Research

Author for correspondence:
A. S. Fokas
e-mail: tf227@cam.ac.uk

An algebraic formula, deep learning and a novel SEIR-type model for the COVID-19 pandemic

A. S. Fokas[1,2,3], N. Dikaios[2] and Y. C. Yortsos[3]

[1]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK
[2]Mathematics Research Centre, Academy of Athens, 11527 Athens, Greece
[3]Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089, USA

ASF, 0000-0002-5881-802X; ND, 0000-0001-9865-0260

The most extensively used mathematical models in epidemiology are the susceptible-exposed-infectious-recovered (SEIR) type models with constant coefficients. For the first wave of the COVID-19 epidemic, such models predict that at large times equilibrium is reached *exponentially*. However, epidemiological data from Europe suggest that this approach is *algebraic*. Indeed, accurate long-term predictions have been obtained via a forecasting model only if it uses an algebraic as opposed to the standard exponential formula. In this work, by allowing those parameters of the SEIR model that reflect behavioural aspects (e.g. spatial distancing) to vary nonlinearly with the extent of the epidemic, we construct a model which exhibits asymptotictly *algebraic* behaviour. Interestingly, the emerging power law is consistent with the typical dynamics observed in various social settings. In addition, using reliable epidemiological data, we solve in a numerically robust way the *inverse problem* of determining all model parameters characterizing our novel model. Finally, using deep learning, we demonstrate that the algebraic forecasting model used earlier is optimal.

1. Introduction

In the framework of Holmdahl & Buckee [1], epidemiological models are broadly divided into two categories: *forecasting* and *mechanistic*. The former fit early-time data with a specific empirical formula, which is then used to predict the time evolution. The limitation of the forecasting models is that they usually remain valid for only a specific, short, period of time, during which the epidemiological situation is unchanged. For example, if a forecasting model is valid during part of a lockdown
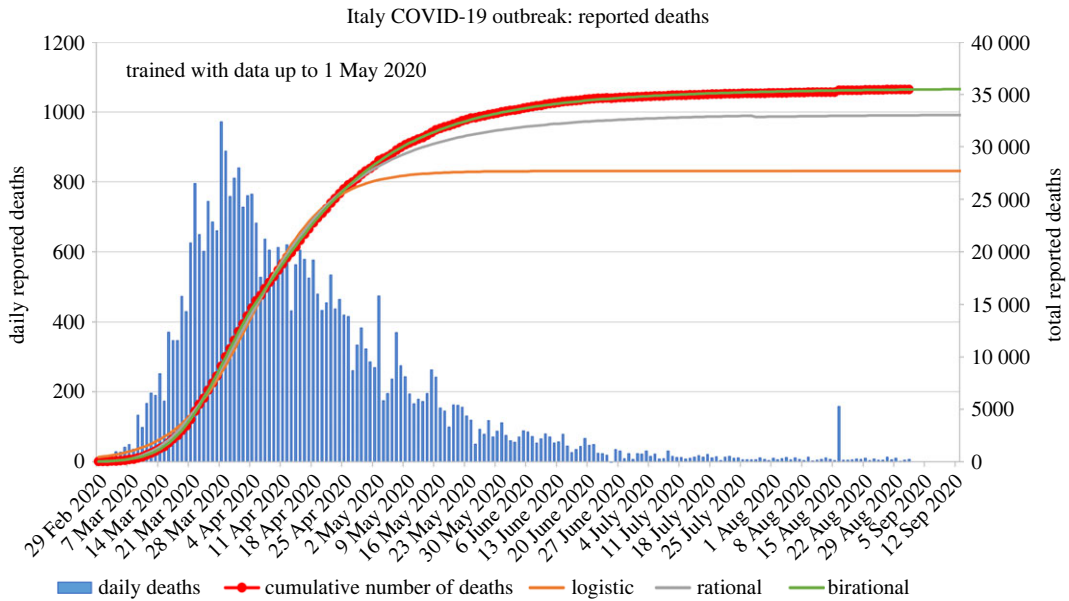
Figure 1. Actual data for the number of reported COVID-19 deaths for Italy (in blue and red) throughout the 'first wave' and predicted results using data until 1 May 2020. The birational formula (in green) is the only one among the three forecasting models (logistic, rational and birational) that matches well the actual data.

period, this model is not expected to make accurate predictions after the lockdown is lifted. As noted in Zhou *et al.* [2], forecasting models are 'not well suited for long-term predictions'. Mechanistic models based on differential equations circumvent this limitation and can be used to make predictions even when the relevant circumstances change. Their limitation is that it is difficult to solve the associated *inverse problem*, namely, to determine the relatively large number of parameters involved in these models from the knowledge of epidemiological data.

The first period of the COVID-19 pandemic, known as 'the first wave', was characterized by the fact that a single viral strain dominated the pandemic. This setting provides an ideal situation for the retrospective study of the various mathematical formulations (both forecasting and mechanistic) used in viral epidemics models. Based on this first wave, we previously introduced [3,4] a novel class of *forecasting* models, which provided accurate long-term predictions for the number of deaths in several European countries caused by the epidemic (figure 1). This was accomplished by using a simple algebraic formula as opposed to the standard exponential formula (logistic model). It is important to emphasize that the logistic formula, as well as similar formulae (see for example [5]) provide accurate fitting for the data of a *fixed* period. In particular, all these formulae can be used to model the data up to 1 May 2020. However, after fixing the parameters of these formulae using the data of a fixed period, these models failed to provide accurate predictions. On the other hand, after 'training' the algebraic formula of for the same period (i.e. after fixing the relevant parameters using data until 1 May 2020), this formula did provide accurate predictions for a period extending longer than 3.5 months, namely until 1 September 2020.

We show typical comparisons in figure 1, where we plot data for deaths during the first wave in Italy, as well as the predictions associated with three models, the logistic and the two novel models of [3,4]. The model routinely used in epidemiology is the logistic defined in (1.1) below; the rational and the birational formulae are given, respectively, by equations (1.2) and (1.3) below:

$$D(t) = \frac{D_F}{1 + \beta e^{-\gamma t}}, \tag{1.1}$$

$$D(t) = \frac{D_F}{1 + \beta(1 + \delta t)^{-\gamma}} \tag{1.2}$$

$$D(t) = \begin{cases} \dfrac{c}{1 + b(1 + dt)^{-k}}, & t \leq T \\ \dfrac{c}{1 + b(1 + dX)^{-k}} - \dfrac{c_1}{1 + b_1(1 + d_1 X)^{-k_1}} + \dfrac{c_1}{1 + b_1(1 + d_1 t)^{-k_1}}, & t > T \end{cases} \tag{1.3}$$

where X is a fixed parameter X usually taken to be T.

Here, $D(t)$ denotes the cumulative number of deaths at time t, with the constant parameters D_F, β, γ and δ determined by 'training' (fitting) the model with the data. Similar considerations are valid for the parameters defining the birational formula (1.3), where the time T corresponds to the inflection point of $D(t)$ (details are given in [3,4]). Equation (1.1) shows that, for large t, the expression $D(t)$ approaches its final value D_F exponentially. In the case of the rational and birational models, the approach is algebraic. In this sense, it is better to refer to the rational formula as *algebraic* [3,4]. All three expressions in (1.1–1.3) are particular solutions of a specific differential equation (ODE) of the Riccati type, uniquely specified by the constant D_F and by a function $a(t)$. The formulae in (1.2) and (1.3) correspond to the case where $a(t)$ is a rational function of t, whereas the logistic model corresponds to the case that $a(t)$ is a constant. The logistic, rational and birational curves are shown along with real data in figure 1 above.

As shown in figure 1, the birational formula closely matches the actual data throughout the time interval considered, in contrast with the other two curves (rational and logistic). It is interesting that even though 'training' occurred during the lockdown period, its predictions remained accurate after the easing of the lockdown conditions. Possible explanations for this unexpected agreement are discussed in [6].

The fact that the correct asymptotic decay is algebraic rather than exponential suggests a social behaviour origin. Indeed, while most, although not all, physical and biological phenomena exhibit exponential decay in the distribution of key variables, or in the behaviour of correlation functions, this is not the case for social phenomena, where one usually observes algebraic dependences and power laws, perhaps because of self-similarity or fractal geometry notions [7]. Indeed, examples of power laws in social settings are ubiquitous, from economics to the distribution of income or of the populations of cities. While such behaviour may also be observed in natural phenomena, e.g. in long-range correlations, or near percolation thresholds in phase transitions [8], the key point regarding epidemiological models is the following: because epidemiological quantities are also affected by social behaviour, e.g. the relaxation of precautions as the epidemic wanes, it is likely that the apparent success of the birational equation is due to such phenomena. Using the *mechanistic* susceptible-exposed-infectious-recovered (SEIR) model and previously reported results in [9], we will demonstrate this assertion quantitatively in the following section.

The most extensively used mechanistic models in epidemiology are variations of the classical SIR model. This is based on the assumption of a constant total population, consisting of three subpopulations, namely, *susceptible*, *infected* and *recovered*, and three implicit conditions, that the control volume where the epidemic occurs is fixed, there is no influx of additional populations, and spatial gradients do not exist. Following chemical reaction engineering process analogies, Ramaswamy *et al.* [10] generalized the SIR model to a spatio-temporal model which relaxed all these restrictions; this led to a set of partial differential equations in terms of population densities. This generalization predicted the onset of spatio-temporal waves, as well as other characteristics inherently absent in the SIR model. In addition, it led to the emergence of a single dimensionless number that characterizes the process, the so-called reproduction number R_0, quantitatively expressed as the ratio of the kinetics of infection to those of recovery [10]. Importantly, R_0 incorporates a combination of physiological and biological parameters, and also reflects measures of social dynamics (e.g. via facial coverings, spatial distancing, areal density, etc.).

If R_0 is constant, SIR predicts that any infection wave wanes asymptotically following an exponential behaviour. In previous work [9], we examined the question of what modifications of the model would lead to an algebraic rather than an exponential asymptotic behaviour, as appears to be the actual case in the first COVID-19 wave. Given that the only relevant parameter in SIR models is R_0, we showed that for an algebraic behaviour to be obtained, it is necessary that beyond a certain time, R_0 must become an increasing function of the infected fraction denoted by i (ratio of the number of infected individuals to the total). The particular dependence was found in [9] to be

$$R_0(i) \propto (1 + a\, i^{1/n})^{-1}, \tag{1.4}$$

where the parameter a is a positive constant, and the exponent n was found to satisfy $n > 2$. Under these conditions, the asymptotic decay is algebraic and follows a power law in time with an exponent equal to $-n/(n-1)$. Equation (1.4) shows that in such cases, as the infection wanes, the reproduction number R_0 is not constant, but rather increases in an algebraic manner with respect to i. The conclusion that the emergence of a power law in time requires that R_0 must increase as the infection wanes, is consistent with the social tendency to relax restrictions (e.g. decrease spatial distancing, increase areal densities, etc.) as the epidemic wanes.

Motivated by the fact that the actual data are best matched using an algebraic dependence and that Luhar *et al.* [9] provide a way to incorporate such behaviour in a simple SIR model, we will extend here this methodology to a more elaborate SEIR-type model, that also includes hospitalized and deceased patients data, neither of which are included in the simpler SIR model of [9].
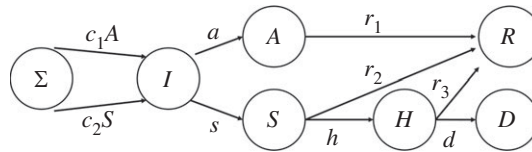
Figure 2. Flow diagram and transmission rates for the different populations in the SEIR model analysed in this work.

The paper is organized as follows: in the Formulation and results section, we present a modification of the SEIR model that accurately gives rise to an algebraic asymptotic behaviour. We then fit the experimental results on the epidemiological data for the deceased of the first wave in Portugal to the algebraic and birational formulae (1.2) and (1.3). These formulae together with the new model are then used to solve the inverse problem and to determine the various parameters of our full mechanistic novel model. Note that, in a certain sense, the mathematical expressions for our forecasting models are optimal, since an elaborate deep-learning algorithm does not seem to yield better results. In solving *the inverse problem*, all parameters of the new SEIR-type model are determined from the knowledge of the deceased and hospitalized data. It is shown that the numerical solution matches well the analytical asymptotic results. Furthermore, as elaborated in the Discussion section, this solution provides a confirmation of the accuracy of the solution of the inverse problem.

2. Materials and methods

2.1. Formulation and results

The SEIR model previously discussed in several publications [11,12] contains the following populations: $\Sigma(t)$, $I(t)$, $A(t)$, $S(t)$, $H(t)$, $R(t)$ and $D(t)$, which denote susceptible, infected, asymptomatic, sick, hospitalized, recovered and deceased individuals, respectively. In this notation, $I(t)$ refers to individuals infected, but not yet infectious, since it is generally assumed that infectiousness requires the elapse of a certain incubation period (typically 5 days) [12]. Then, that individual will either become sick or asymptomatic, both of which are infectious. Importantly, this means that a susceptible individual joins the population $I(t)$ only because of encounters with individuals from populations $A(t)$ or $S(t)$ (but not from $I(t)$). The various entities are interdependent, as indicated in figure 2: if the daily rates at which an infected person becomes either sick or asymptomatic are s or a, respectively, then each day $sI(t)$ or $aI(t)$ individuals leave population $I(t)$ and enter $S(t)$ and $A(t)$, respectively. Likewise, if asymptomatic individuals recover at a daily rate r_1, each day $r_1A(t)$ individuals leave the asymptomatic population and enter the recovered population, $R(t)$. Sick individuals either recover at a rate r_2 or they become hospitalized at a rate h. The hospitalized patients also have two possible outcomes: either they recover at a rate r_3, or they are deceased at a rate d. Finally, susceptible individuals convert to infected, but not yet infectious, populations, because of their encounter with either asymptomatic or sick individuals (we assume that due to special precautions, hospitalized populations cannot infect) at corresponding rates c_1A and c_2S, where c_1 and c_2 denote the respective transmission rates (per person in the respective pools). It is important to note that from all these coefficients, it is only c_1 and c_2 that can be largely affected by social dynamics, the other being largely independent.

2.2. The basic formulation

The previous description leads to the following seven ordinary differential equations:

$$A^{(1)} = aI - r_1A, \tag{2.1}$$

$$S^{(1)} = sI - (h + r_2)S, \tag{2.2}$$

$$H^{(1)} = hS - (r_3 + d)H, \tag{2.3}$$

$$R^{(1)} = r_1A + r_2S + r_3H, \tag{2.4}$$

$$D^{(1)} = dH, \tag{2.5}$$

$$\Sigma^{(1)} = -\Sigma(c_1A + c_2S) \tag{2.6}$$

and

$$I^{(1)} = \Sigma(c_1A + c_2S) - (a + s)I, \tag{2.7}$$

where superscript (n) indicates nth time derivative. Total conservation of individuals leads to

$$I + S + A + R + H + D + \Sigma = T, \tag{2.8}$$

where T is the sum of all individuals in the control volume prior to the pandemic.

Eliminating Σ from equations (2.6) and (2.7) leads to

$$I^{(1)} = [T - (I + S + A + R + H + D)](c_1 A + c_2 S) - (A + s)I. \tag{2.9}$$

We will next proceed with the assumption that the coefficients a, r_1, s, h, r_2, r_3 and d are constants. Then, for arbitrary, and not necessarily constant, values of c_1 and c_2, equations (2.1–2.7) can be reduced to a system of two ODEs. The final result reads as follows:

Populations H, S and I can be expressed in terms of D, via the following equations:

$$H = \frac{D^{(1)}}{d}, \tag{2.10a}$$

$$S = \frac{1}{dh}(D^{(2)} + R_3 D^{(1)}) \tag{2.10b}$$

and

$$I = \frac{1}{dhs}[D^{(3)} + (R_2 + R_3)D^{(2)} + R_2 R_3 D^{(1)}]. \tag{2.10c}$$

Also, equations (2.1) and (2.2) can be written in the form

$$\left. \begin{array}{l} A = \dfrac{a}{r_1} I - \dfrac{A^{(1)}}{r_1} \\[2mm] S = \dfrac{s}{h + r_2} I - \dfrac{S^{(1)}}{h + r_2}. \end{array} \right\} \tag{2.11}$$

and

Proof of equations (2.10): Equation (2.5) is equation (2.10a). Solving equation (2.3) for S and then replacing H via (2.10a) we find (2.10b). Solving (2.2) for I and then replacing S via (21.10b) we find (2.10c). Adding (2.1) and (2.4) and then replacing S, H, I via (2.10) we find,

$$\begin{aligned} (A + R)^{(1)} = \frac{1}{dhs}[aD^{(3)} &+ [a(R_2 + R_3) + sr_2]D^{(2)} \\ &+ [aR_2 R_3 + s(hr_3 + r_2 R_3)]D^{(1)}]. \end{aligned} \tag{2.12}$$

Integrating this equation and simplifying, we obtain

$$A + R = \frac{aD^{(2)}}{dhs} + [a(R_2 + R_3) + sr_2]\frac{D^{(1)}}{dhs} + (FR_2 R_3 - dhs)\frac{D}{dhs} + \xi, \tag{2.13}$$

where ξ is an integration constant. Replacing in (2.1) the function I via (2.10c), we obtain (2.22).

Now, equation (2.9) is

$$I^{(1)} + (a + s)I = [T - (I + S + A + R + H + D)](c_1 A + c_2 S). \tag{2.14}$$

QED

Replacing in (2.6) and (2.7) A and S via (2.11), we find

$$\left. \begin{array}{l} \Sigma^{(1)} = -\Sigma I \left(\dfrac{c_1 a}{r_1} + \dfrac{c_2 s}{h + r_2} \right) + \Sigma \left(\dfrac{c_1 A^{(1)}}{r_1} + \dfrac{c_2 S^{(1)}}{h + r_2} \right) \\[3mm] I^{(1)} = -(a + s)I \left[1 - \dfrac{1}{a + s} \left(\dfrac{c_1 a}{r_1} + \dfrac{c_2 s}{hr_2} \right) \Sigma \right] - \Sigma \left(\dfrac{c_1 A^{(1)}}{r_1} + \dfrac{c_2 S^{(1)}}{h + r_2} \right), \end{array} \right\} \tag{2.15}$$

where R_2 and R_3 are defined by

$$R_2 = r_2 + h, \quad R_3 = r_3 + d. \tag{2.16}$$

In order to make use of the results in Luhar *et al.* [9], we next consider the hypothetical case, $A^{(1)} = S^{(1)} = 0$, in which case equations (2.15) become

$$
\left.
\begin{aligned}
I^{(1)} &= -(a+s)I\left[1 - \frac{1}{a+s}\left(\frac{c_1 a}{r_1} + \frac{c_2 s}{h+r_2}\right)\Sigma\right] \\
\Sigma^{(1)} &= -\Sigma I\left(\frac{c_1 a}{r_1} + \frac{c_2 s}{h+r_2}\right).
\end{aligned}
\right\}
\tag{2.17}
$$

and

These are in a form comparable to the SIR model previously described [9], written here as

$$
\left.
\begin{aligned}
\frac{dI}{d\tau} &= -I\left(1 - \frac{Rs}{1 + x^{1/n}}\right), \quad x = (\lambda - 1)^n \frac{I}{I_1} \\
\frac{d\Sigma}{d\tau} &= -I\Sigma\frac{R}{1 + x^{1/n}}.
\end{aligned}
\right\}
\tag{2.18}
$$

and

The two sets of equations (2.17, 2.18) become self-consistent if we take

$$
\tau = (a+s)t,
$$

and

$$
\frac{1}{a+s}\left(\frac{c_1 a}{r_1} + \frac{c_2 s}{h+r_2}\right) = \frac{R}{1 + x^{1/n}}.
\tag{2.19}
$$

Thus, consistency implies the following dependence for the variables C_1 and C_2:

$$
\begin{aligned}
C_1(I) &\equiv \frac{ac_1}{dhs} = \frac{\rho_1}{1 + \Lambda_1 I^{1/n}} \\
C_2(I) &\equiv \frac{c_2}{dh} = \frac{\rho_2}{1 + \Lambda_2 I^{1/n}},
\end{aligned}
\tag{2.20}
$$

and

where ρ_j, Λ_j, $j = 1,2$, are constants. The subsequent analysis will assume that most of the infection process is via exposure to asymptotic individuals, hence we will take $C_1(I) \gg C_2(I)$.

In order to proceed, we will make use of the following results. Let $\widetilde{A}$ and Q be defined in terms of A and D via

$$
\widetilde{A} = \frac{dhs}{a}A
\tag{2.21a}
$$

and

$$
Q = D^{(3)} + (F + R_2 + R_3)D^{(2)} + (FR_2 + FR_3 + R_2 R_3)D^{(1)} + FR_2 R_3 D - \alpha,
\tag{2.21b}
$$

Then, the functions $\widetilde{A}$ and D satisfy the two ODEs,

$$
\widetilde{A}^{(1)} + r_1\widetilde{A} = D^{(3)} + (R_2 + R_3)D^{(2)} + R_2 R_3 D^{(1)}
\tag{2.22}
$$

and

$$
\frac{Q^{(1)}}{Q} + C_1\widetilde{A} + C_2(D^{(2)} + R_3 D^{(1)}) = 0,
\tag{2.23}
$$

where

$$
F = a + s
$$

$$
\alpha = dhs(T - \xi),
\tag{2.24}
$$

with ξ a constant of integration.

Proof of equations (2.22) and (2.23): It is straightforward to verify that the l.h.s. of (2.14) equals $Q^{(1)}/dhs$, with Q defined in (2.21b). Furthermore, remarkably, equations (2.13) and (2.10) imply that the bracket of the r.h.s. of (2.14) is $-Q/dhs$. Then, (2.14) becomes (2.23). QED

2.3. Large time asymptotic analysis

With $\widetilde{A}$ and D satisfying equations (2.22) and (2.23), C_1 given by (2.20) and taking $C_2 <\!< C_1$, one can show that D approaches asymptotically the constant value D_∞, and the dependence of σ on t for large t is <u>not</u> exponential

$$D \to D_\infty(1 + \sigma),\ t \to \infty,\ \sigma \ll 1 \quad \sigma^{(j+1)} \ll \sigma^{(j)}, \quad j = 0,1,2,\ldots. \tag{2.25}$$

The leading behaviour is given by the following equations:

$$
\left.
\begin{aligned}
\sigma &\sim -\frac{(n-1)^{-1/(n-1)}}{R_2 R_3 D_\infty}\left(\frac{\mu}{\widetilde{\Lambda}_1}\right)^{-n/(n-1)} t^{-1/(n-1)}, \quad \mu = \frac{\rho_1}{r_1}, \\
I &\sim \frac{(n-1)^{-n/(n-1)}}{\mathrm{d}hs}\left(\frac{\mu}{\widetilde{\Lambda}_1}\right)^{-n/(n-1)} t^{-n/(n-1)} \\
\widetilde{A} &\sim \frac{(n-1)^{-n/(n-1)}}{r_1}\left(\frac{\mu}{\widetilde{\Lambda}_1}\right)^{-n/(n-1)} t^{-n/(n-1)}, \quad t \to \infty.
\end{aligned}
\right\} \tag{2.26}
$$

and

Proof of equations (2.26): Substituting (2.25) in equation (2.22), in the definition of Q (equation (2.21b)), and in equation (2.10c), we find the following equations:

$$\widetilde{A}^{(1)} + r_1 \widetilde{A} \sim \widetilde{\sigma}^{(1)}, \tag{2.27a}$$

$$\text{with} \quad \widetilde{\sigma} = R_2 R_3 D_\infty \sigma, \tag{2.27b}$$

$$Q \sim F(v + \widetilde{\sigma}), \quad v = R_2 R_3 D_\infty - \frac{\alpha}{F}, \tag{2.28}$$

$$\text{and} \quad I \sim \frac{\widetilde{\sigma}^{(1)}}{\mathrm{d}hs}, \quad t \to \infty \tag{2.29}$$

It turns out that equation (2.27a) implies

$$\widetilde{A} \sim \frac{\widetilde{\sigma}^{(1)}}{r_1}, \quad t \to \infty. \tag{2.30}$$

Indeed, integrating (2.27a) we find

$$\widetilde{A}\mathrm{e}^{r_1 t} = \text{constant} + \int^t \mathrm{e}^{r_1 \tau}\widetilde{\sigma}^{(1)}(\tau)\mathrm{d}\tau. \tag{2.31}$$

Integration by parts implies

$$\int^t \mathrm{e}^{r_1 \tau}\widetilde{\sigma}^{(1)}(\tau)\,\mathrm{d}\tau = \frac{\mathrm{e}^{r_1 t}\widetilde{\sigma}^{(1)}(t)}{r_1} + \text{const} - \frac{1}{r_1}\int^t \mathrm{e}^{r_1 \tau}\widetilde{\sigma}^{(2)}(\tau)\,\mathrm{d}\tau.$$

Substituting this expression in (2.31) and then multiplying by $e^{-r_1 t}$, we find (2.30).

Letting in equation (2.23) $C_2 = 0$ and then using the expressions for $Q, I, \widetilde{A}$ from (2.28), (2.29) and (2.30), we find

$$\sigma^{(1)}\left[1 + \frac{v + \tilde{\sigma}}{r_1}C_1\right] \sim 0, \quad t \to \infty.$$

Using in this equation, the expression of C_1 from (2.20) we obtain

$$1 + \frac{(v + \widetilde{\sigma})\mu}{1 + \widetilde{\Lambda}_1(\sigma^{(1)})^{1/n_1}} = 0, \quad t \to \infty$$

and

$$1 + v\mu + \widetilde{\Lambda}_1(\widetilde{\sigma}^{(1)})^{1/n_1} + \mu\widetilde{\sigma} = 0. \tag{2.32}$$

Hence, the requirement that σ decays implies

$$v\mu + 1 = 0, \quad \widetilde{\sigma}^{(1)^{1/n_1}} = -\frac{\mu}{\widetilde{\Lambda}_1}\widetilde{\sigma}. \tag{2.33}$$

Integrating (2.33*b*), we find

$$\widetilde{\sigma} = -(n_1 - 1)^{-\frac{1}{n_1 - 1}} \left(\frac{\mu}{\widetilde{\widetilde{\Lambda}}_1}\right)^{-n_1/(n_1-1)} t^{-1/(n_1-1)}, \quad t \to \infty. \tag{2.34}$$

Then, (2.27*b*), (2.29) and (2.30) yield (2.26).

QED

In the subsequent section, we will fit the full equations to the experimental data with goal to determine the various parameters.

We proceed with the assumption $C_2 = 0$. First, we will consider fitting the data to the algebraic birational formula in (1.3).

2.4. Bidirectional long short-term memory network

Bidirectional long short-term memory (BiLSTM) is a powerful generalization of recurrent neural networks that can capture long-term dependencies while at the same time avoiding the problem of vanishing/exploding gradients. The BiLSTM networks are well suited for time series prediction and can potentially completely capture the contextual information of the time series. The BiLSTM network employed in this work was introduced to predict the number of infected for different countries in [13].

3. Results

3.1. An algebraic formula and deep learning

We used data for the cumulative number of deceased, $D(t)$, and the total number of hospitalized patients, $H^c(t)$, from the first wave of COVID-19 in Portugal dating from March to September 2020. During this period, Portugal experienced relatively low numbers of infected, although a larger number of deaths compared with other countries with similar populations, such as Greece. The corresponding data were fitted using the two formulae

$$D(t) = \frac{D_F}{1 + \beta(1 + \delta t)^{-\gamma}}, \quad H^c(t) = \frac{H_F^c}{1 + \tilde{\beta}(1 + \widetilde{\delta}t)^{-\widetilde{\gamma}}}, \tag{3.1}$$

as well as the analogous formulae for the birational model.

It can be shown that the following useful formula is valid:

$$H^c = H + D + R_H, \tag{3.2}$$

and R_H denotes the number of people that recovered after they were hospitalized. We used H^c, as this was the more accurate data reported. The respective parameters, shown in table 1, were determined using data up to the middle of August and the simplex optimization algorithm.

We note that the above formulae provide an optimal fit, given that the use of the efficient deep learning algorithm, BiLSTM network [14], provided similar results. Figure 3 shows the performance of BiLSTM against the logistic and rational models for the number of diseased and hospitalized. Both BiLSTM and the rational had a similar strong correlation to the measured data ($R^2 \sim 0.999$), with the logistic expression resulting into a smaller value $R^2 \sim 0.98$.

3.2. The solution of the inverse problem

We solved the inverse problem by employing our new mechanistic model together with both the rational and the birational formulae.

We used the fact that

$$R_H^{(1)} = r_3 H,$$
$$H = \frac{D^{(1)}}{d}. \tag{3.3}$$

Hence,

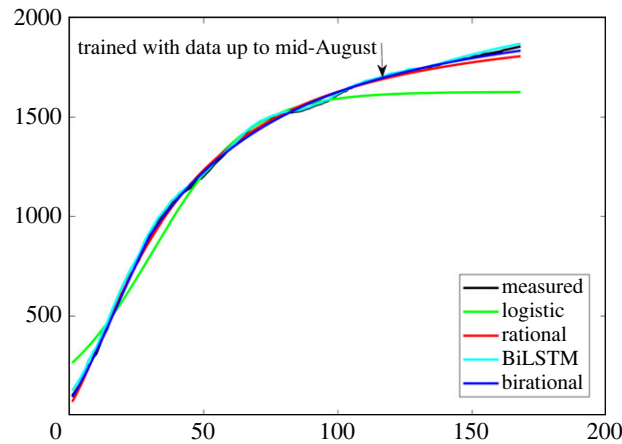$$R_H = \frac{r_3}{d} D + g. \tag{3.4}$$

Figure 3. Comparison of the BiLSTM, the logistic, the rational and the birational models for the number of deceased.

Table 1. Estimated parameters of equation (3.1) when fitted to the number of deceased, $D(t)$, and the total number of hospitalized patients, $H^t(t)$, for the first wave in Portugal. The confidence intervals for each fitting parameter are also presented.

		deceased	hospitalized
logistic	D_F	1626.5 (1576.2,1669.93)	$4.436 \times 10^{+04}$ (4.297×10^4, 4.552×10^4)
	β	5.429 (4.962, 5.867)	47.456 (45.106, 48.711)
	γ	0.056 (0.05, 0.061)	0.045 (0.040, 0.049)
	R2	0.9836 (0.976, 0.989)	0.978 (0.970, 0.986)
rational	D_F	1995.7 (1894.3,2077.3)	5.326×10^4 (5.145×10^4, 5.463×10^4)
	β	41.42 (37.33,44.12)	1.969×10^4 (1.901×10^4, 1.198×10^4)
	δ	0.298 (0.273,0.319)	0.339 (0.322, 0.353)
	γ	1.52 (1.404, 1.614)	2.845 (2.713, 2.926)
	R2	0.988 (0.983, 0.993)	0.987 (0.982, 0.992)
birational	D_{F_1}	2973.9 (2824.8, 3095.6)	$9.390 \times 10 \times 10^{+04}$ (9.076×10^4, 9.64×10^4)
	β_1	39.713 (37.105, 41.62)	760.037 (728.04, 781.36)
	δ_1	0.175 (0.156, 0.189)	0.029 (0.020, 0.038)
	γ_1	1.562 (1.476, 1.664)	4.554 (4.067, 4.890)
	T	30	81
	D_{F_2}	1912.5 (1815.5, 1990.1)	$2.285 \times 10^{+04}$ (2.201×10^4, 2.343×10^4)
	β_2	57.307 (54.4, 60.21)	$1.559 \times 10^{+04}$ (1.509×10^4, 1.611×10^4)
	δ_2	0.3796 (0.36, 0.41)	0.0150 (0.009, 0.020)
	γ_2	1.410 (1.33, 1.467)	9.286 (8.167, 9.971)
	R2	0.999 (0.998, 1)	0.990 (0.986, 0.994)
BiLSTM	R2	0.999 (0.998, 1)	0.998 (0.997, 0.999)

Equation (3.2) became equation (3.5), where we used $R_3 = r_3 + d$.

$$H^c = \frac{1}{d}(D^{(1)} + R_3 D) + g, \quad g \text{ constant.} \tag{3.5}$$

We used equation (3.1) to match (3.5), and thus, we obtained with the aid of the simplex algorithm, the values for parameters R_3, d and g, as shown in table 2.

To determine the remaining parameters, we first introduce the notation

$$X = D^{(2)} + (R_2 + R_3)D^{(1)} + R_2 R_3 D. \tag{3.6}$$

Table 2. Determination of the constant parameters in equation (3.5) and their respective confidence intervals.

	rational	birational
R_3	0.061 (0.055, 0.066)	0.0645 (0.0580, 0.0707)
d	9.232×10^{-04} (8.583×10^{-04}, 1.106×10^{-03})	0.0010 (8.655×10^{-04}, 1.158×10^{-03})
g	-76172 (-72860, -79483)	-77020 (-71932, 80785)

Table 3. Determination of the constant parameters in equation (3.7) and their respective confidence intervals.

	rational	birational
F	0.989 (0.955, 1.022)	0.992 (0.946, 1.036)
R_2	0.116 (0.097, 0.134)	0.131 (0.108, 0.147)
a	5716.6 (5344.4, 6056)	5352.7 (4932.5, 5833.4)
r_1	0.798 (0.701, 0.898)	0.861 (0.732, 0.977)
ρ_1	1.404 (1.22×10^{-04}, 1.57×10^{-04})	1.520–04 (1.370×10^{-04}, 1.64×10^{-04})
$\widetilde{\Lambda}_1$	0.131 (0.123, 0.139)	0.133 (0.124, 0.144)
n	4.002 (3.768, 4.385)	4.026 (3.734, 4.436)

Equations (2.22) and (2.23) read,

$$\widetilde{A}^{(1)} + r_1\widetilde{A} = X^{(1)},$$

$$\frac{X^{(2)} + FX^{(1)}}{X^{(1)} + FX - a} = -C_1\widetilde{A}$$

and

$$C_1 = \frac{\rho_1}{1 + \widetilde{\Lambda}_1 X^{1/n}}. \tag{3.7}$$

Equations (3.7), with the aid of the simplex algorithm, yield the values of the remaining constant parameters shown in table 3.

As expected, the value for n satisfies the constraint $n > 2$.

3.3. Numerical verification

Finally, we verified that the numerical solution of our mechanistic model matches, for large t, the rational formula. Similar considerations are valid for the birational model. For this purpose, we used the values of the various constants (tables 2 and 3), obtained from the experimental data. In this connection, it is interesting to note that from the whole set of equations (2.1)–(2.7), we only need to solve the following smaller set of four differential equations for the variables N, Q, $\widetilde{A}$, M:

$$N^{(1)} = Q + \alpha - FN, \tag{3.8a}$$

$$\frac{Q^{(1)}}{Q} = -C_1\widetilde{A} - C_2M, \tag{3.8b}$$

$$\widetilde{A}^{(1)} = -r_1\widetilde{A} + Q + \alpha - FN \tag{3.8c}$$

and

$$M^{(1)} = Q + \alpha - FN - R_2M, \tag{3.8d}$$

where

$$M(t) = R_3 D^{(1)}(t) + D^{(2)}(t)$$

and

$$N(t) = R_2(R_3 D(t) + D^{(1)}(t)) + R_3 D^{(1)}(t) + D^{(2)}(t). \tag{3.9}$$

We will take $C_2 = 0$ and $C_1 = \rho_1/1 + \widetilde{\Lambda}_1(Q + \alpha - FN)^{1/n}$.
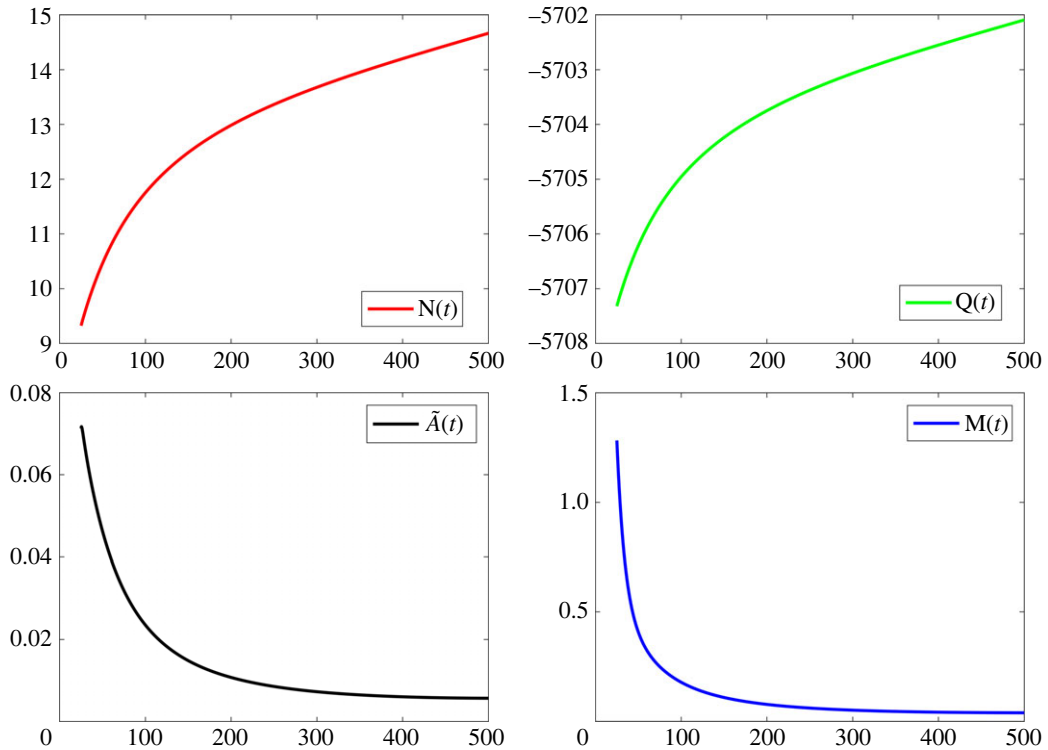
Figure 4. Illustration of the numerical solutions for the functions $M(t)$, $Q(t)$, $\widetilde{A}(t)$ and $N(t)$.

Proof of equations (3.8): The definitions of Q and N imply the identity

$$D^{(3)} + (R_2 + R_3)D^{(2)} + R_2R_3D^{(1)} = Q + \alpha - FN, \tag{3.10}$$

which is equation (3.8a). Also, using the definition of M, equation (2.23) becomes (3.8b). Similarly, using (3.10), equation (2.22) becomes (3.8c). Equation (3.8d) follows from the fact that the l.h.s. of (3.10) can be written in the form $M^{(1)} + R_2M$.

QED

Definition of equations (3.8) at some fixed t: The definition of N, Q, $\widetilde{A}$, M of equations (3.8) at some fixed t, denoted by t_0, imply the following formulae:

$$N(t_0) = R_2(R_3D(t_0) + D^{(1)}(t_0)) + R_3D^{(1)}(t_0) + D^{(2)}(t_0),$$

$$Q(t_0) = D^{(3)}(t_0) + (R_2 + R_3 + F)D^{(2)}(t_0) + (FR_2 + FR_3 + R_2R_3)D^{(1)}(t_0) + FR_2R_3D(t_0) - \alpha$$

$$M(t_0) = R_3D^{(1)}(t_0) + D^{(2)}(t_0). \tag{3.11}$$

Also,

$$\widetilde{A}(t_0) = -\frac{1}{C_1(t_0)}\frac{Q^{(1)}(t_0)}{Q(t_0)}, \tag{3.12}$$

where

$$C_1(t_0) = \frac{\rho_1}{1 + \widetilde{\Lambda}_1[D^{(3)}(t_0) + (R_2 + R_3)D^{(2)}(t_0) + (R_2R_3)D^{(1)}(t_0)]^{1/n_1}}. \tag{3.13}$$

Under the assumption that $C_2 = 0$ and having defined N, Q, $\widetilde{A}$, M at some fixed t_0, we solved equations (3.8) numerically using the Runge–Kutta method and the following expression for $D(t)$,

$$D(t) = \frac{1995.7}{1 + 41.42(1 + 0.298t)^{-1.52}}. \tag{3.14}$$
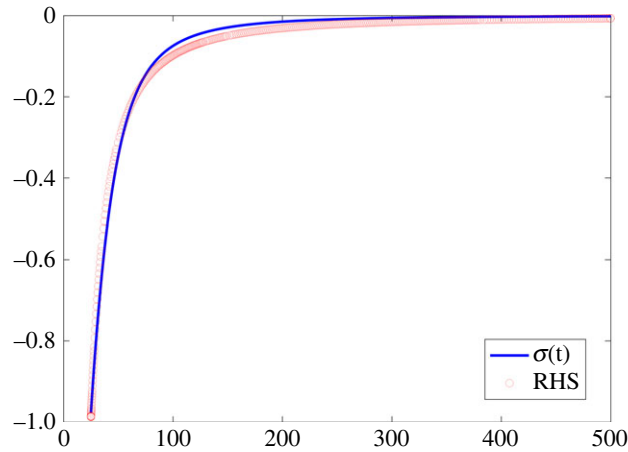
Results are shown in figure 4.

Figure 5. Illustration of the function $\sigma(t)$ and of the r.h.s. of equation (3.17).

The function σ can be determined via

$$D^{(1)}(t) + R_3 D(t) = \frac{N(t) - M(t)}{R_2},\tag{3.15}$$

which yields

$$D_\infty(\sigma^{(1)}(t) + R_3\sigma(t)) = \frac{N(t) - M(t)}{R_2} - R_3 D_\infty.\tag{3.16}$$

The solution of this equation with $\sigma(t_0) = D(t_0)/D_\infty - 1$ is depicted in blue in figure 5. By plotting in the same figure, the equation depicted in red, we find

$$\sigma \sim -\frac{\text{constant}}{t^m}, \quad t \to \infty.\tag{3.17}$$

It becomes clear that for large t, the function $\sigma(t)$ approaches the r.h.s. of (3.17).

The first of equations (2.33) implies the constraint $v\mu + 1 = 0$. Using in this equation the expressions for v and μ given in (2.26) and (2.28) we find the equation

$$\left(R_2 R_3 D_\infty - \frac{\alpha}{F}\right)\left(\frac{\rho_1}{r_1}\right) + 1 = 0.\tag{3.18}$$

Employing the values obtained via the solution of the inverse problem, it can be verified that equation (3.18) is indeed approximately satisfied: using for the constants appearing in the r.h.s. of (3.18) and the numerical values of the constants obtained in tables 2 and 3, we find that the r.h.s. of (3.18) is -0.014 (instead of zero).

4. Discussion

The forecasting models [2,3] suggest that the dynamics of the first wave of COVID-19 approach an equilibrium state algebraically, as opposed to the exponential behaviour normally predicted by the standard SEIR mechanistic models. Motivated by these results, we introduced a novel mechanistic model which for large times does exhibit algebraic behaviour. This model takes into consideration an additional nonlinear mechanism first understood in the context of SIR models in [9,10]. Specifically, in comparison with the standard SEIR models, such as those previously analysed in [11,12], we allowed for the values of the transmission constants to increase as the infection wanes, which is reflecting the social behavioural tendency to be less cautious as the infection diminishes. The specific relationship used fits well the actual data. As a result of this modification, the asymptotic analysis of the equations defining the new model yields the algebraic behaviour predicted by equations (3.8).

It should be emphasized that several authors have previously published papers addressing the need to incorporate sub-exponential/power law dynamics. Some of these papers are mentioned below. The main novelty of our SEIR-type model, in comparison with the earlier very interesting papers, is the introduction of an epidemiologically motivated mechanistic model which allows to determine the

associated large t-dynamics. Indeed, as we hope will become clear in what follows, the previous models concentrated on the early times behaviour. In particular, Chowell *et al.* [15], addressed early reactive behaviour changes associated with the 2014–2015 Ebola epidemic in West Africa, which exhibited a slower than exponential growth pattern. This was modelled by the equation

$$\frac{\mathrm{d}c}{\mathrm{d}t} = rc^p \qquad 0 \le p \le 1, \tag{4.1}$$

which can also be used to model the spread of a range of pathogens, including the epidemics of influenza, Ebola, foot-and-mouth disease, HIV/AIDS, plague, measles and smallpox. Equation (4.1) was also investigated by Viboud *et al.* [16].

An extended version of (4.1), namely, the equation

$$\frac{\mathrm{d}c}{\mathrm{d}t} = rc^p \left(1 - \frac{C^a}{k}\right) \qquad 0 \le p \le 1 \tag{4.2}$$

was investigated in [17].

Several modifications of the standard SIR epidemic model to support early sub-exponential growth dynamics are discussed by Chowell *et al.* [15], including the following: (i) metapopulation models, namely characterizing populations in terms of subpopulations (e.g. based on age, vulnerability, etc.), (ii) incorporating reactive behaviour changes by modelling a time-dependent transmission rate, and (iii) incorporating inhomogeneous mixing through a power law scaling parameter. These considerations motivated the introduction of the SIR type model,

and

$$\left.\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t} &= \frac{-\beta(t)SI}{N}, \\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \frac{\beta(t)SI}{N} - \gamma I \\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \gamma I. \end{aligned}\right\} \tag{4.3}$$

This model can be simply interpreted as an SIR model with a time-varying reproduction number.

The models mentioned above were used in the paper by Chowell *et al.* [18] in order to fit the early epidemiological data in a number of infections.

Another important contribution of our work is that it addresses a well-known difficulty associated with mechanistic models, namely the determination of the values of the various parameters that enter in the models. In this paper, we present a computationally efficient and robust approach for solving this inverse problem. Moreover, we show that knowledge of the numbers of deceased and hospitalized patients is sufficient to determine all the parameters of the model uniquely and accurately.

The numerical solution presented is useful for two reasons. First, the numerical results confirmed the validity of the asymptotic analysis. Indeed, the numerical solution matched well the asymptotic formula, $\sigma \sim t^{-m}$. Second, they confirmed the validity of the solution of the inverse problem. Indeed, the asymptotic analysis gives rise to the constraint $\nu\mu + 1 = 0$. Since ν and μ can be expressed in terms of the model parameters determined via the solution of the inverse problem, this equation provides a check of the accuracy of the solution of the inverse problem.

Concluding, it is worth noting that in addition to SEIR models several other types of models have been introduced in connection with COVID-19 [19–22]. It would be interesting to investigate whether any of these models exhibit large t algebraic behaviour.

References

1. Holmdahl I, Buckee C. 2020 Wrong but useful—what COVID-19 epidemiologic models can and cannot tell us. *N. Engl. J. Med.* **383**, 303–305. (doi:10.1056/NEJMp2016822)

2. Zhou P *et al.* 2020 A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273. (doi:10.1038/s41586-020-2012-7)

3. Fokas AS, Dikaios N, Kastis GA. 2020 Mathematical models and deep learning for predicting the number of individuals reported to be infected with SARS-CoV-2. *J. R. Soc. Interface* **17**, 20200494. (doi:10.1098/rsif.2020.0494)

4. Fokas AS, Dikaios N, Kastis GA. 2021 COVID-19: predictive mathematical formulae for the number of deaths during lockdown and possible scenarios for the post-lockdown period. *Proc. R. Soc. A* **477**, 20200745. (doi:10.1098/rspa.2020.0745)

5. Hogg RV, Klugman SA. 1984 *Loss distributions*, pp. 217–232. Wiley Series in Probability and Statistics. New York, NY: John Wiley & Sons, Inc.

6. Fokas AS, Kastis GA. 2021 SARS-CoV-2: The second wave in Europe. *J. Med. Internet Res.* **23**, e22431. (doi:10.2196/22431)

7. Mandelbrot BB. 1982 *The fractal geometry of nature*. San Francisco, CA: W. H. Freeman and Co.

8. Stauffer D, Aharony A. 1992 *Introduction to percolation theory*. London, UK: Taylor & Francis.

9. Luhar M, Oberai AA, Fokas AS, Yortsos YC. 2022 Accounting for super-spreader events and algebraic decay in SIR models. *Comput. Methods Appl. Mech. Eng.* **401**, 115286. (doi:10.1016/j.cma.2022.115286)

10. Ramaswamy H, Oberai AA, Yortsos YC. 2021 A comprehensive spatial-temporal infection model. *Chem. Eng. Sci.* **233**, 116347. (doi:10.1016/j.ces.2020.116347)

11. Fokas AS, Cuevas-Maraver J, Kevrekidis PG. 2021 Easing COVID-19 lockdown measures while protecting the older restricts the deaths to the level of the full lockdown. *Sci. Rep.* **11**, 5839. (doi:10.1038/s41598-021-82932-8)

12. Fokas AS, Cuevas-Maraver J, Kevrekidis PG. 2020 A quantitative framework for exploring exit strategies from the COVID-19 lockdown. *Chaos Solitons Fractals* **140**, 110244. (doi:10.1016/j.chaos.2020.110244)

13. Fokas AS, Dikaios N, Tsiodras S, Kastis GA. 2022 Simple formulae, deep learning and elaborate modelling for the COVID-19 pandemic. *Encyclopedia* **2**, 679–689. (doi:10.3390/encyclopedia2020047)

14. Graves A, Schmidhuber J. 2005 Framewise phoneme classification with bi-directional LSTM and other neural network architectures. *Neural Netw.* **18**, 602–610. (doi:10.1016/j.neunet.2005.06.042)

15. Chowell G, Sattenspiel L, Bansal S, Viboud C. 2016 Mathematical models to characterize early epidemic growth: a review. *Phys. Life Rev.* **18**, 66–97. (doi:10.1016/j.plrev.2016.07.005)

16. Viboud C, Simonsen L, Chowell G. 2016 A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks. *Epidemics* **15**, 27–37. (doi:10.1016/j.epidem.2016.01.002)

17. Shanafelt DW, Jones G, Lima M, Perrings C, Chowell G. 2018 Forecasting the 2001 foot-and-mouth disease epidemic in the UK. *EcoHealth* **15**, 338–347. (doi:10.1007/s10393-017-1293-2)

18. Chowell G, Viboud C, Simonsen L, Moghadas SM. 2016 Characterizing the reproduction number of epidemics with early subexponential growth dynamics. *J. R. Soc. Interface* **13**, 20160659. (doi:10.1098/rsif.2016.0659)

19. Kaxiras E, Neofotistos G, Angelaki E. 2020 The first 100 days: modeling the evolution of the COVID-19 pandemic. *Chaos Solitons Fractals* **138**, 110114. (doi:10.1016/j.chaos.2020.110114)

20. Kaxiras E, Neofotistos G. 2020 Multiple epidemic wave model of the COVID-19 pandemic: modeling study. *J. Med. Internet Res.* **22**, e20912. (doi:10.2196/20912)

21. Marmarelis VZ. 2020 Predictive modeling of COVID-19 data in the US: adaptive phase-space approach. *IEEE Open J. Eng. Med. Biol.* **1**, 207–213. (doi:10.1109/OJEMB.2020.3008313)

22. Olabode D, Culp J, Fisher A, Tower A, Hull-Nye D, Wang X. 2021 Deterministic and stochastic models for the epidemic dynamics of COVID-19 in Wuhan, China. *Math. Biosci. Eng.* **18**, 950–967. (doi:10.3934/mbe.2021050)