Chaos, Solitons and Fractals 158 (2022) 112097

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Forecasting of a complex phenomenon using stochastic data-based techniques under non-conventional schemes: The SARS-CoV-2 virus spread case

Daniel E. Mendoza [a,b,*], Ana Ochoa-Sánchez [c,d], Esteban P. Samaniego [b,e]

[a] Department of Civil Engineering, University of Cuenca, Av. 12 de Abril sn, CP: 010112 Cuenca, Ecuador
[b] Faculty of Engineering, University of Cuenca, Av.12 de Abril sn, CP: 010112 Cuenca, Ecuador
[c] School of Environmental Engineering, Faculty of Science and Technology, University of Azuay, Cuenca, Ecuador
[d] TRACES, University of Azuay, Cuenca, Ecuador
[e] Department of Water Resources and Environmental Sciences, University of Cuenca, Av. 12 de Abril sn, CP: 010151 Cuenca, Ecuador

ARTICLE INFO

Article history:
Received 12 October 2021
Received in revised form 31 March 2022
Accepted 4 April 2022
Available online 7 April 2022

Keywords:
Autoregressive-with-exogenous-variables
Vector-autoregressive
Non-stationary
Differential-equations
Outbreaks recessions

ABSTRACT

Epidemics are complex dynamical processes that are difficult to model. As revealed by the SARS-CoV-2 pandemic, the social behavior and policy decisions contribute to the rapidly changing behavior of the virus' spread during outbreaks and recessions. In practice, reliable forecasting estimations are needed, especially during early contagion stages when knowledge and data are insipient. When stochastic models are used to address the problem, it is necessary to consider new modeling strategies. Such strategies should aim to predict the different contagious phases and fast changes between recessions and outbreaks. At the same time, it is desirable to take advantage of existing modeling frameworks, knowledge and tools. In that line, we take Autoregressive models with exogenous variables (ARX) and Vector autoregressive (VAR) techniques as a basis. We then consider analogies with epidemic's differential equations to define the structure of the models. To predict recessions and outbreaks, the possibility of updating the model's parameters and stochastic structures is considered, providing non-stationarity properties and flexibility for accommodating the incoming data to the models. The Generalized-Random-Walk (GRW) and the State-Dependent-Parameter (SDP) techniques shape the parameters' variability. The stochastic structures are identified following the Akaike (AIC) criterion. The models use the daily rates of infected, death, and healed individuals, which are the most common and accurate data retrieved in the early stages. Additionally, different experiments aim to explore the individual and complementary role of these variables. The results show that although both the ARX-based and VAR-based techniques have good statistical accuracy for seven-day ahead predictions, some ARX models can anticipate outbreaks and recessions. We argue that short-time predictions for complex problems could be attained through stochastic models that mimic the fundamentals of dynamic equations, updating their parameters and structures according to incoming data.

1. Introduction

Complex dynamical systems are commonly described by a set of differential equations containing the phenomenon's qualitative features [1]. However, while the processes-based models provide essential insights into the dynamics and non-linear interactions involved in complex problems, for some phenomena, the quantity and nature of such interacting factors could significantly complexify the phenomenon's evolution. We are especially interested in cases for which rapid variations might occur. This is likely to happen when the phenomenon is sensitive to highly volatile factors that are prone to fast changes in their states, which makes analysis and prediction challenging [2].

A phenomenon of the kind described above is the case of epidemics' spread. Systems of ordinary differential equations describing the evolution of contagious diseases have been available for a while and still have a great scientific interest [3–5]. The solutions of these deterministic equations provide information about the epidemic advance and its coarse-grained features [6]. An interesting example is provided by [7], who successfully applied a set of ordinary differential equations that consider the effect of asymptomatic patients over the SARS-CoV-2 virus spread [8,9]. However, the COVD-19 pandemic shows us that social behavior and policy decisions could cause serious discrepancies between deterministic models and observations (e.g., [10–18]).

* Corresponding author at: Department of Civil Engineering, University of Cuenca, Av. 12 de Abril sn, CP: 010112 Cuenca, Ecuador.
E-mail address: daniel.mendoza@ucuenca.edu.ec (D.E. Mendoza).

https://doi.org/10.1016/j.chaos.2022.112097
0960-0779/

The COVID-19 pandemic has involved severe human health, social and economic problems. Moreover, during the first months of the virus's spread, the consequences were catastrophic, collapsing the health systems in several places, partly due to the lack of predictive assessments [19,20]. Alternatives were presented for this complex task, such as, for example, the approach proposed by [21], which uses fractal theory combined with fuzzy logic for time series forecasting. Also, [22, 23] have performed spatial analysis and classifications of the contagious dynamics. However, in line with [24], the efforts should continue, and they should tackle short and long time series predictions, as well as spatial analysis to effectively and objectively inform decision-makers about policies and coordination strategies.

Apart from being complex, pandemics are of serious risk at their beginning, thus limiting the time for sophisticated applications or the development of deterministic theories. Therefore, it is desirable to use existing knowledge and modeling frameworks to deal with the problem.

Although originated in a different context, an example of an applied technique that uses an existing framework in an innovative fashion is the "Forecasting at scale" approach proposed by [25–28]. This technique is devised as a flexible tool working in a general mathematical structure, capable of identifying changing points during the training phase for inferring future curve inflections. Other alternatives are, for instance, the well-known Neural Networks, with interesting applications such as the one proposed by Melin et al. [23,29], who combined it with fuzzy logic to predict the virus spread in Mexico, or the approach presented by Namasudra et al. [30], who applied nonlinear neural networks.

As a motivation here, we apply the "Forecasting at scale" technique and the Autoregressive Neural Networks (NNA) for the forecasting of the number of infections variable (as indicated in Fig. 1) in COVID-19 pandemics. Both techniques are implemented as R-packages with the names of "nnetar" and "Prophet" [25,31]. We performed sequential forecasting tasks for several periods of seven days (i.e., seven-day ahead forecasting) using the COVID-19 data sets of Iraq and Iran, which show pronounced recessions and outbreaks in their curves, making them challenging for prediction. Each sequential forecasting was attained by calibrated models using the complete set of past observations relative to the forecasted period. This entails that both the Prophet and the ANN parameters are re-estimated for each forecasted period as more data were progressively integrated into the observations set for the training task. The forecasting results in (Fig. 1) evidences the difficulty in anticipating recessions and outbreaks.

Due to the complexity of predicting COVID-19 cases, the need for testing different stochastic strategies is suggested [32]. Nonetheless, considering the unusual statistical characteristics of a virus' spread, such as the fact that it follows fat-tail distributions that limit the prediction by probabilistic approaches [33], these stochastic models must scape the conventional applications as mentioned by Santosh [34]. In this context, we regard a strategy as unconventional if it relies, at least partly, on the possibility of combining existing knowledge and tools in an innovative manner. The continuous development of these strategies will facilitate fast predictive implementations in future pandemics. Nonetheless, two questions arise before devising such innovations: 1) How to deal with the complex changing conditions of the problem? and 2) When using stochastic approaches, which should be the underlying theory constraining the models?

Regarding the first question, it may be that fast fluctuations of the phenomenon (such as the triggering of outbreaks and sudden recessions) depend on some natural states of the contagious curves or depend significantly on the surrounding conditions in time. Thus, the parameters of any model for the forecasting task should be determined to a great degree by the information closer to the forecast time. For capturing the complex changing conditions, the parameters should update their values progressively. Moreover, the stochastic structures themselves should be updated. As a consequence, the simultaneous updating of both the parameters and stochastic structures should provide the modeling strategy with inherent non-stationary features and greater flexibility for accommodating the incoming data and training it for short-term forecasting tasks. Examples of techniques dealing with shaping the models' parameters in time, or making them dependent on states, are the well-known Generalized-random-walk (GRW) and the State-dependent parameters (SDP) theories.

Regarding the second question, we choose to frame the stochastic structures considering the governing ordinary differential equations of epidemics (in this case, the Susceptible-Infected-Recovered-Deceased SIRD equations). The main idea is to constrain these structures so that they are not entirely oblivious of conceptual or dynamical considerations. The aforementioned pursues, in some sense, the Data-based-modeling's (DBM) philosophy, which in many cases reached high forecasting performances, partly because they are linked to differential

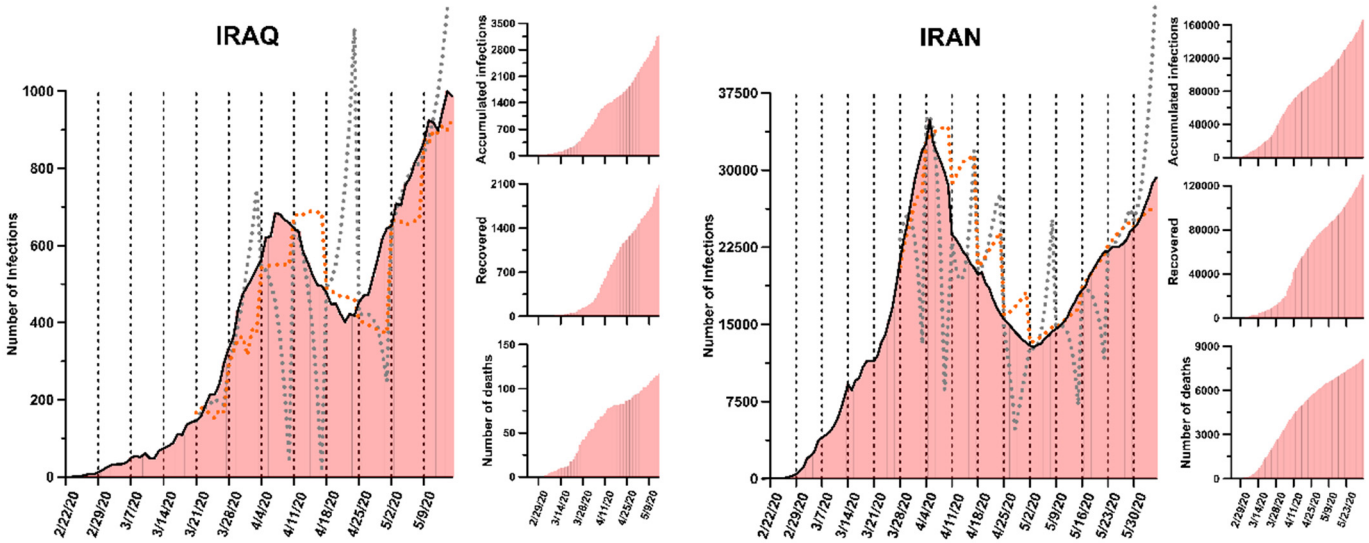

Fig. 1. SARS-CoV-2 contagious curves for Iraq and Iran. Dotted lines identify seven-days periods. Dashes lines are forecasted resulted. Orange for ANN and grey for Prophet model.

equations mechanisms (please refer to [35] book and the examples given therein).

Autoregressions with exogenous variables (ARX) and Vector autoregressions (VAR) are useful stochastic techniques because their structures could be devised to incorporate the form of the differential equations. More precisely, some the SIRD equations are considered to some extent in order to define the structure of the models proposed here [35–38]. Furthermore, the GRW and SDP techniques are merged with the ARX into the Recursive-time series theory, as detailed in Young [35], and implemented in the Matlab package CAPTAIN (which is freely available in http://wp.lancs.ac.uk/captaintoolbox/). It is worth mentioning that the CAPTAIN tool does not provide VAR models. However, under certain assumptions (detailed later in this document), the GRW can be merged with VAR techniques using the available CAPTAIN options.

In addition, some important considerations for the proposed methodology are the following:

1) Since the CAPTAIN tools enables different GRW options – such as Random-Walk (RW), Integrated-Random Walk (IRW) – each ARX-based and VAR-based model will be used considering several combinations of the GRW options.
2) The stochastic structures are analogous to the differential equations and they use the daily rate of deaths and healings as exogenous variables handling the number of infections curve in the ARX model. For the VAR model, the three variables are treated as endogenous variables.
3) We restrict the prediction period to a window of seven days in the future, motivated by the time-average period for incubation of the virus into the human body [39].
4) We hypothesize that the exogenous variables could be sensitive to spread's changes in the previous days (seven days), carrying with them implicit information of recessions and outbreaks in advance. To test this hypothesis and at the same time to unveil the role that different variables have on the problem, we conducted an experiment through the exploration of different model types, which are.
   a) ARX models considering one and two lagged exogenous variables.
   b) Different VAR models' combinations between the endogenous variables.

In summary, we propose a novel modeling strategy for the complex COVID-19 forecasting task based on existing tools and knowledge. The stochastic models are based on analogies to the epidemic's differential equations. Moreover, they simultaneously perform the updating of their parameters and structures for fitting the incoming data. In addition, we explore different ARX and VAR model configurations with different GRW techniques to unveil the usefulness and roles that the variables have on the anticipation of outbreaks and recessions. To the authors' knowledge, all this is a non-conventional strategy, which takes advantage and adapts the existing knowledge and tools to the problem of forecasting a complex phenomenon. This systematic modeling approach might support practical applications and decision making, especially when knowledge and data are insipient during the early contagion stages. Furthermore, the method could help forecast other complex problems.

2. Materials

2.1. Data

We used daily data retrieved by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (https://github.com/CSSEGISandData/COVID-19). From all the collected time-series information, our main interest lies in three different data sets:

a) Accumulated infections (I_a): the daily incremental number of confirmed diagnosed individuals with the virus
b) Accumulated recoveries (H) and deaths (D): the daily incremental number of healed and dead victims of the virus
c) Infection-curve (I): The infected individuals calculated by subtracting *H and D* from the I_a, for each *i-th* day.

Since the (I) data-set describes the change of infections through time (here referred to as the *epidemic-curve*), the modeling process will be focused on emulating this variable. The (H) and (D) data-sets will play the role of the exogenous variables in the ARX models and endogenous variables in the VAR techniques. The I_a information will be used for estimating non-linear functions in the modeling process, as explained in Section 3.4.

Data from Iraq, Iran (Fig. 1) show interesting dynamics in the contagious evolution, which are challenging to be emulated, offering the opportunity for evaluating different models' abilities. Fig. 1 shows the daily infection curves and accumulated infections, recoveries, and deaths of Iraq and Iran populations. Although in different scales in the number of affections, both curves present the typical growing phases and descending parts probably related to the first trigger on the number of infections and the posterior restrictive measures for alleviating the conditions. Interestingly, new outbreaks in both curves are present after the descending phase. These rapidly changing conditions make the forecasting's task a challenge, which motivates the experiments carried out in this work.

3. Methods

3.1. The SIRD model

The *Susceptible-Infected-Recovered-Deceased* (SIRD) model is frequently used for modeling epidemics [3–5]. The model follows the well-known *low-mass-action* principle for describing a dynamical interaction between the four different compartments in a system of *ordinary differential equations*.

$$\dot{S}(t) = -\beta\frac{I(t)S(t)}{N} \quad ; \quad S(0) = S_o \tag{1a}$$

$$\dot{I}(t) = \beta\frac{I(t)S(t)}{N} - \gamma I(t) - \mu I(t) \quad ; \quad I(0) = I_o \tag{1b}$$

$$\dot{H}(t) = \gamma I(t) \quad ; \quad H(0) = H_o \tag{1c}$$

$$\dot{D}(t) = \mu I(t) \quad ; \quad D(0) = D_o \tag{1d}$$

The (N) denotes the total population of individuals involved in the process, while (S), (I) (H), and (D) indicates the *susceptible*, *infected*, *healed*, and *dead* individuals with their respective initial conditions (S_o, I_o, H_o, D_o). The β, γ, and μ are the rates of *infection*, *healings*, and *deaths*, respectively estimated from data. The *healed* (H) and *deaths* (D) can be aggregated into a single variable$R(t) = H(t) + D(t)$; where (R) denotes the *removed* cases from the population.

The model assumes a constant balance in the flux of individuals through the compartments$S(t) + I(t) + R(t) = N$. The latter is proper for epidemics with faster time evolution than the balance between birth and death rates (*vital dynamics process*). Since the transmission speed of the SARS-CoV-2 virus is highly similar to its predecessors (SARS-CoV and MERS-CoV) [40,41], the *vital-dynamics* process could be omitted without significant detriment. The model also assumes that the *removed* cases (R) do not interact with the remaining susceptible (S) population. This property is directly related to the acquired immunity of recovered individuals (at least for a period larger than our short-term forecasting experiment), protecting them from new contagious and limiting the spread of the virus, as has been argued for the SARS-CoV-2 case [42].

The assumptions made by the SIRD justify the study of the SARS-CoV-2 contagious dynamics by analyzing the SIRD mathematical

properties and solutions [43–48]. However, rather than solving the set of differential equations directly, we use them as theoretical support. Thus, we build a restated mathematical expression, which is more convenient here, as detailed below.

When its equivalent balance replaces the component of susceptible individuals$S(t) = N - I(t) - R(t)$, and the Eqs. (1b), (1c), and (1d) are combined in a single mathematical form, we obtain the Eq. (2a).

$$I(t) = a(t)I(t) - R(t);\ \text{with}\ a(t) = \beta\{N - [I(t) + R(t)]\}/N \tag{2a}$$

This expression indicates that the rate of change of infections (I) in an instant (t) is proportional to the present quantity of infected individuals (I), minus the rate of change of the total removed cases ($\dot{R}$). It is worth noticing that (a) is considered a non-stationary parameter, synthesizing the β, I, R, and N variability.

Eq. (2a) can be reformulated in a transfer function form by introducing the differential operator (s) and the auxiliary variable (b) (see Eq. (2b)).

$$I(t) = -\frac{b(t)}{1 + b(t)s}\dot{R}(t);\ \text{with}\ b(t) = -\frac{1}{a(t)} \tag{2b}$$

Furthermore, under certain assumptions, the previous continuous expression can be approximated following a discrete transfer function, which is analog to the continuous transfer function, as shown by Eq. (2c).

$$I(k) = \frac{B(z^{-1})}{A(z^{-1})}\dot{R}(k-\tau_r) + \frac{C(z^{-1})}{A(z^{-1})}e(k);\ e(k)\tilde{N}(0,\sigma^2) \tag{2c}$$

For the sake of brevity, we have focused on the most critical aspects of the theory supporting our work. However, for more details about the analogies and relations between the continuous and discrete forms, the reader should refer to the book of [35]. Here, it is sufficient to mention that Eq. (2c) is the *autoregressive moving average* with *exogenous variables* (AMARX) model. For simplicity, we will use the related *autoregressive with exogenous variables* (ARX) model.

3.2. The ARX approach

We degenerate Eq. (2c) to the expression (3a), also known as autoregressive with exogenous variables (ARX) model.

$$I(k) = \frac{B(z^{-1})}{A(z^{-1})}\dot{R}(k-\tau_r) + \frac{1}{A(z^{-1})}e(k);\ e(k)\tilde{N}(0,\sigma^2) \tag{3a}$$

Such that:

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ldots\ldots a_n z^{-n}$$
$$B(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \ldots\ldots b_m z^{-m}$$

In Eq. (3a), ($\dot{R}$) denotes the daily rate of change of the *removed* individuals (*exogenous variable*), which is approximated by $\dot{R} = R(k) - R(k-1)$. The (e) indicates a *white noise* and *zero-mean* error in the *k-th* daily time step. The (τ_r) is a pure time delay parameter acting on the *exogenous-variable*, which we established according to the restricted minimum period of *seven days* for allowing forecasting tasks. The $A(z^{-1})$ and $B(z^{-1})$ denotes the transfer function's polynomials of the *backward shift operator* (z), which build autoregressive structures of (I), plus regressive forms of ($\dot{R}$) information.

Eq. (3a) supports some of the modeling hypotheses to be tested here. For instance, the parameters n and m, which define the regressive structures, must be identified according to some statistical criteria. Commonly, the identification process uses the total available data, hopefully deriving a suitable model structure that generalizes the main aspects of the phenomenon. However, as mentioned previously, attempting to capture the complex dynamic elements of the epidemic in the early stages, the identification process should evolve according to the updated data retrieved by the monitoring systems.

Eqs. (3a) and (3b) also suggest exploring how sensitive are the exogenous variables ($\dot{H}$) and ($\dot{D}$) in leading the model and how they interact with it. For that purpose, the experimentation must consider the following modeling options. 1) Pure auto-regressive (AR) models, considering the exogenous variable $R = I$. 2) An ARX_H model, assuming the exogenous component as $R = H$. 3) An ARX_D model assuming that $R = D$. 4) An ARX_{H+D} model including the total removed cases $R = H + D$. 5) And finally, an $ARX_{H\&D}$ model considering two independent exogenous variables into it, shown in Eq. (3b).

$$I(k) = \frac{B(z^{-1})}{A(z^{-1})}\dot{H}(k-\tau_h) + \frac{C(z^{-1})}{A(z^{-1})}\dot{D}(k-\tau_d) + \frac{1}{A(z^{-1})}e(k);\ e(k)\tilde{N}(0,\sigma^2) \tag{3b}$$

Such that:

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ldots\ldots a_n z^{-n}$$
$$B(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \ldots\ldots b_m z^{-m}$$
$$C(z^{-1}) = c_0 + c_1 z^{-1} + c_2 z^{-2} + \ldots\ldots c_p z^{-p}$$

As noticed, Eq. (3b) has two exogenous variables. The $B(z^{-1})$ and $C(z^{-1})$ are polynomials acting on ($\dot{H}$) and ($\dot{D}$), while (τ_h) and (τ_d) are their pure time delay parameters. Here, three parameters define the model's structure ($n.m.p$), while (e) denotes a *white noise* and *zero mean* error in the *k-th* daily time step again. The discrete $ARX_{H\&D}$ transfer function does not have a transfer function analogy to the continuous SIRD equations (i.e., not in the sense that ARX of Eq. (3a) has with Eq. (2b). Nonetheless, intending to test whether such a following analogical meaning conceptualized herein has predictive advantages over other alternatives, Eq. (3b) is an interesting contrasting experiment.

It is essential to mention that since parameter (a) in Eq. (2b) is *non-stationary* – because it synthesizes great variability of several factors – this *non-stationarity* should be inserted in the discrete ARX approximations. In that respect, the *Generalized-Random-Walk* (GRW) and the *State-dependent-parameter* (SDP) are adequate complementary techniques for tackling the *non-stationarity* problem. These GRW and SDP techniques are described briefly below using model (3b) ($ARX_{H\&D}$) as a vehicle for explaining in a more general sense how these techniques can deal with the non-stationarity of the ARX's parameters. From the latter, it is only a matter of adding or removing the corresponding elements for the rest of the ARX models,

3.3. The GRW technique for time variable parameters (TVP)

The backward shift operator in Eq. (3b) is lifted, obtaining in this way the expanded form of the $ARX_{H\&D}$ model (see Eq. (4a))

$$I(k) = \mathbf{z}^T(k)\boldsymbol{\rho}(k) + e(k) \tag{4a}$$

where:

$$\mathbf{z}^T(k) = [-I(k-1) - I(k-2).... - I(k-n)\ \dot{H}(k-\tau_h)\ \dot{H}(k-\tau_h-1)\ \dot{H}(k-\tau_h-2)....$$
$$\dot{H}(k-\tau_h-m)\ \dot{D}(k-\tau_d)\ \dot{D}(k-\tau_d-1)\ \dot{D}(k-\tau_d-2)....\dot{D}(k-\tau_d-p)]$$

$$\boldsymbol{\rho}^T(k) = \begin{bmatrix} a_1\ a_2....a_n\ b\ b_0\ b_1\ b_2....b_m\ c_0\ c_1\ c_2....c_p \end{bmatrix} = \begin{bmatrix} \rho_1\ \rho_2\ \rho_3 \cdots \rho_{n+m+p} \end{bmatrix}$$

The elements in ($\mathbf{z}$) are the daily historical observations of the variable I, $\dot{H}$, and$\dot{D}$, while the elements in ($\boldsymbol{\rho}$) vector are the parameters in the linear $ARX_{H\&D}$ model, whose notation is conveniently redefined to

follow an ordinal sequence (ρ_i). In turn, each (ρ_i) parameter is determined by a two dimensional state-vector $\boldsymbol{\rho}_i(k) = [l_i(k)\ d_i(k)]^T$, which follows a *Generalized-Random-Walk* (GRW) process (see eq. 4b) defining their non-stationarity, as follows.

$$\boldsymbol{\rho}_i(k) = \mathbf{A}\boldsymbol{\rho}_i(k-1) + \mathbf{D}\boldsymbol{\eta}_i(k-1) \tag{4b}$$

where:

$$\mathbf{A}_i = \begin{bmatrix} \alpha_i & \beta_i \\ 0 & \gamma_i \end{bmatrix} \ ; \ \mathbf{D}_i = \begin{bmatrix} \delta_i & 0 \\ 0 & \varepsilon_i \end{bmatrix}$$

In the Eq. (4b), each $\boldsymbol{\eta}_i(k) = [\eta_{1,i}(k)\ \eta_{2,i}(k)]^T$ is also a two dimensional zero mean vector, with their noise defined by a diagonal covariance matrix $\mathbf{Q}_{\eta,i}$. Each $\alpha, \beta, \gamma, \delta, \varepsilon$, as well as the elements of $\mathbf{Q}_{\eta,i}$ (or usually the NVR matrix $\mathbf{Q}_{nvr} = \mathbf{Q}_{\eta,i} / \sigma^2$) are the so-called *hyperparameters*, which are *time-invariant*, and are estimated from data through optimization. Such optimization uses likelihood functions based on the innovations obtained from the Kalman-Filter process. Once the *hyperparameters* are optimized, the recursive Kalman-Filter and the Fixed-Interval-Smoothing (KF/FIS) technique gives the desired time-variability property to each parameter in the model. For more details about these techniques, the reader should refer to specialized literature [35].

The GRW process includes several random walk types, from which the most important are: the *random walk* (RW), the integrated random walk (IRW), and smoothed random walk (SRW). The ARX model with constant parameters, the RW and the IRW process, can be set out in the MATLAB-CAPTAIN toolbox, which is available at http://wp.lancs.ac.uk/captaintoolbox/ [49]. Actually, for the reasons explained in the technical notes below, we experiment with different modeling approaches combining the available GRW options to evaluate their effects on the phenomenon.

3.4. The SDP technique

Briefly, the *State-Dependent-Parameter* technique redefines the output variable (I) of Eq. (4a) to make it dependent on some *state-variable* (χ). The non-parametric functions for each parameter (ρ_i), respecting to (χ), can be estimated by the KF/FIS algorithm and its GRW techniques, which are finally parameterized through curve-fitting procedures. Furthermore, the algorithm can be extended for considering different state-variables (χ_i) through the well-known *Back-fitting algorithm* and the *Modified-Dependent-Variables* (MDVs). This work does not provide a detailed explanation of this technique, but the reader could refer to [35,50] for the extended minutia. In summary, the model is defined as follows:

$$I(k) = \mathbf{z}^T(k)\boldsymbol{\rho}(k) + e(k) \tag{5a}$$

$$\boldsymbol{\rho}(k) = \left[\rho_1\{\chi_1(k)\}\ \rho_2\{\chi_2(k)\}.\ldots\rho_{n+m+p}\left\{\chi_{n+m+p}(k)\right\}\right]^T \tag{5b}$$

We conveniently rewrite the linear model (4a) in (5a) for coupling it with the expression (5b), which denotes a vector containing the collection of the parameters (ρ_i). In turn, each element (ρ_i) is a function of a (χ_i) state variable, which is defined in *k-th* time step. Of course, the observation vector ($\mathbf{z}$) is built from autoregressive information of the (I) epidemic curve, plus regressive structures of ($\dot{H}$) and ($\dot{D}$) daily rates observations, of orders *n, m,* and *p*, respectively.

Regarding the (χ) *state variables*, the accumulated infections (I_a) is setting here as the independent variable defining the functions for all the parameters of the autoregressive components (i.e., the parameters $\rho_{1,2,..n}$). The (I_a) is equivalent to ($I + R$) expression, whose variability is included in the parameter (a) of Eq. (2a). Because of the latter, it is a logical assumption to make each parameter in the *A*-polynomial (see Eqs. (3a) and (3b)) dependent on the I_a variable state. Nonetheless, it is worth mentioning that we are implicitly assuming that the variability of I_a absorbs the non-stationarity of (β) in the SIR model, and the population (N) remains constant for simplicity.

Extending the modeling experimentation, we will consider that the set of parameters of both exogenous variables ($\dot{D}$) and ($\dot{H}$) (see the parameters of B and C polynomials in Eq. (3a) and (3b) will act as state variables for themselves. This experimentation is motivated by the possibility that abrupt time changes related to outbreaks or recessions might be more sensible to be detected in the ($\dot{H}$) and ($\dot{D}$) *daily-rates* fast fluctuations. The latter hypothesis was also one of the reasons for including such a *daily-rates* as exogenous variables instead of using the daily soft evolutions of (H) and (D) data.

The SDP is also freely available in the MATLAB-CAPTAIN toolbox at http://wp.lancs.ac.uk/captaintoolbox/ [49]. Therefore, similar to the ARX modeling, we will use the mentioned toolbox for a comprehensive modeling experiment considering different combinations between constant, RW and IRW options, which mainly define the smoothing shape process of the non-parametric functions estimated by the *Back-fitting* algorithm.

The estimated SDP functions' parameterization is usually attained through fitting curve techniques with meaningful physical or conceptual forms. Nonetheless, for the sake of simplicity, we use linear and quadratic interpolations between two and three consecutive points of the non-parametric functions as surrogate parameterizations. Suppose any state variable value $\chi(k)$ falls outside the non-parametric function's domain in the prediction phase. In that case, we use linear and quadratic extrapolations using seven data points from the extremes of the curve. Similar options have been used for other purposes [51].

3.5. The VAR technique

One main limitation of the ARX model is the unidirectional relation between the exogenous variables (D) and (H) and the output variable (I). However, as described by the SIRD equations, the phenomenon's behavior is an interaction between the different compartments, and they have a simultaneous influence on each other in their evolutive processes. In that regard, the autoregressive vector framework can consider such feedback between these variables, treating each one as *endogenous*. In other words, a Vector-Autoregressive (VAR) technique comprises one equation per variable into the system been described [52].

Consider the set of Eqs. (2b), (1c), and (1d). Thereon, following a similar argument as explained in Section 3.1, and discretizing ($\dot{D}$) and ($\dot{H}$), one can arrive at a set of ARX discrete analogical approximations as expressed by Eqs. (6a)–(6c) in the expanded form.

$$I(k) = \mathbf{z}^T(k)\boldsymbol{\rho}(k) + e(k) \tag{6a}$$

$$H(k) = \boldsymbol{\varphi}^T(k)\boldsymbol{\mu}(k) + e_1(k) \tag{6b}$$

$$D(k) = \boldsymbol{\psi}^T(k)\boldsymbol{\xi}(k) + e_2(k) \tag{6c}$$

where

$$\mathbf{z}^T(k) = \left[-I(k-1)-I(k-2).\ldots-I(k-n)\ \dot{R}(k-\tau_r)\ \dot{R}(k-\tau_r-1)\ \dot{R}(k-\tau_r-2).\ldots\dot{R}(k-\tau_r-m)\right]$$

$$\boldsymbol{\varphi}^T(k) = \left[H(k-1)\ I(k-\tau_{I_h})\ I(k-\tau_{I_h}-1)\ I(k-\tau_{I_h}-2).\ldots I(k-\tau_{I_h}-q)\right]$$

$$\boldsymbol{\psi}^T(k) = [D(k-1)\ I(k-\tau_{I_d})\ I(k-\tau_{I_d}-1)\ I(k-\tau_{I_d}-2).\ldots I(k-\tau_{I_d}-s)]$$

$$\boldsymbol{\rho}^T(k) = \left[\rho_1\ \rho_2\ \rho_3\ldots\rho_{n+m}\right]$$

$$\boldsymbol{\mu}^T(k) = \left[1\ \mu_1\ \mu_2\ldots\mu_q\right]$$

$$\xi^T(k) = [1\xi_1\xi_2\ldots\xi_S]$$

The reader is already familiarized with the notation and meaning of Eqs. (6a)–(6c). Nonetheless, it is worth mentioning that the (e), (e_1),

and (e_2) errors are considered white noise processes that may be *contemporaneously* correlated (i.e., correlated just for the *k-th* time step). In the context considered herein, the discrete Eq. (6a) and its similarity with the continuous form 2b was already justified. Also, the regressive structures of the (I) variable in Eqs. (6a) and (6b) surrogate the time-variability nature of (γ) and (μ) parameters in the differential Eqs. (1c) and (1d).

The set of Eqs. (6a)–(6c) interact with each other, allowing recursive predictions. Although the models have no restrictions about the future prediction's horizons, we have limited these to a seven days ahead horizon for comparison purposes with the remaining modeling alternatives.

Eq. (6a) imposes a minimum time delay (τ_r) of one month to ensure the first *k-th* prediction of (I) as a function of$\hat{R}$. After that, (H) and (D) can be estimated through Eqs. (6b) and (6c), respectively. In turn, the last *H and D* estimations are used for a new estimation of (I), continuing in this manner until the desired prediction future step.

For the case herein, we implemented these models (6a) to (6c_ using the TVP concept. Thus, the parameters in the ($\boldsymbol{\rho}$), ($\boldsymbol{\mu}$), and ($\boldsymbol{\xi}$) vectors follow the GRW process (as explained in Section 3.3). The latter modeling perspective is not directly available in the *MATLAB-CAPTAIN toolbox*. However, we implemented an algorithm for including more than one endogenous variable working recursively between them, using the same computational framework.

Since we use the CAPTAIN framework for our purposes, the *hyperparameter* optimization is reached individually for each equation in the system through their error innovations obtained from the *Kalman-Filter* process. The latter rules out the possibility of any *cointegrated* process herein, which is a more complex modeling process that escapes the scope of this work.

Reminding that $R(t) = H(t) + D(t)$, and following one of the objectives herein, we explore the role of each variable inside the modeling through the next experiments: 1) A model assuming that $R = H$ (and renaming τ_r by τ_h), excluding the Eq. (6c) (labeled as VAR_H). 2) A model in which $R = D$ (and renaming τ_r by τ_d), excluding the Eq. (6b) (labeled as VAR_D). 3) A model in which $R = H + D$, excluding Eq. (6c), and considering Eq. (6b) as R (labeled as VAR_{H+D}). 4) A model including Eqs. (6a), (6b), and (6c) simultaneously (labeled as $VAR_{H\&D}$). And finally, 5) an extended model option in which for the vector ($\mathbf{z}$) of Eq. (6a), the variable (R) is separated on individual components of (H) and (D), and the ($\boldsymbol{\varphi}$) and ($\boldsymbol{\psi}$) vectors of Eqs. (6b) and (6c) include historical information of the endogenous variables. The latter model is labeled as $VAR_{H\&D\text{-extended}}$, with specific the vectors of Eqs. (6a), (6b), and (6c) in the following forms:

$$\mathbf{z}^T(k) = [-I(k-1)-I(k-2)\ldots.-I(k-n)\ \dot{H}(k-\tau_h)\ \dot{H}(k-\tau_h-1)\ \dot{H}(k-\tau_h-2)\ldots.$$
$$\dot{H}(k-\tau_h-m)\ \dot{D}(k-\tau_d)\ \dot{D}(k-\tau_d-1)\ \dot{D}(k-\tau_d-2)\ldots.\dot{D}(k-\tau_d-p)]$$

$$\boldsymbol{\varphi}^T(k) = [-H(k-1)-H(k-2)\ldots.-H(k-w)\ I(k-\tau_{l_h})\ I(k-\tau_{l_h}-1)\ I(k-\tau_{l_h}-2)\ldots.I(k-\tau_{l_h}-q)]\ \boldsymbol{\psi}^T(k)$$
$$= [-D(k-1)-D(k-2)\ldots.-D(k-u)\ I(k-\tau_{l_d})\ I(k-\tau_{l_d}-1)\ I(k-\tau_{l_d}-2)\ldots.I(k-\tau_{l_d}-s)]$$

$VAR_{H\&D\text{-extended}}$ is not justified as analogical forms related to the differential equations. Thus, this model contrasts the results against those VAR models defined here as analog to the differential equations.

3.6. Modeling process, considerations and technical notes

So far, we have outlined different techniques, which will be the vehicle for our purposes. Nonetheless, it is necessary to precise the modeling process, as well as some important technical aspects.

For instance, since the problem is complex, we implemented an adaptive modeling strategy, in which the structure of the models (i.e., *m, n, p, q, w, r, s, and* τ, depending on ARX or VAR) changes for best fitting to the historical and new incoming data. As mentioned, this adaptive strategy attempts to emulate the natural conditions of the epidemic evolution. This emulation practice helps evaluate the potential of the models in supporting the decisions through continuous monitoring systems, using the most common data retrieved in the early stages of the epidemic phenomenon.

The first structures of ARX and VAR models were identified using 21 data observations to avoid computational errors. After, the models' structures were updated according to the inclusion of every *seven data* observations sequentially incorporated. The process is repeated until reaching the final *seven-day* string. Each forecasted *seven-day* string was attained using the corresponding models' structures identified previously to the predicted period. In the end, periodical predictions every seven days were obtained and contrasted whit its observations.

Although the central core of the modeling process was summarized in the previous lines, it is essential to mention some other important aspects, which we pointed out in the following paragraphs.

- For the ARX and VAR techniques, the identification of the models' structures was based on the statistical selection from a wide variety of combinations between *m. n, p, q, w, r, s, and* τ parameters, building a model that better fitted the observed data. We use the well-known *Akaike-Information-Criterion* (AIC) as the statistical parameter for selecting the most convenient structure [53,54]. To avoid the explosion of combinations in the identification task, we impose a maximum bound of seven steps for *m. n, p, q, w, r, s, and* τ for the model's parameters. Of course, the latter bound for the ARX models was established after lagging the data for a minimum of seven days, allowing the seven-step-ahead forecasting using historical observations.
- While the forecasting estimations were attained by lagging the exogenous variables a minimum of seven days in the ARX modeling, the forecasted estimations using VAR techniques are attained by integrating sequentially one-step-ahead forecasted values as new inputs. The latter ARX's and VAR's forecasting strategies are known as “DirRec” and “Recursive,” respectively [55]. The latter constitutes an interesting contrasting experiment for evaluating the forecasting estimations using only observations as inputs (in the ARX's models), versus the forecasting using predicted values surrogating observed data (as in VAR's models).
- It is recalled that only the GRW technique was used for modeling the parameters' non-stationarity for the VAR models, while for the ARX kind of models, both GRW and SDP techniques were tried. In all the cases, we explore the modeling of the parameters considering three different combinations: RW, IRW and *stationarity* (SP) (i.e., constant parameter). For the sake of simplicity, we built the combinations in such a way that the same GRW option is established for the entire set of parameters belonging to the same polynomial (i.e., *A*, *B*, or *C* polynomial depending if ARX or VAR model kind).
- The idea of combining the available GRW options for modeling the parameters' non-stationarity is motivated by the fact that a user in practice could not know the best approach a priori in that regard. Therefore, it would be reasonable to test several GRW combinations for the parameters on each of the five ARX and VAR experiments (detailed Sections 3.2 and 3.5). Since different GRW render different forecasting estimations, we select the *50-quantile* (*median*) to synthesize the models' predictive capability (*predictive accuracy*). Additionally, the 25 and 75 quantiles are considered indicators of the variability between the different predictions attained under different GRW combinations. A comparison against the model with the best-fitted approach through all the predicted strings was made to validate the median as a reasonable indicator of the predictive capacity. In the authors' opinion, the adopted modeling strategies, such as the adaptive model's structures and combinations of different GRW and SDP techniques, constitute an unconventional endeavor to objectively tackle such a phenomenon's predictive complexity.

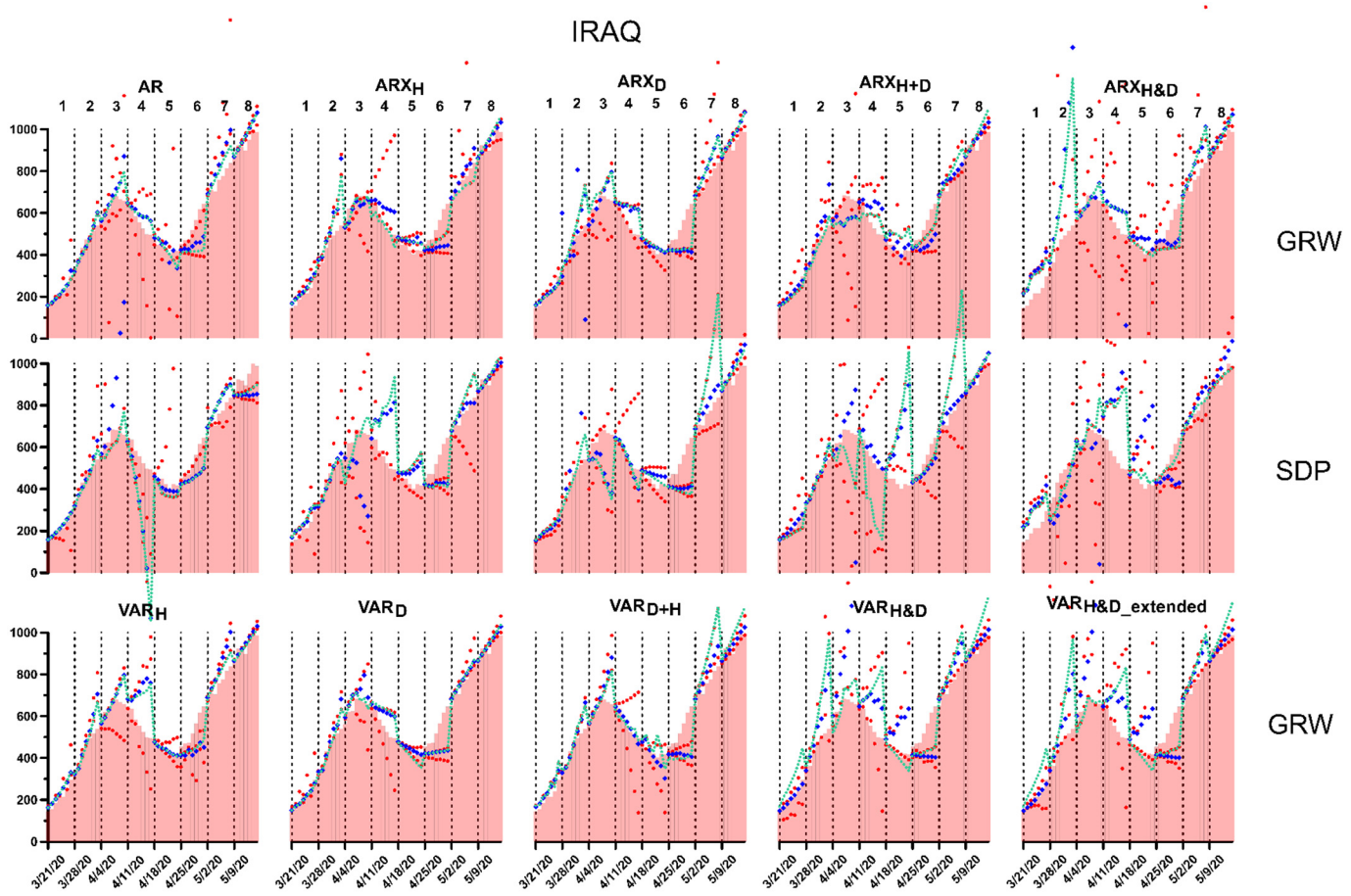

Fig. 2. Seven-days forecasting of Iraq's infection curve. Vertical dotted lines delineate data strings groups of seven days. Blue-dots indicate the *median* estimations calculated from the forecasted combination using different GRW options. Green-lines are the best fitted models for all the strings, following a specific GRW process. Red-dots indicates the 25% percentile and 75% percentile values from all forecasted values.

4. Results and discussion

To analyze the forecasting abilities on different curve phases of the contagious curve, we segmented the data into several strings of seven values each and numbered them, as shown in Figs. 2 and 3. For Iraq (Fig. 2), strings 1 and 2 correspond to the initial growing phase of the epidemics. String 3 contains a maximum, followed by a decrement in the number of cases as shown by string 4. String 5 is the segment where an inflection point occurs before a new outbreak arises in string 6. String 7 and 8 are the growing evolution of the contagious curve of the second outbreak.

For Iran (Fig. 3), string 1 corresponds to the initial growing phase of the epidemics. After a sharp peak, the curve experiments a decrement in the number of cases during the strings 2, 3, and 4. String 5 shows a slight inflection change in the curve, followed by a new increment in the number of cases that continues for the strings 6, 7, 8, 9, and 10. Interestingly, strings 8 and 9 show a slow-down in the contagious during the second increment, accelerating in string 10.

The results referent to the different string data portions are discussed in the following paragraphs for both study cases, Iraq's and Iran's virus spread.

4.1. Predictions for Iraq's case

- As indicated by Table 1a and Table 1b, when using the GRW techniques in both the ARX and VAR models, the median's prediction curve (calculated from the models using different GRW combinations) is statistically similar to the best predictions attained under a specific GRW with the best performance. Since in real-time, modelers would not know an optimum GRW option a priori, the median of all the forecasted values could be considered as a reliable predictive curve. The latter is an alternative against the blindness regarding an optimum GRW combination for the ARX and VAR predictions.
- Table 1a indicates that pure autoregressive (AR) models using GRW techniques are statistically more accurate than those using external variables for the ARX kind. The AR median's prediction curve captures the initial growing epidemics, ameliorating the overshooting effect of the peak (strings 1, 2, 3, and 4 in Fig. 2). More interesting, after the descendent part (strings 4 and 5), the AR model seems to be able to capture the second outbreak that suddenly appears on the curve (string 6) and continuing its growth (strings 7 and 8). The latter supports the hypothesis about the implicit information in the historical observations and their potentials in supporting predictive tasks of complex dynamics features when used as input variables in a model.
- Although not with the best performance, the ARX_{H+D} modeling with GRW techniques seems to shape the peak (strings 2, 3, and 4) and the second outbreak of the curve (string 6), as shown by Fig. 2. According to ARX_H and ARX_D in Fig. 2, the role played individually by the daily rates of H and D variables for representing the latter effects are very similar. Interestingly, the variability of the ARX_D's predictions – represented by the percentile 25 and 75 in Fig. 2 –is narrower than the alternative ARX models. The latter could be related to the identification process (Appendix 1_Table a), in which the

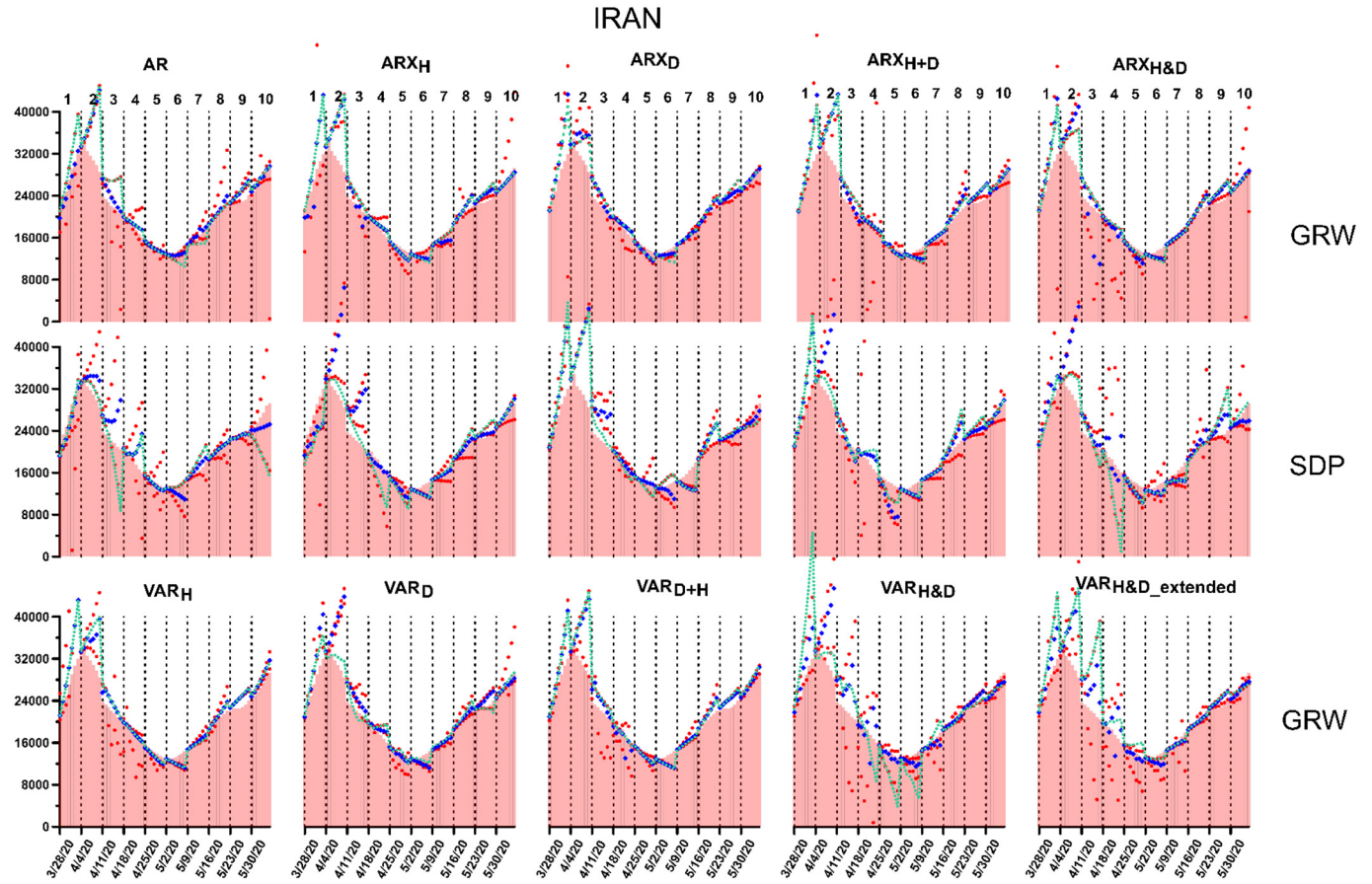

Fig. 3. Seven-days forecasting of Iran's infection curve. Vertical dotted lines indicate data strings groups of seven days. Blue-dots indicate the *median* estimations calculated from the forecasted combination using different GRW options. Green-lines are the best fitted models for all the strings, following a specific GRW process. Red-dots indicates the 25% percentile and 75% percentile values from all forecasted values.

structures' evolution reached by the ARX_D are statistically efficient, especially when included the critical curvatures (i.e., strings 5, 6, and 7) sequentially. Because the number of deaths during an epidemic is probably the most reliable register, the fact that the models' predictivity respond to this variable could be advantageous.

- Splitting the exogenous daily rates for constructing the $ARX_{H\&D}$ does not increase the critical strings' predictive performance in the contagious curve (i.e., strings 3 and 6). The previous is interesting since such a forecasting accuracy is opposite to the fitting in the identification process, reaching efficient values during both, R^2 and AIC parameters (Appendix 1_Table b). Since the fitting is accurate (as shown by R^2 and AIC, respectively), the forecasting failure could be more related to the transfer functions not being in a direct analogy with the continuous forms derived from the differential equations. Because of the latter, the model is over-parameterized.
- It is worth mentioning that although the ARX models using the SDP technique do not attain the best forecasting effects (Table 1a), these show an interesting predictive accuracy during and posterior the second outbreak (i.e., for strings 6, 7, and 8), especially for the AR and for the ARX_{H+D} models. The SDP technique can unravel the implicit information carried by the historical and the exogenous variables, whose non-parametric functions allow the dynamic curve to capture the non-linear changes. Nonetheless, since the ability of SDP relies on the number of observations, it might be that for early stages the models have enough observations to more widely identify the critical states that cause severe changes. Therefore, the model should be tested in later contagious phases, but this is an open future task.

Table 1b
Coefficient of determination (R^2) for Iraq (VAR models).

Parameter's modeling	VAR_H	VAR_D	VAR_{H+D}	$VAR_{H\&D}$	$VAR_{H\&D\text{-}extended}$
GRW-TVP	0.862 (0.888)	0.927 (0.931)	0.911 (0.88)	0.701 (0.772)	0.702(0.763)

The values outside and inside the brackets contain the R2 coefficient for the median and the best model (using a specific GRW combination as specified in Appendix 1_Tables c, d, e), estimated and chosen from all the GRW combinations, respectively.

Table 1a
Coefficient of determination (R^2) for Iraq (ARX models).

Parameter's modeling	AR	ARX_H	ARX_D	ARX_{H+D}	$ARX_{H\&D}$
GRW-TVP	0.934(0.929)	0.888(0.926)	0.910(0.911)	0.906(0.932)	0.608(0.683)
SDP	0.605(0.724)	0.696(0.769)	0.224(0.765)	0.804(0.598)	0.573(0.781)

The values outside and inside the brackets contain the R2 coefficient for the median and the best model (using a specific GRW combination as specified in Appendix 1_Tables a, b), estimated and chosen from all the GRW combinations, respectively.

- When daily rates are separated for constructing the $ARX_{H\&D}$ the predictive capacity does not improve in any sense. The latter supports the hypothesis that stochastic approximations must be analogous to the differential equations describing the phenomenon, regardless of the technique behind the parameter modeling (i.e., GRW or SDP). This could be because ARX stochastic structures following analogies with the differential equations provide the models with conceptual robustness, implicitly helping with the problem of over-parameterization and, consequently, the over-fitting problem.

- On the other hand, as shown in Table 1b, VAR models have similar performances to those attained by the ARX models using GRW techniques. The best model performance is reached when the interacting components are given between the contagious variable (I) and the daily rates of deaths (D), determining the VAR_D model. Also, as indicated by Fig. 2, the prediction's variability using different GRW combinations is narrower for the same VAR_D model than its alternatives. As in ARX models, the low prediction variability could be related to the high statistical fitting and efficiency reached during the identification process (Appendix 1_Table c). The mentioned is notorious even when the critical strings with pronounced curvatures (i.e., strings 5, 6, and 7) are included in the identification process. Since variables (I) and (D) are endogenous in the VAR models, they can be simultaneously predicted, as shown in Appendix 3. This simultaneous prediction capacity is in the best interest of health systems during an epidemic process.
- Fig. 2 reveals that the peak and the second outbreak curvatures (i.e., strings 3, 4, 5, and 6) are captured by VAR_H, VAR_D, and VAR_{H+D}, but those predictions are not as efficient as the ones attained by the ARX models. Although the VAR_D model shows better statistics than its relative VAR models, both exogenous daily rate (H) seems to play an essential role in predicting these complex peaks and outbreaks as shown by the model VAR_H in Fig. 2. Furthermore, other important features such as the descendent section of the contagious curve (strings 4 and 5) are well captured by the VAR_{H+D} model, which is a notorious ability compared with the remaining VAR options. Interestingly, the fitting statistics of VAR_{H+D} structures are lower than those reached by the VAR_H and VAR_D models (Appendix 1_Table d).
- Making the daily rates H and D endogenous variables interacting separately with the contagious (I) variable as specified by $VAR_{H\&D}$ model (see the extended model in Section 3.5) do not increase the predictive capacity (Table 1b and Fig. 2). Moreover, the inclusion of more historical information into each endogenous variable in the $VAR_{H\&D\text{-}extended}$ alternative does not improve any predictive capacity either, despite its good fitting performance (R^2) and efficiency (AIC) during the identification process (Appendix 1_Table e). The latter and the ARX results indicate that better predictions are attained when the model's structures follow similar analogies to the differential equations describing the phenomenon. These forecasting results obey to the overall strategy proposed here, which is sympathetic with the ARX and VAR models. In other words, for the short-term forecasting problem, the ARX and VAR techniques that follow differential equations on their structures, are capable to anticipate rapid curve changes if the parameters (following GRW models) and stochastic structures are updating at the same time.

Table 2b
Coefficient of determination (R^2) for Iran (VAR models).

Parameter's modeling	VAR_H	VAR_D	VAR_{H+D}	$VAR_{H\&D}$	$VAR_{H\&D\text{-}extended}$
GRW-TVP	0.947 (0.942)	0.906 (0.955)	0.916 (0.917)	0.854 (0.803)	0.863(0.76)

The values outside and inside the brackets contains the R2 coefficient for the median and the best model (using a specific GRW combination as specified in Appendix 1_Tables h, i, j), estimated and chosen from al the GRW combinations, respectively.

4.2. Predictions for Iran's case

- Table 2a and Table 2b indicates that the medians attained from the GRW combinations are statistically similar to the best predictions obtained by a specific GRW with the best performance. These results support the stated for the Iraq case about the median predictive curve as a reliable indicator of the model's predictive capacity, surrogating the predictions attained by an optimum model.
- Table 2a reveals an ARX_D model with the best predictions compared to ARX_H and ARX_{H+D}. Fig. 3 graphically confirms the ARX_D's predictive ability in all the forecasting strings. This predictive ability is linked to the fitted statistics attained during the training process (Appendix 2_Table f). This ARX_D model captures the particular curve dynamics, such as the slope changes in the curve from the strings 5 to 6 (Fig. 3). Also, the prediction variability from the different GRW combinations is narrower than the alternative ARX models, which is the same report for the Iraq case.
- Although not with the best prediction accuracy and fitted performance (Table 2a and Appendix 2_Table f, respectively), the AR model reveals abilities to capture the pronounced curvatures in the strings 5, 6, and 7 shown in Fig. 3. The latter ability of the AR model for capturing the curvature changes is even better than the model ARX_D, which confirms that historical data of the contagious curve encodes useful information for predicting new breakouts during the epidemic evolution.
- The statistics shown in Table 2a for the ARX models using SDP techniques are more efficient than those estimated for the Iraq case (Table 2a). Still, Fig. 3 shows serious inaccuracies, especially on strings 1 to 6. However, interesting results are shown for the strings 6 to 10, especially for the ARX_H and ARX_D models. The latter supports the argument derived from the Iraq predictions that although SDP is not helpful in the early stages of the phenomenon, it can be beneficial when the number of observations increases in future phases. The latter has the potential for predicting the strong non-linear behavior in the curves.
- As in Iraq's case, splitting the exogenous H and D daily for the model $ARX_{H\&D}$ does not improve the predictive capacity in both quantitative and qualitative (Table 2a and Fig. 3). Since the fitting statistics are relatively high with the other ARX models (Appendix 2_Table g), the deficient predictions could result from an over-fitting effect. The over-fitting responds to the excess of parameters representing the mathematical structure of the phenomena. Thus, as long as the model follows analogies directly connected with the phenomenon's differential equations, it attains better predictions under a more efficient stochastic form.

Table 2a
Coefficient of determination (R^2) for Iran (VAR models).

Parameter's modeling	AR	ARX_H	ARX_D	ARX_{H+D}	$ARX_{H\&D}$
GRW-TVP	0.885(0.891)	0.889(0.909)	0.947(0.951)	0.907(0.915)	0.882(0.944)
SDP	0.847(0.681)	0.729(0.873)	0.842(0.833)	0.890(0.907)	0.846(0.86)

The values outside and inside the brackets contains the R2 coefficient for the median and the best model (using a specific GRW combination as specified in Appendix 2_Tables f, g), estimated and chosen from al the GRW combinations, respectively.

- On the other hand, VAR models applied over the Iran data have similar statistics to those using ARX models for predictions. Here, the best predictive performance is reached when the predictions are recursively given by an interaction between the components of the contagious curve (I) and the recovered individuals (H), as shown by Table 2b for the VAR_H model. This result is the opposite of the ARX models' results and does not follow the same predictive features of Iraq's modeling case.
- The VAR_H model seems to alleviate the overshooting effect at the peak of the contagious curve. (i.e., strings 1 and 2 Fig. 3). Although VAR_H, VAR_D, and VAR_{H+D} achieved acceptable predictive accuracies (Table 2b), they cannot predict the intense changing slopes in the contagious curve, except the particular accuracy of prediction in string 8 given by the VAR_D model. It is worth mentioning that string 8 is a complicated zone for being adjusted when using VAR_D and VAR_{D+H} techniques (according to Appendix 2_Tables h and i).
- Reminding that ARX models work under the "DirRec" scheme for predictions, while the VAR models estimate the predictions using the "Recursive" scheme, the Iran cases give important insights about the differences in predictions for both types of forecasting approaches. In that regard, although the VAR models provide good predictions according to the statistical parameters, the ARX models show the potential for capturing the slope changes of the curve better than the VAR models. The latter is also replicated in the Iraq data. Thus, one could say that "DirRec" approaches to the forecasting problem attain good results in both quantitative and qualitative aspects of prediction.
- As for Iraq's case, daily rates H and D endogenous variables interacting separately with the contagious (I) variable as specified by $VAR_{H\&D}$ model do not increase the predictive capacity as evidence by Table 2b and Fig. 3. Moreover, more historical information into each endogenous variable in the $VAR_{H\&D\text{-}extended}$ model does not improve any predictive capacity, despite its good fitting performance (R2) and efficiency (AIC) during the identification process shown in Appendix 2_Tables i and j.

5. Summary and conclusions

Based on existing tools and knowledge, we propose a novel modeling strategy for the complex SARS-CoV-2 forecasting task. We used Autoregressive with exogenous variables (ARX) and Vector Autoregressions (VAR) techniques. The stochastic structures of ARX and VAR techniques follow analogies with the epidemic's differential equations, thus making them not entirely oblivious of conceptual or dynamical interpretations. To tackle the problem that entails an accurate forecasting task of rapid variations, the models simultaneously perform the updating of their parameters and structures for fitting the incoming data. In addition, we explore different ARX and VAR model configurations to unveil the usefulness and roles that the variables have in anticipating outbreaks and recessions.

The experiment was addressed using the Iraq and Iraq data. Due to the complexity, the time forecasting horizon was set in seven days, and the parameters in the model's equations were considered non-stationary. The well-known Generalized-Random-Walk (GRW), as well as the State-variable-parameter (SDP) techniques, were used for shaping parameters non-stationary for the ARX models, while for the VAR models, a scheme using the GRW techniques was tried. The use of GRW techniques is motivated because they are compatible with the hypothesis that fast fluctuations in the contagious curve depend significantly on the surrounding previous conditions before the event. The SDP techniques are compatible with the alternative hypothesis that rapid fluctuations in the contagious curve depend on some natural states.

Different GRW and SDP combinations render different forecasting estimations that are finally represented by the median. The investigation limits its applicability to the early stages of the phenomenon's evolution in Iraq and Iran, emulating the problem of data scarcity and uncertainty under high-demand circumstances of predictive skills during stages in which rapid variations can occur (such as sudden recessions and outbreaks). Motivated by this, the authors consider an adaptive fashion of the model's structures simultaneously with the parameters' updating for tackling the rapid dynamic variations of the phenomenon.

The nature of ARX and VAR models allows the evaluation of different forecasting approaches. On the one hand, the ARX models use both observed historical data of exogenous variables and recursive estimations retrieved from the endogenous variable through its autoregressive structures, in an approach known as "DirRec" for the forecasting task. On the other hand, the VAR techniques' feat the interaction between each endogenous variable recursively through one-step-ahead estimations until the desired prediction horizon. According to the general argot, the last forecasting procedure is known as "Recursive".

This experiment suggests that the forecasting task of a short-term horizon is possible for a complex phenomenon with a high dynamical variability when using the ARX models and VAR models with GRW techniques. The results show that the forecasting estimations represented by the medians of all the GRW combinations are similar to the best models among those combinations, thus providing reliability to the criterion used in the strategy. Moreover, the models can capture qualitative features such as the intense curvature changes, especially those minimums related to new outbreaks in the contagious curve. Although SDP techniques applied in the ARX models do not attain good predictions in the early stages of the phenomenon, evidence suggests its capacity for accurately capturing the non-linear qualitative features as observations increase.

The ARX models and GRW techniques seem to be more skillful in predicting strong curvatures than the VAR models using the same GRW approaches. Such ability appears to rely partly on the main difference between "DirRec" and "Recursive" methods, explained above. Interestingly, historical observations of contagious and daily rates of deaths that are closer to rapid curve changes (recessions and outbreaks), seem to play an essential role in the models for anticipating these qualitative features.

One of the most important aspects derived from the experimental modeling is the suggestion that better predictions are attained when stochastic approaches are analogous to the differential equations governing the phenomenon. The reason could be that these analogies implicitly provide the models with conceptual robustness, also control over-parameterizations and over-fittings a priori. Moreover, for predicting a phenomenon of such complexity, it was necessary to consider the time variability of the model's parameters and consider an adaptive structure according to the retrieved information. To the authors' knowledge, the latter is a novel approach for emulating this kind of problem with uncertain conditions in critical stages of its evolution, using the most common and available data.

In summary, the processes proposed for the forecasting strategy are complementary. On the one hand, although the central role of the analogies between ARX and VAR stochastic structures with differential equations is to provide the models with conceptual flavor, this implicitly provides the techniques with efficiency and robustness, ultimately controlling over-parameterizations and over-fitting. On the other hand, the success of GRW techniques suggests that rapid curve variations more likely depend on the surrounding observations when these changes occur, as hypothesized in the introduction. Finally, updating the stochastic structures supports capturing rapid changes that could entail strong discontinuities in the evolution of the infection.

In the end, this strategic stochastic modeling with adaptive structures is approved as a promising alternative for estimating predictions in complex phenomena, such as the case of an epidemic spread, with the potential for anticipating rapid curve changes as such recessions

and outbreaks. This experience could be helpful for supporting decisions in continuous monitoring systems.

Future work could consider the following activities: 1) The evaluation of the SDP abilities, using more information for training the models. The suggestion relies on the fact that although SDP does not help in the early stages of the contagious process, it could be useful in advanced states when the model accounts for more detailed observations for the learning process. 2) The analysis of possible mechanisms behind the stochastic structures. This is suggested because ARX and VAR have analogies to differential equations. Thus, these analogies could be translated into explicit dynamic interpretations. 3) Other alternative techniques should be tested using a similar strategy, such as Autoregressive Neural-Networks or Fractal theory. However, other alternatives might require more extended work about developing theoretical aspects and software for allowing certain advantages that the recursive time series theory already has implemented.

CRediT authorship contribution statement

Conceptualization, D.M., A.O, E.S; methodology, D.M.; software, D.M; formal analysis, D.M, A.O, E.S; writing-original draft preparation, D.M.; writing-review and editing, D.M, A.O, E.S.

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.

Appendix 1

Table a
. Iraq_ARX models: identified structures and fitted statistics.

	AR model					ARX_H model					ARX_D model				
Data portion	n	m	τ	R^2	AIC	n	m	τ_h	R^2	AIC	n	m	τ_d	R^2	AIC
String 1	2	1	10	0.913	4.970	2	3	7	0.502	6.915	1	4	8	0.910	5.204
String 2	1	2	10	0.968	5.366	2	1	10	0.844	6.962	4	3	7	0.940	6.304
String 3	4	6	10	0.992	5.917	1	6	10	0.981	6.600	1	3	8	0.961	7.131
String 4	1	4	11	0.950	8.105	1	6	14	0.937	8.435	1	1	7	0.971	7.429
String 5	1	7	8	0.863	9.298	4	2	8	0.754	9.798	2	1	7	0.955	7.967
String 6	5	3	9	0.664	10.057	3	2	14	0.408	10.514	2	2	7	0.957	7.845
String 7	5	6	8	0.506	10.495	5	6	14	0.518	10.471	1	1	7	0.847	9.035
String 8	1	4	10	0.815	9.488	1	1	13	0.532	10.331	5	5	7	0.969	7.858
Best_GWR	C	RW				C	RW				RW	C			TVP
	RW	C				C	RW				RW	C			SDP

Notations of headers follow Eq. (3a). The model's structures evolve according to the data strings sequentially included in the identification process. R^2 and AIC are the Coefficient-of-determination and Akaike-Information-Criteria, respectively, estimated during the identification process. The last line indicates the best GRW combinations.

Table b
. Iraq_ARX models: identified structures and fitted statistics.

	ARX_{H+D} model					$ARX_{H\&D}$ model						
Data portion	n	m	τ_{h+d}	R^2	AIC	n	m	p	τ_h	τ_d	R^2	AIC
String 1	1	1	9	0.652	6.257	1	1	1	10	9	0.960	4.193
String 2	7	1	7	0.982	5.147	1	7	3	9	8	0.998	3.251
String 3	4	4	10	0.990	5.962	1	7	7	11	10	0.997	5.142
String 4	1	6	14	0.952	8.172	3	2	7	11	10	0.995	6.168
String 5	6	4	8	0.768	9.907	3	7	7	9	7	0.992	6.843
String 6	7	1	12	0.492	10.470	2	7	2	12	7	0.985	7.042
String 7	1	3	10	0.431	10.412	6	7	3	14	7	0.973	7.751
String 8	1	1	7	0.511	10.375	3	2	3	9	7	0.972	7.698
Best_GWR	C	RW				RW	C	C				TVP
	RW	RW				RW	C	C				SDP

Notations of headers follow Eqs. (3a) and (3b). The model's structures evolve according to the data strings sequentially included in the identification process. R^2 and AIC are the Coefficient-of-determination and Akaike-Information-Criteria, respectively, estimated during the identification process. The last line indicates the best GRW combinations.

Table c
. Iraq_VAR models: identified structures and fitted statistics.

	VAR_H model								VAR_D model							
Data portion	n	m	–	q	τ_h	τ_{lh}	R^2	AIC	n	m	–	s	τ_d	τ_{ld}	R^2	AIC
String 1	2	2	1	2	2	4	0.758(0.978)	6.093(2.117)	1	3	1	4	1	2	0.904(0.990)	5.168(−0.997)
String 2	1	3	1	1	2	4	0.989(0.978)	4.347(3.452)	4	3	1	1	4	0	0.988(0.980)	4.662(0.983)
String 3	7	2	1	1	2	5	0.995(0.992)	5.310(3.690)	2	7	1	5	4	0	0.997(0.979)	4.946(2.218)
String 4	2	2	1	5	4	1	0.920(0.997)	8.530(4.426)	5	2	1	3	1	0	0.995(0.983)	5.983(2.425)
String 5	7	7	1	7	6	0	0.915(0.999)	9.064(4.191)	3	2	1	5	1	0	0.980(0.988)	7.255(2.500)
String 6	5	6	1	2	6	2	0.711(0.999)	10.016(4.855)	2	3	1	5	2	0	0.993(0.988)	6.096(2.613)
String 7	1	6	1	4	1	0	0.738(0.999)	9.732(4.670)	2	2	1	6	2	0	0.990(0.990)	6.401(2.595)
String 8	1	5	1	7	1	0	0.731(0.999)	9.895(5.123)	3	2	1	7	1	0	0.992(0.992)	6.359(2.543)
Best_GWR	C	RW	IRW	RW					C	RW	C	IRW				TVP

Notations of headers follow Eqs. (6a)–(6c). The model's structures evolve according to the data strings sequentially included in the identification process. R^2 and AIC are the Coefficient-of-determination and Akaike-Information-Criteria, respectively. Values outside and inside ( ) are the fitting statistics of the equation components (6a) and (6b) for VAR_H, and (6a) and (6c) for VAR_D, estimated during the identification process. The last line indicates the best GRW combinations.

Table d
. Iraq_VAR models: Identified structures and fitted statistics.

	VAR_{H+D} model								$VAR_{H\&D}$							
Data portion	n	m	–	r	τ_r	τ_{lr}	R^2	AIC	–	q	–	s	τ_{lh}	τ_{ld}	R^2	AIC
String 1	2	1	1	1	3	3	0.898(0.982)	5.125(2.292)	1	2	1	4	4	2	0.898(0.978)[0.990]	5.125(2.117)[−0.997]
String 2	4	1	1	1	2	4	0.987(0.982)	4.626(3.753)	1	1	1	1	4	0	0.987(0.978)[0.980]	4.626(3.452)[0.983]
String 3	2	5	1	1	2	7	0.995(0.992)	5.250(4.196)	1	1	1	5	5	0	0.995(0.992)[0.979]	5.250(3.690)[2.218]
String 4	3	1	1	2	1	7	0.927(0.997)	8.440(4.464)	1	5	1	3	1	0	0.927(0.997)[0.983]	8.440(4.426)[2.425]
String 5	4	7	1	6	2	1	0.797(0.999)	9.817(4.566)	1	7	1	5	0	0	0.797(0.999)[0.988]	9.817(4.191)[2.500]
String 6	1	3	1	1	2	5	0.624(0.999)	10.023(5.082)	1	2	1	5	2	0	0.624(0.999)[0.988]	10.023(4.855)[2.613]
String 7	1	5	1	3	2	1	0.571(0.999)	10.193(4.958)	1	4	1	6	0	0	0.571(0.999)[0.990]	10.193(4.670)[2.595]
String 8	1	6	1	4	1	3	0.750(0.999)	9.850(5.323)	1	7	1	7	0	0	0.750(0.999)[0.992]	9.850(5.123)[2.543]
Best_GWR	C	RW	IRW	C					IRW	C	IRW	C				TVP

Notations of headers follow Eqs. (6a)–(6c). The model's structures evolve according to the data strings sequentially included in the identification process. R^2 and AIC are the Coefficient-of-determination and Akaike-Information-Criteria, respectively. Values outside and inside ( ) are the statistics of the equation components (6a) and (6b) for VAR_{H+D}. For including Eqs. (6a), (6b), and (6c) simultaneously in the $VAR_{H\&D}$ the parameters q, s, τ_{lh}, τ_{ld} are coupled to the n, m and τ_r parameters. The values outside ( ), inside ( ), and inside [ ] are the statistics of Eqs. (6a), (6b), and (6c), estimated during the identification process. The last line indicates the best GRW combinations.

Table e
. Iraq_VAR models: Identified structures and fitted statistics.

	$VAR_{H\&D\text{-}extended}$ model												
Data portion	n	m	p	w	q	u	s	τ_h	τ_d	τ_{lh}	τ_{ld}	R^2	AIC
String 1	1	2	2	2	1	1	4	2	1	4	2	0.975(0.979)[0.990]	3.931(2.082)[−0.997]
String 2	1	2	3	1	1	3	1	1	5	4	6	0.995(0.978)[0.987]	3.766(3.452)[0.669]
String 3	4	7	7	2	3	1	5	1	4	1	0	0.998(0.994)[0.979]	3.042(3.591)[2.218]
String 4	2	6	1	4	5	1	3	5	2	3	0	0.998(0.998)[0.983]	4.920(4.248)[2.425]
String 5	3	7	3	3	2	1	5	7	2	2	0	0.998(0.999)[0.988]	5.011(3.970)[2.500]
String 6	4	2	3	3	2	1	5	3	3	0	0	0.998(0.999)[0.988]	5.156(4.690)[2.613]
String 7	2	2	3	3	4	2	7	2	1	2	0	0.997(1.000)[0.992]	5.404(4.470)[2.503]
String 8	2	2	3	4	3	2	5	2	1	0	1	0.997(1.000)[0.992]	5.475(4.886)[2.478]
Best_GWR	RW	C	IRW	C	C	C	C						TVP

Notations of headers follow Eqs. (6a)–(6c). The model's structures evolve according to the data strings sequentially included in the identification process. R^2 and AIC are the Coefficient-of-determination and Akaike-Information-Criteria, respectively. The values in outside ( ), inside ( ) and inside [ ] are the statistics of Eqs. (6a), (6b) and (6c), estimated during the identification process for the extended $VAR_{H\&D\text{-}extended}$ model. The last line indicates the best GRW combinations.

Appendix 2

Table f
. Iran_ARX models: identified structures and fitted statistics.

AR model						ARX_H model					ARX_D model				
Data portion	n	m	τ	R^2	AIC	n	m	τ_h	R^2	AIC	n	m	τ_d	R^2	AIC
String 1	1	4	11	0.976	13.622	1	6	8	0.977	13.715	1	3	12	0.916	14.802
String 2	1	2	11	0.978	14.525	1	4	8	0.978	14.615	2	3	8	0.977	14.685
String 3	2	5	13	0.942	16.078	2	1	12	0.939	15.929	2	6	11	0.986	14.676
String 4	4	2	14	0.924	16.160	2	1	12	0.594	17.715	3	6	12	0.995	13.538
String 5	1	5	8	0.654	17.519	5	1	12	0.540	17.804	3	2	14	0.990	13.896
String 6	1	4	8	0.504	17.712	3	4	9	0.472	17.837	6	5	12	0.983	14.500
String 7	1	4	8	0.586	17.418	2	2	11	0.107	18.157	3	4	8	0.982	14.331
String 8	1	6	11	0.552	18.079	2	7	12	0.503	17.596	2	7	12	0.957	15.142
String 9	1	7	10	0.589	17.299	3	7	12	0.489	17.564	7	5	14	0.861	16.308
String 10	1	7	9	0.566	17.304	4	2	14	0.673	16.976	1	7	7	0.634	17.133
Best_GWR	RW	C				RW	C				RW	RW			TVP
	RW	C				RW	C				RW	C			SDP

Notations of headers follow Eq. (3a). The model's structures evolve according to the data strings sequentially included in the identification process. R^2 and AIC are the Coefficient-of-determination and Akaike-Information-Criteria, respectively, estimated during the identification process. The last line indicates the best GRW combinations.

Table g
. Iran_ARX models: identified structures and fitted statistics.

	ARX_{H+D} model					$ARX_{H\&D}$ model						
Data portion	n	m	τ_{h+d}	R^2	AIC	n	m	p	τ_h	τ_d	R^2	AIC
String 1	2	1	8	0.963	13.931	2	1	1	8	12	0.968	13.852
String 2	1	5	10	0.985	14.304	1	5	1	8	13	0.996	13.170
String 3	4	1	9	0.952	15.812	5	4	1	12	7	0.998	12.795
String 4	2	3	13	0.666	17.606	2	3	7	13	10	0.996	13.480
String 5	1	7	11	0.459	18.041	2	1	6	8	13	0.993	13.678
String 6	3	3	13	0.722	17.164	3	5	2	14	14	0.992	13.803
String 7	3	7	12	0.275	18.125	3	5	2	14	14	0.992	13.673
String 8	1	8	7	0.326	17.903	4	4	7	14	11	0.988	14.048
String 9	2	7	11	0.130	18.073	6	1	4	10	12	0.980	14.363

Table g (*continued*)

	ARX_{H+D} model					$ARX_{H\&D}$ model						
Data portion	n	m	τ_{h+d}	R^2	AIC	n	m	p	τ_h	τ_d	R^2	AIC
String 10	3	8	11	0.821	16.489	4	1	5	12	10	0.981	14.208
Best_GWR	RW	C				RW	C	IRW				TVP
	RW	C				RW	C	C				SDP

Notations of headers follow Eqs. (3a) and (3b). The model's structures evolve according to the data strings sequentially included in the identification process. R^2 and AIC are the Coefficient-of-determination and Akaike-Information-Criteria, respectively, estimated during the identification process. The last line indicates the best GRW combinations.

Table h
. Iran_VAR models: identified structures and fitted statistics.

	VAR_H model								VAR_D model							
Data portion	n	m	–	q	τ_h	τ_{lh}	R2	AIC	n	m	–	s	τ_d	τ_{ld}	R2	AIC
String 1	1	5	1	2	4	1	0.985(0.989)	13.235(11.887)	2	3	1	1	1	3	0.923(0.999)	14.795(6.831)
String 2	5	2	1	1	5	6	0.993(0.996)	13.597(11.673)	2	4	1	5	2	5	0.989(0.997)	13.990(8.511)
String 3	2	7	1	2	2	1	0.933(0.995)	16.333(13.191)	3	4	1	6	1	1	0.975(0.988)	15.266(10.356)
String 4	4	3	1	6	7	2	0.825(0.997)	17.042(13.880)	2	2	1	7	7	1	0.970(0.992)	15.168(10.254)
String 5	3	7	1	2	7	4	0.539(0.995)	17.955(14.655)	7	5	1	4	2	0	0.974(0.990)	15.159(10.678)
String 6	6	5	1	2	5	3	0.555(0.997)	17.797(14.547)	3	4	1	4	1	0	0.981(0.989)	14.500(10.893)
String 7	2	1	1	2	6	2	0.463(0.998)	17.620(14.410)	3	4	1	4	1	0	0.967(0.994)	14.945(10.458)
String 8	3	1	1	3	3	1	0.530(0.999)	17.411(14.294)	2	4	1	4	1	0	0.958(0.994)	15.048(10.439)
String 9	3	1	1	2	7	3	0.708(0.999)	16.862(14.295)	3	3	1	4	6	0	0.946(0.995)	15.215(10.495)
String 10	5	4	1	1	5	6	0.856(0.999)	16.229(14.520)	4	3	1	5	7	0	0.940(0.995)	15.298(10.553)
Best_GWR	RW	RW	IRW	IRW					C	IRW	C	C				TVP

Notations of headers follow Eqs. (6a)–(6c). The model's structures evolve according to the data strings sequentially included in the identification process. R^2 and AIC are the Coefficient-of-determination and Akaike-Information-Criteria, respectively. Values outside and inside ( ) are the statistics of the equation components (6a) and (6b) for VAR_H, and 6a and 6c for VAR_D, estimated during the identification process. The last line indicates the best GRW combinations.

Table i
. Iran_VAR models: identified structures and fitted statistics.

	VAR_{H+D} model								$VAR_{H\&D}$							
Data portion	n	m	–	r	τ_r	τ_{lr}	R^2	AIC	–	q	–	s	τ_{lh}	τ_{ld}	R^2	AIC
String 1	1	5	1	1	3	1	0.966(0.993)	14.055(11.843)	1	2	1	1	1	1	0.966(0.989)[0.999]	14.055(11.887)[6.831]
String 2	2	3	1	6	7	3	0.989(0.997)	13.985(12.045)	1	1	1	5	6	2	0.989(0.996)[0.997]	13.985(11.673)[8.511]
String 3	2	2	1	3	7	7	0.946(0.997)	15.873(13.074)	1	2	1	6	1	1	0.946(0.995)[0.988]	15.873(13.191)[10.356]
String 4	3	2	1	4	7	5	0.596(0.997)	17.795(14.023)	1	6	1	7	2	7	0.596(0.997)[0.992]	17.795(13.880)[10.254]
String 5	3	7	1	1	7	7	0.728(0.996)	17.428(14.747)	1	2	1	4	4	2	0.728(0.995)[0.990]	17.428(14.655)[10.678]
String 6	5	4	1	4	1	1	0.580(0.997)	17.675(14.709)	1	2	1	4	3	1	0.580(0.997)[0.989]	17.675(14.547)[10.893]
String 7	2	1	1	3	4	1	0.381(0.998)	17.763(14.518)	1	2	1	4	2	1	0.381(0.998)[0.994]	17.763(14.410)[10.458]
String 8	1	6	1	2	4	3	0.174(0.999)	18.054(14.461)	1	3	1	4	1	1	0.174(0.999)[0.994]	18.054(14.294)[10.439]
String 9	6	4	1	3	5	0	0.966(0.993)	14.055(11.843)	1	2	1	4	3	6	0.966(0.989)[0.999]	14.055(11.887)[6.831]
String 10	5	4	1	3	5	1	0.989(0.997)	13.985(12.045)	1	1	1	5	6	7	0.989(0.996)[0.997]	13.985(11.673)[8.511]
Best_GWR	RW	RW	RW	RW					IRW	IRW	C	C				TVP

Notations of headers follow Eqs. (6a)–(6c). The model's structures evolve according to the data strings sequentially included in the identification process. R^2 and AIC are the Coefficient-of-determination and Akaike-Information-Criteria, respectively. Values outside and inside ( ) are the statistics of the equation components (6a) and (6b) for VAR_{H+D}. For including Eqs. (6a), (6b), and (6c) simultaneously in the $VAR_{H\&D}$ the parameters q, s, τ_{lh}, τ_{ld} are coupled to the n, m and τ_r parameters. The values outside ( ), inside ( ), and inside [ ] are the statistics of Eqs. (6a), (6b), and (6c), estimated during the identification process. The last line indicates the best GRW combinations for modeling the parameters for the TVP technique.

Table j
. Iran_VAR models: identified structures and fitted statistics.

	$VAR_{H\&D\text{-}extended}$ model												
Data portion	n	m	p	w	q	u	s	τ_h	τ_d	τ_{lh}	τ_{ld}	R^2	AIC
String 1	1	2	2	1	2	2	3	2	7	1	1	0.998(0.989)[0.999]	11.321(11.887)[6.096]
String 2	2	1	6	1	1	2	1	2	1	6	1	0.999(0.996)[0.997]	12.190(11.673)[8.267]
String 3	1	7	7	1	2	4	5	6	2	1	1	0.994(0.995)[0.999]	14.143(13.191)[7.654]
String 4	1	7	7	1	6	2	4	1	2	2	0	0.986(0.997)[0.999]	14.843(13.880)[8.528]
String 5	2	1	6	2	2	1	4	3	6	0	0	0.985(0.996)[0.990]	14.490(14.616)[10.678]
String 6	3	1	5	2	2	3	7	3	6	2	0	0.988(0.997)[0.996]	14.131(14.470)[9.976]
String 7	3	1	4	2	2	1	4	2	7	1	0	0.980(0.998)[0.994]	14.485(14.398)[10.458]
String 8	2	1	5	1	3	1	4	3	6	1	0	0.961(0.999)[0.994]	15.038(14.294)[10.439]
String 9	4	1	3	2	2	1	4	7	5	0	0	0.971(0.999)[0.995]	14.644(14.275)[10.495]
String 10	5	6	3	1	1	2	5	5	1	6	0	0.987(0.999)[0.996]	13.899(14.520)[10.378]
Best_GWR	RW	C	RW	C	RW	C	RW						TVP

Notations of headers follow Eqs. (6a)–(6c). The model's structures evolve according to the data strings sequentially included in the identification process. R^2 and AIC are the Coefficient-of-determination and Akaike-Information-Criteria, respectively. The values in outside ( ), inside ( ) and inside [ ] are the statistics of Eqs. (6a), (6b) and (6c), estimated during the identification process for the extended $VAR_{H\&D\text{-}extended}$ model. The last line indicates the best GRW combinations.

Appendix 3

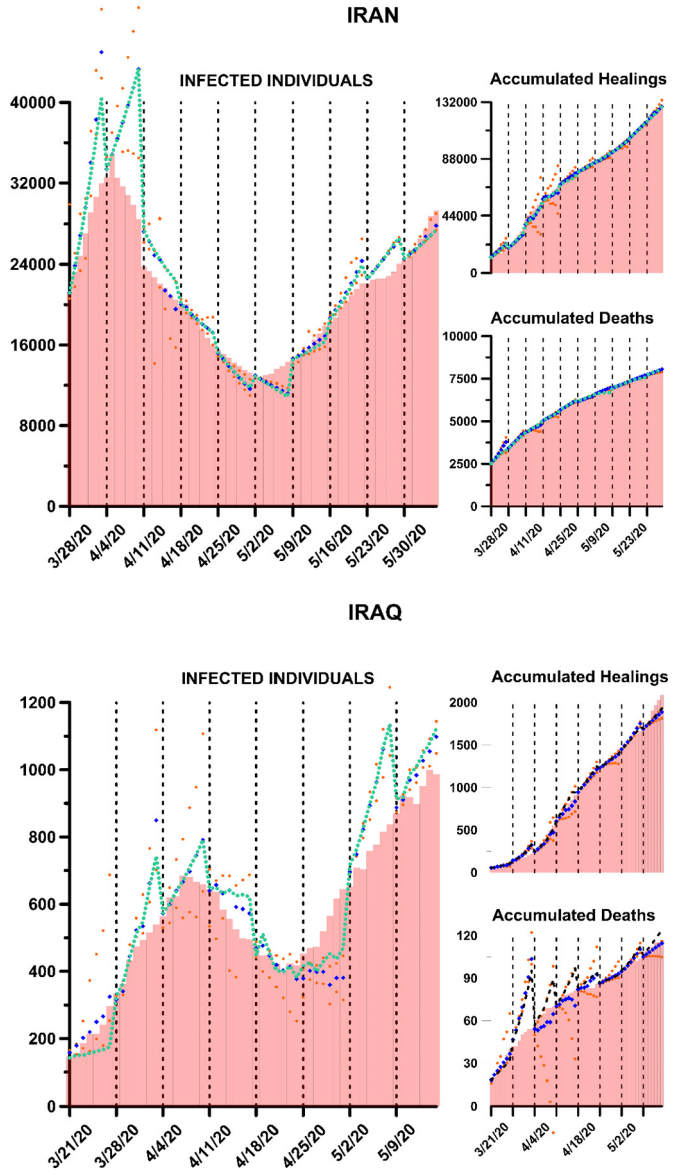

Fig. 4. Examples of VAR included accumulated Healings and Deaths for Iraq and Iran cases.

References

[1] Richardson MJ, Chemero A. Complex dynamical systems and embodiment. The Routledge handbook of embodied cognition; 2014. p. 39–50.

[2] Wang W-X, Lai Y-C, Grebogi C. Data based identification and prediction of nonlinear and complex dynamical systems. Phys Rep. 2016;644:1–76. https://doi.org/10.1016/j.physrep.2016.06.004.

[3] Capasso V. Mathematical structures of epidemic systems. , Vol. 97Springer Science & Business Media; 2008..

[4] Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc R Soc London Ser A. 1927;115(772):700–21.

[5] Newman M. Networks. Oxford university press; 2018..

[6] Adak D, Majumder A, Bairagi N. Mathematical perspective of COVID-19 pandemic: disease extinction criteria in deterministic and stochastic models. Chaos Solitons Fractals. 2021;142:110381.

[7] Sun T, Wang Y. Modeling COVID-19 epidemic in Heilongjiang province, China. Chaos Solitons Fractals. 2020;138:109949.

[8] Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Wang W, Song H, Huang B, Zhu N, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. 2020;395(10224):565–74.

[9] Team EE. Note from the editors: World Health Organization declares novel coronavirus (2019-nCoV) sixth public health emergency of international concern. Eurosurveillance. 2020;25(5):200131e.

[10] Imai N, Dorigatti I, Cori A, Riley S, Ferguson NM. Report 1: estimating the potential total number of novel coronavirus (2019-nCoV) cases in Wuhan City, China; 2020. URL Https://Www. Imperial. Ac. Uk/Mrc-Global-Infectious-Disease-Analysis/News–Wuhan-Coronavirus/Accessed on February, 21, 2020.

[11] Ndiaye BM, Tendeng L, Seck D. Comparative prediction of confirmed cases with COVID-19 pandemic by machine learning, deterministic and stochastic SIR models. ArXiv Preprint ArXiv:2004.13489; 2020..

[12] Roda WC, Varughese MB, Han D, Li MY. Why is it difficult to accurately predict the COVID-19 epidemic? Infect Dis Model. 2020;5:271–81.

[13] Shen M, Peng Z, Xiao Y, Zhang L. Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China. BioRxiv; 2020..

[14] Wang H, Wang Z, Dong Y, Chang R, Xu C, Yu X, Zhang S, Tsamlag L, Shang M, Huang J. Phase-adjusted estimation of the number of coronavirus disease 2019 cases in Wuhan, China. Cell Discov. 2020;6(1):1–8.

[15] Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet. 2020;395(10225):689–97.

[16] You C, Deng Y, Hu W, Sun J, Lin Q, Zhou F, Pang CH, Zhang Y, Chen Z, Zhou X-H. Estimation of the time-varying reproduction number of COVID-19 outbreak in China. Int J Hyg Environ Health. 2020.;113555.
[17] Zhan C, Tse CK, Lai Z, Hao T, Su J. Prediction of COVID-19 spreading profiles in South Korea, Italy and Iran by data-driven coding. PloS One. 2020;15(7):e0234763.
[18] Zhang Y, You C, Cai Z, Sun J, Hu W, Zhou X-H. Prediction of the COVID-19 Outbreak Based on a Realistic Stochastic Model. MedRxiv; 2020.
[19] Barlow P, van Schalkwyk MCI, McKee M, Labonté R, Stuckler D. COVID-19 and the collapse of global trade: building an effective public health response. Lancet Planet Health. 2021;5(2):e102–7.
[20] Lemos DRQ, D'angelo SM, Farias LABG, Almeida MM, Gomes RG, Pinto GP, Cavalcante JN, Feijão LX, Cardoso ARP, Lima TBR, et al. Health system collapse 45 days after the detection of COVID-19 in Ceará, Northeast Brazil: a preliminary analysis. Rev Soc Bras Med Trop. 2020.;53.
[21] 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.
[22] Castillo O, Melin P. A novel method for a covid-19 classification of countries based on an intelligent fuzzy fractal approach. Healthcare. 2021;9(2):196.
[23] Melin P, Monica JC, Sanchez D, Castillo O. Analysis of spatial spread relationships of coronavirus (COVID-19) pandemic in the world using self organizing maps. Chaos Solitons Fractals. 2020;138:109917.
[24] Boccaletti S, Ditto W, Mindlin G, Atangana A. Modeling and forecasting of epidemic spreading: the case of Covid-19 and beyond. Chaos Solitons Fractals. 2020;135:109794.
[25] Taylor SJ, Letham B. Forecasting at scale. Am Stat. 2018;72(1):37–45.
[26] Battineni G, Chintalapudi N, Amenta F. Forecasting of COVID-19 epidemic size in four high hitting nations (USA, Brazil, India and Russia) by Fb-Prophet machine learning model. Appl Comput Inf. 2020. https://doi.org/10.1108/ACI-09-2020-0059.
[27] Indhuja M, Sindhuja P. Prediction of covid-19 cases in India using prophet. Int J Stat Appl Math. 2020.;5(4).
[28] Khayyat M, Laabidi K, Almalki N, Al-Zahrani M. Time series Facebook prophet model and python for COVID-19 outbreak prediction. Comput Mater Continua. 2021:3781–93.
[29] 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. Healthcare. 2020;8(2):181.
[30] Namasudra S, Dhamodharavadhani S, Rathipriya R. Nonlinear neural network based forecasting model for predicting COVID-19 cases. Neural Process Lett. 2021:1–21.
[31] Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. J Stat Softw. 2008;27(3):1–22. https://doi.org/10.18637/jss.v027.i03.
[32] Skiera B, Jürgensmeier L, Stowe K, Gurevych I. How to best predict the daily number of new infections of COVID-19. ArXiv Preprint ArXiv:2004.03937; 2020.
[33] Cirillo P, Taleb NN. Tail risk of contagious diseases. Nat Phys. 2020:1–8.
[34] Santosh KC. COVID-19 prediction models and unexploited data. J Med Syst. 2020;44(9):1–4.
[35] Young PC. Recursive Estimation and Time-series Analysis: An Introduction. Springer Science & Business Media; 2012.
[36] Athanasopoulos G, Poskitt DS, Vahid F. Two canonical VARMA forms: scalar component models Vis-à-Vis the echelon form. Econ Rev. 2012;31(1):60–83.
[37] Box G. Box and Jenkins: time series analysis, forecasting and control. A Very British Affair. London: Palgrave Macmillan; 2013. p. 161–215. https://doi.org/10.1057/9781137291264_6.
[38] Nicholson WB, Matteson DS, Bien J. VARX-L: structured regularization for large vector autoregressions with exogenous variables. Int J Forecast. 2017;33(3):627–51.
[39] Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann Intern Med. 2020;172(9):577–82.
[40] Liu Y, Rocklöv J. The reproductive number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus. J Travel Med. 2021;28(7):taab124. https://doi.org/10.1093/jtm/taab124.
[41] Pecho-Silva S, Arteaga-Livias K, Rodriguez-Morales AJ. Airborne SARS-CoV-2: weighing the evidence for its role in community transmission. J Prev Med Public Health. 2020;53(3):178.
[42] Ju B, Zhang Q, Ge J, Wang R, Sun J, Ge X, Yu J, Shan S, Zhou B, Song S, et al. Human neutralizing antibodies elicited by SARS-CoV-2 infection. Nature. 2020;584(7819):115–9.
[43] Miller JC. A note on the derivation of epidemic final sizes. Bull Math Biol. 2012;74(9):2125–41.
[44] Harko T, Lobo FSN, Mak MK. Exact analytical solutions of the susceptible-infected-recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. Appl Math Comput. 2014;236:184–94.
[45] Caccavo D. Chinese and Italian COVID-19 outbreaks can be correctly described by a modified SIRD model. MedRxiv; 2020.
[46] Fernández-Villaverde J, Jones CI. Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities; 2020.
[47] Ferrari L, Gerardi G, Manzi G, Micheletti A, Nicolussi F, Salini S. Modelling provincial Covid-19 epidemic data in Italy using an adjusted time-dependent SIRD model. ArXiv Preprint ArXiv:2005.12170; 2020.
[48] Miller JC. Mathematical models of SIR disease spread with combined non-sexual and sexual transmission routes. Infect Dis Model. 2017;2(1):35–55.
[49] Taylor CJ, Pedregal DJ, Young PC, Tych W. Environmental time series analysis and forecasting with the captain toolbox. Environ Model Softw. 2007;22(6):797–814. https://doi.org/10.1016/j.envsoft.2006.03.002.
[50] Young PC, Mckenna P, Bruun J. Identification of non-linear stochastic systems by state dependent parameter estimation. Int J Control. 2001;74(18):1837–57. https://doi.org/10.1080/00207170110089824.
[51] Mendoza DE, Samaniego EP, Mora DE, Espinoza MJ, Pacheco EA, Avilés AM. Local rainfall modelling based on global climate information: a data-based approach. Environ Model Softw. 2020;131:104786. https://doi.org/10.1016/j.envsoft.2020.104786.
[52] Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice. OTexts; 2018.
[53] Akaike H. A new look at the statistical model identification. IEEE Trans Autom Control. 1974;19(6):716–23. https://doi.org/10.1109/TAC.1974.1100705.
[54] Akaike H. On entropy maximization principle. Appl Stat. 1977;543:27–41.
[55] Ben Taieb S, Bontempi G, Atiya AF, Sorjamaa A. A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. Expert Syst Appl. 2012;39(8):7067–83. https://doi.org/10.1016/J.ESWA.2012.01.039.