a promising tool. As a model, machine learning seeks out unknown functions, dependencies, or structures between inputs and outputs. Machine-learning methods are applied to forecast the number of confirmed cases and mortality for the upcoming season [23].

## 1.2. Our work and main contributions

Recently, the study of disease dynamics has been much recent interest in the areas of mathematical biology. This paper reviews the role of models in understanding and predicting recurrent outbreak patterns of diseases. We design an approach consisting of two parts: the first one is to provide a recurrent nonautonomous compartmental model for analyzing the outbreak and fitting the data as accurately as possible, and the second one is to provide five machine learning algorithms for getting a more accurate prediction of the outbreak.

In the first part, we propose a nonautonomous recurrent compartmental model governed by ordinary differential equations to explain the behavior of the epidemic. Our dynamic model has time-varying parameters, enabling to capture of fluctuations in the number of population groups due to control measures applied by authorities, population behavioral changes, and/or modifications of the epidemic characteristics. Time-dependent models are more realistic to reflect the actual dynamics of infectious diseases, and at the same time, enable the assessment of the strength of the effects produced by the adopted restrictive measures such as nationwide lockdowns and social distancing. As an advantage, in our model, the parameters are estimated based on the officially reported data and reflect the actual dynamics of infectious diseases. Due to recurrent epidemic waves of COVID-19, which have been

**Table 1**

Findings summary.

The strengths of the paper
1- Developing a recurrent model to discuss the mechanism of recurrent epidemic waves of COVID-19.
2- The accurate analysis of the COVID-19 disease for a long period including 550 days.
3- Developing a dynamic model with time-varying parameters which enables to capture fluctuations in the number of population groups due to control measures applied by authorities.
4- Investigating the tipping phenomena due to parameter drifts in our time-varying parameters model that exhibits a bifurcation-driven tipping or B-tipping in the frozen-in case.
5- Estimating the susceptible and exposed population: Due to the lack of information from the community as well as the incubation period of the COVID-19 disease, data regarding susceptible and exposed people is not available. The proposed method of the article has been able to estimate the population of this group of people.
6- Forecasting the COVID-19 pandemic's trajectories during 120 days in selected 4 countries using dynamic model and five machine learning methods.

observed in many countries, we consider a recurrent model to discuss the mechanism of such recurrent epidemic waves from the viewpoint of mathematical modeling. Our recurrent model is a differential equation of the form

$$\dot{X}(t) = F(X(t), \lambda(t)) \quad (1.2.1)$$

where  $\lambda(t)$  is in general a time-varying parameter varies in a finite-dimensional vector space  $\Lambda$ , and the vector field  $F : \mathbb{R}^{n+1} \times \Lambda \rightarrow \mathbb{R}^{n+1}$  depends on time only due to its dependence on the parameter. We analyze the model dynamically and investigate the long-time behavior of solutions in this model. Various backgrounds of nonautonomous systems were presented in [24] by Kloeden et al. which can be very helpful for studying nonautonomous pandemic models. Due to changes in the contact rate, nonlinearity, and stochasticity caused by time-varying parameters, we observe complex and rich dynamics in our models. If in (1.2.1), the time-dependent parameter  $\lambda(t)$  is replaced by a constant  $\lambda$  in  $\Lambda$ , then we call

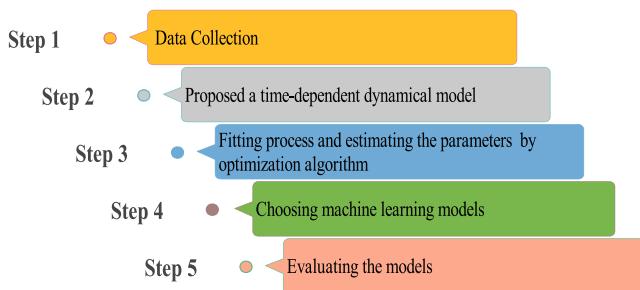
$$\dot{X}(t) = F(X(t), \lambda) \quad (1.2.2)$$

the *parameterized system with frozen parameters*. We refer to stable solutions of (1.2.2) as the *quasi-static attractor* [25].

*Parameter drift* is a phenomenon characterized by time-dependent changes in parameter sets and is gaining much interest in recent years. When deterministic systems with parameter drift witness a dramatic change in performance, a tipping phenomenon occurs [26]. For instance, if the parameter passes through a bifurcation point of the parameterized system, like in our model, where a stable solution loses stability, then the system may tip directly as a result of varying that parameter. Tipping phenomena are found in some areas of research, including climate dynamics [27,28], regime shifts in ecology [29,30], disease spreading in human populations [31,32], and so on.

In this article, we demonstrate that our nonautonomous system admits a pullback attractor where fibers are contained within a positive flow invariant probability simplex. Then, the geometry of the pullback attractor is described using the parameterized system (1.2.2) with frozen parameters. We also investigate tipping phenomena due to parameter drifts in our time-varying parameters model that exhibits a bifurcation in the frozen-in case. In the literature, very few works have been done on the transition behavior of compartmental epidemic models and investigating the tipping phenomena.

In the second part, this research offers a description and comparison of some machine learning models including Radial Base Function (RBF), Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM), Generalized Regression Neural Network (GRNN), and Adaptive Neuro-Fuzzy Inference System (ANFIS). These methods are used for analyzing



**Fig. 1.** Main process of the paper.

and time series forecasting the data of the COVID-19 outbreak. They are used because of their ability to perform high-accuracy simulations with training capabilities and their flexibility as function approximators.

We evaluate the training, learning, and prediction performance of these models. Root mean squared error (RMSE), accuracy, and R-squared are computed to evaluate model performance. According to these metrics and by comparing the model output and observed data, we find that the models with the best performance, fitting, and prediction are ANFIS and RBF. Our main contributions are summarized in [Table 1](#).

### 1.3. Research methodology

This research consists of two main parts, and two widely used epidemiological models from two different categories are compared. A compartmental model that investigates the phenomenon of recurrent outbreaks has been presented in the first part of the paper. In this part, we analyze the model dynamically and discuss the stable solutions of the corresponding autonomous system with frozen parameters. We aim to investigate what kind of tipping phenomena might occur in our model. We then apply some machine-learning techniques to the analysis of the COVID-19 epidemic in the second part. We summarized the structure of the paper in some steps as illustrated in [Fig. 1](#). As shown in [Fig. 1](#), the first step in the article involves collecting COVID-19 data. We use the World Health Organization and GitHub databases to collect data. We analyze several countries' COVID-19 using data published by the World Health Organization (WHO). We also extract information from the GitHub database, which contains statistics on people who were infected, died, and recovered, as well as vaccinated population data. Next, a filtering step is also performed to reduce noise and excess variability in the raw data. We use the fixed-point filter in MATLAB software. Our next step is to normalize filtered data to organize it in our database. In the second step, a dynamic model is designed based on the collected data. A detailed analysis of the model is then performed dynamically. In this stage, the conditions are ready for approximating the parameters and analyzing the features of the model. In the third step, as well as illustrating how the process fits the data, we obtain parameters such as transmission rate, diagnosis rate, treatment rate, and even recurrence rate. This process is done using optimization algorithms. The next step is about providing a prediction package. A dynamic model, along with machine learning algorithms, is used to develop a package for predicting COVID-19. The training and testing data are processed using five different machine learning methods, including LSTM, MLP, ANFIS, GRNN, and RBF. This allows us to compare whether the classic dynamic model means would be best suited for predicting COVID-19 or the selected modern machine learning methods. Finally, some metrics, including RMSE, R-Squared, and accuracy, are used to evaluate the presented models.

## 2. Literature review

This research focuses on two types of mathematical models: (1) dynamic models and (2) forecast models based on Machine Learning. One of the most widely used methods of dynamic models of epidemic spread is compartmental models that can be stochastic or deterministic. These types of models are necessary tools for analyzing the spread dynamics of the pandemic in the population. They include SIR (Susceptible–Infectious–Removed) model [33], the SEIR (Susceptible–Exposed–Infected–Removed) model [34], and so on. The developed compartmental models were proposed to consider the effect of asymptomatic individuals [35–37], isolation [38,39], delay in diagnosis into account [40,41], and other aspects of the spread dynamics of the epidemic [2,3,42–44]. In general, compartmental SIR-type models and their extensions are typically non-linear, which causes model parameter identifications to be difficult. For instance, the transmission rate depends on factors such as social behavior and strain-specific that ensure to be time-dependent. For stochastic models, parameter estimation, the so-called calibration, is a type of statistical inference. Parameter estimation techniques allow exploration of epidemics as they occur [19]. In nonautonomous models, the permanence of disease plays an important role in epidemiology [45–47]. The persistence and extinction of disease were studied for nonautonomous recurrent epidemic dynamical systems [48,49]. A threshold value was established to determine whether the disease is persistent or extinct. It is noticeable that the modeling of biological systems under a nonautonomous environment should be more realistic than the autonomous situation. Particularly, for analyzing the long-term dynamical behavior of an epidemic system, due to changing the parameters of the system with time, nonautonomous models perform better than autonomous models with constant parameters.

Recurrent models are used when periodicity in infectious disease outbreaks occurs. In the coronavirus disease outbreak, recurrent epidemic waves have been observed in many countries. So, to understand the pattern of the outbreak, the recurrent epidemic models are recommended [50–52]. In fact compartmental recurrent epidemic models are an extension of classic forms that have been developed over the last years. It is shown by Nakata et al. [53] that the global dynamics of a class of periodic SEIRS epidemic models are determined by the basic reproduction number using the spectral radius of a linear integral operator. By introduction of random fluctuations, Cai et al. [54] extend a classical SIRS epidemic model from a deterministic framework to one based on stochastic differential equations. The dynamics of a stochastic SIRS epidemic model with saturated incidence are considered by Zhao et al. [55]. The extinction and persistence of epidemics are determined by the stochastic system threshold when the noise is small. The global dynamics of a general SIRS epidemic model based on a ratio-dependent incidence rate, as well as a stochastic differential equation model of the epidemic, are investigated by Cai et al. [56]. Vaccination and nonlinear incidence are considered in a class of nonautonomous SEIRS epidemic models by Zhang et al. [57]. Yang [58] proposed a recurrent SEIRS epidemic model with isolated, exposed class and diagnosed class, relevant to interventions determined by modeling. In [50], the authors investigated nonautonomous recurrent epidemic models with bilinear incidence and disease-induced mortality.

However, in the literature, very few studies have been conducted that consider both aspects simultaneously: time-varying infectivity enabling to capture modifications of the epidemic characteristics, and recurrence due to reinfection.

Since dynamical population prediction models have limitations on predicting the population in the time-series, so machine learning methods are applied for forecasting the likely COVID-19 cases and deaths. In [59], two learning algorithms, namely deep learning and reinforcement learning, were developed to forecast COVID-19. In [60], the authors developed deep learning methods to forecast the COVID-19 spread and compared 5 deep learning algorithms for COVID-19

forecasting. Their results showed the superior performance of the Variational AutoEncoder algorithm compared to the other algorithms. However, the forecasting was performed for two the number of confirmed cases and the number of recovered cases with, the forecasting horizon of 17 days. Deep learning Recurrent Neural Networks (RNN) models proposed to predict the country-wise cumulative confirmed cases, cumulative recovered cases, and cumulative fatalities [61]. Their methods included Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM). In [62], Machine Learning based on LSTM was applied to forecasting the COVID-19 spread. Their results demonstrated vaccination effectiveness in controlling the outbreak. To predict the number of COVID-19 cases in Brazil, India, and Russia, three deep learning models, including CNN, LSTM, and the CNN-LSTM have been developed [63]. The accurate analysis confirmed that the LSTM model had the highest performance. Lalmuanawma et al. in a review article [64] investigated the role of the modern AI and ML technology that were recently implemented to tackle the challenges during the outburst. An overview of the literature was conducted in [65] to examine how AI is being used to fight COVID-19 in the areas of epidemiology, diagnosis, and disease progression. A recurrent neural network-based model was developed in [66] using online data available in the electronic health record (EHR) to predict the outcome of COVID-19 patients without requiring specific feature selection or imputation of missing data. In [67], the authors proposed an Adaptive Neuro-Fuzzy Inference System (ANFIS) to forecast the number of COVID-19 cases in the United Kingdom. The ANFIS integrates an artificial neural network to train a Fuzzy inference system (FIS) based on data. This method has proven its optimal ability to deal with uncertainty and has been widely applied to forecasting systems. In [68], the authors applied ANFIS as a critical rule for automatic COVID-19 detection from chest X-ray images based on the characteristics derived by texture analysis using the gray-level co-occurrence matrix (GLCM) technique.

Our research article combines modern and classical methods to fitting and forecasting the COVID-19 spread and compares five machine learning models (ANFIS, MLP, LSTM, GRNN, and RBF) and dynamic time-dependent model SEIRS for COVID-19 forecasting. Our machine learning models have many attractive features, including handling temporal dependencies in time series data, and their flexibility in modeling nonlinear features. To the best of our knowledge, such a comparison has not been seen in other scientific works. In addition, the accurate analysis and forecasting of the COVID-19 disease for a long period, including 550 days, are one of the features of this article, which cannot be seen in other articles.

### 3. Methods

#### 3.1. Proposed recurrent dynamic epidemiological model

We develop a recurrent compartmental epidemic model with time-varying parameters to fit the released data and provide different analyses of the behavior of epidemics. This model uses real epidemic data to calibrate its parameters to account for fluctuations in the number of population groups because of government control measures. The total population of our model is partitioned into the following stages of the disease:  $S$ , susceptible;  $E$ , exposed individuals;  $I$ , infected population, who have not been isolated, and  $R$ , recovered population.

The model consists of the following differential equations that describe how the population in each compartment evolves:

$$\begin{aligned} \frac{dS(t)}{dt} &= \sigma - \frac{\alpha(t)}{N} S(t)E(t) + \rho(t)R(t) - \mu S(t), \\ \frac{dE(t)}{dt} &= \frac{\alpha(t)}{N} S(t)E(t) - \beta(t)E(t) - \mu E(t) \\ \frac{dI(t)}{dt} &= \beta(t)E(t) - \gamma(t)I(t) - \mu I(t) \\ \frac{dR(t)}{dt} &= \gamma(t)I(t) - \rho(t)R(t) - \mu R(t) \end{aligned} \quad (3.1.1)$$

The parameters of (3.1.1) are time-dependent and positive which may be interpreted as follows:

- $\alpha(t)$  shows the transmission rate (the probability of moving from susceptible group to infected group) due to contact between a susceptible individual and an infected, at time  $t$ .
- $\beta(t)$  indicates diagnosis rate at time  $t$ .
- $\gamma(t)$  is the recovery rate at time  $t$ .
- $\rho(t)$  denotes the recurrence rate at time  $t$  for people who recover and then enter the disease cycle.

The following assumptions are also considered within the model:

- (H1) We consider the probability that the recovered individuals becoming susceptible again.
- (H2) We consider the birth and natural death rates in the system. The natural death is determined by  $\mu = 0.0003$ .
- (H3) According to assumption (H2), we ignore the death resulting from disease.

For simplicity in our calculations, we use the fractions instead and we let  $S$ ;  $E$ ;  $I$ ;  $R$  denote respectively the fraction of the population which is susceptible, the exposed individuals, the infectious individuals and recovered. We also take the next condition:

- (H4) It is assumed that the birth and death rates are equal, that is  $\sigma = \mu$ .

Thus, the model to be studied takes the following form:

$$\begin{aligned} \frac{dS}{dt} &= -\alpha(t)SE + \rho(t)R - \mu S + \mu, \\ \frac{dE}{dt} &= \alpha(t)SE - \beta(t)E - \mu E \\ \frac{dI}{dt} &= \beta(t)E - \gamma(t)I - \mu I \\ \frac{dR}{dt} &= \gamma(t)I - \rho(t)R - \mu R. \end{aligned} \quad (3.1.2)$$

#### 3.1.1. Mathematical analysis of our dynamic model

We carried out the qualitative analysis in this section to learn more about our recurrent dynamic epidemiological model.

##### Lemma 3.1.1. The simplex

$$\Sigma = \{(S, E, I, R) \in \mathbb{R}_+^4 : S(t), E(t), I(t), R(t) \geq 0, \\ S(t) + E(t) + I(t) + R(t) = 1\}$$

is positively flow invariant for model (3.1.2).

**Proof.** Let  $F(t, X)$  be the time-dependent vector field of Eqs. (3.1.2) and  $X \in \Sigma$ . Then, the following holds:

- (i)  $\frac{dS}{dt} \geq 0$  if  $S = 0$ ,  $\frac{dE}{dt} \geq 0$  if  $E = 0$ ,  $\frac{dI}{dt} \geq 0$  if  $I = 0$  and  $\frac{dR}{dt} \geq 0$  if  $R = 0$ ;
- (ii)  $1^T F(t, X) = 0$  holds for every  $X \in \Sigma$  and every  $t \in \mathbb{R}$ .

By conditions (i) and (ii), we get the proof.  $\square$

Among semigroup forms, the two-parameter semigroup or process formalism is the most natural generalization to nonautonomous dynamical systems.

Let  $F(t, X)$  be the time-varying vector field of Eqs. (3.1.2) and take the following initial value problem

$$\frac{dX}{dt} = F(t, X), \quad X(t_0) = X_0.$$

Let us recall that a process is a continuous mapping  $(t, t_0, X_0) \mapsto \phi(t, t_0, X_0) \in \Sigma \subset \mathbb{R}^4$  for  $t, t_0 \in \mathbb{R}$  and  $X_0 \in \Sigma$  with  $t \geq t_0$ , satisfying both the initial value and evolution properties [24].

- (i)  $\phi(t_0, t_0, X_0) = \{X_0\}$  for all  $t_0 \in \mathbb{R}$ , and  $X_0 \in \Sigma \subset \mathbb{R}^4$ ;
- (ii)  $\phi(t_2, t_0, X_0) = \phi(t_2, t_1)\phi(t_1, t_0, X_0))$  for all  $t_0 \leq t_1 \leq t_2$  and all  $X_0 \in \Sigma \subset \mathbb{R}^4$ .

Let us consider the initial condition  $X(t_0) = X_0$ . Therefore, under the assumption that global solutions exist and they are unique, solution  $X(t, t_0, X_0)$  of (3.1.2) defines a continuous time process. Note that the flow associated with (3.1.2) may be defined in the compact and positively invariant set  $\Sigma$  of  $\mathbb{R}^4$ , leading to a complete flow when the initial point  $(S(0), E(0), I(0), R(0))$  belongs to  $\Sigma$ , i.e. solutions are defined for all  $t \in \mathbb{R}^+$  and they are unique for each initial condition  $X(t_0) = X_0$ . By these facts and [Theorem 3.1.2](#), we get the next proposition.

**Proposition 3.1.1.** *For the nonautonomous differential equations defined by (3.1.2), unique solutions generate a process  $\phi$  on  $\Sigma$  which is defined by  $\phi(t, t_0, X_0) = X(t)$ , where  $X(t) = X(t, t_0, X_0)$  is the unique solution starting at the point  $X_0$  at time  $t_0$ .*

### 3.1.2. The existence of pullback attractor

The global pullback attractor  $\mathcal{A}$  with fibers  $A(t)$ , for given  $t$ , is a set of minimal closed sets at time  $-\infty$  that can attract all subsets of  $B \subset \Sigma$ , [24,69], i.e.,  $\lim_{s \rightarrow -\infty} \text{dist}(\phi(t, s, B), A(t)) = 0$  holds for every subset  $B \subset \Sigma$ , and it is invariant, i.e.,  $\phi(t, s, A(s)) = A(t)$  holds for all  $s \leq t$ .

Since our compartmental model (3.1.2) generates a process  $\phi$  on  $\Sigma$ , the following result holds by [24]:

**Theorem 3.1.1.** *Consider the nonautonomous system (3.1.2). Then there exists a pullback attractor  $\mathcal{A}$  with fibers in  $\Sigma$  uniquely determined by*

$$A(t) := \bigcap_{s \leq t} \overline{\bigcup_{r \leq s} \phi(t, r, \Sigma)}. \quad (3.1.3)$$

In the following, we illustrate the geometry of the pullback attractor  $\mathcal{A}$ . Let us consider the nonautonomous system

$$\frac{dX}{dt} = F(X, \lambda(t)), \quad (3.1.4)$$

where  $F$  is the vector field of Eqs. (3.1.2),  $X \in \Sigma$  and  $\lambda(t) = (\alpha(t), \beta(t), \gamma(t))$  is the time-varying parameter. If in (3.1.4) the time-dependent parameter  $\lambda(t)$  is replaced by a constant  $\lambda$ , then we call

$$\frac{dX}{dt} = F(X, \lambda) \quad (3.1.5)$$

the corresponding autonomous ODE with frozen parameters.

In the case of the nonautonomous Eq. (3.1.4), a quasi-static attractor  $A(\lambda)$  frozen at the parameter  $\lambda$  is described as a local attractor of the corresponding autonomous Eq. (3.1.5). Now, for given any  $t$ , we compare the fiber  $A(t)$  of  $\mathcal{A}$  with the corresponding quasi-static attractor  $A(\lambda(t))$ . First, we fix a natural birth rate  $\mu$ .

We fix a parameter value  $\lambda = (\alpha, \beta, \gamma, \rho)$ , and take the parameterized system (3.1.5) with parameter  $\lambda$ . By (3.1.2), we get

$$\frac{dS}{dt} + \frac{dE}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0,$$

which implies  $S(t) + E(t) + I(t) + R(t) = 1$ . So, for simplicity, we can ignore the recovered population and reduce Eqs. (3.1.2) to the following form

$$\begin{aligned} \frac{dS}{dt} &= -\alpha(t)SE + \rho(t)R - \mu S + \mu, \\ \frac{dE}{dt} &= \alpha(t)SE - \beta(t)E - \mu E \\ \frac{dI}{dt} &= \beta(t)E - \gamma(t)I - \mu I. \end{aligned} \quad (3.1.6)$$

Then, we get a reduced form of the parametric Eqs. (3.1.5) corresponding to (3.1.6). To gain the equilibria of this reduced parameterized system, we set

$$-\alpha SE + \rho(1 - S - E - I) - \mu S + \mu = 0$$

$$\alpha SE - (\beta + \mu)E = 0$$

$$\beta E - (\gamma + \mu)I = 0.$$

Thus, the disease-free equilibrium is  $X_0 = (1, 0, 0, 0)$ .

We define the effective reproduction number by

$$R_e = \frac{\alpha}{\beta + \mu}. \quad (3.1.7)$$

By a straightforward computations, the endemic equilibrium is given by  $X_1 = (S^*, E^*, I^*)$ , where

$$\begin{aligned} S^* &= \frac{\beta + \mu}{\alpha}, \\ E^* &= \frac{(\rho + \mu)(\gamma + \mu)(R_e - 1)}{R_e((\beta + \mu)(\gamma + \mu) + \rho(\gamma + \mu) + \rho\beta)}, \\ I^* &= \frac{\beta(\rho + \mu)(R_e - 1)}{R_e((\beta + \mu)(\gamma + \mu) + \rho(\gamma + \mu) + \rho\beta)}. \end{aligned} \quad (3.1.8)$$

Note that the endemic equilibrium exists if  $R_e > 1$ .

**Theorem 3.1.2.** *The disease-free equilibrium  $X_0 = (1, 0, 0)$  of the reduced parameterized system is globally asymptotically stable if  $R_e < 1$  and it is unstable if  $R_e > 1$ .*

**Proof.** The Jacobian matrix of the parameterized system at the equilibrium  $X_0$  is

$$J|_{X_0} = \begin{bmatrix} -(\rho + \mu) & -(\alpha + \rho) & -\rho \\ 0 & \alpha - (\beta + \mu) & 0 \\ 0 & \beta & -(\gamma + \mu) \end{bmatrix}$$

After calculating the eigenvalues of the Jacobian matrix, we have two negative eigenvalues  $-(\rho + \mu)$ ,  $-(\gamma + \mu)$ , and another eigenvalue is  $\alpha - (\beta + \mu)$ . If  $\alpha < \beta + \mu$ , or equivalently  $R_e < 1$ , then all eigenvalues are negative and the equilibrium  $X_0$  is asymptotically stable. If  $\alpha > \beta + \mu$ , or equivalently  $R_e > 1$ , then  $X_0$  becomes unstable.  $\square$

**Theorem 3.1.3.** *If  $R_e > 1$ , the reduced parameterized system admits an endemic equilibrium  $X_1 = (S^*, E^*, I^*)$ , where  $S^*$ ,  $E^*$  and  $I^*$  are given by (3.1.8). In particular, it is asymptotically stable.*

**Proof.** The Jacobian matrix of the reduced parameterized system at  $X_1$  is defined by

$$J|_{X_1} = \begin{bmatrix} -\alpha E^* - (\rho + \mu) & -(\beta + \rho + \mu) & -\rho \\ \alpha E^* & 0 & 0 \\ 0 & \beta & -(\gamma + \mu) \end{bmatrix}$$

Thus, the characteristic equation is given by

$$f(\lambda) = -\lambda(\alpha E^* + \rho + \mu + \lambda)(\gamma + \mu + \lambda) - \alpha E^*((\beta + \mu + \rho)(\gamma + \mu + \lambda) + \rho\beta), \quad (3.1.9)$$

which simplifies to

$$f(\lambda) = -(\gamma + \mu + \lambda)(\lambda(\alpha E^* + \rho + \mu + \lambda) + \alpha E^*(\beta + \mu + \rho)) - \alpha E^* \rho \beta. \quad (3.1.10)$$

**Claim 3.1.1.** *If  $\lambda$  is a real eigenvalue of  $J|_{X_1}$  then  $\lambda$  should be negative.*

**Proof.** By (3.1.10), if  $\lambda > 0$  then  $f(\lambda) < 0$ . So, the Jacobian matrix  $J|_{X_1}$  does not admit a positive eigenvalue. Moreover,

$$f(0) = -\alpha E^*((\gamma + \mu)(\beta + \mu + \rho) + \rho\beta) < 0, \quad (3.1.11)$$

which implies that the Jacobian matrix  $J|_{X_1}$  does not admit 0 as an eigenvalue. This finishes the proof.  $\square$

Now, we verify that  $J|_{X_1}$  has not pure imaginary eigenvalues. For simplicity, we set  $\hat{\gamma} = \gamma + \mu$ ,  $\hat{\beta} = \beta + \mu$  and  $\hat{\rho} = \rho + \mu$ .

**Claim 3.1.2.** *The Jacobian matrix  $J|_{X_1}$  does not admit pure imaginary eigenvalues  $\lambda = \pm i\omega$ .*

**Proof.** By (3.1.8),

$$\alpha E^* = \frac{\alpha(\rho + \mu)(\gamma + \mu)(R_e - 1)}{R_e((\beta + \mu)(\gamma + \mu) + \rho(\gamma + \mu) + \rho\beta)}$$

$$= \frac{\alpha\hat{\gamma}(\alpha - \hat{\beta})}{\alpha(\hat{\beta}\hat{\gamma} + \rho(\hat{\gamma} + \beta))} = \frac{\hat{\beta}\hat{\gamma}(R_e - 1)}{(\hat{\beta}\hat{\gamma} + \rho(\hat{\gamma} + \beta))}. \quad (3.1.12)$$

By (3.1.23), if  $\lambda = i\omega$  is an eigenvalue of  $J$ , then

$$f(i\omega) = -(\hat{\gamma} + i\omega)[(i\omega)(\alpha E^* + \hat{\rho} + i\omega) + \alpha E^*(\hat{\rho} + \beta)] - \alpha E^* \rho \beta = 0.$$

Therefore, the real part of  $f(i\omega)$  should be zero, that is

$$\hat{\gamma}\omega^2 - \hat{\gamma}\alpha E^*(\hat{\rho} + \beta) + \omega^2(\alpha E^* + \hat{\rho}) - \alpha E^* \rho \beta = 0.$$

This yields

$$\omega^2 = \frac{\alpha E^*(\hat{\gamma}(\hat{\rho} + \beta) + \rho\beta)}{\hat{\gamma} + \alpha E^* + \hat{\rho}} \quad (3.1.13)$$

Also, the imaginary part of  $f(i\omega)$  should be zero, that is

$$Im(f(i\omega)) = -\hat{\gamma}\omega(\alpha E^* + \hat{\rho}) - \omega(\alpha E^*(\hat{\rho} + \beta) - \omega^2) = 0. \quad (3.1.14)$$

Since  $\omega \neq 0$ , we get

$$-\hat{\gamma}(\alpha E^* + \hat{\rho}) - (\alpha E^*(\hat{\rho} + \beta) - \omega^2) = 0. \quad (3.1.15)$$

This yields

$$\omega^2 = \alpha E^*(\hat{\gamma} + \hat{\rho} + \beta) + \hat{\gamma}\hat{\rho}. \quad (3.1.16)$$

Combining (3.1.16) and (3.1.13), we get

$$\alpha E^*(\hat{\gamma} + \hat{\rho} + \beta) + \hat{\gamma}\hat{\rho} = \frac{\alpha E^*(\hat{\gamma}(\hat{\rho} + \beta) + \rho\beta)}{\hat{\gamma} + \alpha E^* + \hat{\rho}}. \quad (3.1.17)$$

By (3.1.12), we can substitute  $\alpha E^*$  by  $\frac{\hat{\beta}\hat{\gamma}(R_e - 1)}{(\hat{\beta}\hat{\gamma} + \rho(\hat{\gamma} + \beta))}$  in the above equation.

Also, for simplicity, we set  $S = \hat{\beta}\hat{\gamma} + \rho(\hat{\gamma} + \beta)$  and  $T = \hat{\gamma} + \hat{\rho} + \beta$ . Then, by a straightforward computations, we get

$$\frac{\hat{\beta}(R_e - 1)T + S}{S} = \frac{\hat{\beta}(R_e - 1)(\hat{\gamma}(\hat{\rho} + \beta) + \rho\beta)}{S(\hat{\gamma} + \hat{\rho}) + \hat{\beta}\hat{\gamma}(R_e - 1)}. \quad (3.1.18)$$

This implies that

$$\hat{\beta}\hat{\rho}\hat{\gamma}T\tilde{R}^2 + \hat{\beta}S[T(\hat{\gamma} + \hat{\rho}) - \beta(\hat{\gamma} + \rho)]\tilde{R} + S^2(\hat{\gamma} + \hat{\rho}) = 0, \quad (3.1.19)$$

where  $\tilde{R} = R_e - 1$ . Let,  $Z = \hat{\beta}S[T(\hat{\gamma} + \hat{\rho}) - \beta(\hat{\gamma} + \rho)]$ . Then

$$\begin{aligned} Z &= \hat{\beta}S[(\hat{\gamma} + \hat{\rho} + \beta)(\hat{\gamma} + \hat{\rho}) - \beta(\hat{\gamma} + \rho)] \\ &> \hat{\beta}S[(\hat{\gamma} + \hat{\rho} + \beta)(\hat{\gamma} + \hat{\rho}) - \beta(\hat{\gamma} + \hat{\rho})] \\ &= \hat{\beta}S(\hat{\gamma} + \hat{\rho})^2 > 0. \end{aligned} \quad (3.1.20)$$

Therefore,

$$\tilde{R} = \frac{1}{2}(-Z \pm \sqrt{Z^2 - 4\hat{\beta}\hat{\rho}\hat{\gamma}TS^2(\hat{\gamma} + \hat{\rho})}). \quad (3.1.21)$$

If  $Z^2 - 4\hat{\beta}\hat{\rho}\hat{\gamma}TS^2(\hat{\gamma} + \hat{\rho}) < 0$ , then (3.1.19) has not a real root. Let  $Z^2 - 4\hat{\beta}\hat{\rho}\hat{\gamma}TS^2(\hat{\gamma} + \hat{\rho}) > 0$ . Since

$$Z > \sqrt{Z^2 - 4\hat{\beta}\hat{\rho}\hat{\gamma}TS^2(\hat{\gamma} + \hat{\rho})}, \quad (3.1.22)$$

then by (3.1.22),  $\tilde{R} < 0$  which is impossible, because  $\tilde{R} = R_e - 1 > 0$ . Thus, by these observations, the Jacobian matrix  $J|_{X_1}$  does not admit pure imaginary eigenvalues  $\lambda = \pm i\omega$ .  $\square$

Let  $\lambda = \xi + i\omega$  be an eigenvalue of  $J|_{X_1}$ . Then, by (3.1.10), we get

$$\begin{aligned} f(\xi + i\omega) &= -(\hat{\gamma} + \xi + i\omega)[(\xi + i\omega)(\alpha E^* + \hat{\rho} + \xi + i\omega) + \alpha E^*(\hat{\rho} + \beta)] \\ &\quad - \alpha E^* \rho \beta = 0. \end{aligned} \quad (3.1.23)$$

Thus, both real part and imaginary part of  $f(\xi + i\omega)$  should be zero. That is,

$$\begin{aligned} Real(f(\xi + i\omega)) &= -(\hat{\gamma} + \xi)[\xi(\alpha E^* + \hat{\rho} + \xi) - \omega^2 + \alpha E^*(\hat{\rho} + \beta)] \\ &\quad - \alpha E^* \rho \beta + \omega^2(2\xi + \alpha E^* + \hat{\rho}) = 0, \end{aligned} \quad (3.1.24)$$

and

$$Im(f(\xi + i\omega)) = -(\hat{\gamma} + \xi)[\xi\omega + \omega(\alpha E^* + \hat{\rho} + \xi)]$$

$$- \omega[\xi(\alpha E^* + \hat{\rho} + \xi) + \alpha E^*(\hat{\rho} + \beta) - \omega^2] = 0. \quad (3.1.25)$$

Thus, for  $\omega \neq 0$ , we get,

$$\begin{aligned} Im(f(\xi + i\omega)) &= -(\hat{\gamma} + \xi)(2\xi + \alpha E^* + \hat{\rho}) \\ &\quad - [\xi\alpha E^* + \xi\hat{\rho} + \xi^2 + \alpha E^*(\hat{\rho} + \beta) - \omega^2] = 0. \end{aligned} \quad (3.1.26)$$

We will obtain a common root  $\xi$  for both Eqs. (3.1.24) and (3.1.26). First, by (3.1.24), we have

$$\begin{aligned} \xi^3 + (\alpha E^* + \hat{\rho} + \hat{\gamma})\xi^2 + (\hat{\gamma}(\alpha E^* + \hat{\rho}) - 3\omega^2 + \alpha E^*(\hat{\rho} + \beta))\xi \\ + [\hat{\gamma}\alpha E^*(\hat{\rho} + \beta) + \alpha E^*\rho\beta - \omega^2(\alpha E^* + \hat{\rho} + \hat{\gamma})] = 0, \end{aligned} \quad (3.1.27)$$

and by (3.1.26), we get

$$-3\xi^2 - 2(\alpha E^* + \hat{\rho} + \hat{\gamma})\xi - (\alpha E^*(\hat{\gamma} + \hat{\rho} + \beta) + \hat{\gamma}\hat{\rho} - \omega^2) = 0. \quad (3.1.28)$$

Let us take

$$\begin{aligned} k(\xi) &= \xi^3 + (\alpha E^* + \hat{\rho} + \hat{\gamma})\xi^2 + (\hat{\gamma}(\alpha E^* + \hat{\rho}) - 3\omega^2 + \alpha E^*(\hat{\rho} + \beta))\xi \\ &\quad + [\hat{\gamma}\alpha E^*(\hat{\rho} + \beta) + \alpha E^*\rho\beta - \omega^2(\alpha E^* + \hat{\rho} + \hat{\gamma})], \end{aligned} \quad (3.1.29)$$

and

$$h(\xi) = -3\xi^2 - 2(\alpha E^* + \hat{\rho} + \hat{\gamma})\xi - (\alpha E^*(\hat{\gamma} + \hat{\rho} + \beta) + \hat{\gamma}\hat{\rho} - \omega^2). \quad (3.1.30)$$

The roots of  $h(\xi)$  have the following form

$$\xi = \frac{\alpha E^* + \hat{\rho} + \hat{\gamma} \pm \sqrt{\Delta}}{-3}, \quad (3.1.31)$$

whenever  $\Delta \geq 0$ , where

$$\Delta = (\alpha E^* + \hat{\rho} + \hat{\gamma})^2 - 3(\alpha E^*(\hat{\gamma} + \hat{\rho} + \beta) + \hat{\gamma}\hat{\rho} - \omega^2).$$

But  $\Delta \geq 0$  if

$$\omega^2 \geq (\alpha E^*(\hat{\gamma} + \hat{\rho} + \beta) + \hat{\gamma}\hat{\rho}) - \frac{(\alpha E^* + \hat{\rho} + \hat{\gamma})^2}{3}.$$

If

$$\alpha E^*(\hat{\gamma} + \hat{\rho} + \beta) + \hat{\gamma}\hat{\rho} > \omega^2 \quad (3.1.32)$$

then both roots  $\xi_1$  and  $\xi_2$  of  $h(\xi)$  are negative.

If

$$\alpha E^*(\hat{\gamma} + \hat{\rho} + \beta) + \hat{\gamma}\hat{\rho} < \omega^2 \quad (3.1.33)$$

then  $\xi$  can be positive.

Note that when the reproduction number  $R_e$  tends to 1, then  $E^*$  tends to zero, and the disease-free and endemic equilibria collapse into a single equilibrium, that is  $X_0 = X_1$ . By Theorem 3.1.2, the disease-free equilibrium  $X_0$  at the tipping parameter  $R_e = 1$  admits 0 as an eigenvalue and two other eigenvalues are negative. So, the same holds for endemic equilibrium  $X_1$  at  $R_e = 1$ . Also, the equilibrium point  $X_1$  of the model and its eigenvalues depend continuously on model parameters. By Claim 3.1.2,  $J|_{X_1}$  does not admit pure imaginary complex eigenvalues. These observations show that the Jacobian matrix  $J|_{X_1}$  does not admits complex eigenvalues with positive real parts.

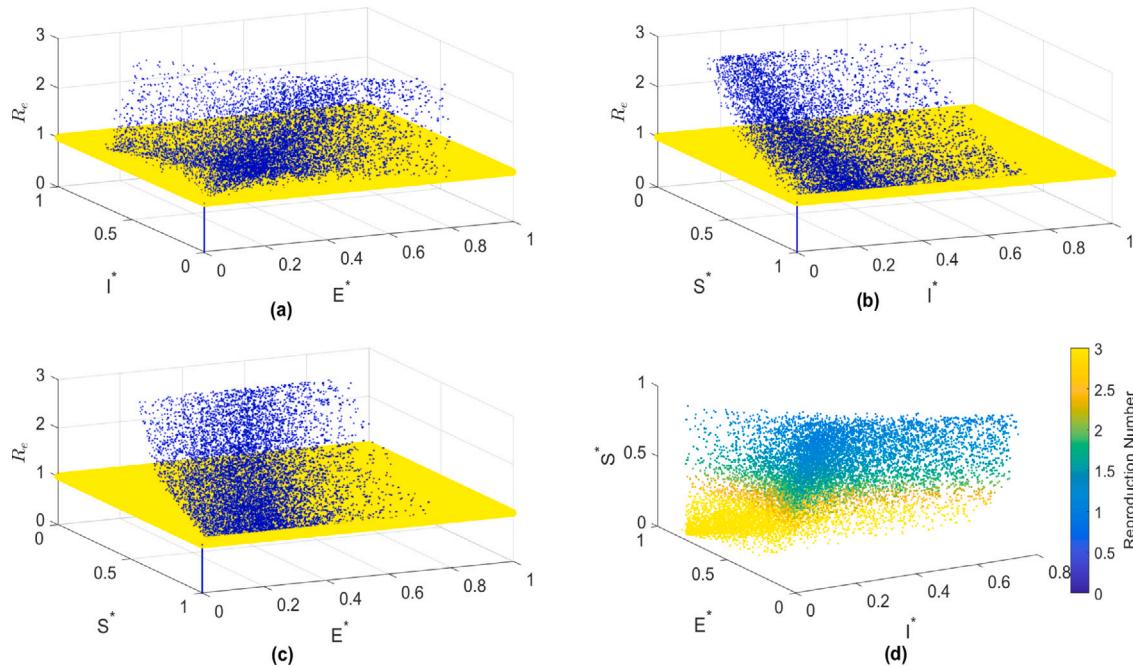
Now, in Eq. (3.1.29), let us take

$$Z = \hat{\gamma}(\alpha E^* + \hat{\rho}) - 3\omega^2 + \alpha E^*(\hat{\rho} + \beta),$$

and

$$W = \hat{\gamma}\alpha E^*(\hat{\rho} + \beta) + \alpha E^*\rho\beta - \omega^2(\alpha E^* + \hat{\rho} + \hat{\gamma}).$$

If  $\alpha E^*(\hat{\gamma} + \hat{\rho} + \beta) + \hat{\gamma}\hat{\rho} > \omega^2$  and  $\omega$  is sufficiently small, then both  $Z$  and  $W$  are positive, hence the Jacobian matrix  $J|_{X_1}$  has not a positive root. In this case,  $\xi$  should be negative. Consequently, by these observations, Claims 3.1.1 and 3.1.2, all eigenvalues of  $J|_{X_1}$  have negative real parts. This implies that the endemic equilibrium should be globally asymptotically stable if  $R_e > 1$ . This finishes the proof.  $\square$



**Fig. 2.** The trajectory of the quasi-static attractor  $A(\lambda(t))$  of the parameterized system versus  $R_e(t) = \frac{\alpha(t)}{\beta(t)+\mu}$ , when  $R_e(t)$  oscillates between 0.2 and 2. Plot (a) illustrates  $A(\lambda(t))$  in  $R_e I^* E^*$ -hyperplane, plot (b) illustrates  $A(\lambda(t))$  in  $R_e S^* I^*$ -hyperplane, plot (c) illustrates  $A(\lambda(t))$  in  $R_e S^* E^*$ -hyperplane, but plot (d) is four dimensional and illustrates  $A(\lambda(t))$  in  $R_e S^* E^* I^*$ -hyperplane. The tipping phenomena are observed.

Here, we have used the effective reproduction number  $R_e$  to assess the effectiveness of virus control strategies and provide a real-time picture of an outbreak. The index  $R_e$  depends on a population's current immune state, therefore, some important events including a mass vaccination campaign or a large scale outbreak can be highly successful in its reducing. [Theorems 3.1.2](#) and [3.1.3](#) show that when  $R_e < 1$ , the spread of the epidemic can be controlled and the number of infected individuals is decreased, however, if  $R_e > 1$  the controllability of the spread of the disease is more difficult.

Additionally, the nonautonomous instability has been described as bifurcation-induced tipping (or B-tipping) and it can be understood in terms of a classical autonomous bifurcation of the stable state in the frozen system. [Fig. 2](#) illustrates the presence of the tipping phenomenon for our model. Generally, when a trajectory starts at one attractor and ends at another, the tipping phenomenon occurs. When the reproduction number  $R_e$  is greater than 1 the quasi-static attractor  $A(t)$  is the endemic equilibrium  $X_1(t)$ . In [Fig. 2](#) plots (a), (b), (c), the reproduction number  $R_e$  oscillates between 0.2 and 2, hence, the quasi-static attractor  $A(t)$  changes from disease-free equilibrium  $X_0(t)$  to endemic equilibrium  $X_1(t)$  and vice-versa. Thus, we observe a dramatic change in the long-term performance of parameterized systems subjected to parameter drift.

**Definition 3.1.1.** We say that [25] the nonautonomous system (3.1.6) undergoes a *bifurcation-driven tipping* or *B-tipping* if  $\lambda(t)$  passes through a bifurcation point of the corresponding parameterized system. This occurs if  $\lambda(t)$  passes through the bifurcation point  $\alpha = \beta + \mu$ , where the quasi-static attractor  $X_0$  loses stability.

The occurrence of epidemic waves indicates that there should be trajectories connecting the endemic equilibrium and the disease-free equilibrium that leads to a bifurcation-driven tipping. By [Theorems 3.1.2](#) and [3.1.3](#), we get the next result.

**Theorem 3.1.4.** *The nonautonomous system (3.1.6) undergoes a bifurcation-driven tipping at  $R_e = 1$ .*

### 3.2. Introduced machine learning models

Machine learning (ML) is a field of inquiry devoted to understanding and building methods that 'learn', that is, methods that leverage data to improve performance on some set of tasks. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so [70]. Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, speech recognition, agriculture, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks. In this section, we review four of the most popular neural network methods. We have used these methods to analyze the data of the COVID-19 outbreak. The architecture of the machine learning methods used in this article is shown in [Fig. 3](#).

#### 3.2.1. Radial basis function neural networks (RBF)

As a part of mathematical modeling, an artificial neural network that uses radial basis functions as activation functions is known as radial basis function networks. Inputs and neuron parameters are combined linearly in the network's output. Depending on their purpose, radial basis function networks can be used to approximate functions, predict time series, and classify, and control systems. The Royal Signals and Radar Establishment researchers, Broomhead and Lowe introduced them in [71]. Three layers are typically present in radial basis function networks: the input layer, a hidden layer with a non-linear activation function, and the linear output layer. It is possible to model the input as a vector of real numbers  $x \in \mathbb{R}^n$ . Then the output of the network is a scalar function of the input vector,  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ , and is given by

$$\phi(x) = \sum_{i=1}^N a_i \rho(\|x - c_i\|) \quad (3.2.1)$$

where  $N$  is the number of neurons in the hidden layer,  $c_i$  is the center vector for neuron  $i$ , and  $a_i$  is the weight of neuron  $i$  in the linear output neuron (see [Fig. 3](#)). The name radial basis function refers to functions that are radially symmetric around a center vector. Each hidden neuron

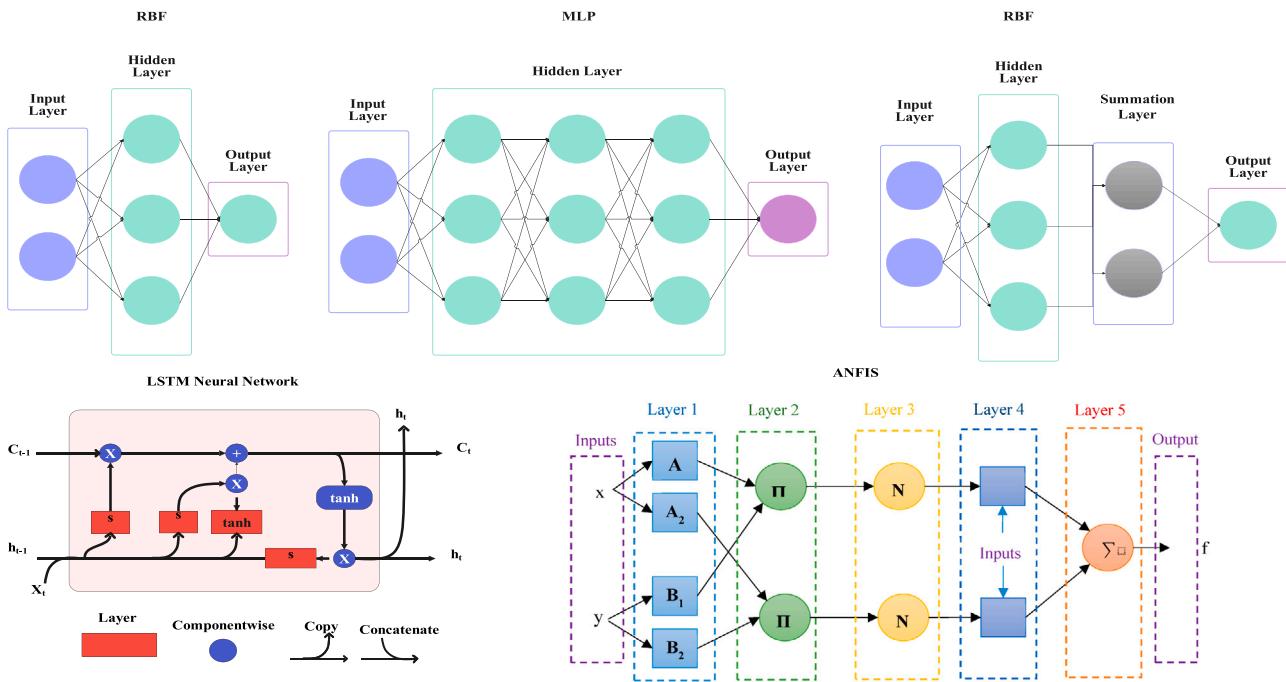


Fig. 3. Architecture of machine learning methods.

is connected to all inputs in the basic form. Radial basis functions are usually Gaussian and the norm Euclidean distance

$$\rho(\|x - c_i\|) = \exp(-\zeta_i \|x - c_i\|^2) \quad (3.2.2)$$

where

$$\lim_{\|x\| \rightarrow \infty} \rho(\|x - c_i\|) = 0. \quad (3.2.3)$$

It means that modifying the parameters of one neuron does not have a great effect on input values that are very far from its center.

### 3.2.2. Multilayer perceptron and neural networks

In artificial neural networks, multilayer perceptrons (MLPs) are fully connected, classes. It is unclear whether the term MLP refers to feedforward ANNs or networks composed of multiple layers of perceptrons (with threshold activation). Nodes in an MLP are arranged in at least three layers: an input layer, a hidden layer, and an output layer (see Fig. 3). With the exception of the input nodes, all nodes are neurons with nonlinear activation functions. In order to train MLPs, backpropagation is utilized as a supervised learning method [72]. Any number of layers can be reduced to a two-layer input–output model as long as all neurons have linear activation functions, that is, linear functions that map weighted inputs to outputs. It has long been known that the sigmoid function is the most common activation function for MLP which is described by

$$y(x_i) = \tanh(x_i), \quad \text{or} \quad y(x_i) = \frac{1}{1 + e^{-x_i}} \quad (3.2.4)$$

### 3.2.3. The general regression neural network

General regression neural network (GRNN) is one of the probabilistic neural networks proposed by Donald F. Specht in [73]. Probabilistic neural networks, such as this one, require a fraction of the training samples necessary for backpropagation neural networks. It is rarely enough to train a backpropagation neural network based on measurements of an operating system. It is therefore especially advantageous to use a probabilistic neural network since it is capable of identifying the underlying function of data even with only a few training samples. For a satisfying fit, the user needs to have relatively little knowledge and can do so without additional input. Therefore, GRNN is an excellent tool for

predicting and comparing system performance. Normal distribution is used in GRNN for probability density functions. Each training sample,  $X_j$ , is used as the mean of a Normal Distribution.

$$Y(X) = \frac{\sum_{i=1}^n Y_i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}, \quad D_i^2 = (X - X_i)^T (X - X_i). \quad (3.2.5)$$

Training samples are compared with prediction points to determine how well they can represent prediction points,  $X$ , by measuring the distance,  $D_j$ , between them. Whenever  $D_j$  is small between the training sample and prediction point,  $\exp\left(-\frac{D_j^2}{2\sigma^2}\right)$  increases. Whenever  $D_j = 0$ ,  $\exp\left(-\frac{D_j^2}{2\sigma^2}\right)$  becomes one, which is the best representation of the point of evaluation. The distance to all the other training samples is bigger. The term  $\exp\left(-\frac{D_j^2}{2\sigma^2}\right)$  becomes smaller as the distance  $D_j$  increases, and thus the other training samples make relatively little contribution to the prediction. A large part of the prediction is based on the term  $Y_j \times \exp\left(-\frac{D_j^2}{2\sigma^2}\right)$ , which is the biggest for the  $j$ th training sample. In Fig. 3 GRNN is shown. A GRNN pattern neuron calculates  $\exp\left(-\frac{D_i^2}{2\sigma^2}\right)$  (normal distribution centered at every training sample) for each training sample. Signals from pattern neurons  $i$  are weighted by the training samples  $Y_i$  going into the denominator neurons. The weights on the signals going into the Numerator Neuron are one. GRNN predicts each point based on a sample from the training data.

### 3.2.4. Long short-term memory neural network

Long short-term memory (LSTM) is a type of artificial neural network used in the field of artificial intelligence and deep learning [74]. There are feedback connections in LSTMs, which differ from feedforward neural networks. In addition to processing single data points (such as images), recurrent neural networks (RNNs) are also capable of processing entire sequences of data (such as speech or video recording). One of the most cited neural networks of the 20th century is the LSTM [75]. LSTMs are named based on the analogy that standard RNNs have both “long-term memory” and “short-term memory”. In the network, the weights and biases of connections change once per training

episode, similar to how synaptic strength changes in the brain store long-term memories; in the network, activation patterns change one-time step, just as short-term memories are stored moment-to-moment by the brain's electric firing patterns. There are four basic components of an LSTM unit: a cell, an input gate, an output gate, and a forget gate (See Fig. 3). During arbitrary time intervals, the cell remembers value, and three gates control the flow of information.

In the equations below, the lowercase variables represent vectors. Input and recurrent connections are weighted by matrices  $W_q$  and  $U_q$ , where  $q$  represents either input gate  $i$ , output gate  $o$ , forget gate  $f$ , or memory cell  $c$ , depending on the activation being calculated. We use vector notation to show mathematical relations. For an LSTM cell with a forget gate, the compact equations are as follows:

$$\begin{aligned} f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f) \\ i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i) \\ o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o) \\ \tilde{c}_t &= \tanh(W_c x_t + U_c h_{t-1} + b_c) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \quad (3.2.6)$$

where  $\sigma$  is a sigmoid function, the initial values are  $c_0 = 0$  and  $h_0 = 0$  and the operator  $\odot$  denotes the Hadamard product (element-wise product). The subscript  $t$  indexes the time step. Also,  $x_t \in \mathbb{R}^d$  denotes the input vector to the LSTM unit,  $f_t \in (0, 1)^h$  is forgotten gate's activation vector,  $i_t \in (0, 1)^h$  denotes input/update gate's activation vector,  $o_t \in (0, 1)^h$  denotes output gate's activation vector,  $h_t \in (-1, 1)^h$  denotes hidden state vector also known as output vector of the LSTM unit,  $\tilde{c}_t \in (-1, 1)^h$  denotes cell input activation vector,  $c_t \in \mathbb{R}^h$  denotes cell state vector, and  $W \in \mathbb{R}^{h \times d}$ ,  $U \in \mathbb{R}^{h \times d}$ , and  $b \in \mathbb{R}^h$  denote weight matrices and bias vector parameters which need to be learned during training. Also, the superscripts  $d$  and  $h$  refer to the number of input features and several hidden units, respectively. Active functions are sigmoid function and hyperbolic tangent function.

### 3.2.5. Adaptive Neuro-Fuzzy inference system (ANFIS) approach

Adaptive neuro-fuzzy inference systems (ANFIS) are based on Takagi-Sugeno fuzzy inference systems and are a type of artificial neural network. The technique was developed in the early 1990s [76]. It combines the advantages of neural networks and fuzzy logic in one system because it integrates both systems. IF-THEN fuzzy rules form the inference system, which is flexible, adaptive, and capable of approximating nonlinear functions. Therefore, ANFIS is considered as a universal estimator. When parameters are obtained by a genetic algorithm, ANFIS can be more efficient. It is used in an intelligent situational energy management system.

Two parts can be identified in the network structure, namely the part with the preconditions and the part with the consequences. Specifically, the architecture consists of five layers (See Fig. 3). The first layer takes the input values and determines the associated membership functions. It is commonly referred to as the fuzzification layer. The membership degrees of each function are computed based on the parameter set of the premise, namely {a,b,c}. The second layer is responsible for generating the firing strengths for the rules. Due to its task, the second layer is called the "control layer". The third layer is responsible for normalizing the calculated firing strengths by dividing each value by the total firing strength. The fourth layer takes as input the normalized values and the follow-up parameters {p,q,r}. The values returned from this layer are the de-normalized values, and these values are passed to the last layer to get the final output.

### 3.3. Performance evaluation indices

A machine learning pipeline begins with data preparation and training and ends with evaluating the trained model's performance. Machine learning models are classified as adaptive or non-adaptive according to their ability to generalize data that has not yet been seen. Our model would be more predictive if we used different performance metrics for performance evaluation before applying it to unseen data. If ML models are not evaluated using various metrics before they are applied to unseen data, problems can arise, as well as poor predictions. Consequently, it does not generalize well with unseen data because the model does not learn and remember.

It is possible to classify machine learning tasks, similar to performance metrics, as either regression or classification. We mention the most common metrics related to both problems, along with information about model performance. It is also important to remember that metrics are different from loss functions. Model performance is measured by loss functions. A machine learning model is trained using them (using an optimization technique like gradient descent), and the parameters are usually differentiable. A model's performance (during training and testing) is tracked and measured using metrics and does not need to be differentiable. It is also possible to use the performance metric as a loss function (perhaps with some additional regularization), such as MSE, if it is differentiable for some tasks.

Regression models have a continuous output. Therefore, we need a metric based on calculating some sort of distance between the prediction and the actual outcome. Some important regression metrics are Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and  $R^2$  (R-Squared).

There is much research in the field of classification problems. It is almost impossible to find a manufacturing or industrial environment without use cases. Models with classification output generate discrete datasets, so we must find a metric that is capable of comparing discrete classes. There are various classification metrics that evaluate the accuracy or inaccuracy of a model, but they all evaluate it differently. Some important metrics to evaluate classification models are Accuracy, Confusion Matrix (not a metric but fundamental to others), Precision and Recall, F1-score, and AU-ROC.

Here, we use RMSE, Accuracy, and R-Squared that define as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{N}}, \quad (3.3.1)$$

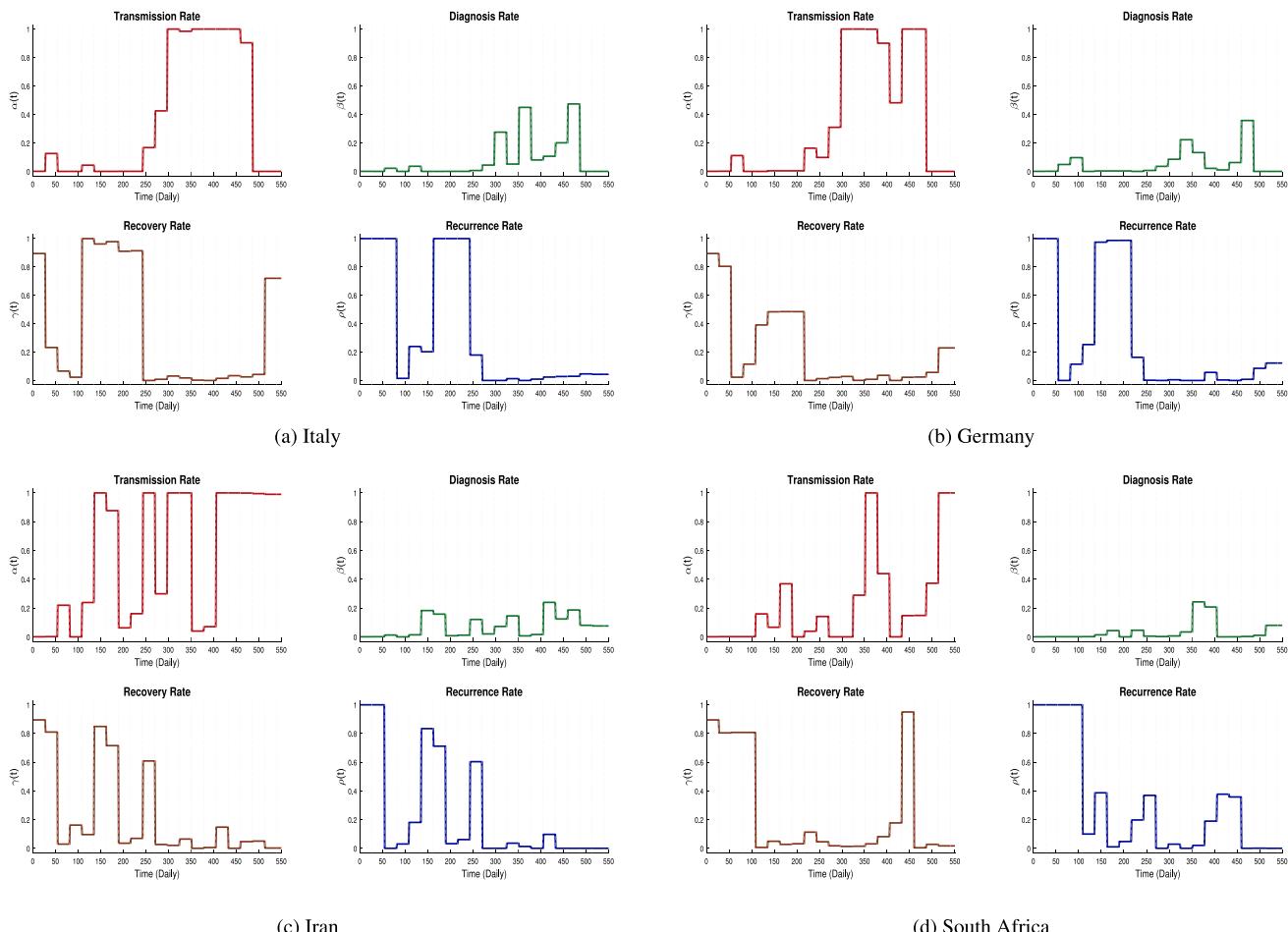
$$Accuracy = 1 - \sqrt{\frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{\sum_{i=1}^N (Predicted_i)^2}}, \quad (3.3.2)$$

$$R - Squared = 1 - \frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{\sum_{i=1}^N (Actual_i)^2}. \quad (3.3.3)$$

## 4. Numerical analysis

### 4.1. Datasets

A dataset from WHO was used, which included daily COVID-19 reports and global geographical distributions. In the first step, different countries' data was examined, and then countries from different continents were selected based on their important features. Therefore, the targeted countries in this article are Italy and Germany from Europa, Iran from Asia, and South Africa from the African continent. To understand the evolution of the COVID-19 outbreak, we used 550 daily data of national cumulative infection numbers from Germany, Italy, Iran, and South Africa from Jan 22, 2020, to Sep 24, 2021, and divided the data into two, which are training set (439 days), and test set (111 days). This duration has many features, including different types of variants and their consequences, like several peaks of COVID-19 that can make it difficult to analyze.



**Fig. 4.** Parameters for the recurrent model (3.1.1). A panel plot of transmission rate, diagnosis rate, recovery rate, and recurrence rate for four countries Italy, plot (a), France, plot (b), Iran, plot (c), and South Africa, plot (d).

First, we evaluate all 550 days for analysis with the help of the dynamic model, and then we use the data in two categories (trained and tested data) for the forecasting process. Finally, we can compare the results of the dynamical model with machine learning methods in prediction scenarios and fitting processes.

#### 4.2. Estimating parameters of dynamic model

Here, we consider an optimization algorithm for estimating the parameters of recurrent dynamical model (3.1.1) including transmission rate,  $\alpha(t)$ , diagnosis rate,  $\beta(t)$ , recovered rate,  $\gamma(t)$ , and recurrence rate,  $\rho(t)$ . First, we assume that all parameters are piecewise constant. We divided all-time duration into 20-time intervals. The results of parameters for four targeted countries are shown in Fig. 4. It is essential to note that data are available for only two groups of the population: recovered and infected individuals. This is a big challenge for the fitting process. The algorithm used for parameter estimation must be able to succeed in this challenge. Therefore, susceptible and exposed individuals are approximated by the optimization algorithm.

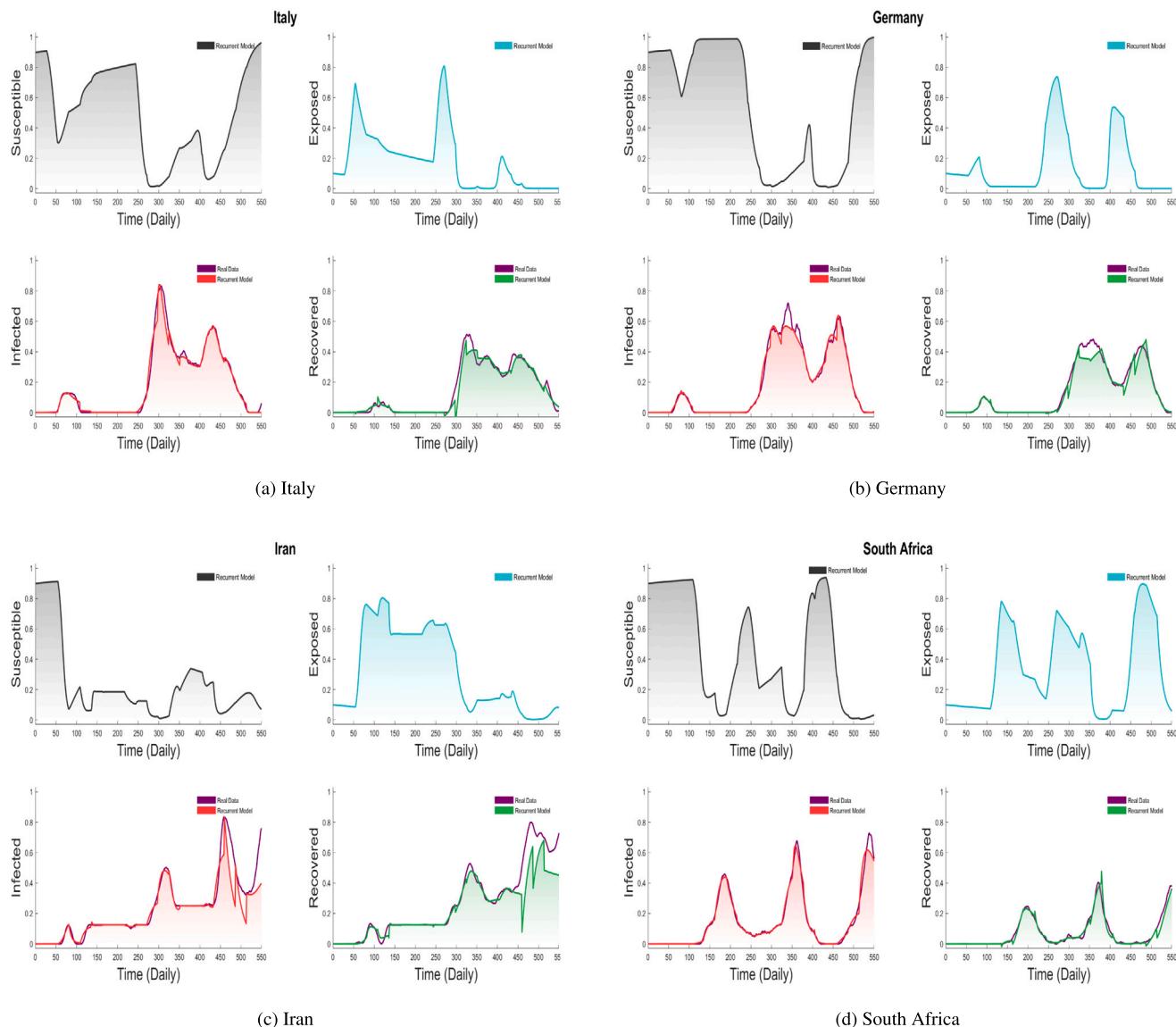
As can be seen in Fig. 4, all the countries studied had low transmission rates at the beginning of the COVID-19 pandemic. This period is significant for the decision-makers to be able to control the disease by applying non-pharmaceutical interventions like lockdown, wearing a mask, and so on. Also, given that all countries are facing a new pandemic, it is natural that the recurrence rate is high. Because not only hospitals and medical staff are not prepared in terms of information and equipment, but also the body's immune system is weak against the new virus. The recurrence rate decreases over time, which can be due

to people's awareness of COVID-19 and compliance with its protocols. Also, due to the change of the virus to a new variant and the weak performance of hospitals in critical situations, the diagnosis rate of the disease is always low. For this reason, it is very difficult to control the epidemic of COVID-19.

After estimating parameters, we solve the recurrent dynamical model (3.1.1) with initial conditions to get a numerical solution. In this part, we compare the results of our model with the actual data of COVID-19 in two compartments including recovered and infected cases in Fig. 5. Also, we approximated daily cases of susceptible and exposed individuals in targeted countries in Fig. 5. It can be seen that the model and algorithm used in this article have a good performance in the fitting process. With respect to all possible combinations of any three individual populations of the model, the phase portrait diagrams of stability analysis is shown for Italy, Germany, Iran, and South Africa in Figs. 6, 7, 8, and 9. A red point indicates the starting point and a blue point indicates the final destination.

### *4.3. Prediction process*

In different papers, Data-driven approaches have been developed for modeling COVID-19, based on statistical and neural network methods, to predict cases, hospitalizations, recovered, and deaths, and to instigate the effect of some control factor on the epidemic. We have compared two categories of methods for predicting COVID-19 in this section; the classic one including the recurrent dynamic model, and the machine learning methods including GRNN, RBF, LSTM, MLP, and ANFIS.



**Fig. 5.** Trajectory of the COVID-19 epidemic in targeted countries using recurrent model (3.1.1). Daily results are investigated: Comparison between model output and observed data for (a) Italy, (b) France (c) Iran and (d) South Africa for all susceptible, exposed, infected, and recovered individuals.

For evaluating the performance of our model in the prediction scenario, we estimate the parameters of our recurrent model just for initial 439 days and then use the subsequent 111 days of actual data to evaluate the model's accuracy and conformity. The results of this study evaluate the effectiveness of dynamic models for predicting COVID-19. We also evaluated a similar process using machine learning methods. By implementing neural network methods including LSTM, GRNN, RBF, ANFIS, and MLP, we have given 439 days of initial data as training to the neural networks and obtained 111 days of prediction from each of them. We obtain the results for recovered and infected compartments. The forecast for the infected population and recovered population can be seen in Figs. 10, and 11 respectively.

#### 4.4. Performance evaluation outcome

As can be seen in Table 2, we applied RMSE evaluation for all countries in two zones of training data and testing data for two compartments of recovered and infected individuals. Six methods have been compared. Based on the output of RMSE, the number corresponding to the best method is bolded in each row of the table to distinguish it from the rest of the numbers. For example, for Iran, the performance of

MLP in the trained zone is better than all methods for both infected and recovered. Also, in tested data, ANFIS is the best in both compartments. For convenience, we calculated a general measure at the end of Table 2 so that we can have a better comparison under RMSE. We calculated two means for the performance of different methods in two zone of trained data and tested data. Based on these two criteria, we can rank the methods in terms of their performance. According to RMSE, in the zone of trained data, model ranking from the best method to the weakest one is MLP, ANFIS, RBF, Dynamic system (DS), GRNN, and LSTM, respectively. Also, in the zone of tested data, ANFIS and RBF are superior in predicting compared to other machine learning models. Models ranking from good performance to the lowest in entire scenarios is ANFIS, RBF, LSTM, MLP, GRNN, and DS, respectively.

Similar to Table 2, the same results can be derived from Tables 3 and 4. In Tables 3 and 4, we observe the numbers of evaluation metrics are close to 1 which indicate a better fit between prediction and actual value. These numbers tell us how well our model is making all its predictions on a scale of zero to one. According to Table 3, model ranking from the best method to the weakest one for trained data is ANFIS, MLP, RBF, LSTM, DS, and GRNN. In addition, this ranking for tested data is ANFIS, RBF, MLP, LSTM, GRNN, and DS. This shows

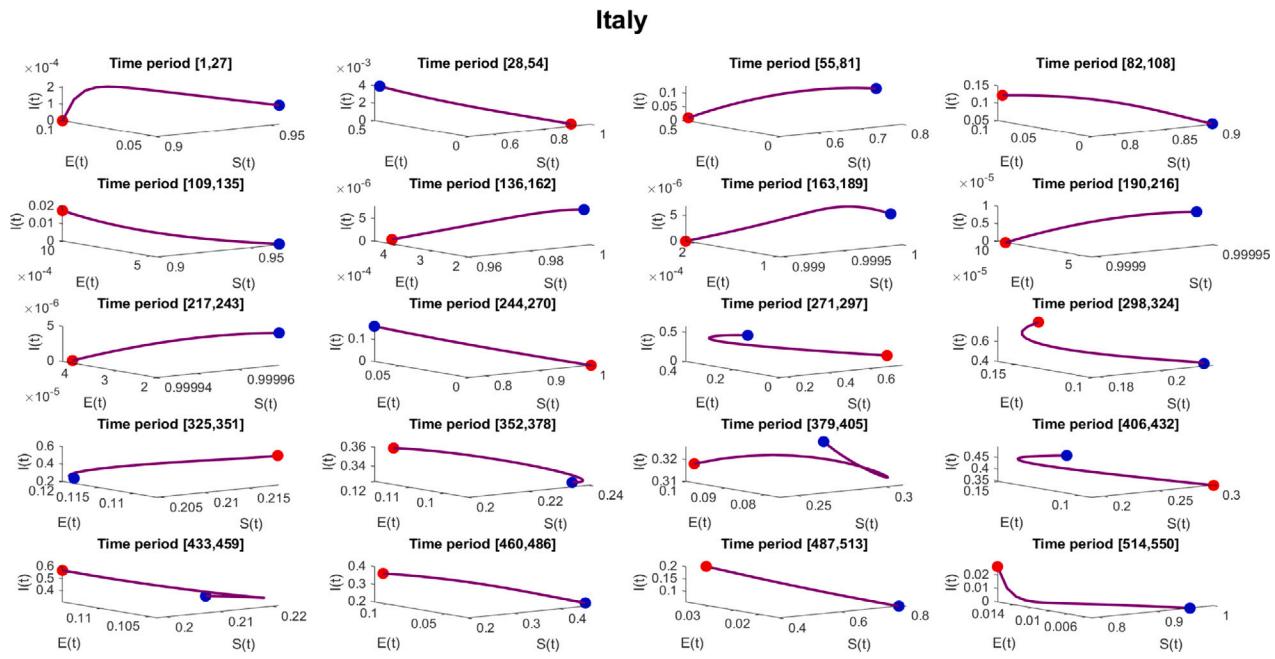


Fig. 6. An illustration of phase portraits based on all possible combinations for Italy. The red point marks the starting point, and the blue point marks the finish.

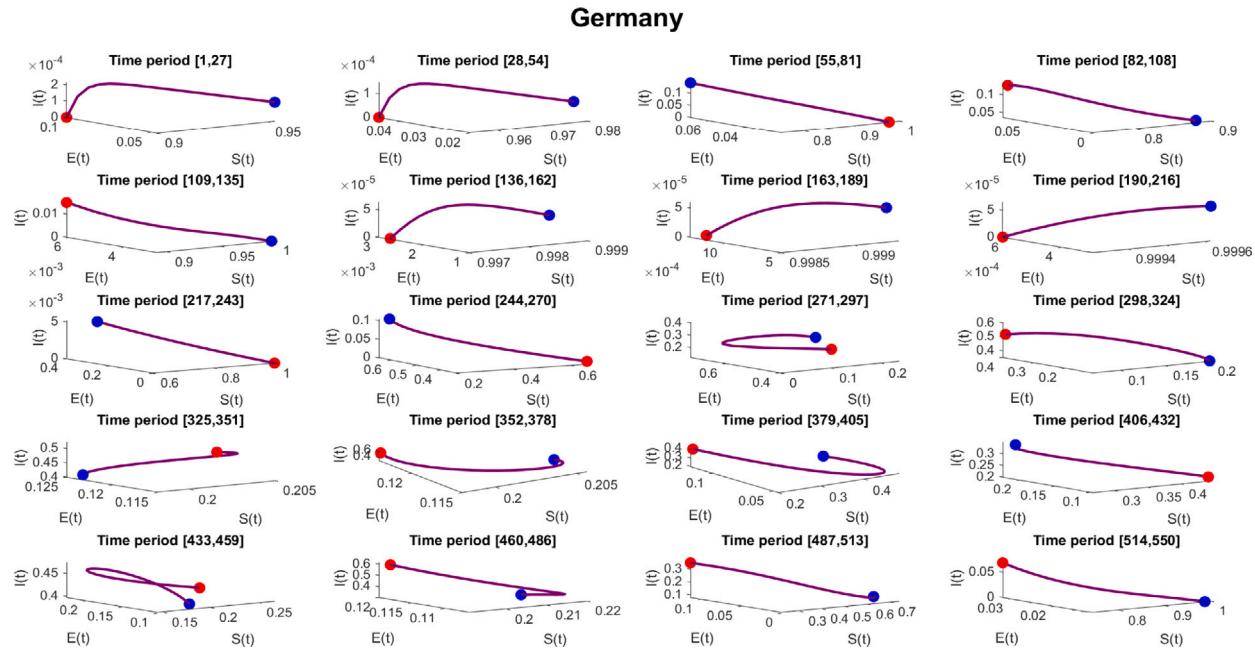


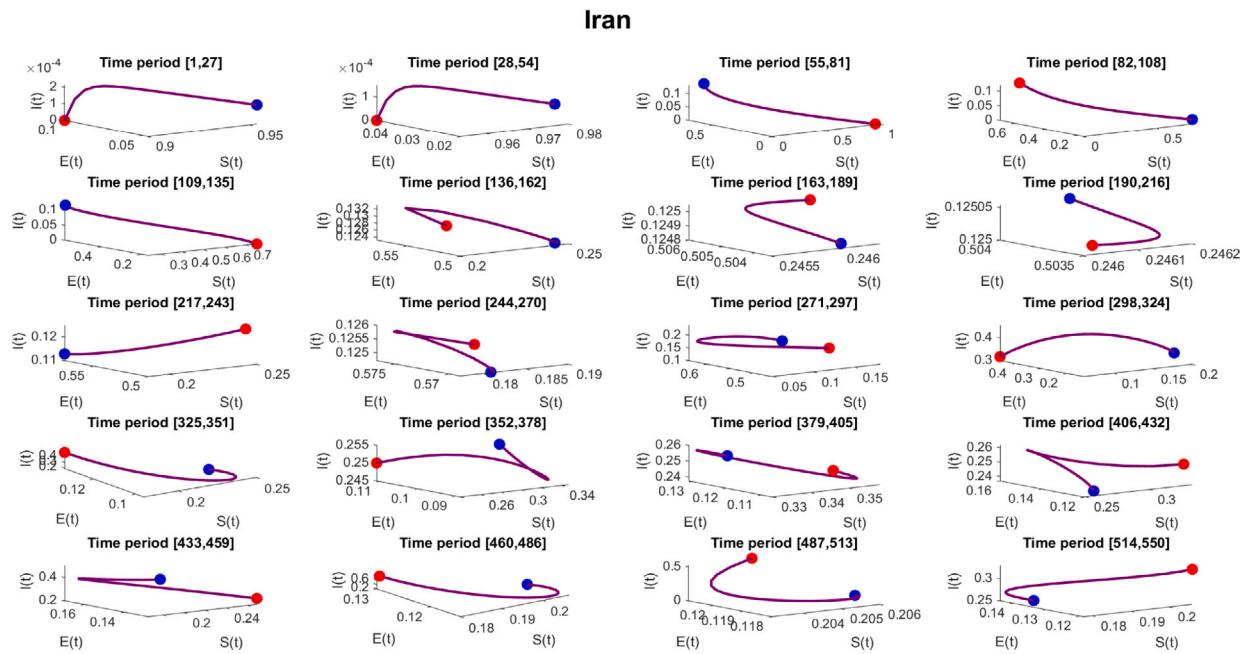
Fig. 7. An illustration of phase portraits based on all possible combinations for Germany. The red point marks the starting point, and the blue point marks the finish.

that the DS method is not suitable for prediction. Similarly, based on Table 4, the ranking of methods for trained data is ANFIS, RBF, MLP, LSTM, DS, and GRNN. In addition, this ranking for tested data is ANFIS, RBF, MLP, LSTM, GRNN, and DS. This shows that ANFIS has a higher application with accurate prediction and fitting.

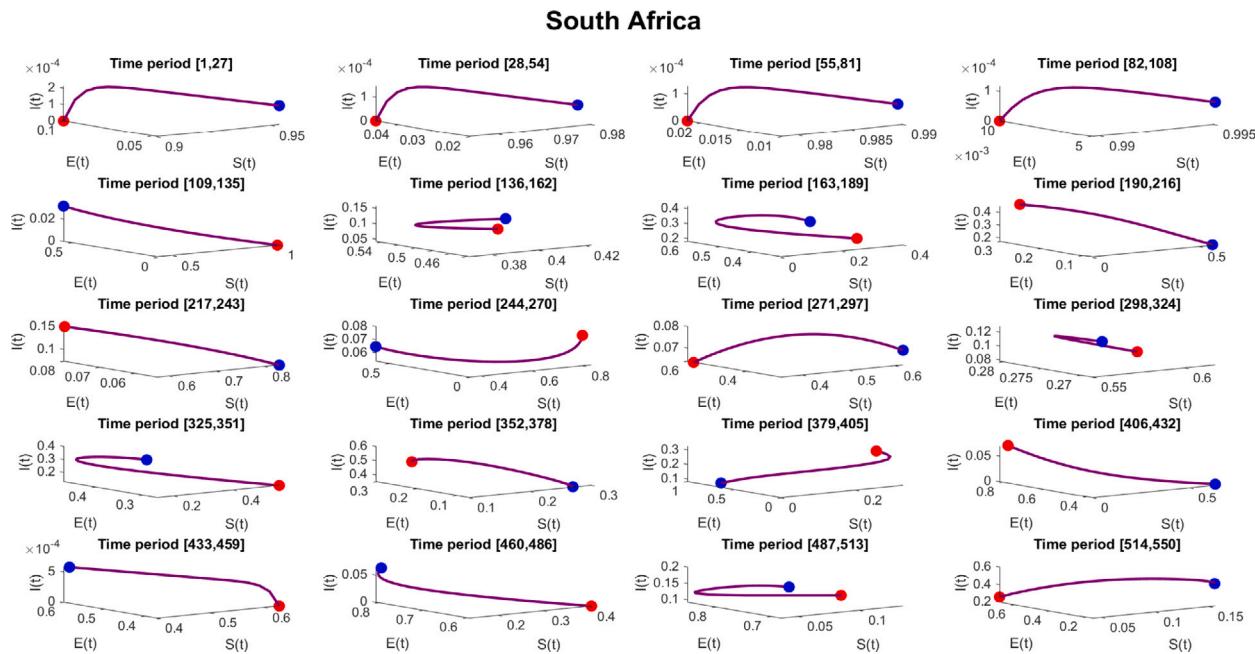
## 5. Results and discussions

This research has developed an approach to combine two widely used epidemiological models from two different categories. In the first part, we have proposed a nonautonomous recurrent compartmental model governed by ordinary differential equations to explain the

behavior of the epidemic. It has been conducted to consider both time-varying infectivity enabling to capture modifications of the epidemic characteristics and recurrence due to reinfection. For the nonautonomous system, we have considered the corresponding parameterized system with frozen parameters and illustrated its stable solutions which refer to them as quasi-static attractors. We have investigated the occurrence of parameter drift due to time-dependent changes in parameter sets. When deterministic systems with parameter drift witness a dramatic change in performance, a tipping phenomenon occurs. Here, we have shown that our nonautonomous system admits a pullback attractor in which fibers are contained in a positive flow invariant probability simplex. Then, the geometry of the pullback attractor was described using the parameterized system with frozen parameters. Also,



**Fig. 8.** An illustration of phase portraits based on all possible combinations for Iran. The red point marks the starting point, and the blue point marks the finish.



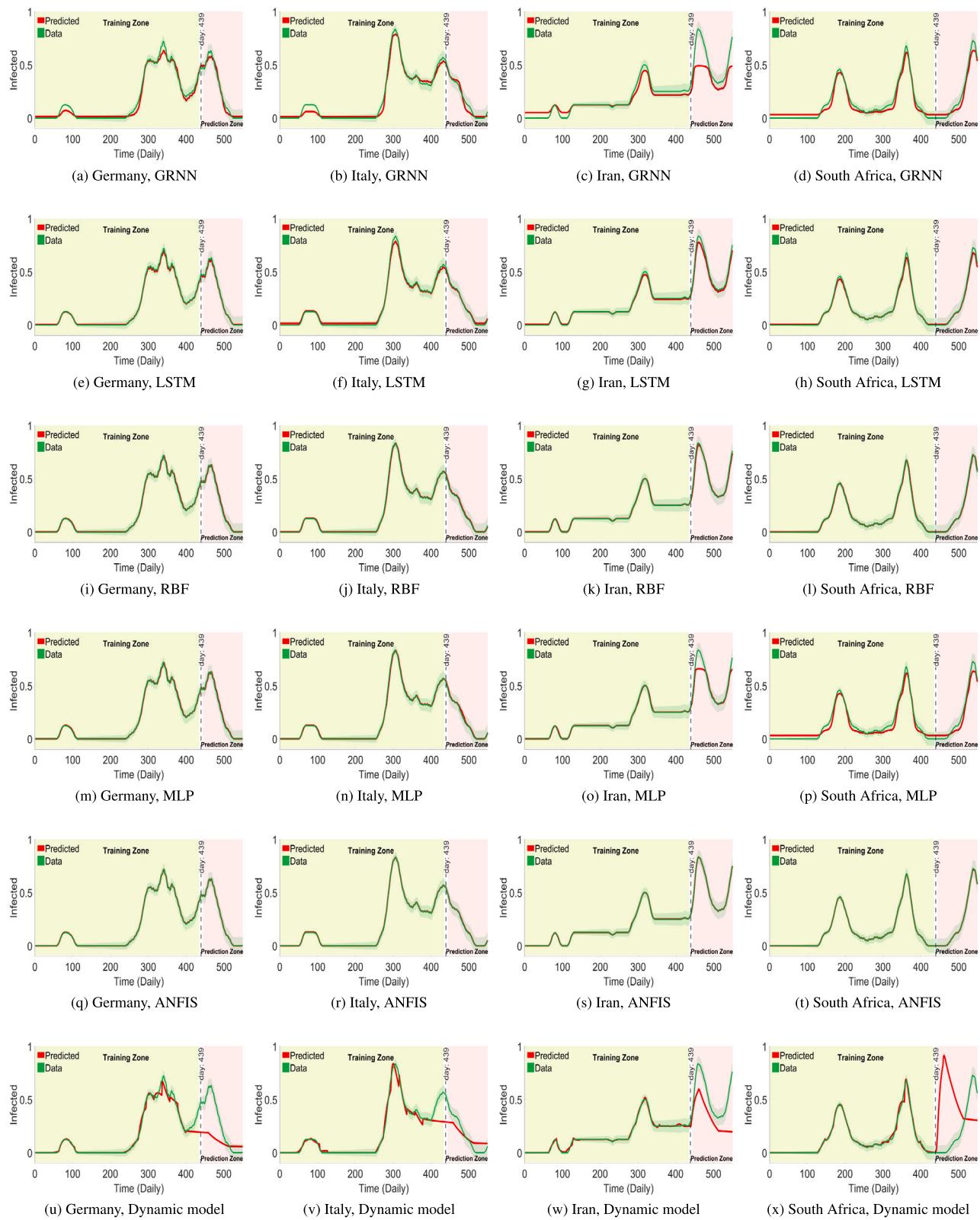
**Fig. 9.** An illustration of phase portraits based on all possible combinations for South Africa. The red point marks the starting point, and the blue point marks the finish.

we have shown that our model undergoes a bifurcation in the frozen-in case. In the second part, we have selected some Machine Learning (ML) techniques to handle past data and predict the spread of COVID-19. The predictive methods included ANFIS, RBF, LSTM, MLP, and GRNN. Generally, ML models are deployed for feature prediction involving risk and epidemic analysis. In Section 3, we have carried out a profound analysis to predict the COVID-19 outbreak using these ML techniques. As an important goal, this study has examined the predictive accuracy of the selected ML algorithms and compared their performance with the proposed dynamic model. We have reported the results of this goal in the fourth section. Furthermore, we have applied our models to investigate COVID-19 cases to targeted countries including Italy, Germany, Iran, and South Africa. For our dynamic model, the parameters have been estimated based on the officially reported data and

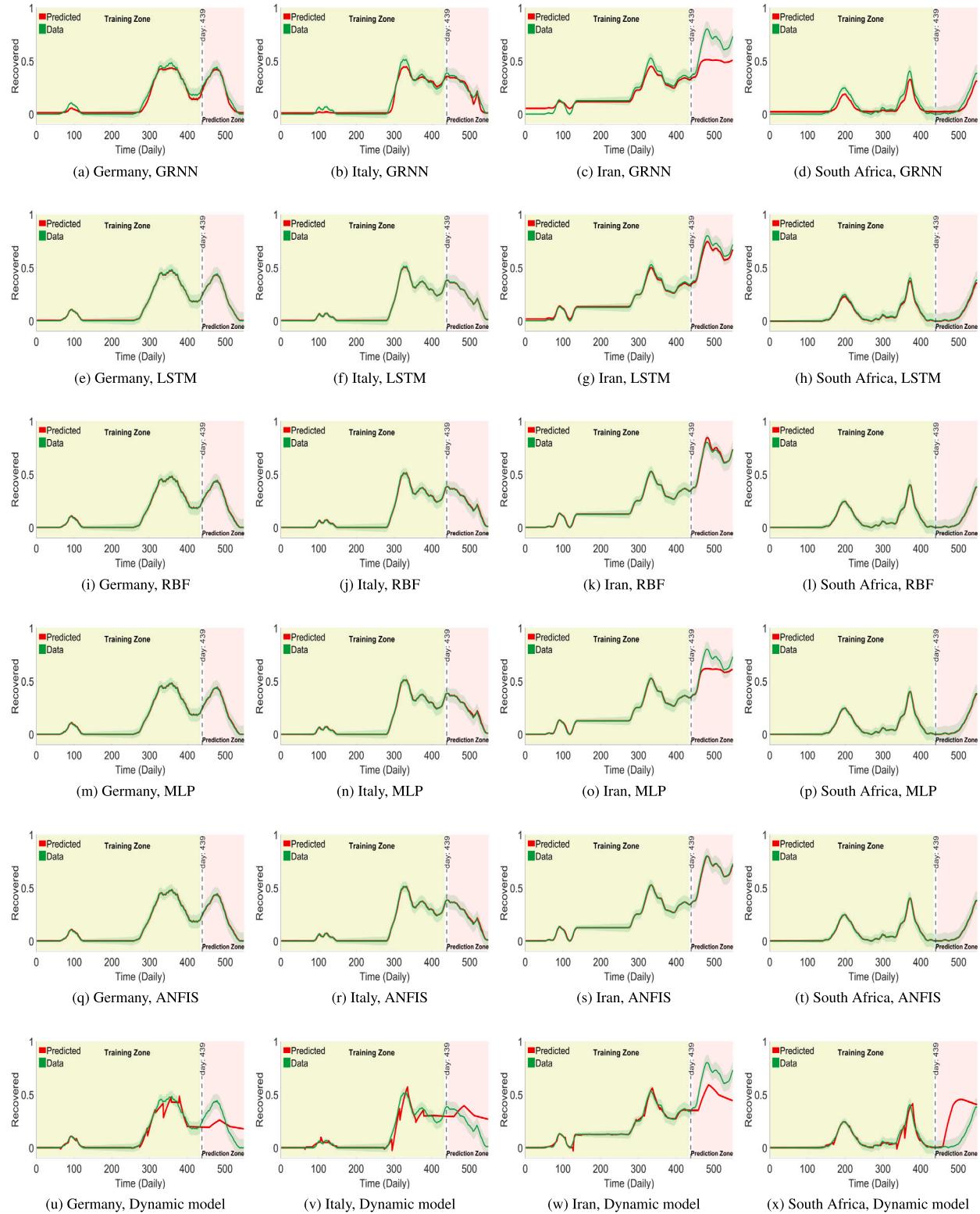
reflect the actual dynamics of infectious diseases. It causes the model evaluates the effect of the implementation of protocols such as nationwide lockdowns, social distancing measures, and so on. Some metrics including RMSE, R-Squared, and accuracy, were used to evaluate the presented models. Our findings suggest that the models with the best performance, fitting, and prediction are ANFIS and RBF.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.



**Fig. 10.** Trajectory of the COVID-19 prediction in Germany, Italy, Iran, and South Africa. Prediction scenario for infected individuals are done by recurrent dynamical model and machine learning methods including LSTM, GRNN, RBF, ANFIS and MLP.



**Fig. 11.** Trajectory of the COVID-19 prediction in Germany, Italy, Iran, and South Africa. Prediction scenario for recovered individuals are done by recurrent dynamical model and machine learning methods including LSTM, GRNN, RBF, ANFIS and MLP.

**Table 2**  
RMSE evaluation.

Country	Data type	Compartments	GRNN	RBF	LSTM	MLP	ANFIS	DS
Iran	Trained	Recovered	0.0351	0.0049	0.2409	<b>0.0046</b>	0.0047	0.0077
		Infected	0.0324	0.0045	0.2803	<b>0.0043</b>	0.0044	0.0073
	Tested	Recovered	0.1660	0.0226	0.0527	0.0726	<b>0.0171</b>	0.3294
		Infected	0.1795	0.0198	0.0230	0.0572	<b>0.0099</b>	0.3162
Germany	Trained	Recovered	0.0281	0.00743	0.3144	0.00742	<b>0.00741</b>	0.0416
		Infected	0.0245	0.0043	0.2237	<b>0.0042</b>	0.0043	0.0140
	Tested	Recovered	0.0274	0.0118	<b>0.0108</b>	0.0119	0.0119	0.2428
		Infected	0.0206	<b>0.00780</b>	0.0170	0.0082	0.00781	0.1809
Italy	Trained	Recovered	0.0294	0.0083	0.3325	0.0083	<b>0.0082</b>	0.0176
		Infected	0.0251	0.0050	0.2336	<b>0.0048</b>	0.0049	0.0147
	Tested	Recovered	0.0294	<b>0.0088</b>	0.0146	0.0089	0.0101	0.3503
		Infected	0.0275	<b>0.0080</b>	0.0139	0.0101	0.0094	0.1378
South Africa	Trained	Recovered	0.0336	0.008453	0.2466	0.0086	<b>0.008452</b>	0.0119
		Infected	0.0343	0.005497	0.1285	0.0056	<b>0.005496</b>	0.0103
	Tested	Recovered	0.0469	<b>0.0116</b>	0.0217	0.0207	0.0117	0.3376
		Infected	0.0433	<b>0.00533</b>	0.0213	0.0060	0.00534	0.1541
Mean of trained data			0.0303	0.0060	0.2501	<b>0.00597</b>	0.00598	0.0156
Mean of tested data			0.0675	0.0119	0.0218	0.0244	<b>0.0104</b>	0.2561

**Table 3**  
Accuracy evaluation.

Country	Data type	Compartments	GRNN	RBF	LSTM	MLP	ANFIS	DS
Iran	Trained	Recovered	0.8056	0.9756	0.9326	0.9769	<b>0.9767</b>	0.9613
		Infected	0.8440	0.9797	0.9553	<b>0.9808</b>	0.9805	0.9673
	Tested	Recovered	0.5910	0.9583	0.8957	0.8561	<b>0.9690</b>	-0.3450
		Infected	0.6242	0.9699	0.9631	0.9045	<b>0.9845</b>	0.0794
Germany	Trained	Recovered	0.8964	0.9738	<b>0.9823</b>	0.9738	0.9738	0.8446
		Infected	0.8691	0.9783	0.9325	<b>0.9786</b>	0.9784	0.9288
	Tested	Recovered	0.9203	0.9665	<b>0.9689</b>	0.9661	0.9662	-0.3532
		Infected	0.9240	<b>0.97229</b>	0.9359	0.9707	0.97228	-0.5780
Italy	Trained	Recovered	0.9011	0.9728	0.9706	0.9727	<b>0.9732</b>	0.9417
		Infected	0.8690	0.9751	0.9627	<b>0.9759</b>	0.9757	0.9259
	Tested	Recovered	0.8735	<b>0.9631</b>	0.9413	0.9627	0.9575	0.2894
		Infected	0.8884	<b>0.9689</b>	0.9478	0.9606	0.9634	0.5552
South Africa	Trained	Recovered	0.8250	0.95978	0.9568	0.9586	<b>0.95979</b>	0.9429
		Infected	0.6014	0.95187	0.8714	0.9501	<b>0.95188</b>	0.9081
	Tested	Recovered	0.8473	<b>0.9655</b>	0.9325	0.9372	0.9654	-99.9182
		Infected	0.6346	<b>0.9651</b>	0.8433	0.9603	0.9650	-81.9747
Mean of trained data			0.8264	0.9708	0.9455	0.9709	<b>0.9712</b>	0.9275
Mean of tested data			0.7879	0.9661	0.9285	0.9397	<b>0.9679</b>	-22.7806

**Table 4**  
R-squared evaluation.

Country	Data type	Compartments	GRNN	RBF	LSTM	MLP	ANFIS	DS
Iran	Trained	Recovered	0.9695	0.9994	<b>0.9996</b>	0.9994	0.9994	0.9985
		Infected	0.9795	0.9995	0.9981	0.999627	<b>0.999629</b>	0.9989
	Tested	Recovered	0.9114	0.9983	0.9910	0.9830	<b>0.9990</b>	0.6448
		Infected	0.9237	0.9990	0.9987	0.9922	<b>0.9997</b>	0.7589
Germany	Trained	Recovered	0.9901	0.9993	<b>0.9996</b>	0.9993	0.9993	0.9781
		Infected	0.9851	0.99953	0.9959	<b>0.99954</b>	0.99953	0.9950
	Tested	Recovered	0.9938	0.9988	<b>0.9990</b>	0.9988	0.9988	0.5295
		Infected	0.9945	<b>0.999224</b>	0.9963	0.9991	0.999223	0.5810
Italy	Trained	Recovered	0.9908	0.99926	0.9991	0.99926	<b>0.99928</b>	0.9966
		Infected	0.9846	0.9993	0.9985	<b>0.99942</b>	0.99941	0.9946
	Tested	Recovered	0.9838	<b>0.99855</b>	0.9961	0.99852	0.9980	-1.0492
		Infected	0.9881	<b>0.9989</b>	0.9969	0.9983	0.9986	0.7135
South Africa	Trained	Recovered	0.9744	0.998385	0.9982	0.9982	<b>0.998386</b>	0.9967
		Infected	0.9098	0.99768	0.9863	0.9975	<b>0.99769</b>	0.9918
	Tested	Recovered	0.9814	<b>0.998847</b>	0.9813	0.9985	0.998842	0.0124
		Infected	0.9241	<b>0.99885</b>	0.9813	0.9985	0.99884	0.0134
Mean of trained data			0.9729	0.99904	0.9969	0.99903	<b>0.99907</b>	0.9937
Mean of tested data			0.9626	0.99880	0.9925	0.9958	<b>0.99887</b>	0.2755

**Data sharing**

We gathered used data from the following publicly available data sources: WHO (<https://www.who.int/countries/irn/>).

**Acknowledgments**

This research was supported by a grant (No. 33104) from Ferdowsi University of Mashhad, Iran.

## References

- [1] Q. Lin, S. Zhao, D. Gao, Y. Lou, S. Yang, S.S. Musa, M.H. Wang, Y. Cai, W. Wang, L. Yang, et al., A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action, *Int. J. Infect. Dis.* 93 (2020) 211–216.
- [2] J. Hu, G. Qi, X. Yu, L. Xu, Modeling and staged assessments of the controllability of spread for repeated outbreaks of COVID-19, *Nonlinear Dynam.* 106 (2) (2021) 1411–1424.
- [3] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, M. Colaneri, et al., A SIDARTHE model of COVID-19 epidemic in Italy, 2020, arXiv preprint arXiv:2003.09861.
- [4] T. Li, Y. Guo, Modeling and optimal control of mutated COVID-19 (Delta strain) with imperfect vaccination, *Chaos Solitons Fractals* 156 (2022) 111825.
- [5] L. Simonsen, The global impact of influenza on morbidity and mortality, *Vaccine* 17 (1999) S3–S10.
- [6] S.-W. Huang, Y.-W. Hsu, D.J. Smith, D. Kiang, H.-P. Tsai, K.-H. Lin, S.-M. Wang, C.-C. Liu, I.-J. Su, J.-R. Wang, Reemergence of enterovirus 71 in 2008 in taiwan: dynamics of genetic and antigenic evolution from 1998 to 2008, *J. Clin. Microbiol.* 47 (11) (2009) 3653–3662.
- [7] S. Chiba, R. Kogasaka, M. Akihara, K. Horino, T. Nakao, Recurrent attack of rotavirus gastroenteritis after adenovirus-induced diarrhoea, *Arch. Dis. Child.* 54 (5) (1979) 398–400.
- [8] B.F. Finkenstädt, O.N. Björnstad, B.T. Grenfell, A stochastic model for extinction and recurrence of epidemics: estimation and inference for measles outbreaks, *Biostatistics* 3 (4) (2002) 493–510.
- [9] A.L. Lloyd, Estimating variability in models for recurrent epidemics: assessing the use of moment closure techniques, *Theor. Popul. Biol.* 65 (1) (2004) 49–65.
- [10] R. Olinky, A. Huppert, L. Stone, Seasonal dynamics and thresholds governing recurrent epidemics, *J. Math. Biol.* 56 (6) (2008) 827–839.
- [11] J. Verdasca, M.T. Da Gama, A. Nunes, N. Bernardino, J. Pacheco, M. Gomes, Recurrent epidemics in small world networks, *J. Theoret. Biol.* 233 (4) (2005) 553–561.
- [12] D.J. Earn, P. Rohani, B.M. Bolker, B.T. Grenfell, A simple model for complex dynamical transitions in epidemics, *Science* 287 (5453) (2000) 667–670.
- [13] C.T. Bauch, D.J. Earn, Transients and attractors in epidemics, *Proc. R. Soc. B* 270 (1524) (2003) 1573–1578.
- [14] P.E. Fine, The interval between successive cases of an infectious disease, *Am. J. Epidemiol.* 158 (11) (2003) 1039–1047.
- [15] R. Sonabend, L.K. Whittles, N. Imai, P.N. Perez-Guzman, E.S. Knock, T. Rawson, K.A. Gaythorpe, B.A. Djaafara, W. Hinsley, R.G. FitzJohn, et al., Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in England: a mathematical modelling study, *Lancet* 398 (10313) (2021) 1825–1835.
- [16] G. Giordano, M. Colaneri, A. Di Filippo, F. Blanchini, P. Bolzern, G. De Nicolao, P. Sacchi, P. Colaneri, R. Bruno, Modeling vaccination rollouts, SARS-CoV-2 variants and the requirement for non-pharmaceutical interventions in Italy, *Nat. Med.* 27 (6) (2021) 993–998.
- [17] S. Moore, E.M. Hill, M.J. Tildesley, L. Dyson, M.J. Keeling, Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study, *Lancet Infect. Dis.* 21 (6) (2021) 793–802.
- [18] S. Bekiros, D. Kouloumpou, SBDiEM: a new mathematical model of infectious disease dynamics, *Chaos Solitons Fractals* 136 (2020) 109828.
- [19] A. Mumtaz, O.M. Otunuga, Parameter identification for a stochastic SEIRS epidemic model: case study influenza, *J. Math. Biol.* 79 (2) (2019) 705–729.
- [20] R. Ikram, A. Khan, M. Zahri, A. Saeed, M. Yavuz, P. Kumam, Extinction and stationary distribution of a stochastic COVID-19 epidemic model with time-delay, *Comput. Biol. Med.* 141 (2022) 105115.
- [21] N. Sweilam, S. Al-Mekhlafi, D. Baleanu, A hybrid stochastic fractional order Coronavirus (2019-nCoV) mathematical model, *Chaos Solitons Fractals* 145 (2021) 110762.
- [22] T. Khan, G. Zaman, Y. El-Khatib, Modeling the dynamics of novel coronavirus (COVID-19) via stochastic epidemic model, *Results Phys.* 24 (2021) 104004.
- [23] T. Hastie, R. Tibshirani, J.H. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Vol. 2, Springer, 2009.
- [24] P.E. Kloeden, M. Rasmussen, Nonautonomous Dynamical Systems, 176, American Mathematical Soc., 2011.
- [25] P. Ashwin, S. Wieczorek, R. Vitolo, P. Cox, Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system, *Phil. Trans. R. Soc. A* 370 (1962) (2012) 1166–1184.
- [26] B. Kaszás, U. Feudel, T. Tél, Tipping phenomena in typical dynamical systems subjected to parameter drift, *Sci. Rep.* 9 (1) (2019) 1–12.
- [27] T.M. Lenton, Early warning of climate tipping points, *Nature Clim. Change* 1 (4) (2011) 201–209.
- [28] M. Ghil, Mathematical theory of climate sensitivity, *Interactions* 55 (2012).
- [29] T. Andersen, J. Carstensen, E. Hernandez-Garcia, C.M. Duarte, Ecological thresholds and regime shifts: approaches to identification, *Trends Ecol. Evol.* 24 (1) (2009) 49–57.
- [30] K. Siteur, M.B. Eppinga, A. Doelman, E. Siero, M. Rietkerk, Ecosystems off track: rate-induced critical transitions in ecological models, *Oikos* 125 (12) (2016) 1689–1699.
- [31] Z. Mukandavire, S. Liao, J. Wang, H. Gaff, D.L. Smith, J.G. Morris Jr., Estimating the reproductive numbers for the 2008–2009 cholera outbreaks in Zimbabwe, *Proc. Natl. Acad. Sci.* 108 (21) (2011) 8767–8772.
- [32] L. Li, C.-H. Wang, S.-F. Wang, M.-T. Li, L. Yakob, B. Cazelles, Z. Jin, W.-Y. Zhang, Hemorrhagic fever with renal syndrome in China: Mechanisms on two distinct annual peaks and control measures, *Int. J. Biomath.* 11 (02) (2018) 1850030.
- [33] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. R. Soc. A* 115 (772) (1927) 700–721.
- [34] K. Dietz, Epidemiologic interference of virus populations, *J. Math. Biol.* 8 (3) (1979) 291–300.
- [35] B. Tang, X. Wang, Q. Li, N.L. Bragazzi, S. Tang, Y. Xiao, J. Wu, Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions, *J. Clin. Med.* 9 (2) (2020) 462.
- [36] A. Kronichler, D. Kresse, S. Yoon, K.H. Lee, M. Effenberger, J.I. Shin, Asymptomatic patients as a source of COVID-19 infections: A systematic review and meta-analysis, *Int. J. Infect. Dis.* 98 (2020) 180–186.
- [37] S. Chowdhury, M. Forkan, S.F. Ahmed, P. Agarwal, A.S. Ali, S. Muyeen, Modeling the SARS-CoV-2 parallel transmission dynamics: asymptomatic and symptomatic pathways, *Comput. Biol. Med.* 143 (2022) 105264.
- [38] Z. Liu, P. Magal, O. Seydi, G. Webb, Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data, 2020, arXiv preprint arXiv:2002.12298.
- [39] Z. Memon, S. Qureshi, B.R. Memon, Assessing the role of quarantine and isolation as control strategies for COVID-19 outbreak: a case study, *Chaos Solitons Fractals* 144 (2021) 110655.
- [40] X. Rong, L. Yang, H. Chu, M. Fan, Effect of delay in diagnosis on transmission of COVID-19, *Math. Biosci. Eng.* 17 (3) (2020) 2725–2740.
- [41] L. Pellis, F. Scarabel, H.B. Stage, C.E. Overton, L.H. Chappell, E. Fearon, E. Bennett, K.A. Lythgoe, T.A. House, I. Hall, et al., Challenges in control of COVID-19: short doubling time and long delay to effect of interventions, *Philos. Trans. R. Soc. B* 376 (1829) (2021) 20200264.
- [42] B.M. Castro, Y.D.A. de Melo, N.F. Dos Santos, A.L. da Costa Barcellos, R. Choren, R.M. Salles, Multi-agent simulation model for the evaluation of COVID-19 transmission, *Comput. Biol. Med.* 136 (2021) 104645.
- [43] S.P. Brand, J. Ojal, R. Aziza, V. Were, E.A. Okiro, I.K. Kombe, C. Mburu, M. Ogero, A. Agweyu, G.M. Warimwe, et al., COVID-19 transmission dynamics underlying epidemic waves in Kenya, *Science* 374 (6570) (2021) 989–994.
- [44] S.L. Chang, N. Harding, C. Zachreson, O.M. Cliff, M. Prokopenko, Modelling transmission and control of the COVID-19 pandemic in Australia, *Nature Commun.* 11 (1) (2020) 1–13.
- [45] T. Zhang, Z. Teng, Permanence and extinction for a nonautonomous SIRS epidemic model with time delay, *Appl. Math. Model.* 33 (2) (2009) 1058–1071.
- [46] T. Tuong, D.H. Nguyen, N. Dieu, K. Tran, Extinction and permanence in a stochastic SIRS model in regime-switching with general incidence rate, *Nonlinear Anal. Hybrid Syst.* 34 (2019) 121–130.
- [47] K. Guo, W. Ma, Permanence and extinction for a nonautonomous Kawasaki disease model with time delays, *Appl. Math. Lett.* 122 (2021) 107511.
- [48] T. Zhang, Z. Teng, Global behavior and permanence of SIRS epidemic model with time delay, *Nonlinear Anal. RWA* 9 (4) (2008) 1409–1424.
- [49] S. Li, S. Guo, Permanence and extinction of a stochastic SIS epidemic model with three independent Brownian motions, *Discrete Contin. Dyn. Syst. Ser. B* 26 (5) (2021) 2693.
- [50] T. Kuniya, Recurrent epidemic waves in a delayed epidemic model with quarantine, *J. Biol. Dyn.* 16 (1) (2022) 619–639.
- [51] X.-C. Duan, J.-F. Yin, X.-Z. Li, Global Hopf bifurcation of an SIRS epidemic model with age-dependent recovery, *Chaos Solitons Fractals* 104 (2017) 613–624.
- [52] H. Cao, D.X. Yan, A. Li, Dynamic analysis of the recurrent epidemic model, *Math. Biosci. Eng.* 16 (5) (2019) 5972–5990.
- [53] Y. Nakata, T. Kuniya, Global dynamics of a class of SEIRS epidemic models in a periodic environment, *J. Math. Anal. Appl.* 363 (1) (2010) 230–237.
- [54] Y. Cai, Y. Kang, M. Banerjee, W. Wang, A stochastic SIRS epidemic model with infectious force under intervention strategies, *J. Differential Equations* 259 (12) (2015) 7463–7502.
- [55] Y. Zhao, D. Jiang, The threshold of a stochastic SIRS epidemic model with saturated incidence, *Appl. Math. Lett.* 34 (2014) 90–93.
- [56] Y. Cai, Y. Kang, W. Wang, A stochastic SIRS epidemic model with nonlinear incidence rate, *Appl. Math. Comput.* 305 (2017) 221–240.
- [57] L. Zhang, X. Fan, Z. Teng, Global dynamics of a nonautonomous SEIRS epidemic model with vaccination and nonlinear incidence, *Math. Methods Appl. Sci.* 44 (11) (2021) 9315–9333.
- [58] W. Yang, Dynamical behaviors and optimal control problem of an seirs epidemic model with interventions, *Bull. Malays. Math. Sci. Soc.* 44 (5) (2021) 2737–2752.
- [59] R.L. Kumar, F. Khan, S. Din, S.S. Band, A. Mosavi, E. Ibeke, Recurrent neural network and reinforcement learning model for COVID-19 prediction, *Front. Public Health* 9 (2021).
- [60] A. Zeroual, F. Harrou, A. Dairi, Y. Sun, Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study, *Chaos Solitons Fractals* 140 (2020) 110121.

- [61] K. ArunKumar, D.V. Kalaga, C.M.S. Kumar, M. Kawaji, T.M. Brenza, Forecasting of COVID-19 using deep layer recurrent neural networks (RNNs) with gated recurrent units (GRUs) and long short-term memory (LSTM) cells, *Chaos Solitons Fractals* 146 (2021) 110861.
- [62] E.A. Rashed, S. Kodera, A. Hirata, COVID-19 forecasting using new viral variants and vaccination effectiveness models, *Comput. Biol. Med.* 149 (2022) 105986.
- [63] L. Xu, R. Magar, A.B. Farimani, Forecasting COVID-19 new cases using deep learning methods, *Comput. Biol. Med.* 144 (2022) 105342.
- [64] S. Lalmuanawma, J. Hussain, L. Chhakchhuak, Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review, *Chaos Solitons Fractals* 139 (2020) 110059.
- [65] H.B. Syeda, M. Syed, K.W. Sexton, S. Syed, S. Begum, F. Syed, F. Prior, F. Yu Jr., Role of machine learning techniques to tackle the COVID-19 crisis: systematic review, *JMIR Med. Inform.* 9 (1) (2021) e23811.
- [66] 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 Digit. Health* (2022).
- [67] K.T. Ly, A COVID-19 forecasting system using adaptive neuro-fuzzy inference, *Finance Res. Lett.* 41 (2021) 101844.
- [68] A. Al-Ali, O. Elharrouss, U. Qidwai, S. Al-Maaddeed, ANFIS-Net for automatic detection of COVID-19, *Sci. Rep.* 11 (1) (2021) 1–13.
- [69] J. Merker, B. Kunsch, G. Schulte, Nonlinear compartment models with time-dependent parameters, *Mathematics* 9 (14) (2021) 1657.
- [70] T.M. Mitchell, T.M. Mitchell, *Machine Learning*, Vol. 1, McGraw-hill, New York, 1997.
- [71] D.S. Broomhead, D. Lowe, Radial Basis Functions, Multi-Variable Functional Interpolation and Adaptive Networks, Technical Report, Royal Signals and Radar Establishment, Malvern (United Kingdom), 1988.
- [72] M.W. Gardner, S. Dorling, Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences, *Atmos. Environ.* 32 (14–15) (1998) 2627–2636.
- [73] D.F. Specht, et al., A general regression neural network, *IEEE Trans. Neural Netw.* 2 (6) (1991) 568–576.
- [74] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (8) (1997) 1735–1780.
- [75] J. Schmidhuber, A. Blog, The most cited neural networks all build on work done in my labs.
- [76] J.-S. Jang, ANFIS: adaptive-network-based fuzzy inference system, *IEEE Trans. Syst. Man Cybern.* 23 (3) (1993) 665–685.