



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

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



Ziegler and Nichols meet Kermack and McKendrick: Parsimony in dynamic models for epidemiology

Michael Nikolaou*

Chemical and Biomolecular Engineering Department, University of Houston, Houston TX 77204-4004, USA

ARTICLE INFO

Article history:

Received 16 April 2021

Revised 30 October 2021

Accepted 26 November 2021

Available online 1 December 2021

Keywords:

COVID-19

Epidemics

Ziegler-Nichols

SIR

First-order-plus-time-delay

ABSTRACT

The COVID-19 crisis popularized the importance of mathematical modeling for managing epidemics. A celebrated pertinent model was developed by Kermack and McKendrick about a century ago. A simplified version of that model has long been used and became widely popular recently, even though it has limitations that its originators had clearly articulated and warned against. A basic limitation is that it unrealistically assumes zero time to recovery for most infected individuals, thus underpredicting the peak of infectious individuals in an epidemic by a factor of as much as about 2. One could avoid this limitation by returning to the original comprehensive model, at the cost of higher complexity. To remedy that, we blend Ziegler-Nichols modeling ideas, developed for automatic controller tuning, with Kermack-McKendrick ideas to develop novel model structures that predict infectious peaks accurately yet retain simplicity. We illustrate these model structures with computer simulations on real epidemiological data.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

The COVID-19 crisis popularized the importance of mathematical modeling for managing infectious disease epidemics (Adam 2020; Giordano et al., 2020; Jewell et al., 2020; Kucharski et al., 2020; Wang et al., 2020). Concepts such as R_0 , herd immunity, or flattening the curve entered the global vernacular shortly after the viral epidemic started its ominous exponential spread (Tufekci 2020). Underlying these concepts is seminal modeling work originating in the 1920s, most notably by Kermack and McKendrick (1927) (K-M) who contributed a remarkably general model to the mathematical theory of epidemics. That model partitions a fixed-size population into compartments S (susceptible to infection), I (infectious as a result of infection), and R (the remaining fraction of the population, comprising individuals removed from the I compartment by recovery or death or already being immune to the disease from the outset¹) and keeps track of the corresponding population fractions, s , i , r , over time by bal-

ancing loading and discharge (removal) rates for each compartment (Fig. 1). The result is a set of three general nonlinear integrodifferential equations (IDEs, APPENDIX A).

In developing their general IDEs, K-M monitor at time t the fraction $v_{t,\theta}$ of individuals in compartment I who have been infectious for a time period θ since their infection at time $t - \theta$, and express their removal rate from I as $\psi_\theta v_{t,\theta}$. They further simplify their IDEs by assuming that $\psi_\theta = \psi = \text{const.}$, which they combine with constant infectivity, ϕ , to get the celebrated SIR model (Anderson and May 1979; Anderson et al., 1992; Diekmann et al., 1995; Murray 2002; Keeling and Rohani 2008; Brauer and Castillo-Chavez 2012; Brauer 2017), comprising the equations

$$\dot{s}(t) = -\beta s(t)i(t) \quad (1)$$

$$\dot{i}(t) = (\beta s(t) - \gamma)i(t) \quad (2)$$

$$s(t) + i(t) + r(t) = 1 \quad (3)$$

where β, γ are constants (counterparts of the K-M parameters ϕ, ψ) related to infectivity and removal rates, and the ordinary differential equation (ODE) $\dot{r}(t) = \gamma i(t)$ may trivially be used in place of Eq. (3). The SIR Eqs. (1)-(3), as well as the general IDEs from which they were derived, lead K-M to state two important principles about (a) how epidemics spread ("No epidemic can occur if the [susceptible] population density is below this threshold value.") and (b) how epidemics end ("... from a particular relation

* Corresponding author at: Chemical & Biomolecular Engineering Department, University of Houston, 4226 MLK Blvd, Houston TX 77004, USA.

E-mail address: Nikolaou@uh.edu

¹ It is worth emphasizing two issues regarding the interpretation of R in the SIR acronym: (a) Even though R includes the notion of recovery, it does not stand for recovery alone, a fact already stipulated in the original K-M paper: "Each infected person runs through the course of his sickness, and finally is removed from the number of those who are sick, by recovery or by death.", and (b) R may also include individuals who are initially neither susceptible to the disease nor infectious, e.g. by natural or induced immunity.

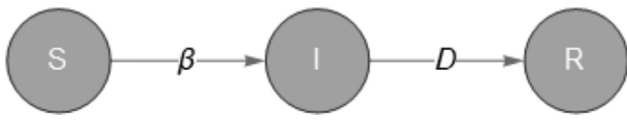


Fig. 1. The SIR model structure. Note that $D \triangleq 1/\gamma$.

between the population density, and the infectivity, recovery, and death rates.”). In fact, K-M showed that the corresponding values of both the threshold value for epidemic spread and the long-term values of the three fractions, s, i, r , of a population that just went through an epidemic are robust, namely not affected by the particular form of the kernels (infectious duration distributions) involved in the IDEs referred to above. The simplicity, intuitive appeal, and descriptive power of the SIR model made it enormously successful over the years, either as a stand-alone or as a module in numerous combinations and permutations of basic compartments and associated dynamics, as discussed by Hethcote (1994) who figuratively refers to about an order of magnitude below a myriad of epidemic models.

Recent use of the simple SIR model to guide the management of COVID-19 has brought to the forefront the importance of the above two K-M principles through two basic insights, with profound implications for designing interventions against this epidemic (Nikolaou 2021, and references therein): (a) to contain the spread of COVID-19 without any intervention, the susceptible fraction of the population would have to drop below 40%, and (b) to achieve this level of immunity in the population (above 60%) by contacting the disease (in the absence of a vaccine) would have meant that about 90% of the population would have to be infected throughout the epidemic. With mortality rate of the infected at about 0.5%, the immunity building option just described would have meant fatalities for about 0.5% of the population, an inordinately large number. This vital insight was provided by a simple model that is strictly speaking “wrong” yet utterly useful (Box 1979).

As important as the standard SIR model is, there is also a subtle shortcoming in Eqs. (1)–(3) that had not been fully appreciated before (Kemper 1980) but emerges naturally when one considers “Flattening the Curve” (Qualls et al., 2017; Ferguson et al., 2020)): As the infectious fraction, i , goes through a peak value i^* to subsequently decline asymptotically to zero, i^* is not accurately characterized by the SIR model. The values of i^* produced by the SIR and by the full IDEs may differ by as much as a factor of about 2, as pointed out by Nikolaou (2021), the essence of whose arguments are summarized below (Section 2.2). In the same fashion, the rise time to i^* suggested by the SIR model can also be commensurately inaccurate (Nikolaou 2021). These shortcomings are of conceptual and practical importance when designing epidemic management strategies to lower the infectious fraction peak, i^* , and thus lessen the burden on resources for treatment of infected patients in need of intensive care. More importantly, the issue extends to the numerous compartment-based model variants spawned by the SIR model for management of various kinds of infectious disease epidemics (Hethcote 1994) and even concerns efforts to capture the dynamics of epidemics by sophisticated model structures that may embed SIR modules (Horrocks and Bauch 2020).

It turns out that these shortcomings emerge from uncritically following the assumption of constant ψ_θ (in the original K-M notation), which intuitively corresponds to an exponential distribution of infectious duration (Fig. 2). An exponential distribution suggests that most infectious individuals are removed from I in zero time after infection. This is a realistically untenable proposition, which is not merely inaccurate but has important consequences, as we discuss next.

It can be easily demonstrated (Nikolaou 2021) that as the infectious duration probability density function (PDF) peaks away from $\theta = 0$ (corresponding to the cumulative density function (CDF) approaching a step, Fig. 2) the response of the original system of IDEs quickly approaches that of a system of delay differential equations (DDEs) as shown in Fig. 3. For the latter, Nikolaou (2021) developed a simple counterpart of the standard SIR model using Padé approximations for replacement of Eq. (2) of the SIR model by an equally simple ODE.

Nevertheless, situations arise (Wallinga and Lipsitch 2007) where the infectious duration CDF, albeit far from exponential, is still not close enough to a step and, as a result, DDEs (and, accordingly, their Padé approximations by ODEs mentioned above) may not be a reasonably accurate approximation of the general IDEs. The purpose of this publication is to develop a simple approximation for such a case. Specifically, we borrow from the field of automatic control, where a high-order system may be fruitfully approximated by a first-order-plus-time-delay (FOPTD) model. Ziegler and Nichols (Z-N) showed that the FOPTD model is convenient both to obtain from data and to use in PID controller tuning following the celebrated tuning rules developed by these authors (Ziegler and Nichols 1993). Applying the Z-N idea to the K-M IDEs, we develop a continuum of simple models which can accurately approximate the original general IDEs by simple ODEs that retain the simple SIR model structure and feature only one additional parameter easily obtainable from epidemiological data on infectious duration CDFs.

In the rest of the paper, we first provide a summary of relevant background facts and an overview of subsequent developments. We then show our main results. We use a case study based on relevant data easily available in literature to illustrate these results. Finally, we present a broader context, conclusions, and recommendations for further study.

2. Background and overview

2.1. The basic model for general infectious duration distributions

For a fixed-size population with initially susceptible, infectious, and immune fractions $s(0), i(0), r(0)$, respectively, one can stack the values of s, i, r and monitor their changes over discrete time by performing corresponding mass balances, as shown in Fig. 4 Nikolaou 2021). For infectious duration CDF, $\mathcal{F}(\theta)$, remaining invariant throughout an epidemic, the model resulting from Fig. 4 comprises Eqs. (1), ((3), and the IDE

$$r(t) = r(0) + \int_0^t (1 - r(0) - s(t - \theta)) \underbrace{\mathcal{F}'(\theta)}_{\{(\theta)\}} d\theta \quad (4)$$

(Cushing 1977; Hethcote 2000).

2.2. Models for specific infectious duration distributions

For $\mathcal{F}(\theta) \triangleq F(\gamma\theta) \triangleq F(\frac{\theta}{D}) = 1 - \exp(-\frac{\theta}{D})$ in Eq. (4), combination with Eqs. (1), (3) immediately yields the SIR model Eqs. (1)–(3) with

$$\gamma \triangleq 1/D \quad (5)$$

representing the average infectious duration and $\frac{\theta}{D}$ or, equivalently, $\theta\gamma$ representing dimensionless time. Similarly, for CDF being the delayed Heaviside (step) function $\mathcal{F}(\theta) \triangleq F(\gamma\theta) \triangleq F(\frac{\theta}{D}) = H(\frac{\theta}{D} - 1)$, as shown in Fig. 2, one can immediately get the algebraic delay equation (ADE) (Hethcote 2000)

$$r(t) = r(0) + (1 - r(0) - s(t - D))H(t - D) \quad (6)$$

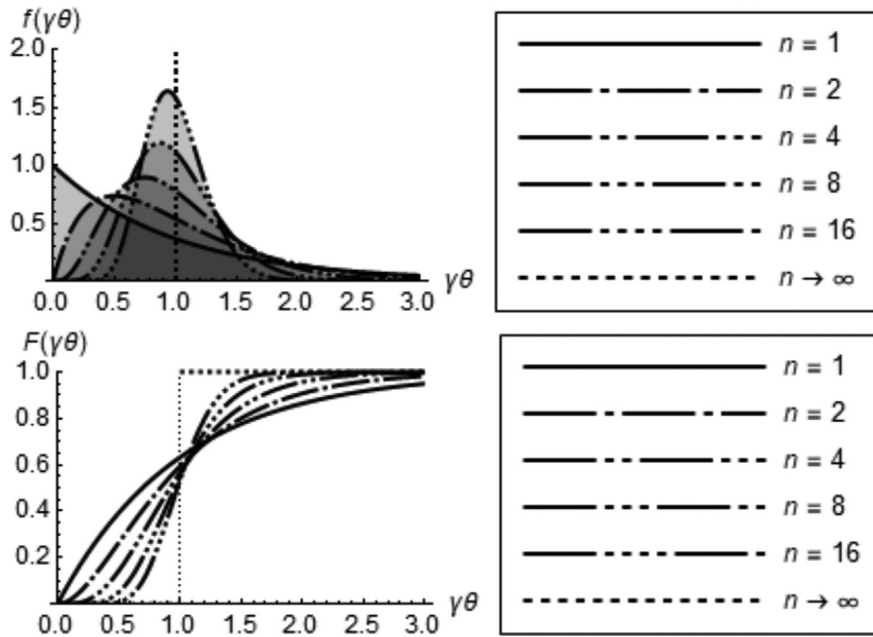


Fig. 2. Sample cumulative distribution functions $\mathcal{F}(\theta) \triangleq F(\gamma\theta)$ (top) and corresponding probability distribution functions $f(\theta) = \mathcal{F}'(\theta) \triangleq f(\gamma\theta) = F'(\gamma\theta)$ (bottom) for discharge time from the I compartment of a population, θ ($\gamma\theta$ in dimensionless form). Curves follow the formulas $F(\gamma\theta) = 1 - \frac{\Gamma(n, \gamma\theta)}{\Gamma(n, 0)}$ and $f(\gamma\theta) = F'(\gamma\theta)$, (see APPENDIX B). The exponential distribution with $F(\gamma\theta) = 1 - e^{-\gamma\theta}$ (top) and $f(\gamma\theta) = e^{-\gamma\theta}$ (bottom) corresponds to $n = 1$, whereas the impulse distribution with $F(\gamma\theta) = H(\gamma\theta - 1)$ (unit step at $\gamma\theta = 1$, bottom) and $f(\gamma\theta) = \delta(\gamma\theta - 1)$ (unit impulse at $\gamma\theta = 1$, top) correspond to $n = \infty$.

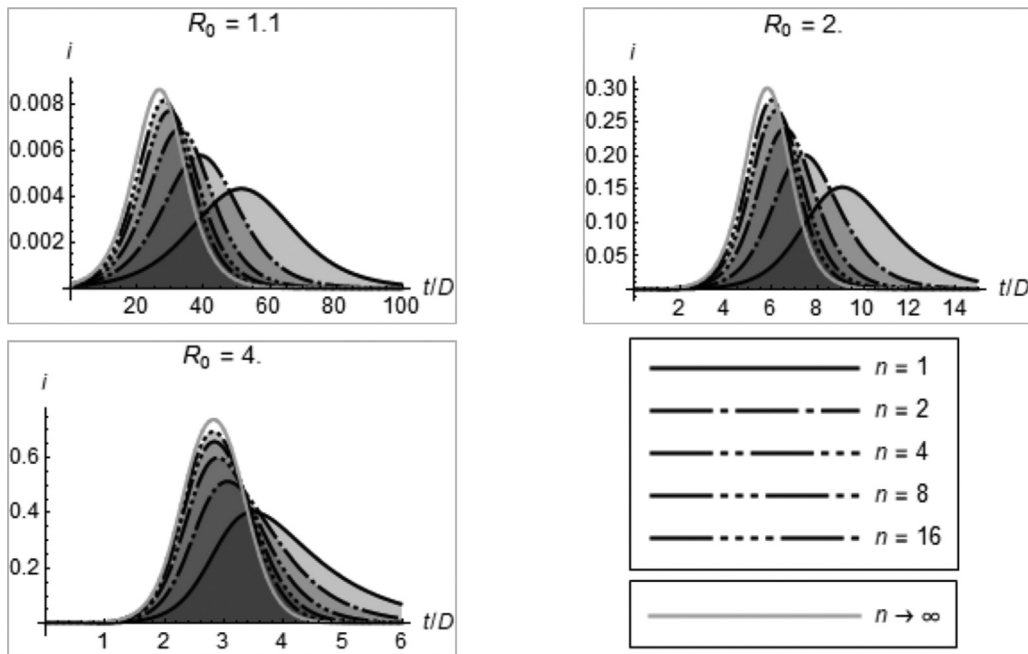


Fig. 3. Response of the infectious fraction, $i(t)$, according to the model of Eqs. (1), (4), and (3) for distributions shown in Fig. 2. Note that the distribution for $n = 4$ is closer to the exponential distribution ($n = 1$) than to the impulse distribution ($n = \infty$) in Fig. 2, yet the response of $i(t)$ for $n = 4$ is a lot closer to the response for $n = \infty$ rather than to that for $n = 1$.

which, along with Eqs. (1) and (3), constitutes a full dynamic model for s, i, r .

It was recently shown, using first- or second-order Padé approximations in the Laplace domain Nikolaou 2021), that the d-SIR model, comprising the above Eqs. (6), (1), and (3), can be well approximated by ODEs, namely Eqs. (1), (3), and either of the following two ODEs:

$$\dot{i}(t) = 2(\beta s(t) - \gamma)i(t) \tag{7}$$

or

$$\dot{i}'(t) = 12\gamma^2 \left(\frac{\beta}{\gamma} s(t)i(t) - i(t) - \frac{1}{2\gamma} i'(t) \right) \tag{8}$$

In addition to the exponential or step CDF discussed above, a variety of physically meaningful CDFs between these two extremes have been considered for F (Byrne et al., 2020). While these intermediate CDFs have a few minor differences, they all follow simple profiles, as shown in Fig. 2. It is for CDF profiles qualitatively

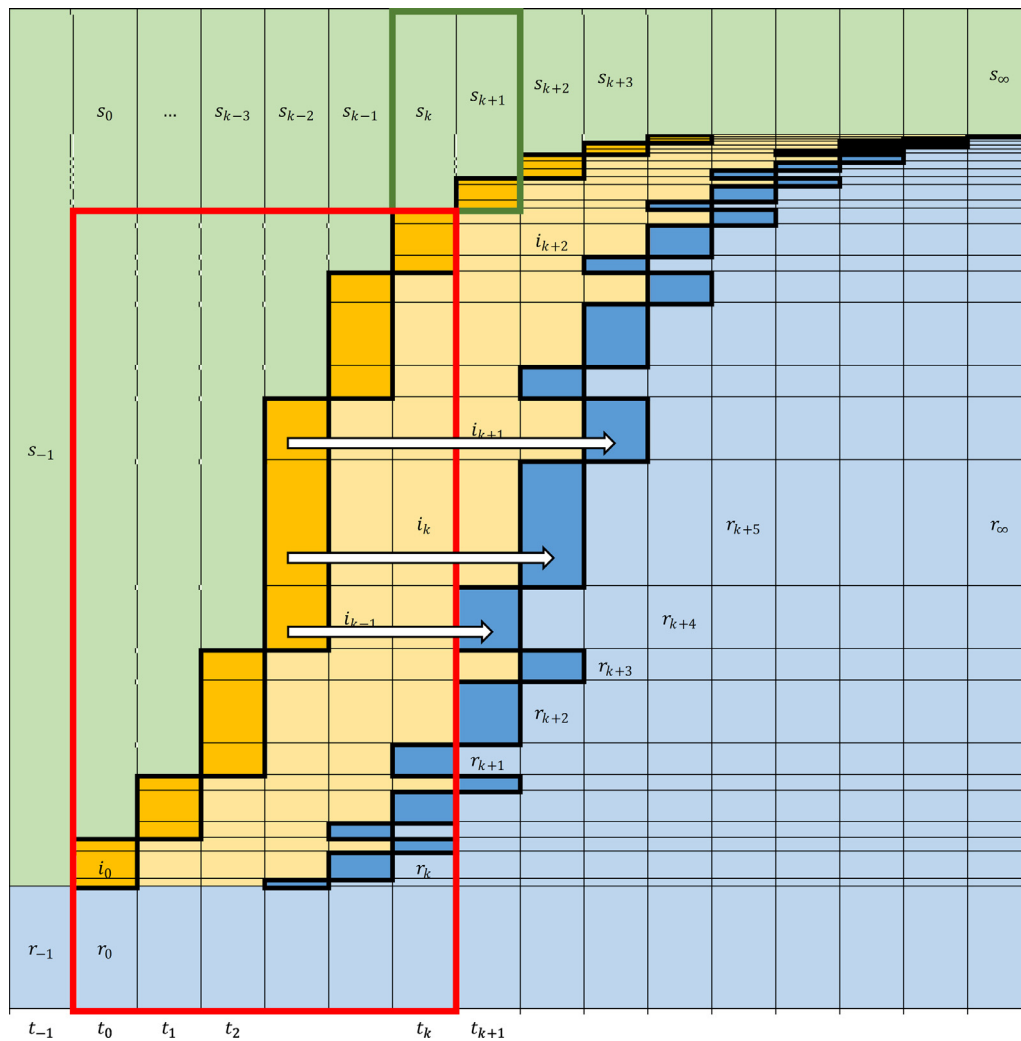


Fig. 4. Schematic of time-varying susceptible (*s*, green), infectious (*i*, orange), and removed (*r*, blue) fractions of a fixed-size population after an initial infection, i_0 , at discretized time t_0 (Nikolaou 2021). Each new part of the infectious fraction i (thick-framed orange rectangles) moves to the removed fraction, r (thick-framed blue rectangles) piecewise in a number of time steps, following a certain distribution of infectious duration (age). The population eventually reaches a steady state at $s_{\infty}, r_{\infty} = 1 - s_{\infty}$, and $i_{\infty} = 0$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

following such patterns that we show in the next section how to create simple approximations using Z-N ideas.

In fact, we show that the entire span of IDE models resulting from CDFs ranging from exponential to step (e.g. as shown in Fig. 2) can be approximated by a family of models similar, in form, to the SIR model Eqs. (1)-(3) with the simple addition of a corresponding time delay to Eq. (2). The delay renders the counterpart of Eq. (2) a DDE.

We further show that approximation can remarkably simplify that DDE to the simple ODE

$$\dot{i}(t) = \alpha(\beta s(t) - \gamma)i(t) \tag{9}$$

where the parameter α takes values in the interval $[1, 2]$ and approximately produces the range of responses for $i(t)$ shown in Fig. 3. More specifically, at the lower end of that interval ($\alpha = 1$) Eq. (9) trivially yields Eq. (2) of the standard SIR model (corresponding to exponential CDF) whereas at the upper end of that interval ($\alpha = 2$) Eq. (9) trivially yields the first-order Padé-SIR model comprising Eqs. (1), (7), and (3). Behavior between these two extremes can be captured by knowledge of this single parameter.

It should be stressed, as mentioned in Introduction, that the threshold value for epidemic spread and long-term values sug-

gested by all models discussed coincide. It is in transients where significant differences emerge.

2.3. The basic reproductive ratio

Making time dimensionless in the above models, as

$$\eta \triangleq t/D \tag{10}$$

or, equivalently, as $\eta \triangleq t\gamma$, trivially results in retention of only one parameter in the corresponding equations, namely $R_0 \triangleq \frac{\beta}{\gamma} \triangleq \beta D$. That parameter, implicitly suggested in the K-M paper and explicitly introduced as such later MacDonald 1957; Heesterbeek 2002; Brauer 2017) is called “the basic reproductive ratio” and is widely considered “one of the most critical epidemiological parameters” (Dietz 1993; Keeling and Rohani 2008). In fact, making time dimensionless in Eqs. (1) and (2) immediately yields

$$s'(\eta) = -R_0 s(\eta) i(\eta) \tag{11}$$

$$i'(\eta) = (R_0 s(\eta) - 1) i(\eta) \tag{12}$$

This suggests that the dynamics of the SIR model is essentially governed by a single parameter, R_0 (hence this parameter's paramount

importance, as already mentioned). This observation corroborates the use of dimensionless time in Fig. 2 and Fig. 3, a convention that will also be followed in the sequel.

The value of R_0 depends on many factors, both natural and resulting from human interventions, and is typically estimated from epidemiological data (Dietz 1993). Nikolaou (2021) presented a detailed analysis of the systematic errors that emerge in predictions by the SIR model if R_0 is estimated from data collected in the early exponential growth of an epidemic and is combined with estimates of the average infectious duration, $\frac{1}{\gamma}$. It is also noted that numerous refinements of R_0 may be considered in more elaborate models (the proverbial “alphabet soup” of models such as SEIR, SIR, and many others) that account for factors such as (a) compartments and corresponding interactions different from the standard S, I, and R (Hethcote 1994), (b) distinct subpopulations based on age, social contact structure, or other clustering factors (Anderson and May 1991; Anderson et al., 1992; Hethcote 1996; Ferguson et al., 2006; Keeling and Rohani 2008)), (c) variation in space (Keeling and Rohani 2008)), or (d) combinations thereof, which collectively lead to diverse stratification patterns (Kestenbaum 2019). For all such models, the term capturing discharge from a corresponding compartment typically corresponds to an exponential CDF of time from loading to discharge (even in advanced modeling exercises purporting to develop sophisticated tools such as automated algorithmic discovery (Horrocks and Bauch 2020)) with all limitations already elaborated on.

In the ensuing discussion we will consider that R_0 takes values above 1, because the “Threshold Theorem” of K-M, referred to in Introduction, asserts that an epidemic spreads iff $R_0 > 1$.

3. Main results

The main idea in the subsequent developments, previewed in summary in Section 2.2, is that the family of CDFs shown in Fig. 2 can be thought of, in approximation, as the unit-step response of a FOPTD model – the celebrated Z-N idea developed for control systems. We show next how this idea can be exploited to yield a dynamic model for s, i, r that approximates Eqs. (1), (4), and (3), and how further simplification by subsequent approximations may arise to span a broad spectrum of dynamics entailed by the K-M IDEs.

3.1. Z-N for K-M

Taking Laplace transforms in Eq. (4) with

$$\tilde{v}(q) \triangleq \mathcal{L}[\underbrace{r(t) - r(0)}_{v(t)}], \tilde{w}(q) \triangleq \mathcal{L}[\underbrace{1 - s(t) - r(0)}_{w(t)}] \quad (13)$$

and using the convolution theorem yields

$$\tilde{v}(q) = \tilde{w}(q)\tilde{f}(q) \quad (14)$$

where q is the Laplace domain variable in place of the most commonly used s (reserved here for the susceptible fraction of the population).

Because $f(t) = \mathcal{F}'(t)$ and $\mathcal{F}(t)$ can be thought of as a unit step response, as mentioned above, one can write the simple approximation

$$\tilde{f}(q) = \frac{\exp(-t_d q)}{(\tau q + 1)} \Leftrightarrow \tilde{\mathcal{F}}(q) = \frac{\exp(-t_d q)}{q(\tau q + 1)} \quad (15)$$

(Fig. 5) which, combined with Eq. (14), yields

$$\tau v'(t) + v(t) = w(t - t_d) \quad (16)$$

where $v(t) \triangleq r(t) - r(0)$.

Therefore, Eqs. (13) and (16) imply that the dynamics of s, i, r are captured by Eqs. (1), (3), and the simple DDE

$$\tau r'(t) + r(t) = 1 - s(t - t_d) \quad (17)$$

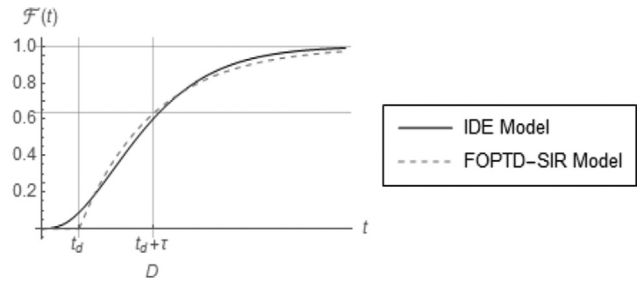


Fig. 5. Approximation of $\mathcal{F}(t)$ by the CDF $(1 - \exp(-\frac{t-t_d}{\tau}))H(t - t_d)$, the response of a system comprising a first-order element (of steady-state gain 1 and time constant τ) followed by a time delay (t_d) element to a unit-step-change on the system's input at time $t = 0$. Note that the average corresponding to the above CDF is the area between the CDF and 1 for $t \geq 0$, which can be shown to be $D = t_d + \tau$ (APPENDIX C).

3.2. The FOPTD-SIR model

Substituting $r(t)$ by $1 - s(t) - i(t)$ in Eq. (17) and using Eq. (1) immediately yields

$$i'(t) = \beta s(t)i(t) - \frac{1}{\tau}(i(t) + s(t) - s(t - t_d)) \quad (18)$$

The above Eq. (18) generalizes Eq. (2) of the standard SIR model in a direct and simple manner, through replacement of the term $-\gamma i(t)$ (corresponding only to an exponential distribution of infectious duration) by $-\frac{1}{\tau}(i(t) + s(t) - s(t - t_d))$ which can approximately account for a variety of infectious duration distributions. The resulting FOPTD-SIR model comprises Eqs. (1), (18), and (3). Note that Eq. (2) is trivially recovered from Eq. (18) for $t_d = 0$. Similar to Eqs. (11) and (12), which render the standard SIR model dimensionless, thus linking its dynamics to R_0 alone, the FOPTD-SIR model can be made dimensionless as well. This can be done by considering the natural counterpart of the standard SIR model parameter γ (inverse of the average of infectious duration when the latter follows an exponential distribution). This counterpart is the average, D , of infectious duration when the latter follows the CDF shown in Fig. 5 can be easily shown (APPENDIX C) to be

$$D = t_d + \tau \quad (19)$$

where t_d and τ are the dead time and time constant of the first-order lag, respectively. Making time dimensionless in Eqs. (1) and (18), and defining R_0 as

$$R_0 \triangleq \beta D \quad (20)$$

immediately yields

$$s'(\eta) = -R_0 s(\eta)i(\eta) \quad (21)$$

$$i'(\eta) = R_0 s(\eta)i(\eta) - \alpha(i(\eta) + s(\eta) - s(\eta - \eta_d)) \quad (22)$$

where

$$\alpha \triangleq 1 + \frac{t_d}{\tau} = \frac{D}{\tau} \geq 1 \quad (23)$$

and

$$\eta_d \triangleq \frac{t_d}{D} = \frac{\alpha - 1}{\alpha}, 0 \leq \eta_d \leq 1 \quad (24)$$

The above suggests that the dynamics of the FOPTD-SIR model is essentially governed by a single additional parameter, α , compared to R_0 of standard SIR dynamics. This is a remarkable simplification of dynamics representation by approximation, compared, for example, to representations based on full account of infectious duration distributions (Wallinga and Lipsitch 2007).

3.3. Properties of the FOPTD-SIR model

It is worth exploring the form of the FOPTD-SIR model for different values of α , given infectious duration distribution profiles such as shown in Fig. 2.

An exponential distribution of infectious duration ($n = 1$ in Fig. 2) corresponds to $t_d = 0$ and $\tau = D$, yielding $\alpha = 1$. This is the lowest possible value of α , corresponding to the standard SIR model, as immediately evident from Eqs. (21) and (22), which revert to Eqs. (11) and (12), respectively.

An impulse distribution of infectious duration (with unit step CDF, $n \rightarrow \infty$ in Fig. 2) corresponds to $t_d = D$ and $\tau = 0$, yielding $\alpha = \infty$, which by simple manipulation turns Eqs. (22) to the algebraic delay equation

$$i(\eta) = s(\eta - 1) - s(\eta) \tag{25}$$

thus resulting in a D-SIR model comprising Eqs. (11), (25), and (3). Incidentally, the above Eq. (25) can also be directly deduced from the graph in Fig. 4, with the infectious duration CDF being a step rather than S-shaped. In fact, we will make the case in section 3.6 that the behaviors of FOPTD-SIR models for $2 \leq \alpha \leq \infty$ virtually coincide, thus leaving the $1 \leq \alpha \leq 2$ as the range of relevant values for α .

The properties of the FOPTD-SIR model can be studied using methods for DDEs well studied in biology (Gopalsamy 1992; Kuang 1993). Even though DDEs – possessing infinite spectra and derivative discontinuities, among other features – may be considered more difficult to analyze than ODEs (Cushing 1977, p. 5) the corresponding theory “does not present substantial additional difficulties” (Bellen and Zennaro 2003, p. 6). Therefore, we will only present next two important properties of the FOPTD-SIR model, to confirm that this model is in agreement with the standard SIR model: The corresponding value for herd immunity suggested by the Threshold Theorem, in section 3.4; and the long-term values of s, i, r after an epidemic, which capture the total number of infections throughout an epidemic, hence the number of potential fatalities. To establish a difference between FOPTD-SIR and standard SIR, an analytical expression for characterization of the maximum of the infectious fraction along with the exponential growth rate leading to it will be developed through introduction of a further simplified SIR-like model, in section 3.6.

3.4. The threshold theorem and herd immunity for the FOPTD-SIR model

Applying standard stability theory for the FOPTD-SIR model linearized around the stationary point $\{\bar{s}, \bar{i} = 0, \bar{r} = 1 - \bar{s}\}$ using Taylor series yields (APPENDIX D) that the threshold value of \bar{s} , above which the epidemic is guaranteed to spread (i.e. the population does not possess herd immunity), is

$$\bar{s} = \frac{1}{\beta(\tau + t_d)} = \frac{1}{\beta D} \triangleq \frac{1}{R_0} \tag{26}$$

as anticipated by general K-M results.

3.5. Long-term values of s, i, r for the FOPTD-SIR model

It can be shown (APPENDIX E) that at the end of an epidemic ($i(\infty) = 0$) that started at $s(0) = 1 - \epsilon \approx 1$, $i(0) = \epsilon \approx 0$, and $r(0) = 0$, the total fraction of infected throughout the epidemic is

$$r(\infty) = 1 + \frac{W[-R_0 \exp(-R_0)]}{R_0} = 1 - s(\infty) \tag{27}$$

according to the FOPTD model, where W is the Lambert function (Kesisoglou et al., 2021). These values are the same as produced by the standard SIR model (Keeling and Rohani 2008).

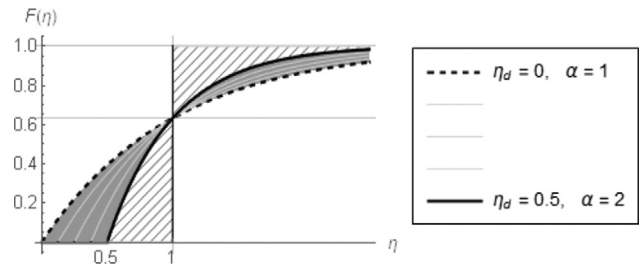


Fig. 6. Family of infectious duration CDFs of the form $F(\eta) = (1 - \exp(-\frac{\eta - \eta_d}{1 - \eta_d}))H(\eta - \eta_d)$ for $0 \leq \eta_d \leq 0.5$ or $1 \leq \alpha \leq 2$ (gray area). The dynamics of FOPTD-SIR or α -SIR models resulting from CDFs with $0.5 \leq \eta_d \leq 1$ (hatched area) are similar.

3.5.1. The α -SIR model

While Eq. (18) is a useful generalization of Eq. (2) (in that it accounts for a variety of infectious duration distributions other than exponential), it is in the form of a DDE, which, as we discussed, might be less appealing than an ODE for analytical computations. To get the analytical computation benefits of a simple ODE from Eq. (18), assuming no significant compromise of accuracy, one can show (APPENDIX G) that the following approximation of Eq. (18) can be used, which immediately turns out to be Eq. (9) presented in dimensionless form:

$$i'(\eta) \approx \alpha(R_0 s(\eta) - 1)i(\eta) \tag{28}$$

According to its definition in Eq. (23), the parameter α in the above Eq. (28) takes values in $[1, \infty]$, as the CDFs shown in Fig. 2 traverse the range from exponential ($n = 1$) to unit step ($n = \infty$). Fig. 3 makes the case that the resulting dynamics for a contiguous family of these CDFs, as the index n increases above a certain value towards ∞ , is tightly clustered very close to the dynamics resulting from the CDF corresponding to $n = \infty$, i.e. to the dynamics of the D-SIR model presented above. The D-SIR model has been analyzed in detail by Nikolaou (2021), who also introduced an approximate ODE version of the D-SIR model using first- and second-order Padé approximations, eventually leading to Eqs. (7) or (8), respectively. Simple inspection of the above Eqs. (28) and (7) immediately indicates that Eq. (28) for $\alpha = 2$ is exactly Eq. (7) (in dimensionless time). Therefore, the relevant range of values for the parameter α , over which appreciable differences can be seen in the dynamics of respective models, is

$$1 \leq \alpha \leq 2 \tag{29}$$

The corresponding CDFs for FOPTD-SIR models in dimensionless time are shown in Fig. 6. This figure indicates that all dynamics resulting from dimensionless dead time $0.5 \leq \eta_d \leq 1$ ($2 \leq \alpha \leq \infty$ by Eq. (24)) remain essentially indistinguishable, thus leading to Eq. (29); dimensionless dead times $0 \leq \eta_d < 0.5$ result in a range of distinguishable dynamics.

Eq. (28), combined with Eqs. (1)/(11), and (3), offers an exceptionally simple parametrization of an SIR-like model, which we call α -SIR, for brevity. That model approximates a wide range of epidemic spread dynamics, without necessarily involving the specific CDF of infectious duration, resorting instead to its FOPTD approximation.

Eq. (28) trivially suggests that the threshold value for s resulting from the α -SIR model is exactly the same as produced by the standard SIR model, for the same average of the infectious duration distribution. It is also easy to show (APPENDIX E) that the long-term values of s, i, r resulting from the α -SIR model are exactly the same as those produced by the standard SIR model. These findings confirm that α -SIR model provides the same insight as the standard SIR model regarding (a) herd immunity, and (b) projections of total number of infected throughout an epidemic and consequent

number of deaths. However the α -SIR model (therefore, by close approximation, the FOPTD model as well) provide better insight on how to “Flatten the Curve”, by providing more relevant representation of the infectious fraction peak, i^* , and the initial exponential rate of rise of i towards i^* . This is discussed next.

3.5.2. Peak of infectious fraction

It can be easily shown (Kermack and McKendrick 1927) that the peak value of the infectious fraction in the standard SIR model is $i_{SIR}^* = s(0) - \frac{\ln(R_0 s(0))}{R_0} - \frac{1}{R_0}$. A similar expression cannot be derived for the peak value suggested by the FOPTD-SIR model. However, the α -SIR model can be used to derive the approximate expression

$$i_{FOPTD-SIR}^* \approx i_{\alpha-SIR}^* = \alpha \left(s(0) - \frac{\ln(R_0 s(0))}{R_0} - \frac{1}{R_0} \right) = \alpha i_{SIR}^* \quad (30)$$

(APPENDIX F). The above approximation suggests that as α increases, the infectious fraction peak, $i_{FOPTD-SIR}^*$, of the FOPTD-SIR model can be as high as about twice the infectious fraction peak, i_{SIR}^* , of the SIR model (Nikolaou 2021) as demonstrated in Fig. 3.

3.5.3. Exponential growth rate of infectious fraction

All models discussed in this manuscript produce approximately exponential growth of the infectious fraction at early stages of a spreading epidemic. Expressions for these initial growth rates can be easily written following Taylor-series linearization of each model around a steady state (equilibrium point) comprising $s = \bar{s}$, $i = 0$, and $r = 1 - \bar{s}$. Taking the α -SIR model with $1 \leq \alpha \leq 2$ to represent the entire range of models considered here (from standard SIR for $\alpha = 1$ to D-SIR for $\alpha = 2$), linearizing Eq. (28) around the equilibrium point $\{\bar{s}, 0, 1 - \bar{s}\}$ immediately yields

$$\frac{i(\eta)}{i(0)} \approx \exp[\alpha(R_0 \bar{s} - 1)\eta] \quad (31)$$

for $i(0) = \epsilon \approx 0$ and $s(0) = \bar{s} - \epsilon \approx \bar{s}$. The above Eq. (31) implies that, for a certain $R_0 \bar{s}$, the standard SIR model, with a growth rate $R_0 \bar{s} - 1$, may underpredict the initial growth rate of the infectious fraction of the FOPTD-SIR model by as much as a factor of 2.

It is also of practical significance to examine the values of i^* predicted by the entire range of α -SIR models when values of R_0 are estimated from available data on early exponential growth of the infectious fraction. (The average infectious duration D can be estimated relatively precisely from epidemiological data (Keeling and Rohani 2008, section 2.1.1), thus time can be easily made dimensionless.) Eq. (31) indicates that the corresponding dimensionless-time growth rate is

$$r = \alpha(R_0 \bar{s} - 1) \quad (32)$$

Solving Eq. (32) for $R_0 \bar{s}$ and substituting into Eq. (30) for $s(0) \approx \bar{s} = 1$ yields

$$i_{\alpha-SIR}^* = \frac{r - \alpha \ln\left(\frac{r}{\alpha} + 1\right)}{\frac{r}{\alpha} + 1} \quad (33)$$

The ratio of $i_{\alpha-SIR}^*$ ($1 < \alpha \leq 2$) over i_{SIR}^* of the standard SIR model ($\alpha = 1$) is shown in Fig. 7. That figure indicates that as α moves from 1 towards 2 discrepancies between $i_{\alpha-SIR}^*$ and i_{SIR}^* increase, except on a single line where the $i_{\alpha-SIR}^*/i_{SIR}^*$ surface intersects the plane at 1.

3.5.4. Applicability range of approximations

While all approximations considered above invariably include approximation errors, they must also conform to the fundamental requirements $s'(\eta) \leq 0$, $r'(\eta) \geq 0$, and $0 \leq s(\eta), i(\eta), r(\eta) \leq 1$. We examine below the implications of these inequalities for the α -SIR model.

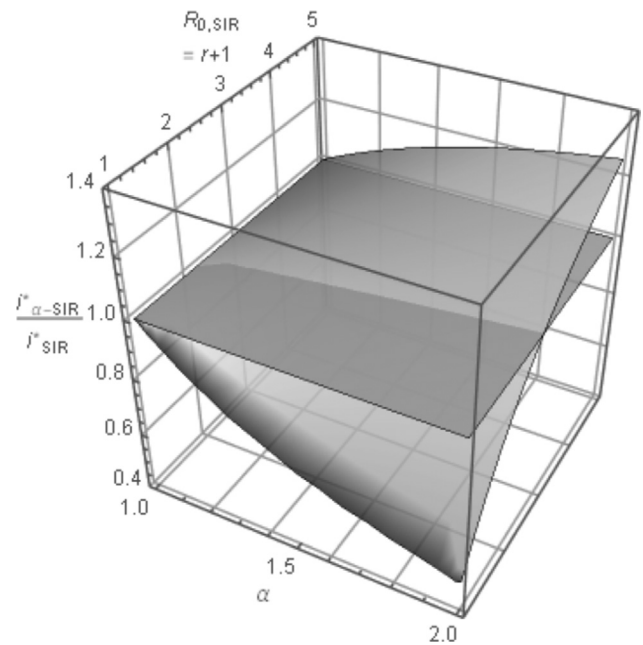


Fig. 7. Ratio of the maximum infectious fraction produced by the α -SIR model, $i_{\alpha-SIR}^*$, over the maximum infectious fraction produced by the standard SIR model, i_{SIR}^* , each tuning its corresponding value of R_0 based on the same rate of exponential growth of the infectious fraction.

The inequality $s'(\eta) \leq 0$ is trivially satisfied by Eq. (1)/(11), as long as $s(\eta) \geq 0$, $i(\eta) \geq 0$.

Satisfying the inequality $r'(\eta) \geq 0$ is more interesting, as Eqs. (28), (1)/(11), and (3), imply $r'(\eta) = -s'(\eta) - i'(\eta) = (\alpha - (\alpha - 1)R_0 s(\eta))i(\eta) \geq 0$, which requires

$$R_0 s(\eta) \leq \frac{\alpha}{\alpha - 1} = \frac{D}{t_d} \quad (34)$$

with $1 \leq \alpha \leq 2$. For $s(0) \approx 1$ the above inequality is guaranteed for $R_0 \leq 2$, which covers a range of infection spreads of practical significance, particularly if containment measures are taken. Nevertheless, $R_0 > 2$ in many practical situations as well. This would lead to $r'(\eta) < 0$ for initial times η , which is clearly infeasible. That situation can be handled by the simple α -SIR model adjustment

$$r'(\eta) = i(\eta) \max(0, \alpha - (\alpha - 1)R_0 s(\eta)) \quad (35)$$

and combination of the above Eq. (35) with Eqs. (1)/(11), and (3). The resulting dynamics will quickly revert to the original α -SIR model dynamics as $s(\eta)$ decreases and makes $\alpha - (\alpha - 1)R_0 s(\eta) \geq 0$ in Eq. (35).

The inequalities $0 \leq s(\eta), i(\eta), r(\eta) \leq 1$ can be handled in a similar manner. For example, Eq. (30) suggests that $0 \leq \alpha(s(0) - \frac{\ln(R_0 s(0))}{R_0} - \frac{1}{R_0}) \leq 1$.

A full investigation of the limits of the above investigations (e.g. for very large R_0) is beyond the scope of the current analysis, and will be pursued elsewhere. Nevertheless, the analysis presented above already covers important aspects of the proposed approximations. Additionally, the case study presented next indicates that the α -SIR model works remarkably well for a wide range of values of R_0 .

4. Case study

We use literature data analyzed by Anderson et al. (1986) to illustrate the concepts developed in the previous sections. The analysis presented therein is based on use of the standard SIR model for predictions of corresponding compartment sizes, even though

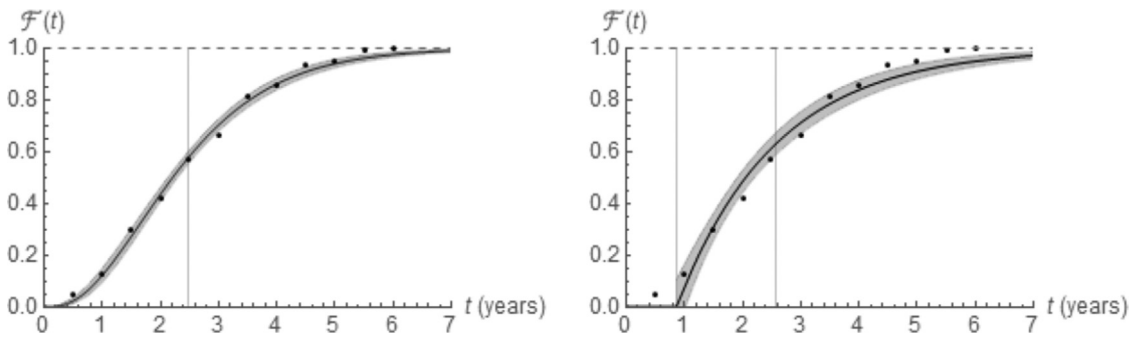


Fig. 8. Fitting of HIV epidemiological data (Anderson et al., 1986) by the CDFs $\mathcal{F}(t) = 1 - \Gamma(n, \frac{t}{D})/\Gamma(n, 0)$ with $n = 3.02 \pm 0.24$, $D = 2.47 \pm 0.04$, $R^2 = 0.999$ (left), $\mathcal{F}(t) = (1 - \exp(-\frac{t-t_d}{\tau}))H(t - t_d)$ with $t_d = 0.86 \pm 0.08$, $\tau = 1.72 \pm 0.13$, $R^2 = 0.996$ (right). 95% confidence bands for mean predictions are shown. All calculations and plotting performed on Mathematica (Wolfram Research Inc. 2021).

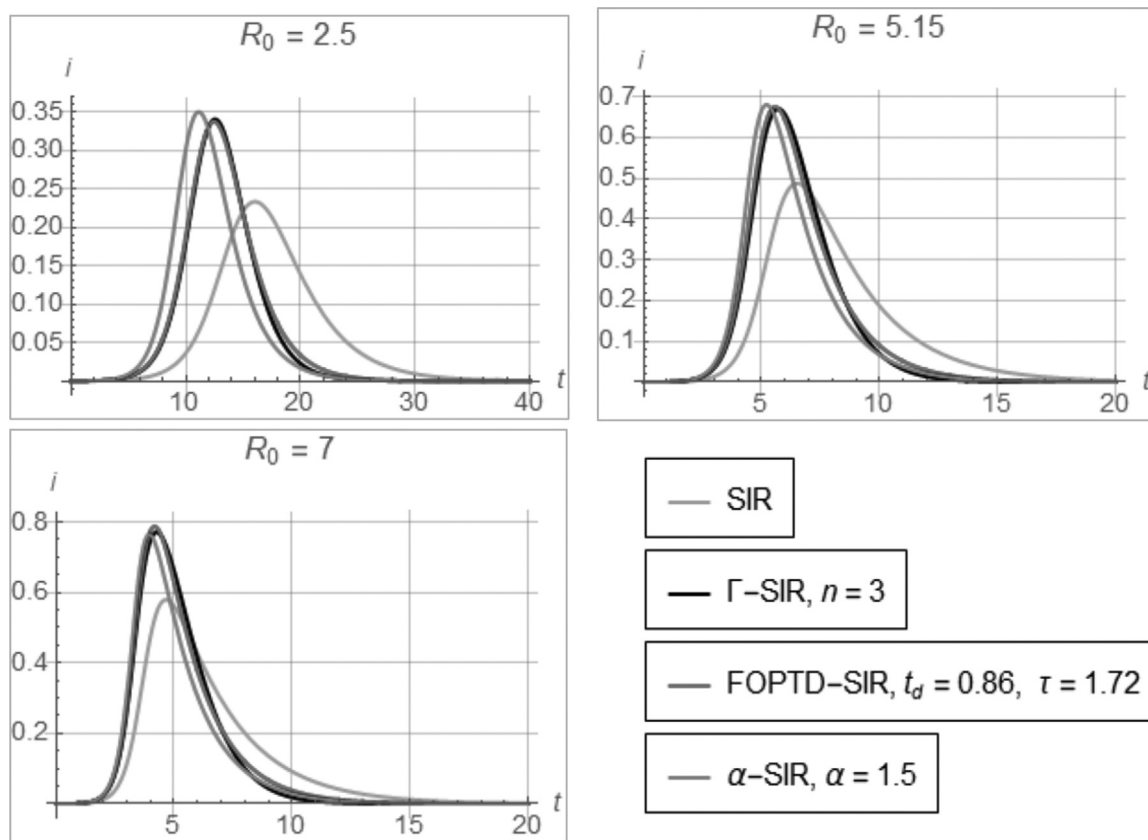


Fig. 9. Simulation results for cases (a)-(d) corresponding to different values of R_0 in Case Study.

data on a corresponding CDF is also shown. The data refer to the spread of HIV infection and development of AIDS. Specifically, the data shown in Fig. 8 refer to the infectious duration CDF governing the incubation period, namely the time from infection of previously susceptible individuals (entering the I compartment) to appearance of full-blown AIDS symptoms (individuals leaving the I and entering the R compartment). Note that entering the R compartment for this case refers to mere removal from the I compartment rather than recovery or death. The latter issue, as gravely important as it is, is not the focus here; if of interest, that issue can be modelled by augmenting this SIR model (Keeling and Rohani 2008, section 2.2); for a potential structure in series with SIR see Fig. 5 in Lipsitch et al. (2003). The data set of this case study is deliberately selected to correspond to an infectious duration CDF that is not close to a step, so that the approach developed here can be meaningfully applied.

Excellent data fit by the Gamma and FOPTD CDFs is shown in Fig. 8. It is noted that $n = 3$ for the Gamma CDF Fig. 2), suggesting that the neither the SIR model ($n = 1$) nor the Padé SIR model ($n = \infty$) is the most appropriate for this case, according to Fig. 3. Rather, the full IDE model (Eqs. (1), ((4), (3) featuring the data-fitted Gamma distribution in Fig. 8) or the corresponding FOPTD-SIR and the α -SIR counterpart are more appropriate. We illustrate these claims in the following simulation results for different values of R_0 :

- a The SIR model, Eqs. (1)-(3), with $\gamma = \frac{1}{2.5} \text{ y}^{-1}$.
- b The Γ -SIR model, Eqs. (1), (4), (3), with $D = 2.5 \text{ y}$, $n = 3$.
- c The FOPTD-SIR model, Eqs. (1), (18), (3), with $t_d = 0.86 \text{ y}$, $\tau = 1.72 \text{ y}$.
- d The α -SIR model, Eqs. (1), (28)/(9), (3), with $D = \frac{1}{\gamma} = 2.5 \text{ y}$, $\alpha = 1.5$.

The values of R_0 selected correspond to those used in Anderson et al. (1986) for rapid spread of the disease ($R_0 = 5.15, 7$) or somewhat tamed ($R_0 = 2.5$). Because $R_0 > 2$ Eq. (35) is used in all simulations.

A number of interesting patterns are evident in Fig. 9:

- There is obvious discrepancy between the profiles produced by the standard SIR model and the rigorous Γ -SIR model. This discrepancy is most important for the peaks of the infectious fraction, i^* , indicated by each of the two respective models, with the peaks rigorously produced by the Γ -SIR model being about 35 – 50% higher than their SIR counterparts. This is in agreement with Eq. (30) which suggests $i_{\text{FOPTD-SIR}}^*$ would be about 50% higher than i_{SIR}^* for $\alpha = 1.5$. The significance of this discrepancy for managing epidemics was already emphasized in Introduction.
- There is remarkable agreement between the infectious fraction profiles produced by the rigorous Γ -SIR model and the FOPTD-SIR model. This agreement underscores the value of the FOPTD-SIR model, which accurately captures model dynamics by mere addition of one easily interpretable parameter to the standard SIR model.
- Although somewhat less accurately matching the profiles of i rigorously produced by the FOPTD-SIR model, the α -SIR model is still clearly superior to the standard SIR model, while retaining simplicity. This advantage is afforded by the α -SIR model with only one additional parameter, α , whose value in the interval $[1, 2]$ qualitatively communicates the behavior of the α -SIR model between two extremes: The standard SIR model ($\alpha = 1$) and the d-SIR model ($\alpha = 2$).

5. Conclusions and discussion

Combining ideas from two seminal contributions in their respective fields of epidemiology (K-M) and automatic control (Z-N), we developed simple model structures that capture the basic dynamics of epidemics described by general, rigorous integrodifferential equations. The merit of the proposed modeling approach was illustrated via computer simulations on actual epidemiological data from the literature.

While the results presented were developed for a three-compartment, fixed-size population, they bear relevance to the broader class of compartment-based models that can capture the dynamics of epidemics for any infectious agent. Extensions, study of their properties, and testing on actual epidemiological data would certainly be of interest for future investigations.

The models proposed here are parsimonious, in that they retain a simple structure with the addition of only one parameter, which accepts intuitive interpretation. As such, the proposed models are easy to include in related software for practitioners or researchers (Obadia et al., 2012).

Of course, the proposed models are, by design, empirical, aiming at usefulness rather than ultimate accuracy of representation. Indeed, the message about the utility of models of such nature was emphasized in both publications from which we drew here: After they use their SIR model to fit data from a plague outbreak, Kermack and McKendrick (1927) state that

deductions as to the actual values of the various constants should not be drawn. It may be said, however, that the calculated curve, ..., conforms roughly to the observed figures.

Similarly, in developing a model useful for controller design, Ziegler and Nichols (1993) observe that

The difficulty of dealing mathematically with processes involving a series of lags ... is very great indeed. An approximate description of the characteristics of a process is given by values of ... two quantities.

... True, these two are only a rough measure of the entire reaction curve, ... but they give enough of the story to allow a prediction ...

There is a similar message here, which resonates with the well known dictum permeating all science and engineering: “All models are wrong, but some are useful” (Box 1979). We hope that the models proposed here will be useful.

Author statement

None.

Declaration of Competing Interest

None.

Acknowledgments

The Institute of Allergy and Infectious Diseases of the National Institutes of Health under award (grant number R01AI140287) supported the research reported in this publication. 100% of the project costs were financed with Federal money. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Funding sources had no involvement in study design; in the collection, analysis and interpretation of data, nor in the writing of the report nor in the decision to submit the article for publication.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.compchemeng.2021.107615.

Appendix A. The Kermack and McKendrick (1927) IDEs

$$v(t) = s'(t) = -s(t) \left[\int_0^t A(\theta) v(t-\theta) d\theta + A(t)i(0) \right] \quad (36)$$

$$r'(t) = \int_0^t C(\theta) v(t-\theta) d\theta + C(t)i(0) \quad (37)$$

$$i(t) = \int_0^t B(\theta) v(t-\theta) d\theta + B(t)i(0) \quad (38)$$

$$B(\theta) \triangleq \exp \left[-\int_0^\theta \psi(a) da \right] \quad (39)$$

$$A(\theta) \triangleq \phi(\theta)B(\theta) \quad (40)$$

$$C(\theta) \triangleq \psi(\theta)B(\theta) \quad (41)$$

Appendix B. Distribution functions in Fig. 2

To derive the formulas for $F(\theta)$ in Fig. 2, note that the inverse Laplace transform of $\tilde{F}(q) = \frac{1}{q(\frac{1}{n\gamma}q+1)^\pi}$ is the response of multiple first-order capacities in series to a unit-step input, which is well known to be $1 - \frac{\Gamma(n, n\gamma\theta)}{\Gamma(n, 0)}$ where $\Gamma(a, z) = \int_z^\infty t^{a-1} e^{-t} dt$.

Appendix C. Derivation and graphical interpretation of Eq. (19)

Given the infectious duration CDF

$$\mathcal{F}(t) = \begin{cases} 1 - \exp\left(-\frac{t-t_d}{\tau}\right), & t \geq t_d \\ 0, & t < t_d \end{cases}$$

(Fig. 5) the infectious duration average, D , is

$$D \triangleq \int_{-\infty}^{\infty} t\mathcal{F}'(t)dt = \int_{t_d}^{\infty} \frac{t}{\tau} \exp\left(-\frac{t-t_d}{\tau}\right) dt = t_d + \tau$$

In addition

$$D \triangleq \int_{t \rightarrow -\infty}^{t \rightarrow \infty} t\mathcal{F}'(t)dt = \int_{F \rightarrow 0^+}^{F \rightarrow 1^-} t dF$$

which suggests that D is the area between $\mathcal{F}(t)$, $t \geq 0$ and 1.

Appendix D. Stability analysis of the FOPTD-SIR model

Linearization of the Eqs. (1) and (18) around $\{\bar{s}, \bar{i} = 0\}$ yields

$$\Delta s'(t) = -\beta\bar{s}\Delta i(t) \tag{42}$$

and

$$\Delta i'(t) = \left(\beta\bar{s} - \frac{1}{\tau}\right)\Delta i(t) + \frac{1}{\tau}(\Delta s(t - t_d) - \Delta s(t)) \tag{43}$$

where Δ denotes deviation from the stationary state.

Taking Laplace transforms of Eqs. (42) and (43) yields

$$\begin{aligned} q\Delta\tilde{i}(q) - \Delta i(0) &= \left(\beta\bar{s} - \frac{1}{\tau}\right)\Delta\tilde{i}(q) + \frac{1}{\tau}(\exp(-t_d q) - 1)\Delta\tilde{s}(q) \\ &= \left(\beta\bar{s} - \frac{1}{\tau}\right)\Delta\tilde{i}(q) + \frac{1}{\tau}(\exp(-t_d q) - 1) \\ &\quad \times \frac{-\beta\bar{s}\Delta\tilde{i}(q) + \Delta s(0)}{q} \\ \Rightarrow (q^2 - \left(\beta\bar{s} - \frac{1}{\tau}\right)q + \frac{\beta\bar{s}}{\tau}(\exp(-t_d q) - 1))\Delta\tilde{i}(q) \\ &= \frac{\Delta s(0)}{\tau}(\exp(-t_d q) - 1) + \Delta i(0)q \end{aligned}$$

Therefore the characteristic equation is

$$\begin{aligned} q^2 - \left(\beta\bar{s} - \frac{1}{\tau}\right)q + \frac{\beta\bar{s}}{\tau}(\exp(-t_d q) - 1) &= 0 \Rightarrow q^2 - \left(\beta\bar{s} - \frac{1}{\tau}\right)q \\ + \frac{\beta\bar{s}}{\tau}\left(-t_d q + \frac{t_d^2 q^2}{2!} - \frac{t_d^3 q^3}{3!} + \dots\right) &= 0 \Rightarrow \left(1 + \frac{t_d^2}{2!}\right)q \\ + \left(-\beta\bar{s} - \frac{\beta\bar{s}}{\tau}t_d + \frac{1}{\tau}\right) + \frac{\beta\bar{s}}{\tau}\left(-\frac{t_d^3 q^2}{3!} + \dots\right) &= 0 \end{aligned}$$

By the Routh-Hurwitz theorem, instability occurs if not all coefficients of the above polynomial in q have the same sign. Because $1 + \frac{t_d^2}{2!} > 0$, it follows that

$$-\beta\bar{s} - \frac{\beta\bar{s}}{\tau}t_d + \frac{1}{\tau} < 0$$

guarantees instability, which immediately yields the threshold value in Eq. (26).

Appendix E. Long-term values of s, i, r for FOPTD and α -SIR

Combining Eq. (1) with Eq. (18) and taking Laplace transforms, \mathcal{L} , of both sides of the resulting equation (q being the Laplace variable rather than the usual s , which is the susceptible fraction here) with $i(t) = 0, s(t) = \bar{s}$ for $t < 0$ and $i(0) = \varepsilon, s(0) = \bar{s} - \varepsilon$ yields

$$\begin{aligned} i'(t) &= -s'(t) - \frac{1}{\tau}(i(t) + s(t) - s(t - t_d)) \\ \Rightarrow q\mathcal{L}[i] - \varepsilon &= -q\mathcal{L}[s] + \bar{s} - \varepsilon - \frac{1}{\tau}(\mathcal{L}[i] + \mathcal{L}[s] - \exp(-t_d q)\mathcal{L}[s]) \\ \Rightarrow \mathcal{L}[i] &= \frac{1}{\tau q + 1}(-\tau q - 1 + \exp(-t_d q))\mathcal{L}[s] + \frac{\varepsilon\bar{s}}{\tau q + 1} \end{aligned}$$

Now, manipulating Eq. (1), taking Laplace transforms, and substituting $\mathcal{L}[i]$ by the above expression yields

$$\frac{d \ln(s(t))}{dt} = -\beta i(t) \Rightarrow q\mathcal{L}[\ln s] - \ln(\bar{s} - \varepsilon)$$

$$\begin{aligned} &= \frac{\beta}{\tau q + 1}(\tau q + 1 - \exp(-t_d q))\mathcal{L}[s] - \frac{\beta\tau\bar{s}}{\tau q + 1} \\ \Rightarrow \lim_{q \rightarrow 0} q\mathcal{L}[\ln s] - \ln(\bar{s} - \varepsilon) \\ &= \beta \lim_{q \rightarrow 0} \frac{1 - \exp(-t_d q)}{q} \lim_{q \rightarrow 0} q\mathcal{L}[s] - \beta\tau\bar{s} \\ \Rightarrow \ln s(\infty) - \ln(\bar{s} - \varepsilon) &= \beta t_d s(\infty) - \beta\tau\bar{s} \end{aligned}$$

by the Final Value Theorem. For $\bar{s} - \varepsilon \approx \bar{s} = 1$ the last equation yields

$$\begin{aligned} \ln s(\infty) &\approx \beta(\tau + t_d)s(\infty) - \beta(\tau + t_d) = \beta D(s(\infty) - 1) \\ \Rightarrow \frac{1}{R_0} \ln s(\infty) - s(\infty) &\approx -1 \\ \Leftrightarrow \frac{1}{R_0} \ln(1 - r(\infty)) + r(\infty) &\approx 0 \end{aligned}$$

which is the standard equation for the total fraction of the population infected from the beginning to the end of the epidemic (Kermack and McKendrick 1927).

In the final step the last equation can be trivially transformed to

$$(-R_0 s(\infty)) \exp(-R_0 s(\infty)) = -R_0 \exp(-R_0(\infty))$$

which is of the form $x e^x = a$, thus immediately leading to the solution $x = W(a)$, where W is the Lambert function (Kesisoglou et al., 2021).

The α -SIR model also reaches the same result: Dividing Eq. (11) by Eq. (28) and rearranging yields

$$\begin{aligned} \int_{s(0)}^s \frac{R_0 s - 1}{R_0 s} ds &= -\alpha \int_{i(0)}^i di \Rightarrow s(\eta) - s(0) - \frac{1}{R_0} \ln\left(\frac{s(\eta)}{s(0)}\right) \\ &= \alpha(i(0) - i(\eta)) \end{aligned}$$

Taking the limit as $\eta \rightarrow \infty$ yields the result.

Appendix F. Peak of the infectious fraction for FOPTD-SIR and α -SIR

When reaching its peak at time t^* , the infectious fraction, i , satisfies $i'(t^*) = 0$, which, using Eq. (18), yields

$$\begin{aligned} \beta s(t^*)i(t^*) - \frac{1}{\tau}(i(t^*) + s(t^*) - s(t^* - t_d)) &= 0 \Rightarrow \beta s(t^*)i(t^*) \\ - \frac{1}{\tau}(i(t^*) + s'(t^*)t_d) &\approx 0 \Rightarrow \beta s(t^*)i(t^*) \\ - \frac{1}{\tau}(i(t^*) - \beta s(t^*)i(t^*)t_d) &\approx 0 \Rightarrow s^* = \frac{1}{\beta(\tau + t_d)} \stackrel{def}{=} \frac{1}{R_0} \end{aligned}$$

Now, combining Eq. (1) with the approximation

$$i'(t) \approx \beta s(t)i(t) - \frac{1}{\tau}(i(t) - \beta s(t)i(t)t_d) \tag{44}$$

of Eq. (18) yields

$$\begin{aligned} \frac{ds}{dt} &\approx \frac{-\beta s}{\beta s - \frac{1}{\tau}(1 - \beta s t_d)} \Rightarrow \left(1 + \frac{t_d}{\tau}\right) - \frac{1}{\beta\tau s} ds \approx -di \\ \Rightarrow \left(1 + \frac{t_d}{\tau}\right)(s - s_0) - \frac{1}{\beta\tau} \ln\left(\frac{s}{s_0}\right) &\approx -(i - i_0) \\ \Rightarrow i^* &\approx i_0 - \left(1 + \frac{t_d}{\tau}\right)(s^* - s_0) - \frac{1}{\beta\tau} \ln\left(\frac{s^*}{s_0}\right) \\ \Rightarrow i^* &\approx i_0 + \alpha\left(-\frac{1}{R_0} + s_0 - \frac{1}{R_0} \ln(R_0 s_0)\right) \end{aligned}$$

where $R_0 \triangleq \beta D = \beta(\tau + t_d)$.

For the case of no prior immunity, we have $s_0 \approx 1$ and the above equation yields

$$i^* \approx \alpha\left(-\frac{1}{R_0} + 1 - \frac{1}{R_0} \ln(R_0)\right)$$

Appendix G. The α – SIR model

Eq. (44) can be written as

$$\begin{aligned} i'(t) &\approx \beta \left(s(t) - \frac{1}{\beta\tau} (1 - \beta s(t)t_d) \right) i(t) \\ &= \beta \underbrace{\left(1 + \frac{t_d}{\tau} \right)}_{\alpha} \left(s(t) - \frac{1}{\beta(\tau + t_d)} \right) i(t) \stackrel{\text{def}}{=} \alpha \beta \left(s(t) - \frac{1}{R_0} \right) i(t) \\ &= \frac{\alpha}{D} (R_0 s(t) - 1) i(t) \Rightarrow i'(\eta) \approx \alpha (R_0 s(\eta) - 1) i(\eta) \end{aligned}$$

References

- Adam, D., 2020. The simulations driving the world's response to COVID-19. *Nature* 580, 316–318.
- Anderson, R., May, R., 1991. *Infectious Diseases of Humans. Dynamics and Control*. Oxford University Press.
- Anderson, R., Medley, G., May, R., Johnson, A., 1986. A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. *IMA J. Math. Appl. Med. Biol.* 3 (4), 229–263.
- Anderson, R.M., Anderson, B., May, R.M., 1992. *Infectious Diseases of humans: Dynamics and Control*. Oxford University Press.
- Anderson, R.M., May, R.M., 1979. Population biology of infectious diseases: Part I. *Nature* 280 (5721), 361–367.
- Bellen, A., Zennaro, M., 2003. *Numerical Methods for Delay Differential Equations*. Oxford.
- Box, G.E.P. (1979). Robustness in the strategy of scientific model building. *Robustness in Statistics* R. L. Launer and G. N. Wilkinson, Academic Press: 201–236.
- Brauer, F., 2017. Mathematical epidemiology: past, present, and future. *Infect. Dis. Modell.* 2 (2), 113–127.
- Brauer, F., Castillo-Chavez, C., 2012. *Mathematical Models in Population Biology and Epidemiology*. Springer.
- Byrne, A.W., McEvoy, D., Collins, A.B., Hunt, K., Casey, M., Barber, A., Butler, F., Griffin, J., Lane, E.A., McAloon, C., Brien, K., Wall, P., Walsh, K.A., More, S.J., 2020. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ Open* 10 (8), e039856.
- Cushing, J.M., 1977. *Integrodifferential Equations and Delay Models in Population Dynamics*. Springer-Verlag.
- Diekmann, O., H. Heesterbeek and H. Metz (1995). The legacy of Kermack and McKendrick. *Epidemic Models: Their Structure and Relation to Data*. D. Mollison.
- Dietz, K., 1993. The estimation of the basic reproduction number for infectious diseases. *Stat. Methods Med. Res.* 2 (1), 23–41.
- Ferguson, N.M., Cummings, D.A.T., Fraser, C., Cajka, J.C., Cooley, P.C., Burke, D.S., 2006. Strategies for mitigating an influenza pandemic. *Nature* 442 (7101), 448–452.
- Ferguson, N.M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunubá, Z., Cuomo-Dannenburg, G., Dighe, A., Dorigatti, I., Fu, H., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Okell, L.C., Emsland, S.v., Thompson, H., Verity, R., Volz, E., Wang, H., Wang, Y., Walker, P.G., Walters, C., Winskill, P., Whittaker, C., Donnelly, C.A., Riley, S., Ghani, A.C., 2020. Report 9 - Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand. Imperial College doi:10.25561/77482.
- Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., Colaneri, M., 2020. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nat. Med.*
- Gopalsamy, K., 1992. *Stability and Oscillations in Delay Differential Equations of Population Dynamics*. Springer, Netherlands.
- Heesterbeek, J.A., 2002. A brief history of R_0 and a recipe for its calculation. *Acta Biotheor.* 50 (3), 189–204.
- Hethcote, H.W., 1994. A Thousand and One Epidemic models. *Frontiers in Theoretical Biology*. S. A. Levin. Berlin. Springer-Verlag, pp. 504–515.
- Hethcote, H.W., 1996. Modeling heterogeneous mixing in infectious disease dynamics. In: Isham, V., Medley, G. (Eds.), *Models For Infectious Human Diseases: Their Structure and Relation to Data*. Publications of the Newton Institute, pp. 215–238.
- Hethcote, H.W., 2000. The mathematics of infectious diseases. *SIAM Rev.* 42 (4), 599–653.
- Horrocks, J., Bauch, C.T., 2020. Algorithmic discovery of dynamic models from infectious disease data. *Sci. Rep.* 10 (1), 7061.
- Jewell, N.P., Lewnard, J.A., Jewell, B.L., 2020. Predictive mathematical models of the COVID-19 pandemic: underlying principles and value of projections. *JAMA* 323 (19), 1893–1894.
- Keeling, M.J., Rohani, P., 2008. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press.
- Kemper, J.T., 1980. On the identification of superspreaders for infectious disease. *Math. Biosci.* 48 (1), 111–127.
- Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A* 115, 700.
- Kesisoglou, I., Singh, G., Nikolaou, M., 2021. The Lambert function should be in the engineering mathematical toolbox. *Comput. Chem. Eng.* 148, 107259.
- Kestenbaum, B., 2019. *Epidemiology and Biostatistics*. Springer.
- Kuang, Y., 1993. *Delay Differential Equations With Applications in Population Dynamics*. Academic Press.
- Kucharski, A.J., Russell, T.W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R.M., Sun, F., Jit, M., Munday, J.D., Davies, N., Gimma, A., van Zandvoort, K., Gibbs, H., Hellewell, J., Jarvis, C.I., Clifford, S., Quilty, B.J., Bosse, N.I., Abbott, S., Klepac, P., Flasche, S., 2020. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect. Dis.* 20 (5), 553–558.
- Lipsitch, M., Cohen, T., Cooper, B., Robins James, M., Ma, S., James, L., Gopalakrishna, G., Chew Suok, K., Tan Chorgh, C., Samore Matthew, H., Fisman, D., Murray, M., 2003. Transmission dynamics and control of severe acute respiratory syndrome. *Science* 300 (5627), 1966–1970.
- MacDonald, G., 1957. *The Epidemiology and Control of Malaria*. Oxford Univ. Pr, London.
- Murray, J.D., 2002. *Mathematical Biology: I. An Introduction*. Springer.
- Nikolaou, M., 2021. Revisiting the Standard for Modeling the Spread of Infectious Diseases. *Nature Scientific Reports Under Review*.
- Obadia, T., Haneef, R., Boëlle, P.-Y., 2012. The R_0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC Med. Inform. Decis. Mak.* 12 (1), 147.
- Qalls, N.L., Levitt, A.M., Kanade, N., Wright-Jegede, N., Dopson, S., Biggerstaff, M., Reed, C., Uzicanin, A., 2017. Community mitigation guidelines to prevent pandemic influenza – United States, 2017. *MMWR. Recommendations and Reports: Morbidity and Mortality Weekly Report*, 66. Recommendations and reports.
- Tufekci, Z., 2020. This Overlooked Variable Is the Key to the Pandemic. *The Atlantic* (September 30).
- Wallinga, J., Lipsitch, M., 2007. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc. R. Soc. B* 274 (1609), 599–604.
- Wang, C., Liu, L., Hao, X., Guo, H., Wang, Q., Huang, J., He, N., Yu, H., Lin, X., Pan, A., Wei, S., Wu, T., 2020. Evolving Epidemiology and Impact of Non-Pharmaceutical Interventions on the Outbreak of Coronavirus Disease 2019 in Wuhan, China. *medRxiv* 2020.2003.2003.20030593.
- Wolfram Research Inc, 2021. *Mathematica*. Version 12.3.1, Champaign, IL.
- Ziegler, J.G., Nichols, N.B., 1993. Optimum settings for automatic controllers. *J. Dyn. Syst. Measur. Control-Trans. ASME* 115 (2B), 220–222.