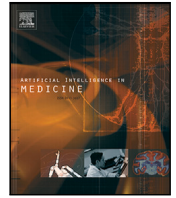




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

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



Research paper

Fuzzy-SIRD model: Forecasting COVID-19 death tolls considering governments intervention

Amir Arslan Haghray, Sehraneh Ghaemi^{*}, Mohammad Ali Badamchizadeh

Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran



ARTICLE INFO

Keywords:

Interval type 2 fuzzy logics
Fuzzy logic systems
Epidemiologic time series
Compartmental models
COVID-19 pandemic

ABSTRACT

Modeling the trend of contagious diseases has particular importance for managing them and reducing the side effects on society. In this regard, researchers have proposed compartmental models for modeling the spread of diseases. However, these models suffer from a lack of adaptability to variations of parameters over time. This paper introduces a new Fuzzy Susceptible–Infectious–Recovered–Deceased (Fuzzy-SIRD) model for covering the weaknesses of the simple compartmental models. Due to the uncertainty in forecasting diseases, the proposed Fuzzy-SIRD model represents the government intervention as an interval type 2 Mamdani fuzzy logic system. Also, since society's response to government intervention is not a static reaction, the proposed model uses a first-order linear system to model its dynamics. In addition, this paper uses the Particle Swarm Optimization (PSO) algorithm for optimally selecting system parameters. The objective function of this optimization problem is the Root Mean Square Error (RMSE) of the system output for the deceased population in a specific time interval. This paper provides many simulations for modeling and predicting the death tolls caused by COVID-19 disease in seven countries and compares the results with the simple SIRD model. Based on the reported results, the proposed Fuzzy-SIRD model can reduce the root mean square error of predictions by more than 80% in the long-term scenarios, compared with the conventional SIRD model. The average reduction of RMSE for the short-term and long-term predictions are 45.83% and 72.56%, respectively. The results also show that the principle goal of the proposed modeling, i.e., creating a semantic relation between the basic reproduction number, government intervention, and society's response to interventions, has been well achieved. As the results approve, the proposed model is a suitable and adaptable alternative for conventional compartmental models.

1. Introduction

Epidemic diseases have arisen in different periods of human civilization, and with all the sufferings they have inflicted on human beings, they have changed the course of history. The Black Death and the Spanish flu are Two examples of such epidemics [1]. Black death occurred in the Middle Ages when it was impossible to record data extensively. However, there are scattered monographs on this subject. Especially in the post-disease environment, its impact on thoughts, economy, and society has been recorded better. Although the importance of recording data was more widely understood in the late 19th and early 20th centuries, during the Spanish flu, governments sought to conceal disease statistics due to the condition arising from the conflict between countries in the First World War [2]. However, compared with the Black Death, information about the subsequent effects of this pandemic is more available and precise. With all our advances in the early 21st century, we are witnessing another pandemic, COVID-19. Starting from Wuhan city in Hubei Province of the People's Republic of

China, COVID-19 has spread to many other countries [3]. By observing the basic reproduction number (R_0) and the Case Fatality Rate (CFR) of the COVID-19 disease, an alarm has been sounded for a global crisis. The emergency committee of the World Health Organization (WHO) declared a global health emergency on 30 January 2020. Unlike previous pandemics, many countries actively identified cases and reported death tolls. Excavating the connection of this information with the economic, social, and demographic parameters makes governments more ready to deal with the conditions created during and after the pandemics.

Most of the efforts done to analyze the COVID-19 data include predicting the trend of infections and deaths. In [4], the authors have proposed a hybrid approach for forecasting the COVID-19 time series by combining the fractal dimension and fuzzy logic. In this paper, the complexity of the time series dynamic has been measured using the concept of fractal dimension. Also, the uncertainties of the prediction process have been managed using fuzzy logic. The inputs of the fuzzy system are linear and nonlinear fractal dimensions. The

^{*} Corresponding author.

E-mail addresses: haghray@tabrizu.ac.ir (A.A. Haghray), ghaemi@tabrizu.ac.ir (S. Ghaemi), mbadamchi@tabrizu.ac.ir (M.A. Badamchizadeh).

time series of both confirmed cases and death tolls are under study in this paper. In another study, ensemble neural networks have been used to make predictions about the COVID-19 time series, and fuzzy logic has been used to aggregate the responses of the neural networks-based predictors [5]. As a powerful tool for managing uncertainties, fuzzy logic has been used to make the final decision in the forecasting process. This study also uses the confirmed cases and deaths data for simulations. In [6], deep learning methods and the Bayesian optimization (BO) algorithm have been combined to predict the confirmed cases of COVID-19. The Bayesian optimization algorithm has been used to select parameters of three deep learning models; multi-head attention, long short-term memory (LSTM), and convolutional neural network. The results achieved by the proposed approaches have been compared with many other methods, too. According to the authors of this paper and based on the performance measures, the LSTM-BO has outperformed other models. In [7], Auto-Regressive Integrated Moving Average (ARIMA) and artificial neural networks have been utilized for predicting the COVID-19 time series. Another model, which has been utilized for COVID-19 time series prediction, is the Gompertz model [8, 9]. In [10], support vector machine (SVM), LSTM, Gated Recurrent Units (GRU), and Bi-LSTM models have been studied for forecasting the COVID-19 time series for several countries. Machine learning and cloud computing have been combined to analyze and predict outbreak behavior in different countries [11]. Also, the Gaussian mixture model has been utilized for predicting the COVID-19 time series [12].

In [13], to have a better understanding of COVID-19 dynamics, a sequential genetic algorithm-based probabilistic cellular automata has been used. This article contains rich studies on the COVID-19 trend in many countries. The authors claim that the motivation of the proposed method was to develop a data-driven, generalized, and spatial framework that can be used to estimate relevant epidemiological parameters. Dendritic Neural Regression (DNR) approach has been used for the COVID-19 time series prediction [14]. In this paper, DNR has been improved using a combination of the S metric selection algorithm and the scale-free local search. Also, different algorithms have been used in this paper to predict the disease trend in many countries, and it contains a rich comparative study. In [15], wavelet-coupled random vector functional link networks have been used for modeling and forecasting COVID-19 spread. In [16], the performance of the countries against COVID-19 has been studied using machine learning algorithms and weighted stochastic imprecise data envelopment analysis. Many other methods have also been utilized for forecasting of COVID-19 trend, which fuzzy time series [17], convolution-LSTM [18], auto-regressive integrated moving average [19], seasonal auto-regressive integrated moving average [20], and grey rolling model [21] are among them.

Compartmental models are traditional tools for modeling the disease spread in epidemiologic studies. In [22], a new time-varying Susceptible-Infected-Recovered-Deceased (SIRD) model has been proposed for modeling the COVID-19 outbreak in Italy. Given that the parameters of the conventional SIRD model can change over time by increasing knowledge about the disease, social reaction, and government intervention, they can be estimated on a functional basis. In another paper having a similar approach to the previous one, the SIRD model parameters are exponential functions of time [23]. The results achieved by this method are compatible with the COVID-19 time series of Italy and China. Another time variable SIRD model is proposed in [24]. The proposed method in this paper assumes parameters of the model as a function of time and determines these functions intending to minimize the model error. From a different point of view in [25], lockdowns during epidemic diseases have been studied. This paper seeks to find the optimal period of lockdown using the three-phase maturation SIRD model. Based on the reported results in this paper, the optimal lockdown time for China, India, and Italy would be 73, 69, and 88 days, respectively.

Almost in all recent studies concerning the enhancement of the Compartmental models, the model parameters have been assumed

time-varying. However, most of these newly proposed methods have solved a curve-fitting problem instead of improving the compartmental models from a fundamental point of view. The main feature of the compartmental models is that they are constructed based on the essential properties of infectious diseases. In the case of emerging diseases, as a strength, these properties can be extracted using system identification techniques and collected data. So, any effort to improve these methods should not overshadow their essential motivations.

A vast majority of the studies concerning the extraction of the infectious disease parameters using compartmental models use the confirmed cases data. However, it should be noted that the reported data for confirmed cases is inaccurate due to the variety of testing policies. Also, most COVID-19 diagnosis tests are done on the people visiting the hospitals [26]. In other words, inconsistencies between confirmed cases and the death toll are probable. So there can be a considerable change in their proportion. That is why using either the hospitalized case data or the death tolls will lead to more accurate, but still underestimated, results [27].

This paper proposes a new fuzzy-SIRD model for overcoming deficiencies of the conventional SIRD model. The biggest weakness of the conventional SIRD models is that the epidemiological parameters of infectious diseases are assumed to be invariant over time, while epidemics occur in a complex system named society with complicated behavior. Therefore, it is essential to model the society's response and changes in the epidemiological parameters caused by it. Government intervention is the main link between society's response and epidemic disease. That is why this paper aims to model the effects of government intervention on society's response and, as a result, on infectious disease parameters.

Fuzzy systems allow the construction of mathematical models using verbal expressions. Based on this feature of fuzzy systems, in the proposed Fuzzy-SIRD model, an Interval Type 2 Fuzzy System (IT2 FLS) models the government intervention using fuzzy IF-THEN rules. Type 2 fuzzy logic has been chosen because of the capability of this approach for handling uncertainties. The efficiency of the proposed model has been evaluated by predicting the COVID-19 time series. For this purpose, seven countries have been selected, and the time series of death data have been used for modeling in two different time horizons. The model parameters have been selected by the Particle Swarm Optimization (PSO) algorithm aiming to minimization of the Root Mean Square Error (RMSE). The results approve the efficiency of the proposed model for predicting the COVID-19 death toll time series.

The rest of the paper is organized as follows: Section 2 introduces some preliminaries about this work, including the factors affecting disease outbreaks, the conventional SIRD model, and interval type 2 fuzzy logic systems. The proposed model is introduced in Section 3. The simulation results are reported in Section 4. And finally, the paper conclusions are provided in Section 5.

2. Preliminaries

This section discusses some preliminaries of the proposed Fuzzy-SIRD model and aims to clarify the motives for proposing a new SIRD-based model. Also, it introduces the mathematical basis on which the new model has been constructed over it.

2.1. Factors affecting disease outbreaks

Many studies have been done to predict the course of epidemic diseases and subsequent deaths. However, most of them rely solely on reported data without considering the influential factors. Another issue that makes the results of studies less reliable is the use of confirmed cases of infection, which is a function of disease diagnosis and testing policies. However, more attention should be paid to the effective parameters for controlling disease rather than the course of time series. Decisive factors in the spread of diseases and subsequent deaths can be divided into several categories such as economic, social, demographic, etc. Some of these factors that can be considered to study are as follows:

- Purchasing power parity (PPP)
- Expenditure on health per capita
- The moral stance of the society (Individualistic or communitarian)
- Government intervention
- Age structure (Population over the age of 65)
- Residential population density

Purchasing power parity is an economic indicator that makes it possible to compare the absolute purchasing power of the countries' currencies. The potential importance of economic factors, like PPP, has been also pointed out in other studies concerning the COVID-19 pandemic [28]. Another economic factor that directly affects the death tolls of epidemic diseases is the expenditure on health per capita. Some researchers have claimed that higher expenditure on health by countries causes lesser case fatality [29]. The pursuit of the public interest by the individuals in a society largely depends on their moral stance. It can be argued that compared with individualistic, in communitarian societies, people pay more attention to advice on social behaviors [30], such as keeping a social distance and using face masks. The accuracy of this claim can be determined by measuring the impact of the government intervention to control the disease by imposing restrictions. Another important factor concerning the death rate due to epidemics is the demographic parameters of the countries [31]. From the first days of the COVID-19 pandemic, it has been reported that ages over 65 constitute the most at-risk population facing the disease [32]. Thus, the population pyramid of the countries would play a decisive role in COVID-19 and other disease death tolls. Another issue affecting the prevalence of the disease is residential population density [33]. Numerous other influential factors can be mentioned, too.

However, the proposed enhanced compartmental models had not been able to assume the impact of these parameters, and a more general method is needed for this purpose. It can be argued that these factors affect epidemic diseases more than their intrinsic characteristics. So any attempt to develop epidemiologic models must aim to model these factors or the outcomes of their impacts. This paper tries to assume the government intervention directly and some other parameters indirectly for modeling and predicting epidemics. The Fuzzy-SIRD approach proposed in this paper can model the effects of these parameters.

2.2. SIRD model

One of the well-known and common models, which has been used in several studies on mathematic modeling of infectious diseases' dynamics and has achieved good results, is the Susceptible–Infectious–Recovered–Deceased (SIRD) model. This model is defined by a set of differential equations as below: [34]:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta IS}{N}, \\ \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I - \mu I, \\ \frac{dR}{dt} &= \gamma I, \\ \frac{dD}{dt} &= \mu I. \end{aligned}$$

n which $S(t)$, $I(t)$, $R(t)$, $D(t)$, and N represent the susceptible, infectious, recovered, deceased, and total population, respectively. Based on this definition it is obvious that $S(t) = N - I(t) - R(t) - D(t)$. Furthermore, β , γ , and μ are infection, recovery, and mortality rates, respectively. These parameters can be affected by many factors. Some of these factors have been reviewed in Section 2.1.

2.2.1. Closed-form solution of SIRD model

First we assume that $C(t) = R(t) + D(t)$ denote the removed individuals, and $\eta = \gamma + \mu$. Then the solution of the SIRD model can be achieved as follows [22].

$$S(t) = S^0 (1 + \kappa)^\rho (1 + \kappa e^{(\beta-\eta)t})^{-\rho} \tag{1}$$

$$I(t) = I^0 (1 + \kappa)^\rho (1 + \kappa e^{(\beta-\eta)t})^{-\rho} e^{(\beta-\eta)t} \tag{2}$$

$$C(t) = N - S^0 (S^0 + I^0)^\rho (S^0 + I^0 e^{(\beta-\eta)t})^{-\rho} \tag{3}$$

Where the zero superscripts indicate the initial value of the variables, $\kappa = \frac{I^0}{S^0}$ and $\rho = \frac{\beta}{\beta-\eta}$. Also, based on the solution for $C(t)$, $R(t)$, and $D(t)$ can be calculated using the γ and μ coefficients. The closed-form solution of the SIRD model gives us a deeper insight into its capabilities. Considering the formulas (1), (2), and (3), we get that the flexibility and adaptability of this model are very low in medium and long-term periods while facing changes in parameters. So it is important to update the conventional SIRD model to adapt to the variation of epidemiological parameters over time.

2.3. Interval type 2 fuzzy logic systems

Fuzzy systems allow the rapid construction of mathematical models from linguistic expressions. That is why they are widely used in various applications. In recent years, type 2 fuzzy systems have been used by many researchers in different studies [36]. The main strength of type 2 fuzzy systems is their ability to handle uncertainty. Interval type 2 fuzzy systems require lesser computational effort while maintaining the capability for handling uncertainties. In this paper, an Interval Type 2 Fuzzy Logic System (IT2 FLS) has been used to model the effects of government intervention on the basic reproduction number of the COVID-19 disease. Fig. 1 represents the overall structure of an IT2 FLS [35]. The first step in an IT2 FLS is the fuzzification of crisp inputs. This is done using interval type 2 fuzzy sets, defined using their upper and lower membership functions. In most applications, the fuzzy sets chosen for this purpose are interval type 2 singleton fuzzy sets. The second step in an IT2 FLS is evaluating fuzzy rules and calculating their outcomes. Every fuzzy IF-THEN rule, in the rule-base of an IT2 FLS, can be represented as below:

$$\begin{aligned} R^l : & \text{ IF } x_1 \text{ is } \tilde{F}_{1,l} \text{ and } \dots \text{ and } x_p \text{ is } \tilde{F}_{p,l}, \\ & \text{ THEN } y \text{ is } \tilde{G}_l \quad l = 1, \dots, M \end{aligned}$$

In which the l shows the index of the rules and p indicates the index of the inputs. M represents the number of rules. Furthermore, $\tilde{F}_{i,l}$ and \tilde{G}_l are input and output sets, respectively. To evaluate the fuzzy rules, first, the upper and lower firing strengths of the rules are calculated using the formulas below:

$$\bar{f}_l(x) = \prod_{i=1}^p \bar{\mu}_{\tilde{F}_{i,l}}(x_i) \tag{4}$$

$$\underline{f}_l(x) = \prod_{i=1}^p \underline{\mu}_{\tilde{F}_{i,l}}(x_i) \tag{5}$$

Now the result of the l th fuzzy rule can be concluded as the interval type 2 fuzzy set B_l , with upper and lower membership functions as below:

$$\bar{\mu}_{B_l}(x) = \bar{f}_l(x) * \bar{\mu}_{G_l}(x) \tag{6}$$

$$\underline{\mu}_{B_l}(x) = \underline{f}_l(x) * \underline{\mu}_{G_l}(x) \tag{7}$$

After evaluating the rules, we must aggregate them to achieve the final result. This aggregation is typically done by using the centroid method. However, there are also other methods like Center of Sets (CoSets), Center of Sum (CoSum), Height, and Modified Height. To calculate the centroid of B_l s, we must calculate the Join of these sets:

$$\tilde{Y} = \prod_{l=1}^M \tilde{B}_l \tag{8}$$

\tilde{Y} is the output fuzzy set. In this step, we must perform a type-reduction on the output set, which can be done using multiple algorithms such as KM, EAISC, WM, NT, etc.

$$Y = \{ (x, \mu_Y(x)) \mid \forall x \in X \} \tag{9}$$

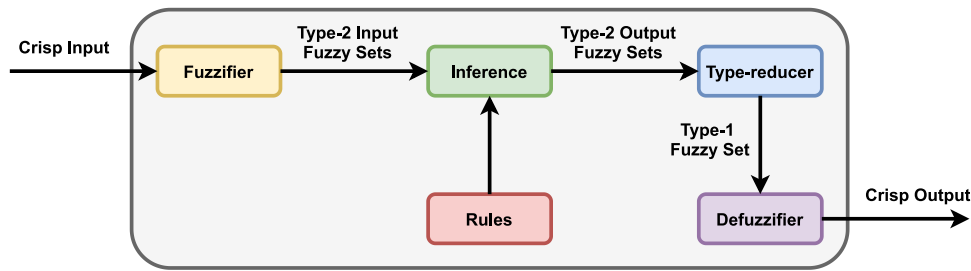


Fig. 1. Block diagram representation of the Interval Type 2 Fuzzy Logic Systems [35].

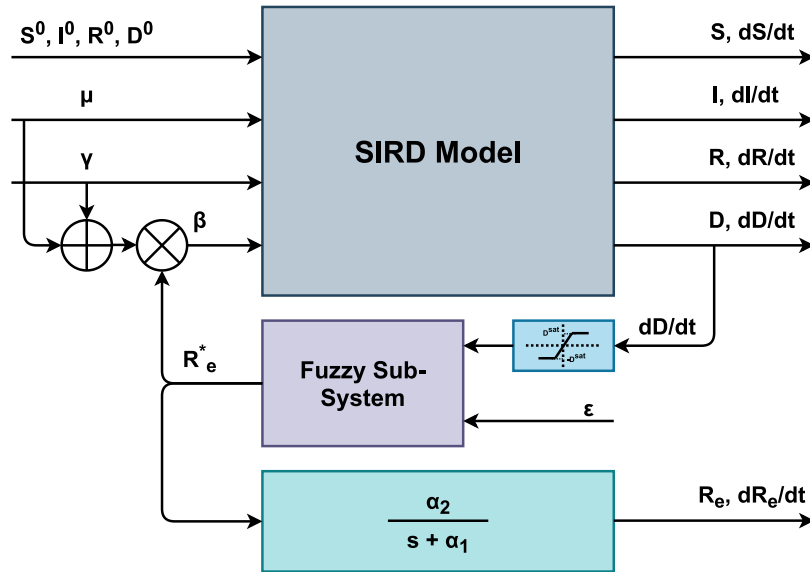


Fig. 2. Block diagram representation of the proposed Fuzzy-SIRD model.

The output of the type reduction algorithm is an all one fuzzy set, which can be demonstrated as an interval in which the membership degree is equal to 1.

$$[y_l, y_r] \tag{10}$$

Finally, the type-reduced fuzzy set must be defuzzified. The crisp output is calculated as the center of the interval above.

$$y = \frac{y_l + y_r}{2} \tag{11}$$

3. Fuzzy-SIRD model

This section proposes a new Fuzzy-SIRD model, which can efficiently model the effects of government intervention in the outbreak of infectious diseases. Governments enforce preventive restrictions based on the death toll to control epidemics. Such an intervention primarily affects the effective reproduction number, R_e , of the infectious diseases. Therefore, the behavior of governments and its effectiveness will have a great impact on the outbreak trend. So, it is important to consider government intervention while modeling the spread of diseases.

The basis of the proposed model is the conventional SIRD model. The proposed approach introduces a new variable for representing the effective reproduction number of the disease. The Fuzzy-SIRD model can be described by a set of differential equations as below:

$$\frac{dS}{dt} = -\frac{\beta IS}{N}, \tag{12}$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I - \mu I, \tag{13}$$

$$\frac{dR}{dt} = \gamma I, \tag{14}$$

$$\frac{dD}{dt} = \mu I, \tag{15}$$

$$\frac{dR_e}{dt} = -\alpha_1 R_e + \alpha_2 \psi \left(\epsilon, \frac{dD}{dt}, D^{sat} \right). \tag{16}$$

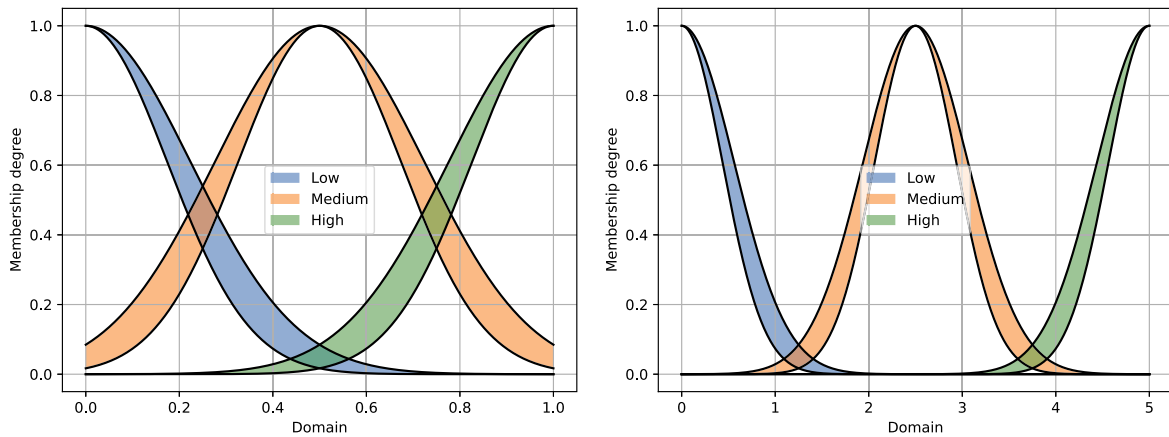
The first four equations are the same as the conventional SIRD model. In the fifth equation, α_1 and α_2 are model coefficients that indicate society's response to government intervention. ψ is the sub-system modeling the government intervention, and has three inputs ϵ (efficiency factor), $\frac{dD}{dt}$ (death rate), and D^{sat} (saturation factor of death rate). The block diagram of the proposed Fuzzy-SIRD model has been demonstrated in Fig. 2. According to the simple SIRD model, the effective reproduction number can be calculated by the equation $R_e = \frac{\beta}{\gamma + \mu}$ [37].

3.1. Fuzzy sub-system

The ψ function consists of an interval type 2 Mamdani fuzzy system with the rule-base represented in Table 1. The rules in this table have been designed to express the impact of the government intervention on the effective reproduction number, R_e , according to the parameters ϵ and $sat \left(\frac{dD}{dt}, D^{sat} \right)$. For example, let us assume the first rule of Table 1.

$$\begin{array}{l} \text{IF } \epsilon \text{ IS LOW AND} \\ \text{sat} \left(\frac{dD}{dt}, D^{sat} \right) \text{ IS LOW THEN} \\ R_e^* \text{ IS HIGH} \end{array} \tag{17}$$

In this situation, the efficiency of the government intervention is low, caused by factors like economic condition, state of social cohesion in society, etc. Also, the death rate is low, meaning there is little increment in the number of deceased people. In these circumstances, the



(a) Fuzzy sets defining the inputs. Low: mean=0, std.=0.2, spread of std.=0.05; Medium: mean=0.5, std.=0.2, spread of std.=0.05; High: mean=1, std.=0.2, spread of std.=0.05. (b) Fuzzy sets defining the output. Low: mean=0, std.=0.1, spread of std.=0.025; Medium: mean=0.5, std.=0.1, spread of std.=0.025; High: mean=1, std.=0.1, spread of std.=0.025.

Fig. 3. The gaussian interval type 2 fuzzy sets with uncertain standard deviation (std.) values used for defining the fuzzy sub-system.

government does not apply restrictions, so the effective reproduction number of the disease will increase.

Also, the interval type 2 fuzzy sets used for defining the system have been represented in Fig. 3. Fig. 3(a) demonstrates the fuzzy sets defining both inputs of the system, ϵ , and $sat\left(\frac{dD}{dt}, D^{sat}\right)$. Fig. 3(b) represents the fuzzy sets defining the output of the system. As can be seen, the Gaussian membership functions have been used for constructing interval type 2 fuzzy sets. The main reasons for such a choice are the smoothness of the Gaussian membership functions and their non-zero membership grades over the universe of discourse [38]. However, as an alternative, triangular, or more generally trapezoidal membership functions can also be used with similar performance [39].

An important issue concerning fuzzy systems is selecting the number of fuzzy sets and fuzzy rules while defining them. Of course, using more fuzzy sets and more fuzzy rules will improve the results achieved by the system. But it will cause a significant increase in execution time and an extra effort for adjusting the parameters of the fuzzy sets. Therefore, there is always a tradeoff between number of the rules and computational effort. In Section 4, we will discuss why we have used three sets for each input of the fuzzy system and why consequently, we have defined nine fuzzy IF-THEN rules.

The parameters of this model, which must be extracted using time series, are:

- Initial infectious population I^0
- Initial recovered population R^0
- Initial deceased population D^0
- Initial effective reproduction number R_e^0
- Recovery rate γ
- Fatality rate μ
- Coefficients α_1 and α_2
- Saturation factor for death rate D^{sat}
- Government intervention efficiency ϵ

3.2. Response of society

One of the main features of the proposed model is modeling the society's response to government intervention by assuming it as a first-order linear system. As it has been shown in Fig. 2, the system representing the response of society can be defined in the frequency domain as the linear system below:

$$H(s) = \frac{\alpha_2}{s + \alpha_1} \tag{18}$$

Table 1
The rule-base of the fuzzy sub-system.

ϵ	$sat\left(\frac{dD}{dt}\right)$	R_0^*
LOW	LOW	HIGH
LOW	MEDIUM	HIGH
LOW	HIGH	MEDIUM
MEDIUM	LOW	HIGH
MEDIUM	MEDIUM	MEDIUM
MEDIUM	HIGH	LOW
HIGH	LOW	MEDIUM
HIGH	MEDIUM	LOW
HIGH	HIGH	LOW

The response of such a system for a Heaviside input function would be:

$$y(t) = \frac{\alpha_1}{\alpha_2} (1 - e^{-\alpha_2 t}) \tag{19}$$

Such an assumption helps enhancement of the SIRD model in different ways. In practice, when governments try to regulate the basic reproduction number of infectious diseases, society pursues government planning under the influence of governance tools. However, this adaptation to government planning would not be ideal. It means that the goal will not be fully achieved, and society will gradually adapt to the new planning. The proposed model can demonstrate all these conditions. In other words, the behavior of society is expressed by α_1 and α_2 . The larger α_2 , the faster society will adapt. Also, the ratio of α_1 and α_2 reflects the difference in social behavior and government expectations. It can be said that, in addition to being a tool for modeling the behavior of society, this system and the two parameters are affected by other sociological issues.

3.3. Parameters tuning

Generally, we can use two approaches for determining the model parameters; using infection time series or deceased time series. According to epidemiology experts, while we have encountered an epidemic, it is not possible to identify all the infected cases accurately with a fixed criterion. So, the infected cases time series will not be a good choice for accurately determining the epidemic disease trend. Unlike many other studies in this paper, the deceased people time series has been used to identify the model parameters. It should also be noted that delays in data aggregation and reporting cause little fluctuations in their natural course. Therefore in this study, a 7-day moving average filter has been

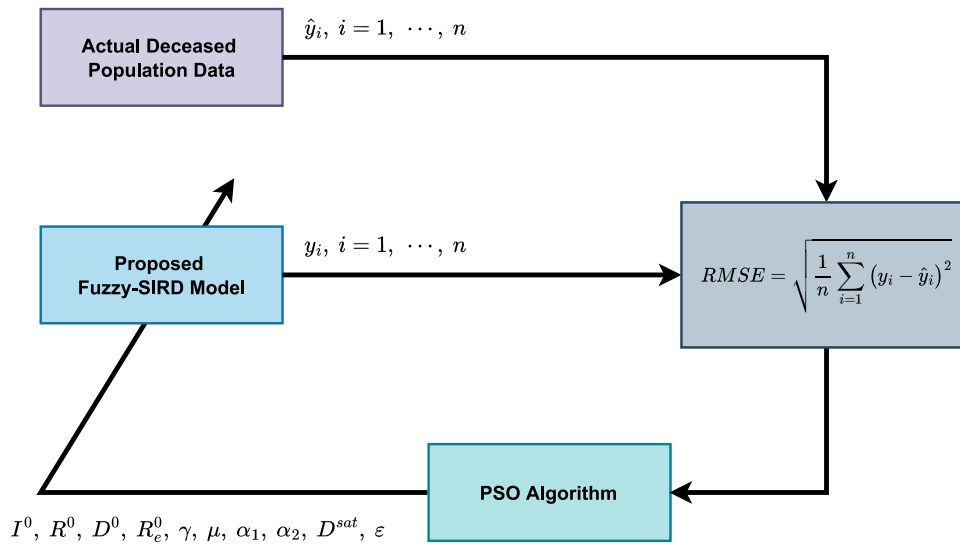


Fig. 4. Block diagram representation of the relation between the PSO algorithm and the proposed Fuzzy-SIRD model.

used. The error function to identify the model parameters has been defined as below:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (20)$$

In which, n is the number of samples, y_i represents the actual deceased population until the i th day, and \hat{y}_i shows the deceased population until the i th day achieved by the model.

After defining the error function, an optimization algorithm can be used to select the parameters for minimizing the error. In this paper, the PSO algorithm has been used to accomplish this task. Each particle's solution vector has been defined as below:

$$P = [I^0 \ R^0 \ D^0 \ R_e^0 \ \gamma \ \mu \ \alpha_1 \ \alpha_2 \ D^{sat} \ \epsilon] \quad (21)$$

Fig. 4 demonstrates the relation between the PSO algorithm and the proposed Fuzzy-SIRD model.

The only constraint on these parameters is their positiveness. An important issue that should be considered while searching for model parameters using the PSO algorithm is consciously creating the initial population of particles. It means that, for faster convergence, the initial particles should be properly dispersed in the search space. The formulas below have been used for creating the initial population:

$$I^0 = I_*^0 + 0.1 \times I_*^0 (rand - 0.5) \quad (22)$$

$$R^0 = R_*^0 + 0.1 \times R_*^0 (rand - 0.5) \quad (23)$$

$$D^0 = D_*^0 + 0.1 \times D_*^0 (rand - 0.5) \quad (24)$$

$$R_e^0 = 2.0 + 4.0 \times (rand - 0.5) \quad (25)$$

$$\gamma = 0.05 + 0.1 \times (rand - 0.5) \quad (26)$$

$$\mu = 0.05 + 0.1 \times (rand - 0.5) \quad (27)$$

$$\alpha_1 = 0.5 + 1.0 \times (rand - 0.5) \quad (28)$$

$$\alpha_2 = 0.5 + 1.0 \times (rand - 0.5) \quad (29)$$

$$D^{sat} = 50.0 + 100.0 \times (rand - 0.5) \quad (30)$$

$$\epsilon = 0.5 + 1.0 \times (rand - 0.5) \quad (31)$$

In which I_*^0 , R_*^0 , and D_*^0 represent the infectious, recovered, and deceased population in the starting time index of our data, used for fitting model parameters. Also, the $rand$ term indicates a random number generated in the interval $[0, 1)$.

3.4. Stability analysis of Fuzzy-SIRD model

We will study the stability of the proposed Fuzzy-SIRD model in this section. The equilibrium point of the system can be achieved by equating the five Eqs. (12)–(16) to zero. As a result, the equilibrium point would be:

$$I = 0, \quad R_e = \frac{\alpha_2}{\alpha_1} \psi \left(\epsilon, \left. \frac{dD}{dt} \right|_{I=0}, D^{sat} \right) \quad (32)$$

The meaning of $I = 0$ in the equilibrium point is that the number of infected people is zero. In practice, we know that this means the end of the epidemic. Now let us study the stability of this system. To prove the stability of the Fuzzy-SIRD system, first, we divide the system into two sub-systems and analyze the stability of these sub-systems separately. The only state variable of the first sub-system is R_e , and the state variables of the second sub-system are $X_1 = [S \ I \ R \ D]^T$. Also, it must be noticed that the function $\psi \left(\epsilon, \frac{dD}{dt}, D^{sat} \right)$ is a positive real number in the interval $[\psi_{min}, \psi_{max}]$. So the dynamic of the first sub-system can be written as below:

$$\frac{dR_e}{dt} = -\alpha_1 R_e + \alpha_2 \psi \left(\epsilon, \frac{dD}{dt}, D^{sat} \right) \quad (33)$$

The equilibrium point of this system is $\bar{R}_e = \frac{\alpha_2}{\alpha_1} \psi \left(\epsilon, \frac{dD}{dt}, D^{sat} \right)$. So, we assume both conditions when the $R_e > \bar{R}_e$ and $R_e < \bar{R}_e$.

$$R_e > \bar{R}_e \longrightarrow \frac{dR_e}{dt} < 0 \quad (34)$$

$$R_e < \bar{R}_e \longrightarrow \frac{dR_e}{dt} > 0 \quad (35)$$

This shows that the first sub-system is stable and tracks the $\psi \left(\epsilon, \frac{dD}{dt}, D^{sat} \right)$ value. The ψ function itself, is depended on two constant values ϵ and D^{sat} , and a variable $\frac{dD}{dt} = I$. So if the state variable I eventually converges to a constant value, the R_e would also converge to a constant value. Next, we must prove the stability of the second sub-system consisting of $S, I, R,$ and D state variables. Considering the Eqs. (14) and (15), proving the stability of the S and I is enough to prove the stability of the overall system. Before starting the proof, we should recall the property of the system under study. The initial conditions for the state variables are always positive numbers. It can also be proved that they will always remain positive. In this regard, we will study three cases, ($I > 0, S = 0$), ($I = 0, S > 0$), and ($I = 0, S = 0$). The first case is when $I > 0$ and $S = 0$:

$$\frac{dS}{dt} = 0 \quad (36)$$

$$\frac{dI}{dt} = -\gamma I - \mu I \quad (37)$$

In practice, this is the case when there is no susceptible person, and the number of infected people decays over time because both γ and μ are real positive numbers. When the state variable I converges to zero, its derivative will also tend to zero. The second case is when $I = 0$ and $S > 0$:

$$\frac{dS}{dt} = 0 \quad (38)$$

$$\frac{dI}{dt} = 0 \quad (39)$$

In practice, this condition occurs when there is no infected person anymore. Naturally, when there is no infected person, the number of susceptible people would not be changed. Finally, the last case is when $I = 0$ and $S = 0$:

$$\frac{dS}{dt} = 0 \quad (40)$$

$$\frac{dI}{dt} = 0 \quad (41)$$

In practice, this situation occurs when all people have been infected, and the results of all infected cases have been determined. We have proved that the state variables I and S , with real positive values, cannot pass the $I = 0$ and $S = 0$ lines to negative values.

Considering these three conditions for S and I , we can prove the stability of the system by applying the Lyapunov stability theory conditions only for $S \geq 0$ and $I \geq 0$. Let us assume the Lyapunov candidate function as below:

$$V(S, I) = \frac{1}{2}c_1I^2 + \frac{1}{2}c_2S^2 + c_3IS \quad (42)$$

In which c_1 , c_2 , and c_3 are greater than zero. The derivative of the Lyapunov candidate would be:

$$\begin{aligned} \dot{V}(S, I) &= \frac{c_1\beta I^2 S}{N} - c_1\gamma I^2 - c_1\mu I^2 - \frac{c_2\beta IS^2}{N} \\ &\quad - \frac{c_3\beta I^2 S}{N} + \frac{c_3\beta IS^2}{N} - c_3\gamma IS - c_3\mu IS \\ &= (c_1 - c_3) \frac{\beta I^2 S}{N} + (c_3 - c_2) \frac{\beta IS^2}{N} \\ &\quad - c_1(\gamma + \mu) I^2 - c_3(\gamma + \mu) IS \end{aligned}$$

For achieving stability conditions, the coefficients of the Lyapunov candidate must be selected in a way that $c_1, c_2, c_3 > 0, c_3 > c_1$ and $c_2 > c_3$. So we proved that the system states starting from any feasible initial condition ($S > 0$ and $I > 0$) will reach to the region $I = 0$, in which the Eqs. (12)–(15) are zero there, and subsequently the R_e will converge to the constant value $\psi(\epsilon, 0, D^{sat})$. As it can be seen, the stability of the second sub-system is unaffected by the first sub-system, which determines the value of the parameter β . As proved, the state variable R_e (and consequently β) never diverges. Therefore the overall system is stable.

4. Simulations

The Fuzzy-SIRD model has been used for modeling the epidemic trend of COVID-19 disease in seven countries to evaluate its capability and efficiency. These seven countries, which are as follows, are among the populated and developed or under-development countries: USA, Brazil, Germany, United Kingdom, Iran, Russia, and Italy. Matching the model parameters has been done using two different time intervals, one short (75 days) and the other longer (150 days), starting from 2020-01-22. Also, a 7-day moving average has been used to eliminate the fluctuations due to delays or errors in the reports. The predictions with the proposed model in scenarios 1 and 2 are in 10 and 20-day intervals. So in the first case, 65 days of data have been used for the learning phase and ten days for prediction. And in the second

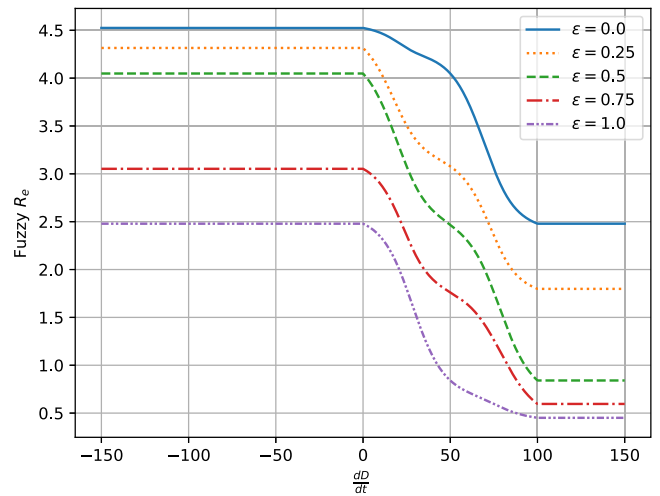


Fig. 5. The output plot of the fuzzy sub-system for different ϵ values, $D^{sat} = 100$, and $\frac{dD}{dt}$ in the interval $[-150, 150]$.

case, 130 days of data have been used for learning and 20 days for prediction. It should be noted that the data used in this study have been taken from DATAHUB (<https://datahub.io/core/covid-19>, LastAccessed:2022-03-29). DATAHUB has collected these data from many sources like World Health Organization (WHO), European Centre for Disease Prevention and Control (ECDC), US Centers for Disease Control and Prevention (CDC), Worldometers, COVID Tracking Project, etc.

The model parameters have been achieved using the PSO algorithm. The number of particles and iterations have been assumed 100 and 200, respectively. For each case, the optimization process has been repeated 100 times. The results achieved from the Fuzzy-SIRD model have been compared with the simple SIRD model. To have a deeper insight into the fuzzy sub-system of the proposed model, its output for different ϵ values (0.0, 0.25, 0.5, 0.75, 1.0) has been depicted in Fig. 5. In this figure, the D^{sat} has been assumed to be 100, and the $\frac{dD}{dt}$ is in the interval $[-150, 150]$. As it can be seen, when the efficiency of the government intervention is at the most, the fuzzy R_e lays in the interval $[0.4507, 2.479]$. And when the efficiency is the least, the fuzzy R_e varies in the interval $[2.479, 4.524]$. Furthermore, by approaching the ϵ to 0.5 from both upper and lower values, the length of the R_e interval is increased. The largest interval for effective reproduction number is $[0.841, 4.048]$, when $\epsilon = 0.5$. It should be noted that the effective reproduction number of diseases varies in different countries, depending on the policies applied to control the disease and its success. So the existence of the factor ϵ , which expresses the possible diversity of the effective reproduction number, has critical importance. In other words, this parameter, along with the D^{sat} , makes the algorithm adaptable for different countries with different policies. The success of a government in implementing preventive policies is reflected in the value of ϵ . When the $\epsilon = 1$, the maximum and minimum experienceable effective reproduction number will be the lowest, and as the ϵ decreases, these values increase. This increment in the maximum and minimum effective reproduction number indicates a lower success rate in government intervention. The D^{sat} affects the R_e in another way. It specifies at which point the R_e would reach its minimum. Also, the output plane of the fuzzy sub-system has been depicted in Fig. 6.

The achieved root mean square errors for predicting the COVID-19 disease trend in seven countries and two scenarios have been represented in Table 2. As can be seen, in all cases, the error achieved by the proposed model is less than the conventional SIRD model. The last column of this table shows the decrement of RMSE in the proposed model compared with the conventional SIRD model in percent. According to

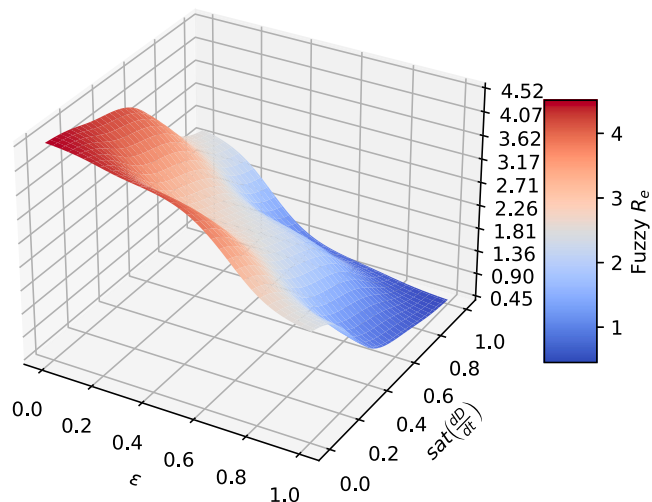


Fig. 6. The 3D representation of the fuzzy sub-system's output plane.

Table 2

The RMSE error achieved by Fuzzy-SIRD and SIRD models for each country in both time horizons.

Country	Scenario	SIRD	Fuzzy-SIRD	Change
USA	Scenario 1	3510.8128	1414.4021	-59.71%
	Scenario 2	2253.4547	348.9075	-84.52%
Brazil	Scenario 1	1066.4176	431.3594	-59.55%
	Scenario 2	5987.1328	915.4509	-84.71%
Germany	Scenario 1	21.5559	6.2062	-71.21%
	Scenario 2	153.3652	20.0909	-86.90%
Iran	Scenario 1	61.1203	41.1801	-32.62%
	Scenario 2	2203.4382	251.8999	-88.57%
Italy	Scenario 1	29.1475	23.0834	-20.80%
	Scenario 2	353.1337	77.0946	-78.17%
Russia	Scenario 1	194.3770	79.4394	-59.13%
	Scenario 2	170.8443	166.4179	-2.59%
United Kingdom	Scenario 1	56.9077	46.8063	-17.75%
	Scenario 2	341.1123	59.8646	-82.45%

Table 3

The parameters of Fuzzy-SIRD and SIRD models achieved for the US COVID-19 time series.

Parameter	Scenario 1		Scenario 2	
	Fuzzy-SIRD	SIRD	Fuzzy-SIRD	SIRD
I^0	959977.4761	767955.2043	1003654.3044	12660.6912
R^0	1013044.4219	2281038.7976	965102.5878	1439051.1596
D^0	73882.1078	73901.7487	72753.6761	83068.4621
R_e^0	0.9470	0.0718	0.7588	0.1803
γ	0.4448	0.0912	0.1035	0.1176
μ	0.0018	0.0022	0.0019	0.0602
α_1	0.1131	-	0.0199	-
α_2	0.1131	-	0.0199	-
D^{sat}	593.7055	-	976.3858	-
ϵ	0.4372	-	0.7108	-

Table 2, we can deduce that the conventional SIRD has a very poor performance in long-term predictions. However, the proposed model outperforms the conventional SIRD model in both cases. In many cases, using the Fuzzy-SIRD model, learning and prediction errors have been reduced by over 80%. The detailed description of the solutions achieved has been separately reported in corresponding tables for any case.

Also, it must be noted that, in order to select the number of the fuzzy sets and fuzzy rules used for defining the fuzzy system, three different simulations have been done. In each simulation, a fuzzy system has

been defined using a different number of fuzzy sets and, as a result, a different number of fuzzy rules:

- System 1: 3 fuzzy sets per input and 9 IF-THEN rules
- System 2: 4 fuzzy sets per input and 16 IF-THEN rules
- System 3: 5 fuzzy sets per input and 25 IF-THEN rules

These systems have been used for the prediction of the COVID-19 time series for the US in the first scenario. Considering the results achieved by the second system, increasing the number of rules to 16 has improved the results by 1.7112% with respect to the first system. Compared with the first system, increasing the number of rules to 25 in the third system has improved the results by 2.1347%. This slight improvement in the results is while the execution time has increased by 80.8447% and 168.1486% for 16-rules and 25-rules systems, respectively. Based on this observation, we decided to use three fuzzy sets per input and nine IF-THEN rules in the fuzzy system to have a good performance along with reasonable execution time. In this case, the main reason that performance of the system is high, while there are a few rules, is the use of type 2 fuzzy systems, which are more adaptable to uncertainties.

Table 3 represents the parameters achieved by the PSO algorithm for Fuzzy-SIRD and SIRD models in both 75 days and 150 days scenarios for the United States. The actual, learned, and predicted deceased population and the corresponding error achieved by the proposed Fuzzy-SIRD model have been depicted in Fig. 7 in the left-most plot. The middle plot represents the infectious and recovered population over the time horizon.

The plot placed on the right side represents the achieved fuzzy and real effective reproduction numbers. It must be noted that the real effective reproduction number is the response of society to the fuzzy effective reproduction number, which is directly affected by government interventions. Fig. 8 demonstrates the real, learned, and predicted deceased population.

Also, in the right-side plot, infectious and recovered populations achieved by the conventional SIRD model have been represented. Both these figures are the representation of results achieved for the 75-day time horizon. Corresponding results for the 150-day time horizon have been represented in Figs. 9 and 10. Also, Fig. 11 demonstrates the convergence diagram of the PSO algorithm used for identifying the system parameters for the US case. This convergence diagram has been achieved by calculating the average convergence characteristics of the algorithm over all 100 runs.

As it can be seen in Table 2, considering the United States of America, for both scenarios, the Fuzzy-SIRD model has reached a better prediction of death tolls. In the second scenario, the enhancement of the results while using the Fuzzy-SIRD model is more noticeable and even better than in the first scenario. Prediction of such a complex and intricate time series affected by many environmental parameters gets harder by extending the time horizon. This situation has two causes. The first reason is the lack of ideal performance while running the optimization algorithm for finding the model parameters. For handling this problem, as it has been said, the optimization problem has run 100 times, and the best results have been reported. The second reason is the coincidence of the forecast period with an event that has a remarkable impact on the course of the data. In this situation, as the time horizon of the data used for learning becomes longer, the proposed model tries to digest the affecting event and thus reduces the error in long-term prediction. Based on Fig. 7, it can be claimed that the effective reproduction number experiences a gradual increase in the first forty days, and after that, a shock enters it. However, the long-term results give us a more precise and deeper insight. As it can be seen in Fig. 9, after a slight decrease in effective reproduction number, a gradual increase has been experienced until it approximately reaches 1.2. Then after experiencing a peak, the R_e decreases, but a second peak is on the way.

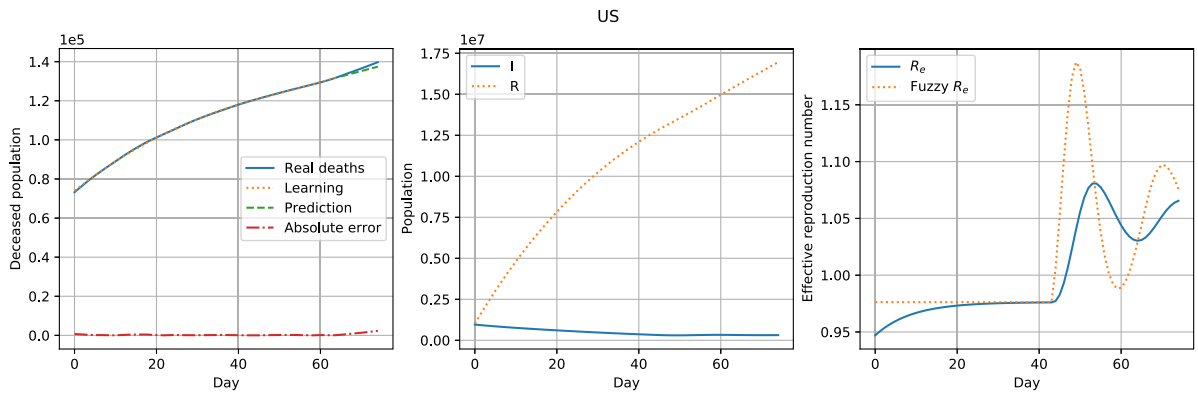


Fig. 7. The results achieved by Fuzzy-SIRD model in the learning and prediction phases, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the short time horizon for the US.

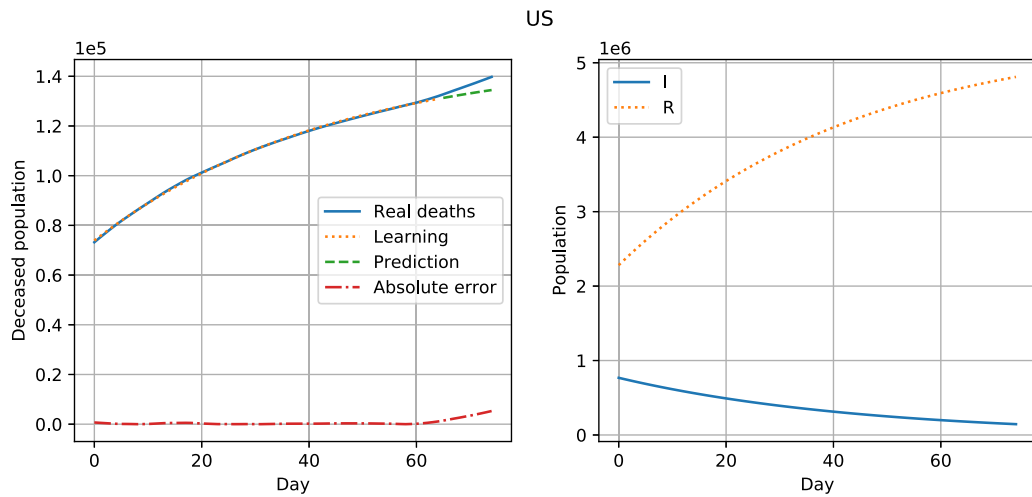


Fig. 8. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the short time horizon for the US.

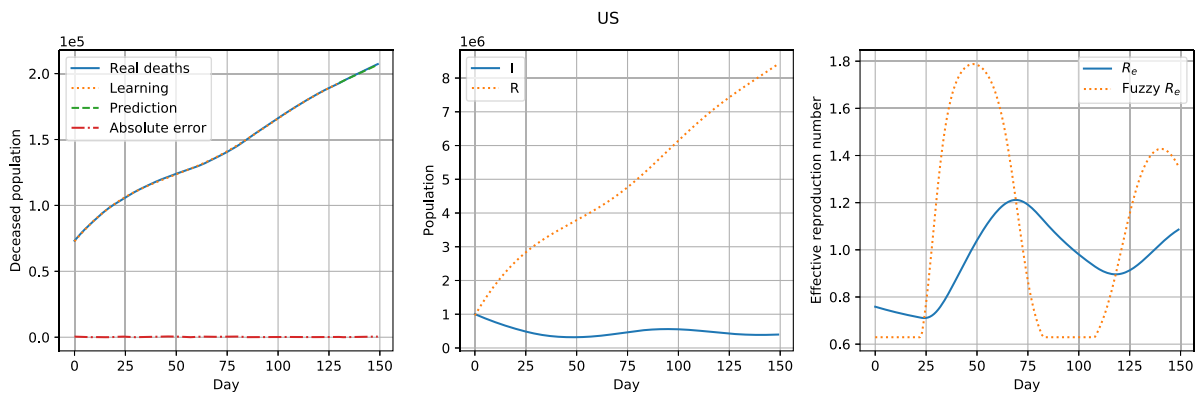


Fig. 9. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the long time horizon for the US.

The $D^{sat} = 976.3858$, achieved in the second scenario, shows that intervention of the US government facing COVID-19 reaches its maximum when the daily death toll increases to about a thousand or more. Furthermore, $\epsilon = 0.7108$ shows that the success of the US government intervention was moderate, and because of this, the effective reproduction number R_e could be in the range $[0.5842, 3.4048]$.

The results achieved for the COVID-19 time series of Germany have been represented in Figs. 12, 13, 14, and 15. The parameters of Fuzzy-SIRD and SIRD models for both scenarios have been reported

in Table 4. One of the best results achieved by the proposed model is for Germany, in which the predictions and actual values have only a little difference. The results show that the German government has successfully controlled the disease, and the R_e value was lower than 1 for a considerably long time. However, as it can be seen, the effective reproduction number has increased over time, and approximately on the 100th day of the second scenario, it has passed the value 1, meaning that the disease will spread in an exponential form again. As the fuzzy

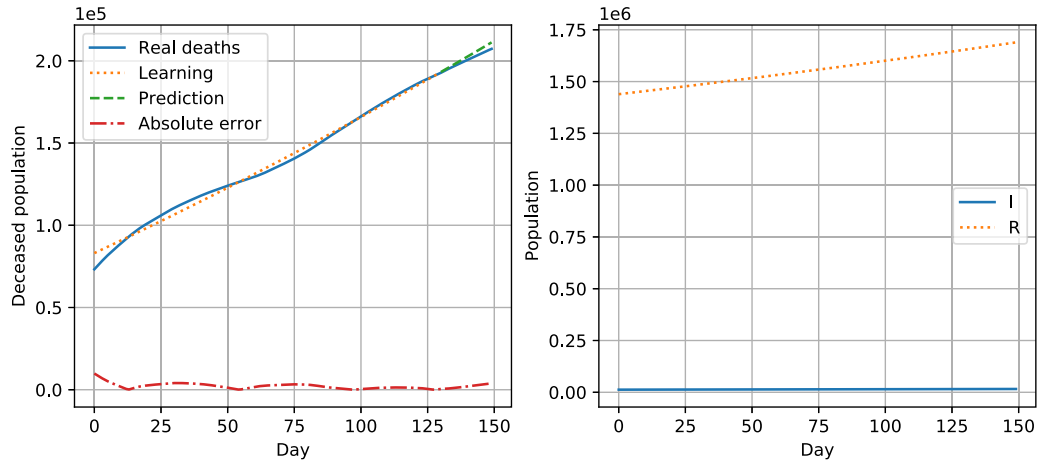


Fig. 10. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the long time horizon for the US.

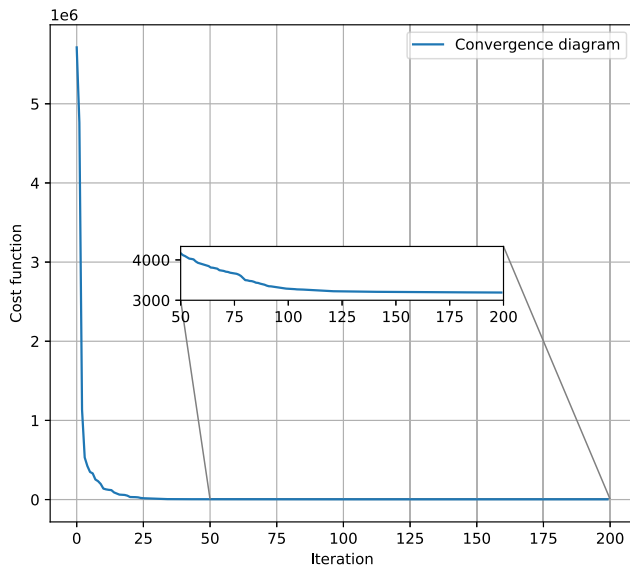


Fig. 11. The convergence diagram of the PSO algorithm (average of 100 runs) for identifying the parameters of the proposed system for the US in a 150-day time horizon.

R_e shows in Fig. 14, government intervention can successfully control the disease, and R_e may experience its peak at about 1.5.

In the case of Germany, D^{sat} is 40.4283 in the second scenario. It means that the government in Germany shows the most reaction when the daily death toll reaches about 40 or more. We can compare this number with the achieved one for the US by normalizing them with respect to these countries' populations. The normalized D^{sat} for Germany is 4.8568×10^{-7} , and for the US is 2.9632×10^{-6} . The achieved number for Germany is 6.1012 times less than the US, which shows that in the considered time interval compared with the US government, the German government was more sensitive to its citizens. Also, the government intervention efficiency factor, ϵ , for Germany is 1.0, which is the maximum efficiency. Based on the achieved ϵ factor, the effective reproduction number could be in the range [0.4507, 2.479], which is less than the achieved one for the US with $\epsilon = 0.7108$.

The rest of the simulations have been provided in the Appendix section. It must be noted that the simulations have been done by Python using the PyIT2FLS [40], Numpy [41], and SciPy [42] libraries. Also, representation of the data has been done using the Matplotlib [43] library.

US

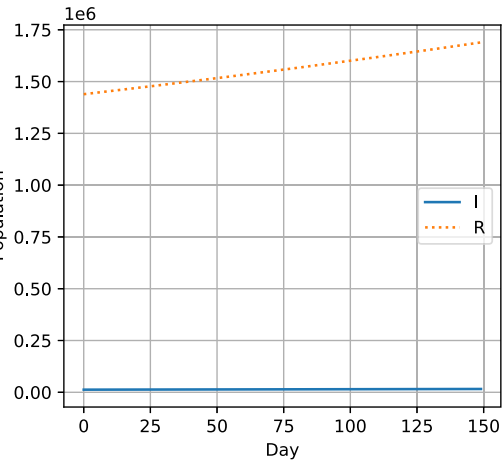


Table 4

The parameters of Fuzzy-SIRD and SIRD models achieved for the Germany COVID-19 time series.

Parameter	Scenario 1		Scenario 2	
	Fuzzy-SIRD	SIRD	Fuzzy-SIRD	SIRD
I^0	1983.6087	64324.9844	8311.0249	1548.0295
R^0	181336.6979	156181.2473	203670.9836	24998.4027
D^0	7047.6271	7081.7298	7059.6980	7174.2583
R_e^0	0.0954	0.0828	0.0786	0.0565
γ	0.0205	0.1263	0.0434	0.0434
μ	0.0534	0.0014	0.0119	0.0497
α_1	0.0853	-	0.0090	-
α_2	0.0853	-	0.0090	-
D^{sat}	15.0644	-	40.4283	-
ϵ	1.0	-	1.0	-

5. Conclusion

In this paper, a new Fuzzy-SIRD model has been proposed for epidemiologic studies. The proposed innovative approach aims to model the effects of government intervention on the effective reproduction number of the disease. Considering that the reports of the infected people are not accurate enough, the death toll time series has been used for extracting the system parameters. Using the infected population as data for tuning the model parameters may lead the models to predict the governments' policies of test and diagnosis rather than disease spread. Some simulations have been done on the COVID-19 time series to test the efficiency of the proposed model. Seven countries have been studied with two scenarios, and the results achieved by the proposed and conventional SIRD models have been compared. In all cases, the proposed Fuzzy-SIRD model is much better than the conventional SIRD model. Compared with the other method, the results show lesser root mean square error, which approves the efficiency and preciseness of the proposed method. In the short-term scenario, the minimum decrement of prediction RMSE is 17.75% in the UK case, and the maximum decrement is 71.21% in the case of Germany. Utilizing the proposed Fuzzy-SIRD model causes an average decrement of RMSE equal to 45.83%. As it can be seen, the proposed method yields a significant reduction in prediction errors. This decrement is even larger in long-term time intervals. In long-term case studies, the proposed Fuzzy-SIRD model causes an average decrement of RMSE equal to 72.56%. The maximum decrement of RMSE in long-term scenarios belongs to Iran. In the case of Iran, the RMSE of long-term predictions has been reduced by 88.57%. Also, the minimum decrement of RMSE in long-term scenarios belongs to Russia, 2.59%. Among the different countries, the COVID-19 trend in Russia shows great cohesion with the conventional SIRD

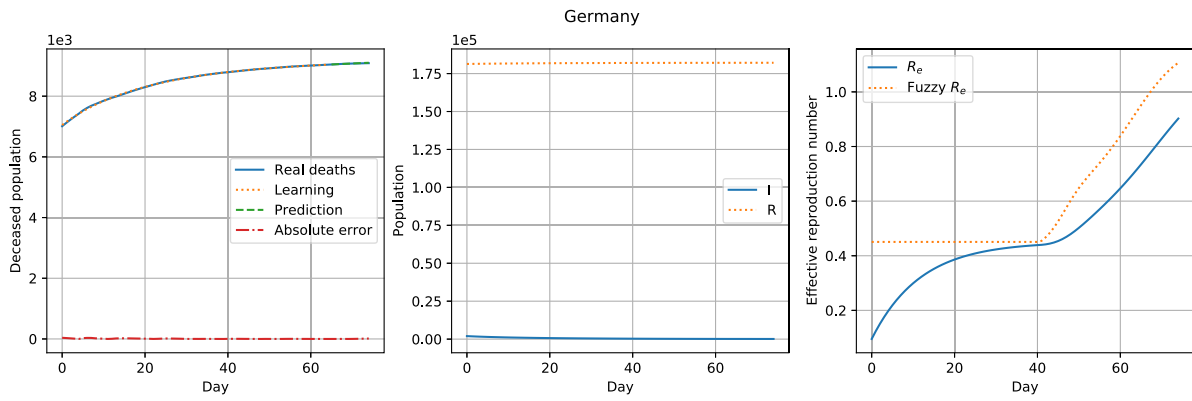


Fig. 12. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the short time horizon for Germany.

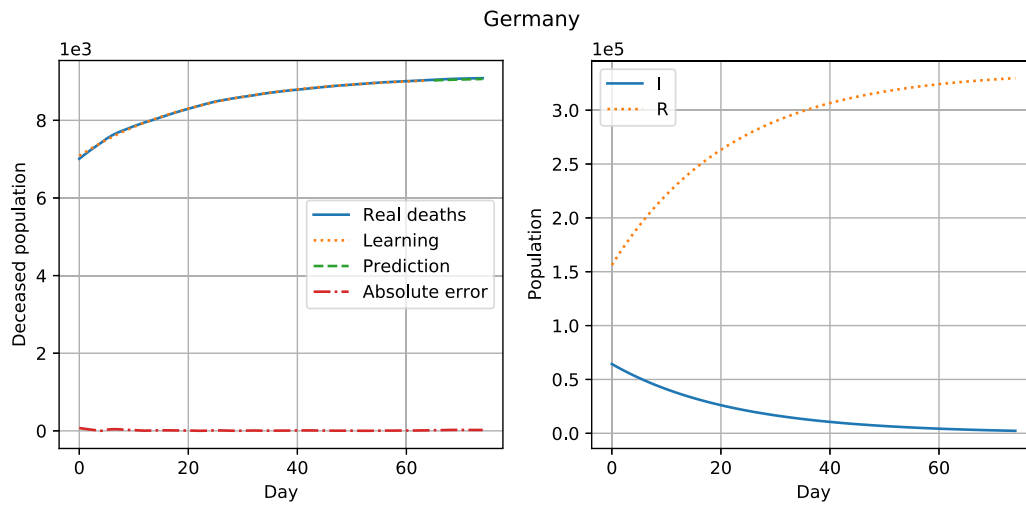


Fig. 13. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the short time horizon for Germany.

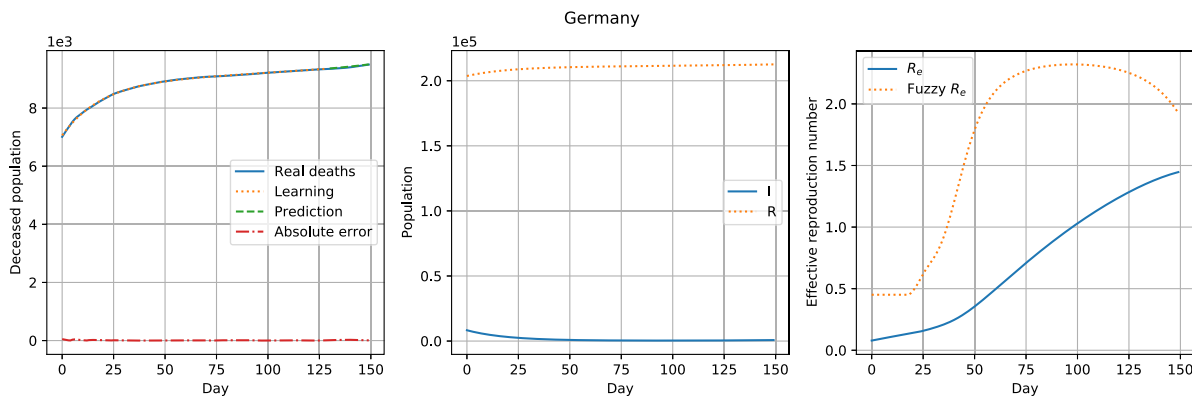


Fig. 14. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the long time horizon for Germany.

model, but even in this case, the proposed model causes more precise predictions.

The proposed approach is able to model the government intervention and its effects on epidemiological parameters of the disease, such as effective reproduction number. The achieved R_e and the fuzzy R_e , which the latter is a measure of government intervention, have been plotted for all seven countries and both scenarios. The variation of these two parameters over time has been interpreted in all cases. Such interpretations give us an insight into the actions done by governments

and the state of disease transmission. Furthermore, the parameters of the proposed model, such as α_1 and α_2 , give us an estimation of society's response to government interventions. For example, considering the optimal parameters achieved for Iran and the US in the long-term scenario, we can compare the society's response in these two countries. For Iran and the US, α_2 has been achieved as 0.0292 and 0.0199, respectively. It means that in the US, society responds faster than in Iran. According to Fig. 9, there is a three-week gap between Fuzzy R_e and R_e for the US. While this gap for Iran is about four weeks, as can

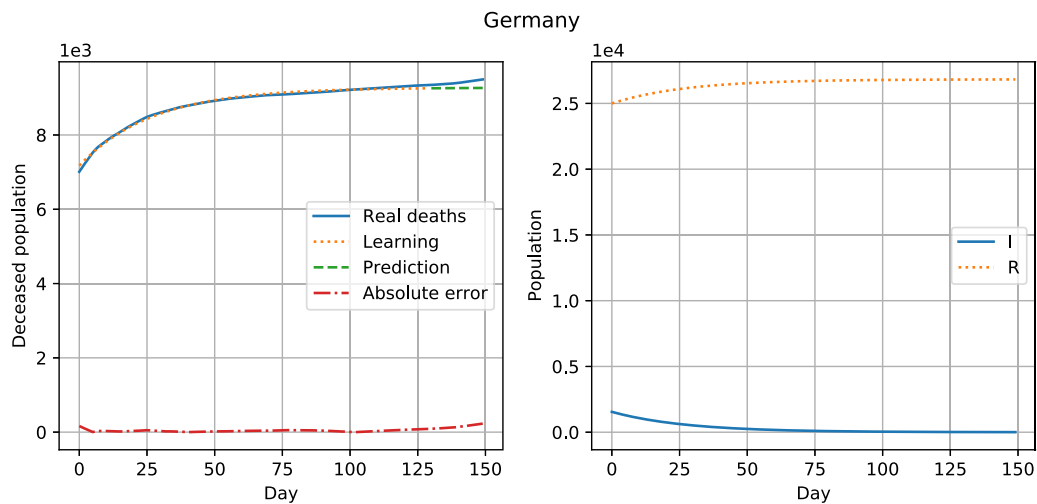


Fig. 15. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the long time horizon for Germany.

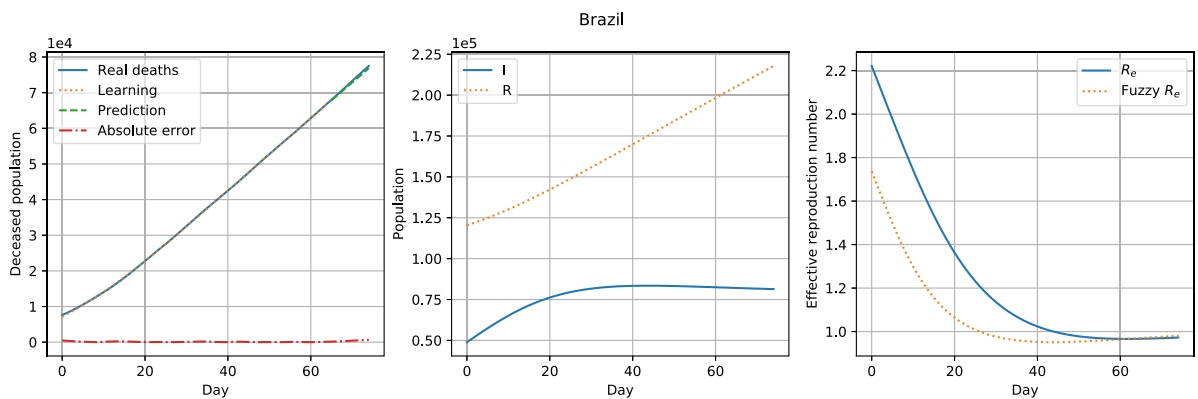


Fig. 16. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the short time horizon for Brazil.

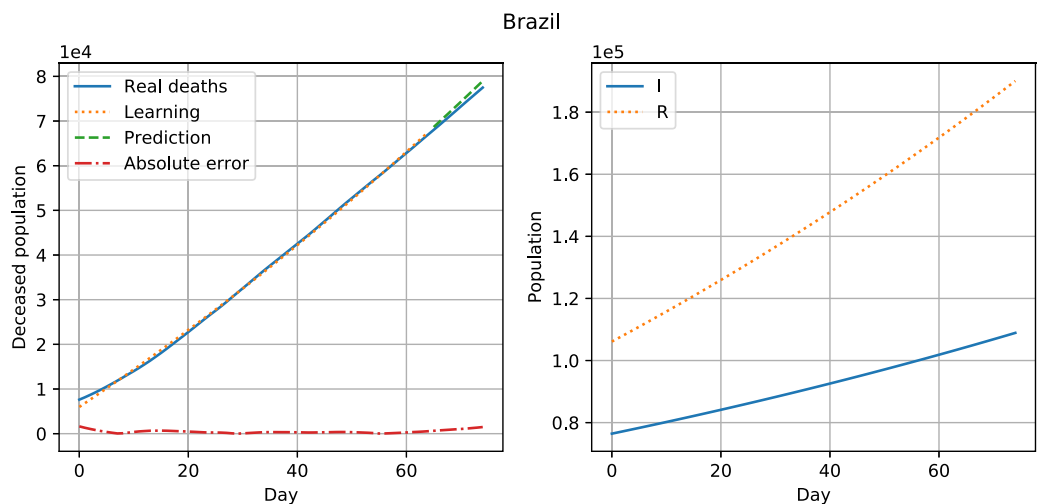


Fig. 17. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the short time horizon for Brazil.

be seen in Fig. 22. That is why the proposed model can be of particular importance in the study of epidemics.

Despite all the advantages listed for the model proposed in this paper, the Fuzzy-SIRD model is not able to adapt to the condition

after starting the vaccination. Therefore, the idea of this paper can be generalized for models in which vaccination is also included. However, variety in vaccine types, the time required to provide immunity by different vaccines, the degree of immunity after injection of each dose, the

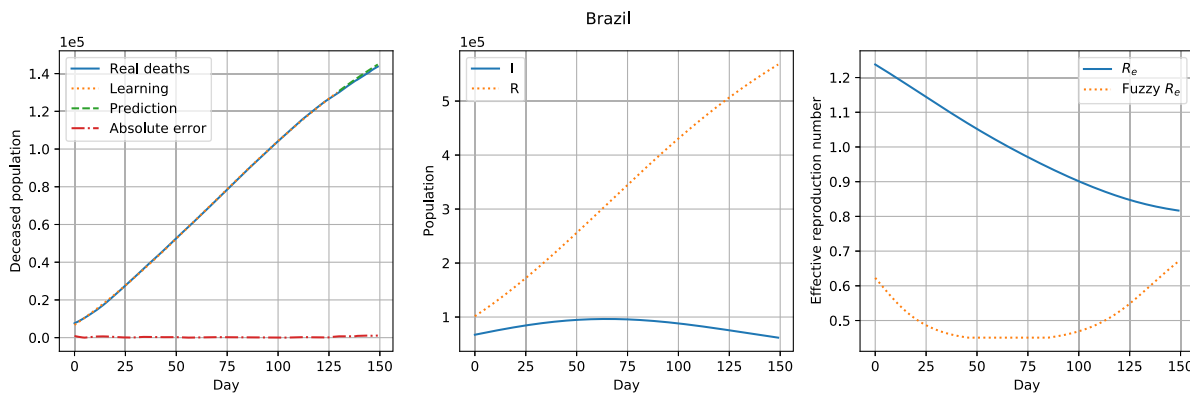


Fig. 18. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the long time horizon for Brazil.

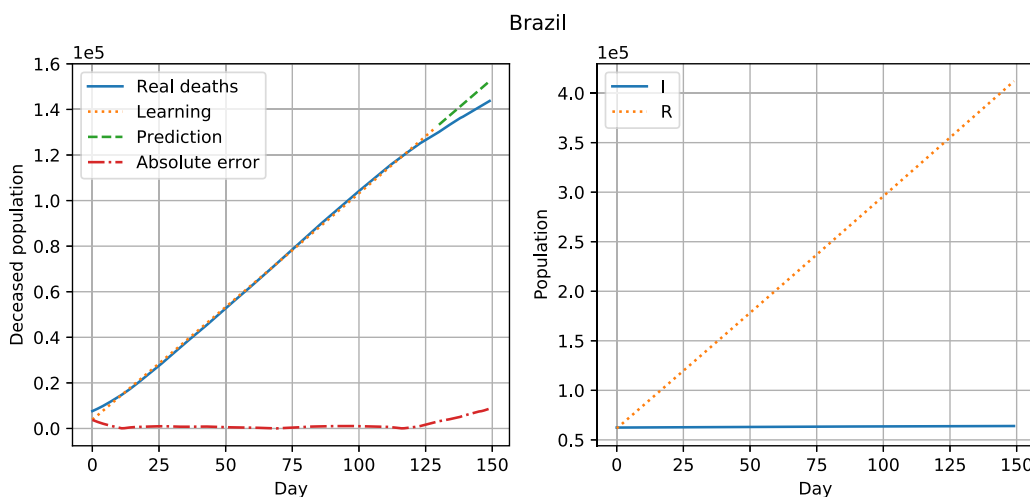


Fig. 19. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the long time horizon for Brazil.

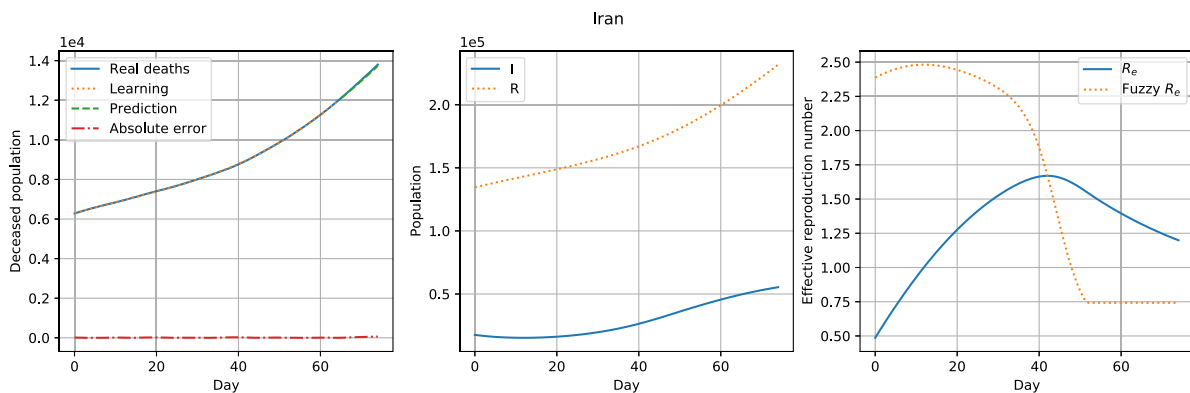


Fig. 20. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the short time horizon for Iran.

reduction of immunity after a specific time, the effectiveness of various vaccines, the rate of transmission prevention in vaccinated individuals, etc. make this task very difficult and unimaginably complex.

In conclusion, it can be said that the proposed approach has achieved the goal of modeling the impact of government intervention on epidemics. It has also significantly improved the conventional

compartmental models. With the advent of machine learning-based models, compartmental models are receiving less attention. Considering the proposed method, we predict that compartmental models will gain their place again in epidemiological time-series studies. It seems that by the development of hybrid fuzzy and classic models, in addition to increasing the precision of the classical models, we will be able to

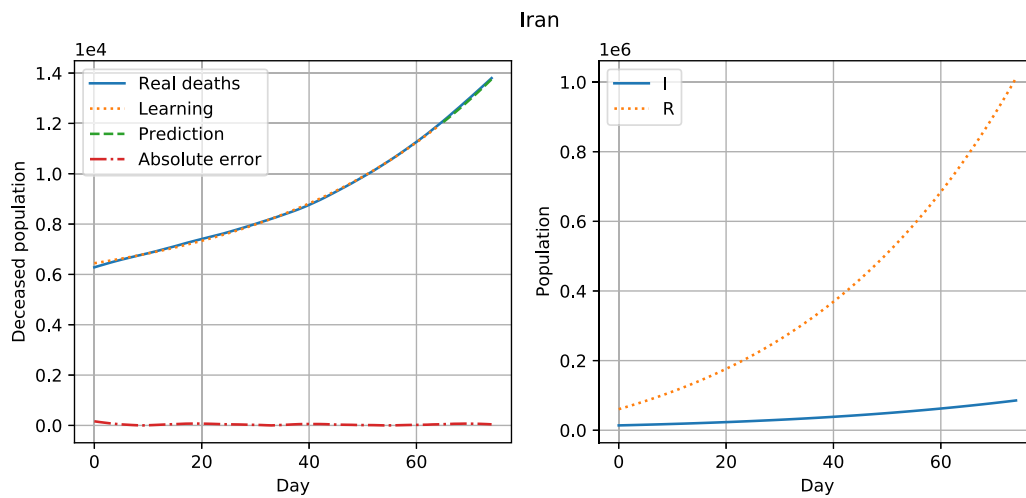


Fig. 21. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the short time horizon for Iran.

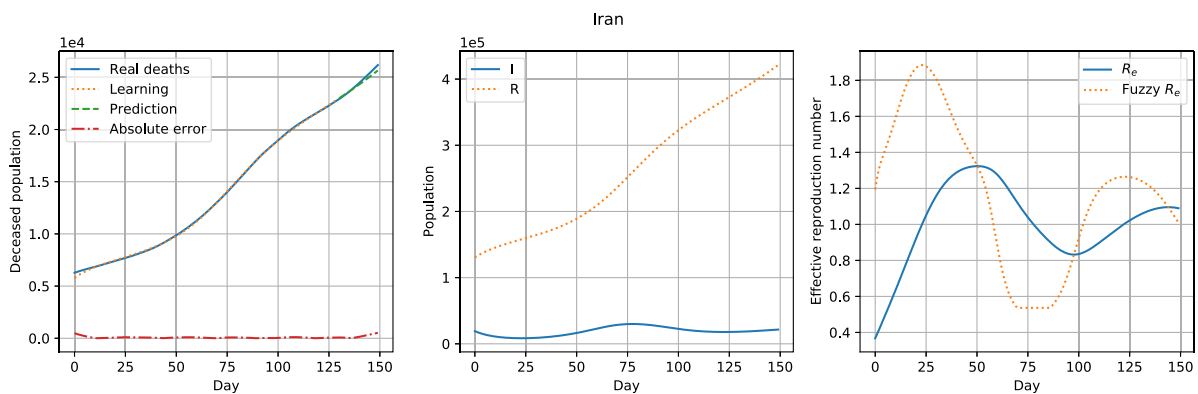


Fig. 22. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the long time horizon for Iran.

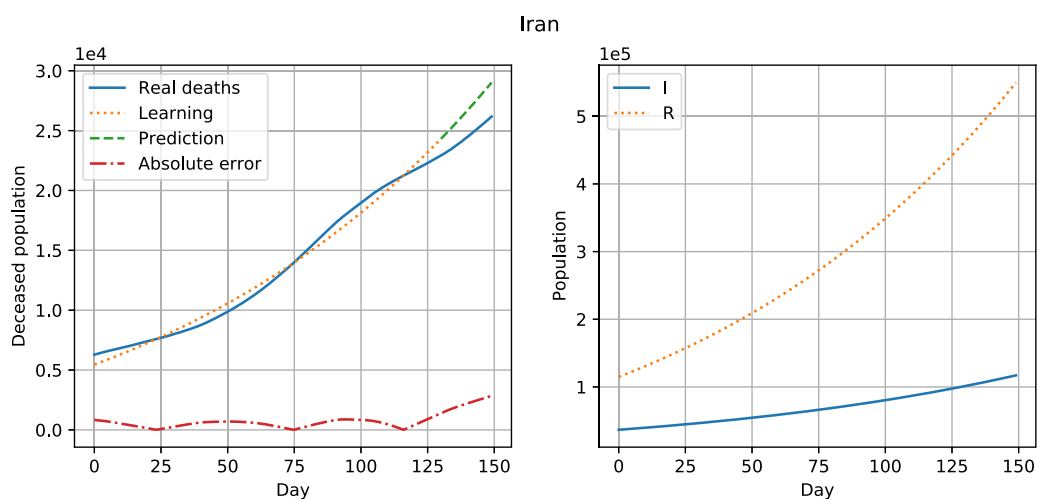


Fig. 23. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the long time horizon for Iran.

express field data as human experiences. In future work, the factors introduced in Section 2.1 can be described using fuzzy IF-THEN rules, and their effects can be considered in the proposed model. Such an effort can improve the precision of the proposed model and give the researcher a deeper insight into the complex society's response facing the epidemic disease.

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.

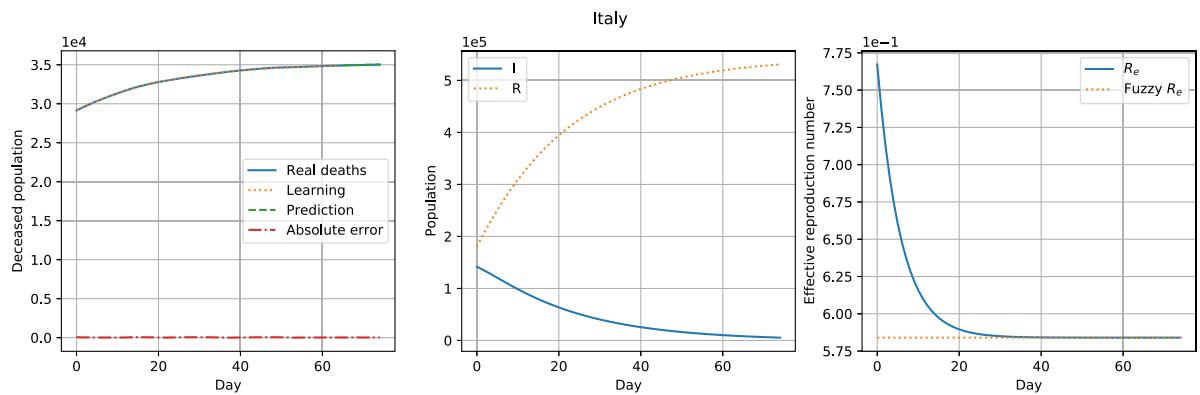


Fig. 24. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the short time horizon for Italy.

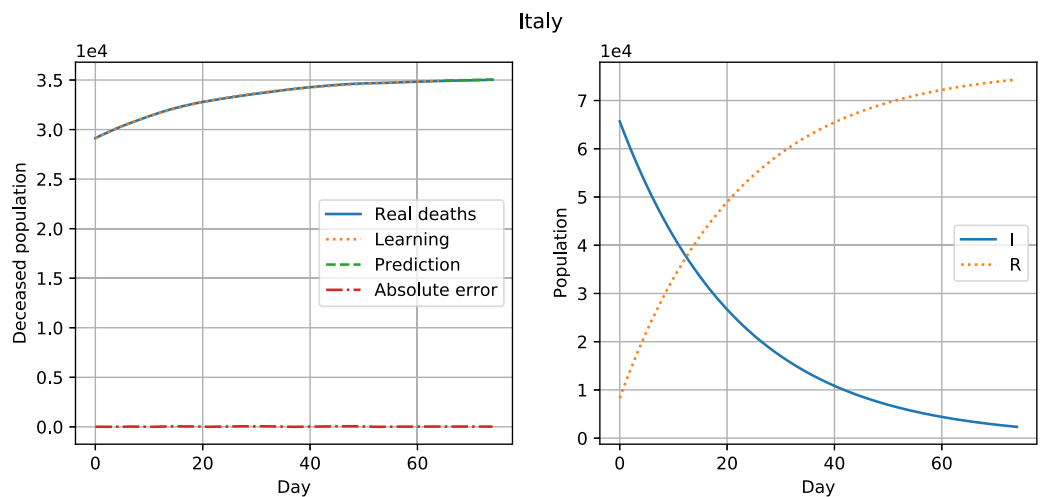


Fig. 25. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the short time horizon for Italy.

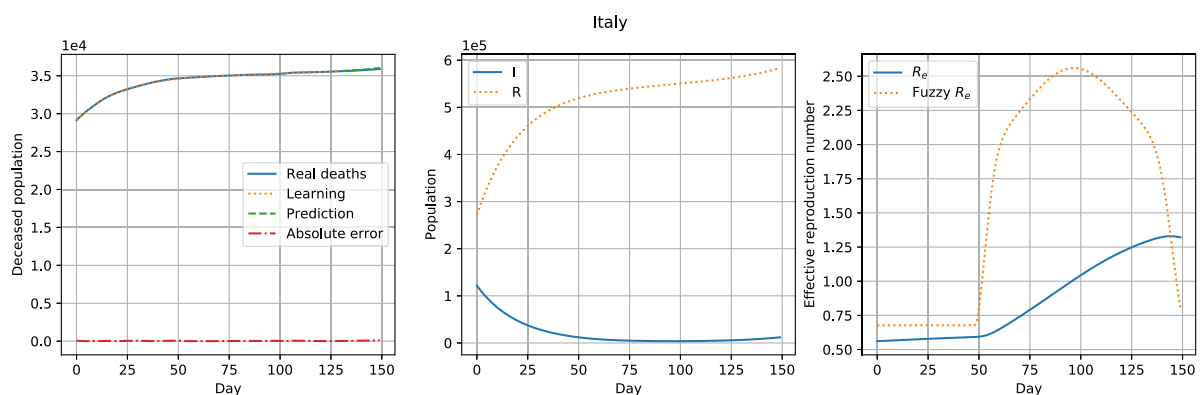


Fig. 26. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the long time horizon for Italy.

Appendix

The results for Brazil have been represented in Figs. 16, 17, 18, and 19, in the same manner as for the US. The optimal parameters for both models in both scenarios achieved for Brazil have been reported in Table 5. Similar to the US, the results achieved by the Fuzzy-SIRD model in both scenarios are better than conventional SIRD. By extending the time horizon, the RMSE error for the SIRD model has increased more than fivefold, but for the Fuzzy-SIRD model, it has only been approximately doubled. Unlike the United States, effective

reproduction number experiences a slight reduction over time in both short and long scenarios.

Iran is the next country where the simulations have been done based on its time series. The output plots have been demonstrated in Figs. 20, 21, 22, and 23. Table 6 represents the model parameters achieved for this country. Like Germany, in which the results of both scenarios were consequent, for Iran, the R_e and fuzzy R_e in Fig. 22 are almost a continuation of Fig. 20. In 150 days of the assumed time horizon, Iran has experienced two peaks in the disease transmission rate. In the first peak, which has approximately occurred 6th week

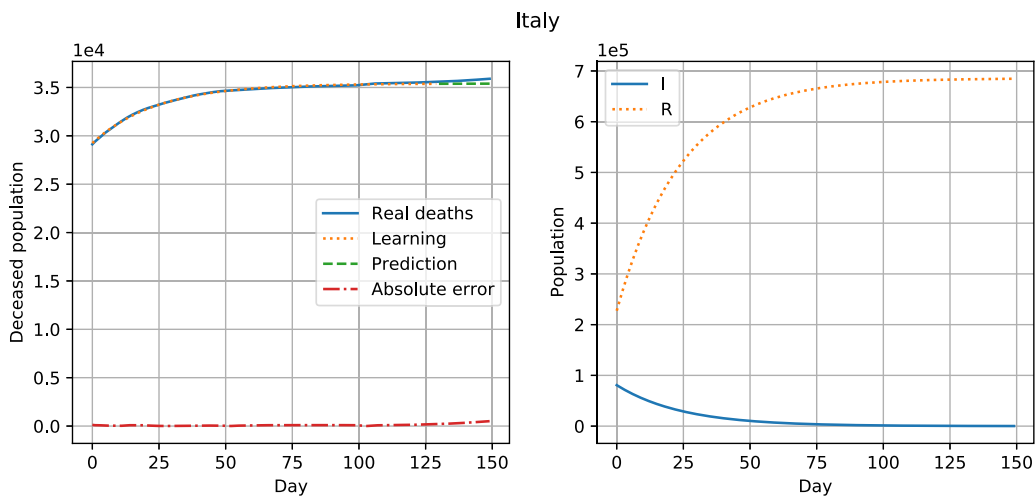


Fig. 27. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the long time horizon for Italy.

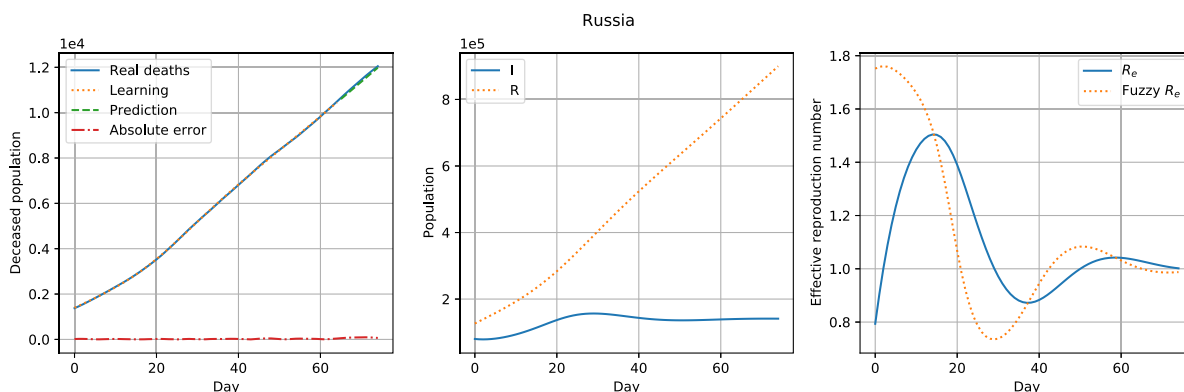


Fig. 28. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the short time horizon for Russia.

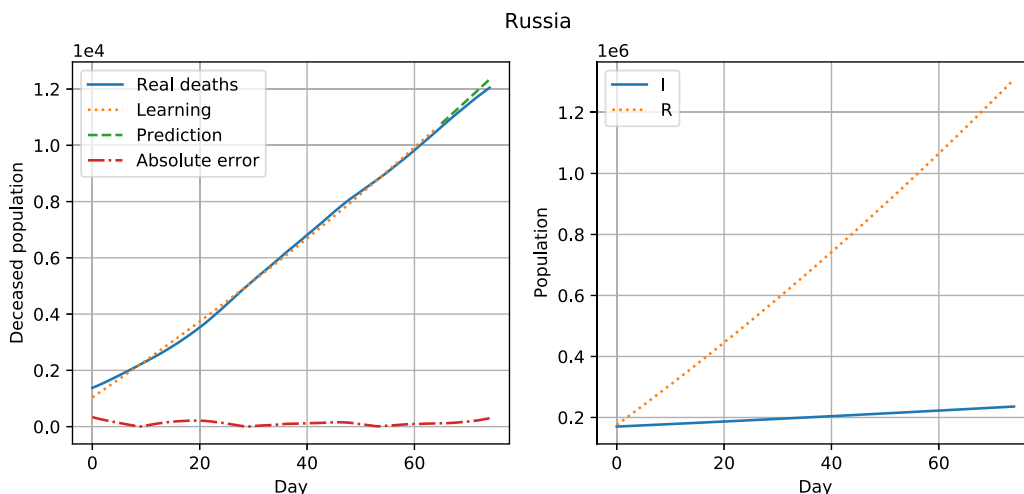


Fig. 29. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the short time horizon for Russia.

of the time horizon, the effective reproduction number has reached higher than 1.3. But through the actions done by the government with a considerably large delay of about 30 days, the disease has been brought under control. Also, the second peak, milder than the first, occurred on approximately the 140th day of the time horizon. Although the data have been accompanied by fluctuations, and the conventional SIRD

model has completely failed in the second scenario, the proposed model has provided an acceptable result.

Figs. 24, 25, 26, and 27 depict the outputs achieved for Italy. Also, Table 7 demonstrates the corresponding model parameters. In Italy, according to the results, the effective reproduction number starts from values lower than 1, and gradually in 100 days, the R_e reaches its peak

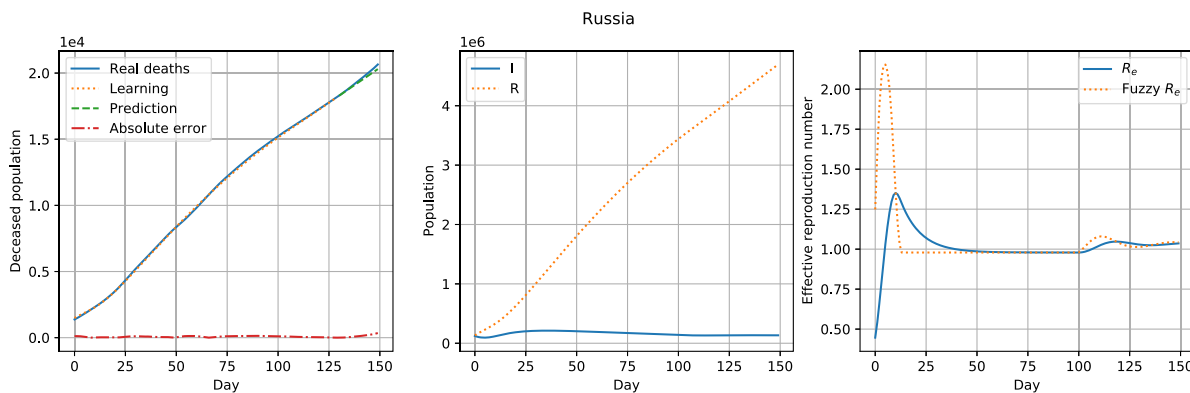


Fig. 30. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the long time horizon for Russia.

Table 5

The parameters of Fuzzy-SIRD and SIRD models achieved for the Brazil COVID-19 time series.

Parameter	Scenario 1		Scenario 2	
	Fuzzy-SIRD	SIRD	Fuzzy-SIRD	SIRD
I^0	48912.9733	76443.9333	67210.2989	62426.2085
R^0	120356.9097	106098.9387	101487.8183	61574.8858
D^0	7134.6536	5977.7386	6751.3856	3853.9964
R_e^0	2.2212	0.0279	1.2377	0.0532
γ	0.0170	0.0124	0.0371	0.0372
μ	0.0122	0.0108	0.0110	0.0157
α_1	0.1011	-	0.0058	-
α_2	0.1011	-	0.0058	-
D^{sat}	2277.9255	-	1027.3267	-
ϵ	1.0	-	1.0	-

Table 6

The parameters of Fuzzy-SIRD and SIRD models achieved for the Iran COVID-19 time series.

Parameter	Scenario 1		Scenario 2	
	Fuzzy-SIRD	SIRD	Fuzzy-SIRD	SIRD
I^0	17622.7468	13900.6365	19119.3638	36702.9093
R^0	134520.9409	60503.7486	130548.5783	114682.1795
D^0	6294.6422	6444.9966	5794.8825	5453.1067
R_e^0	0.4865	0.3467	0.3677	0.0523
γ	0.0452	0.3179	0.1041	0.0420
μ	0.0035	0.0024	0.0071	0.0023
α_1	0.0256	-	0.0292	-
α_2	0.0256	-	0.0292	-
D^{sat}	129.9242	-	202.7566	-
ϵ	0.5769	-	0.8222	-

Table 7

The parameters of Fuzzy-SIRD and SIRD models achieved for the Italy COVID-19 time series.

Parameter	Scenario 1		Scenario 2	
	Fuzzy-SIRD	SIRD	Fuzzy-SIRD	SIRD
I^0	141688.4272	65669.5624	121514.8229	80724.1999
R^0	180681.9331	8287.8052	271737.0489	227410.2261
D^0	29191.1873	29117.5745	29070.2029	29245.5362
R_e^0	0.7672	0.0062	0.5613	0.1975
γ	0.1067	0.0470	0.1063	0.2339
μ	0.0018	0.0042	0.0024	0.0031
α_1	0.1748	-	0.0064	-
α_2	0.1748	-	0.0064	-
D^{sat}	2.9185	-	30.1318	-
ϵ	0.7636	-	0.6520	-

above 1.25. Approximately after the 100th day of the time horizon, COVID-19 spreads in an exponential form.

Table 8

The parameters of Fuzzy-SIRD and SIRD models achieved for the Russia COVID-19 time series.

Parameter	Scenario 1		Scenario 2	
	Fuzzy-SIRD	SIRD	Fuzzy-SIRD	SIRD
I^0	80077.4312	170118.6521	122759.6347	2857.1751
R^0	126455.4159	175410.9377	131940.8516	15.5306
D^0	1393.2686	1039.8632	1483.2160	502.3461
R_e^0	0.7935	0.0815	0.4466	0.3642
γ	0.0796	0.0758	0.1908	0.3069
μ	0.0011	0.0008	0.0008	0.0608
α_1	0.1219	-	0.1015	-
α_2	0.1219	-	0.1015	-
D^{sat}	186.3529	-	109.5131	-
ϵ	0.7595	-	0.4363	-

Table 9

The parameters of Fuzzy-SIRD and SIRD models achieved for the United Kingdom COVID-19 time series.

Parameter	Scenario 1		Scenario 2	
	Fuzzy-SIRD	SIRD	Fuzzy-SIRD	SIRD
I^0	57509.3925	418725.9331	307269.0862	181130.1661
R^0	9715.1647	202470.8131	182486.6248	124728.4758
D^0	28845.8782	28915.9224	28887.7614	28571.8191
R_e^0	0.8197	0.4616	0.8145	0.0746
γ	0.1353	0.4757	0.1358	0.1134
μ	0.0083	0.0011	0.0015	0.0031
α_1	0.0138	-	0.0126	-
α_2	0.0138	-	0.0126	-
D^{sat}	92.2009	-	53.9868	-
ϵ	1.0	-	1.0	-

The outputs achieved by Fuzzy-SIRD and SIRD models in both scenarios for Russia have been represented in Figs. 28, 29, 30, and 31. The parameters obtained by the PSO algorithm for these two models have been reported in Table 8. Based on the results achieved for the COVID-19 time series of Russia, in the first two weeks of the time horizon, a quick increase in R_e has occurred, which is simultaneous with the decrement of fuzzy R_e . Since the fuzzy R_e is a measure of government intervention and R_e theoretically is following it, it can be said that the Russian government has successfully controlled the disease and reduced the spread speed by decreasing the fuzzy R_e . Until the 100th day of the time horizon, the R_e for Russia is less than one, but over time a small increment has been experienced.

The results achieved for the United Kingdom have been demonstrated in Figs. 32, 33, 34, and 35. Table 9 presents the model parameters for different cases and models. In the first scenario, the effective reproduction number is almost constant value over the time horizon. However, in the second scenario, R_e experiences a rise after a slight

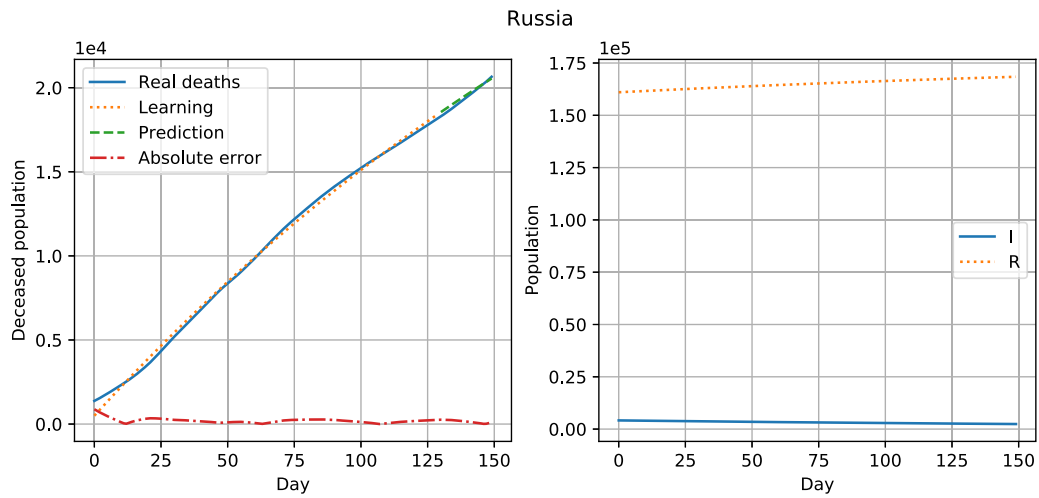


Fig. 31. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the long time horizon for Russia.

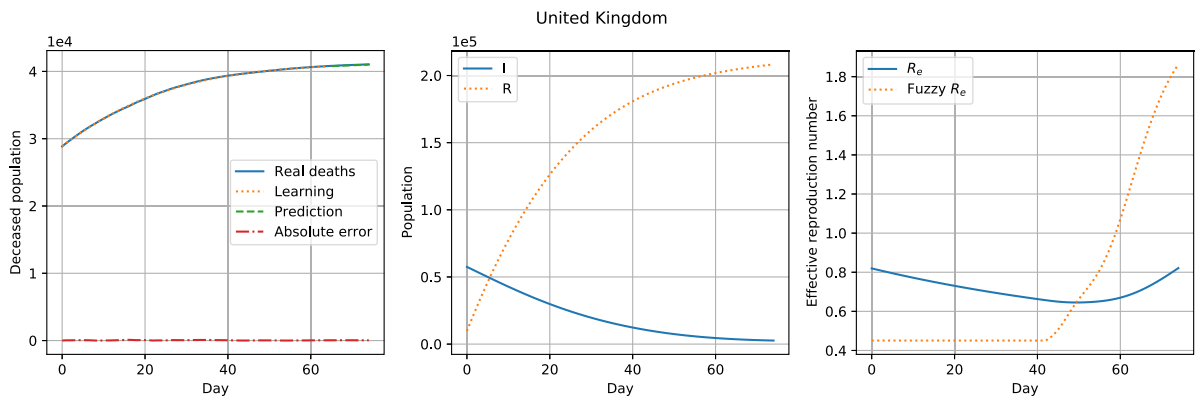


Fig. 32. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the short time horizon for UK.

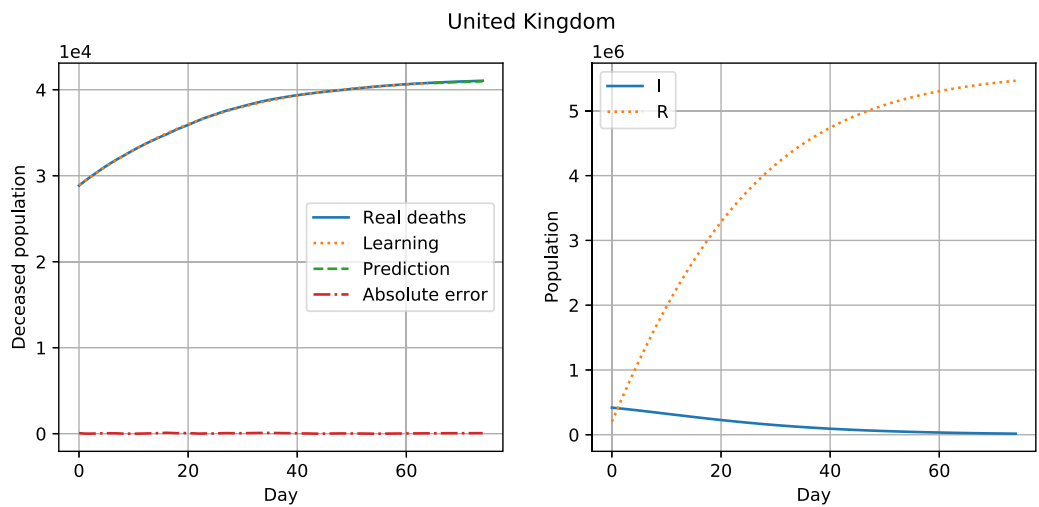


Fig. 33. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the short time horizon for UK.

decrement. The maximum value for R_e is approximately 1.3, which has also started its drop in the last two weeks of the time horizon. The fuzzy effective reproduction number starts a fast increment about the 50th

day of the time horizon, which shows that the UK government has lifted some restrictions. After the 100th day, it seems that the government has again intervened in the favor of lowering the R_e .

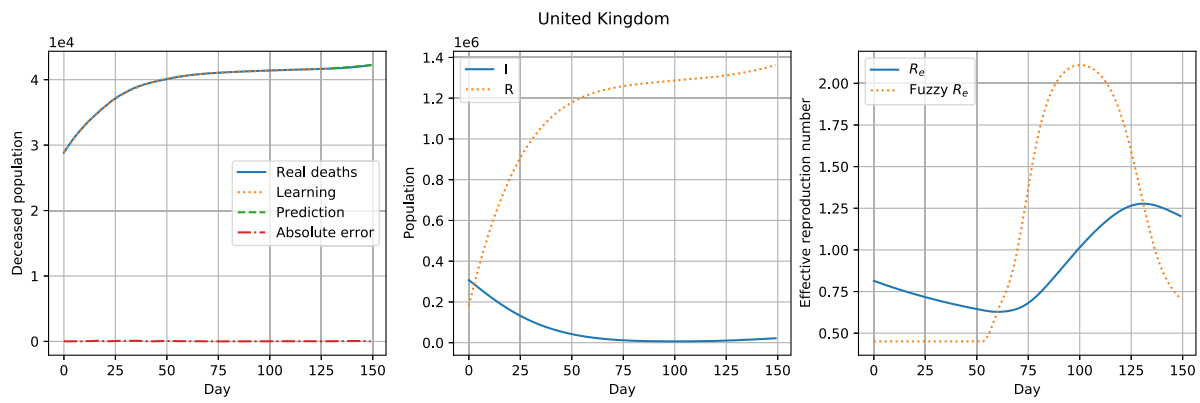


Fig. 34. The results achieved by Fuzzy-SIRD model in the learning and prediction phase, the corresponding error, infectious and recovered population, R_e , and output of the fuzzy system (fuzzy R_e) over the long time horizon for UK.

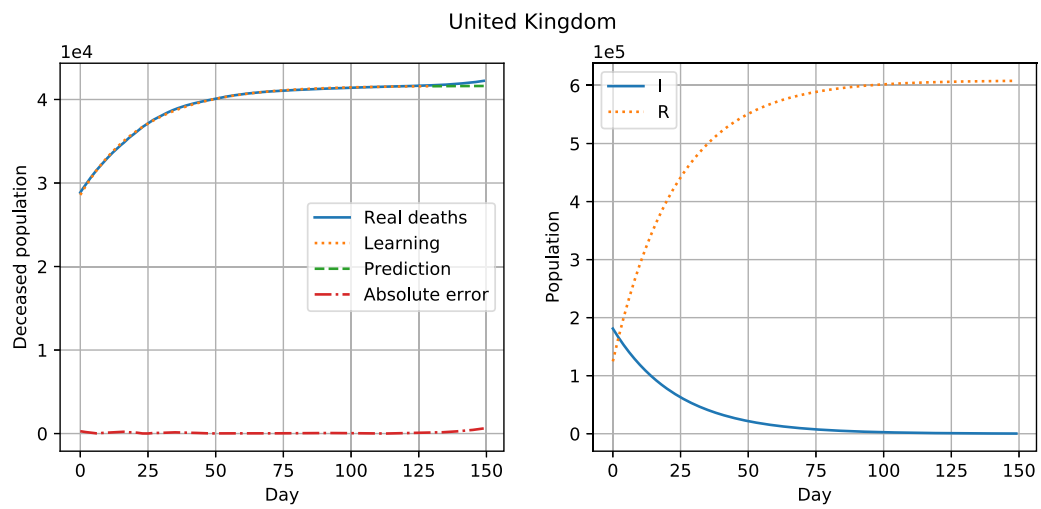


Fig. 35. The results achieved by SIRD model in the learning and prediction phases, infectious and recovered population over the long time horizon for UK.

References

[1] Huremović D. Brief history of pandemics (pandemics throughout history). In: Psychiatry of pandemics. Springer; 2019, p. 7–35.

[2] Martini M, Gazzaniga V, Bragazzi NL, Barberis I. The Spanish influenza pandemic: a lesson from history 100 years after 1918. *J. Prevent. Med. Hygiene* 2019;60(1):E64.

[3] Yuki K, Fujiogi M, Koutsogiannaki S. COVID-19 pathophysiology: A review. *Clin. Immunol.* 2020;108427.

[4] Castillo O, Melin P. Forecasting of COVID-19 time series for countries in the world based on a hybrid approach combining the fractal dimension and fuzzy logic. *Chaos Solitons Fractals* 2020;140:110242.

[5] Melin P, Monica JC, Sanchez D, Castillo O. Multiple ensemble neural network models with fuzzy response aggregation for predicting COVID-19 time series: the case of Mexico. In: *Healthcare*. 8, (2):Multidisciplinary Digital Publishing Institute; 2020, p. 181.

[6] Abbasimehr H, Paki R. Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization. *Chaos Solitons Fractals* 2021;142:110511.

[7] Leila M, Mozghan S, Sadat SM. Exponentially increasing trend of infected patients with Covid-19 in Iran: A comparison of neural network and arima forecasting models. *Iran. J. Publ. Health* 2020;49:92–100.

[8] Jia L, Li K, Jiang Y, Guo X, Zhao T. Prediction and analysis of coronavirus disease 2019. *arxiv e-prints*. 2020, arXiv preprint arXiv:2003.05447.

[9] Castorina P, Iorio A, Lanteri D. Data analysis on coronavirus spreading by macroscopic growth laws. *Internat J Modern Phys C* 2020;31(07):2050103.

[10] Shahid F, Zameer A, Muneeb M. Predictions for COVID-19 with deep learning models of LSTM, GRU and bi-LSTM. *Chaos Solitons Fractals* 2020;140:110212.

[11] Tuli S, Tuli S, Tuli R, Gill SS. Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing. *Internet Things* 2020;11:100222.

[12] Singhal A, Singh P, Lall B, Joshi SD. Modeling and prediction of COVID-19 pandemic using Gaussian mixture model. *Chaos Solitons Fractals* 2020;138:110023.

[13] Ghosh S, Bhattacharya S. A data-driven understanding of COVID-19 dynamics using sequential genetic algorithm based probabilistic cellular automata. *Appl Soft Comput* 2020;96:106692.

[14] Dong M, Tang C, Ji J, Lin Q, Wong K-C. Transmission trend of the COVID-19 pandemic predicted by dendritic neural regression. *Appl Soft Comput* 2021;111:107683.

[15] Hazarika BB, Gupta D. Modelling and forecasting of COVID-19 spread using wavelet-coupled random vector functional link networks. *Appl Soft Comput* 2020;96:106626.

[16] Aydin N, Yurdakul G. Assessing countries' performances against COVID-19 via WSIDEA and machine learning algorithms. *Appl Soft Comput* 2020;97:106792.

[17] Kumar N, Susan S. Particle swarm optimization of partitions and fuzzy order for fuzzy time series forecasting of COVID-19. *Appl Soft Comput* 2021;107611.

[18] Yudistira N, Sumitro SB, Nahas A, Riama NF. Learning where to look for COVID-19 growth: Multivariate analysis of COVID-19 cases over time using explainable convolution-LSTM. *Appl Soft Comput* 2021;109:107469.

[19] Hernandez-Matamoros A, Fujita H, Hayashi T, Perez-Meana H. Forecasting of COVID19 per regions using ARIMA models and polynomial functions. *Appl Soft Comput* 2020;96:106610.

[20] ArunKumar K, Kalaga DV, Kumar CMS, Chilkoor G, Kawaji M, Brenza TM. Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered

- and deaths) for top-16 countries using statistical machine learning models: Auto-regressive integrated moving average (ARIMA) and seasonal auto-regressive integrated moving average (SARIMA). *Appl Soft Comput* 2021;103:107161.
- [21] Ceylan Z. Short-term prediction of COVID-19 spread using grey rolling model optimized by particle swarm optimization. *Appl Soft Comput* 2021;107592.
- [22] Calafiore GC, Novara C, Possieri C. A time-varying SIRD model for the COVID-19 contagion in Italy. *Annu Rev Control* 2020.
- [23] Caccavo D. Chinese and Italian COVID-19 outbreaks can be correctly described by a modified SIRD model. 2020, MedRxiv.
- [24] Ferrari L, Gerardi G, Manzi G, Micheletti A, Nicolussi F, Biganzoli E, et al. Modelling provincial covid-19 epidemic data in Italy using an adjusted time-dependent SIRD model. 2020, arXiv preprint arXiv:2005.12170.
- [25] Lalwani S, Sahni G, Mewara B, Kumar R. Predicting optimal lockdown period with parametric approach using three-phase maturation SIRD model for COVID-19 pandemic. *Chaos Solitons Fractals* 2020;138:109939.
- [26] Gnanvi J, Salako KV, Kotanmi B, Kakai RG. On the reliability of predictions on Covid-19 dynamics: a systematic and critical review of modelling techniques. *Infect Dis Model* 2021.
- [27] Holmdahl I, Buckee C. Wrong but useful—what covid-19 epidemiologic models can and cannot tell us. *N Engl J Med* 2020;383(4):303–5.
- [28] Imtyaz A, Haleem A, Javaid M. Analysing governmental response to the COVID-19 pandemic. *J Oral Biol Craniofacial Res* 2020;10(4):504–13.
- [29] Khan JR, Awan N, Islam M, Muurlink O, et al. Healthcare capacity, health expenditure, and civil society as predictors of COVID-19 case fatalities: A global analysis. *Front Publ Health* 2020;8:347.
- [30] Etzioni A. The responsive community: A communitarian perspective. *Am Sociol Rev* 1996;1–11.
- [31] Goldstein JR, Lee RD. Demographic perspectives on the mortality of COVID-19 and other epidemics. *Proc Natl Acad Sci* 2020;117(36):22035–41.
- [32] Cao Y, Hiyoshi A, Montgomery S. COVID-19 case-fatality rate and demographic and socioeconomic influencers: worldwide spatial regression analysis based on country-level data. *BMJ Open* 2020;10(11):e043560.
- [33] Sun Z, Zhang H, Yang Y, Wan H, Wang Y. Impacts of geographic factors and population density on the COVID-19 spreading under the lockdown policies of China. *Sci Total Environ* 2020;746:141347.
- [34] Bailey NT, et al. The mathematical theory of infectious diseases and its applications. (2nd edition). Charles Griffin & Company Ltd 5a Crendon Street, High Wycombe, Bucks HP13 6LE.; 1975.
- [35] Mendel J, Hagrass H, Tan W-W, Melek WW, Ying H. Introduction to type-2 fuzzy logic control: theory and applications. John Wiley & Sons; 2014, p. 87.
- [36] Jana DK, Roy K, Dey S. Comparative assessment on lead removal using micellar-enhanced ultrafiltration (MEUF) based on a type-2 fuzzy logic and response surface methodology. *Separ Purif Technol* 2018;207:28–41.
- [37] Macêdo AM, Brum AA, Duarte-Filho GC, Almeida FA, Ospina R, Vasconcelos GL. A comparative analysis between a SIRD compartmental model and the richards growth model. 2020, MedRxiv.
- [38] Wu D. Twelve considerations in choosing between Gaussian and trapezoidal membership functions in interval type-2 fuzzy logic controllers. In: 2012 IEEE International conference on fuzzy systems. IEEE; 2012, p. 1–8.
- [39] Sadollah A. Introductory chapter: which membership function is appropriate in fuzzy system? In: Fuzzy logic based in optimization methods and control systems and its applications. IntechOpen; 2018.
- [40] Haghrah AA, Ghaemi S. PyIT2FLS: A new python toolkit for interval type 2 fuzzy logic systems. 2019, arXiv preprint arXiv:1909.10051.
- [41] Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. *Nature* 2020;585(7825):357–62.
- [42] Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* 2020;17(3):261–72.
- [43] Hunter JD. Matplotlib: A 2D graphics environment. *IEEE Ann Hist Comput* 2007;9(03):90–5.