

A flexible COVID-19 model to assess mitigation, “reopening”, virus mutation and other changes

Sergio Bienstock, PhD (Retired)

Past Affiliations: Deutsche Bank, New York, NY, USA; Salomon Brothers Inc, New York; AT&T Bell Labs, Murray Hill, NJ; Harvard-Smithsonian Center for Astrophysics, Cambridge, MA; Harvard University, Cambridge, MA

Abstract

The COVID-19 epidemic which began in China last year has expanded worldwide. A flexible SEIRD epidemiological model with time-dependent parameters is applied to modeling the pandemic. The value of the effective reproduction ratio is varied to quantify the impact of quarantines and social distancing on the number of infections and deaths, on their daily changes. and on the maxima in these daily rates expected during the epidemic. The effect of changing R_{eff} is substantial. It ought to inform policy decisions around resource allocation, mitigation strategies and their duration, and economic tradeoffs. The model can also calculate the impact of changes in infectiousness or morbidity as the virus mutates, or the expected effects of a new therapy or vaccine assumed to arrive at a future date. The paper concludes with a discussion of a potential endemic end of COVID-19, which might involve times of about 100 years.

Introduction

The epidemic of novel coronavirus disease (COVID-19) that began in China in late 2019 has expanded rapidly to over 220 countries and all U.S. states, and upended the lives and livelihoods of much of the world’s population. Over ten million cases and more than 500,000 deaths have been reported worldwide, of which 2.7 million and 128,000 respectively in the

United States as of July 1, 2020 (1). The pandemic and associated community mitigation measures, have had a large negative economic impact as well. The virus could trim global economic growth by 3.0% to 6.0% in 2020, with a partial recovery expected in 2021 (2). Unemployment in the United States temporarily surged to levels not seen since the 1930s. Further, the epidemic is expected to exacerbate global poverty. It might push millions of people in the developing world, such as in India, Sub-Saharan Africa and South Asia, into extreme poverty (3).

Planning for the pandemic, keeping supply chains open, caring for the sick, and searching for effective therapies and vaccines have diverted the attention of healthcare professionals, researchers, planners, and essential workers around the world. Institutions such as the CDC in the United States are developing and refining COVID-19 scenarios designed to inform decisions by public health and other government officials (4). These aim at helping to evaluate the effects of mitigation strategies such as quarantines and social distancing, as well as with hospital resource allocation. Each scenario is based on a set of numerical values for the biological and epidemiological characteristics of COVID-19. These values — called *parameter values* — can be used in mathematical and statistical models to estimate the possible effects of the epidemic in different regions. They are being updated and augmented over time, as more is learned about the epidemiology of COVID-19.

Here I outline a differential equation-based SEIRD model with time-varying parameters and apply it to COVID-19 epidemic data. I discuss some of the potential implications of community mitigation strategies and their weakening in order to reopen the economy. This modeling approach is able to follow the entire development of an epidemic, from inception to eradication or endemic equilibrium, if desired. The model helps quantify tradeoffs in terms of

additional cases and deaths in a given region and their maximum rates of increase, expected to result from a weakening of social distancing, or “reopening”. It can also be used to assess the impact of changes in a pathogen’s infectiousness or morbidity over time.

Model and parameters

SIR and SEIR compartmental epidemiological models have been studied for nearly one hundred years. The term “compartmental” may have originated by analogy to trains or ships, essential modes of transportation in the first part of the twentieth century. In the non-autonomous SEIRD model discussed here, a population is divided into five non-overlapping classes, known as *compartments*:

- S, susceptible hosts;
- E, exposed hosts, presumed to be latently infected but not yet infectious;
- I, infectious hosts;
- R, hosts recovered from the exposed and infectious population and,
- D, hosts deceased due to the infection

This approach leads to a system of five ordinary differential equations in the compartmental variables, with time-dependent model parameters as coefficients. Shown in the Appendix, the equations describe the time evolution of each of the above-mentioned population groups. Key inputs to the model, defined in the Appendix, are the pathogen’s *basic reproduction number*, R_0 , or a closely related parameter, the *effective reproduction ratio*, R_{eff} . From these the transmission rate from Susceptible to Exposed is calculated. Other required data are the transition rates from each group to the next. An E host can either get infectious (move to I) or recover (move to R), and an I host can either move to R (recover) or move to D (die from the disease). All Recovered hosts are assumed to be immune, or “removed” from the epidemic, although it is possible to

relax this assumption. Hosts in all groups but D can also die for reasons unrelated to the infection, at an average rate. Finally, a gross birth rate for the Susceptible group is specified, and often set to the same value as the death rate unrelated to the infection (“steady state”). Such a SEIRD model would be described as an *open-population model in steady state*. All of the model parameters can vary with time, which affords the model a great deal of flexibility.

The starting point of the calculation (“time zero”) is specified by the fraction of the population, I, infectious at that time (typically very small, e.g. one per million) while S is nearly 1.0 for a new epidemic. Any other starting point is also possible. Expressed as fractions of the overall population, $S + E + I + R + D$ should add up to 1.0 at all times for steady-state models. This provides a check on the accuracy of the numerical solution. All calculations were performed in the **R** Statistical Environment (5). The differential equations were easily integrated out to 100 years using the `ode` function in **R** (6) (7) (8) in order to explore the long-term behavior of the solutions. For practical uses, the focus is on the first year or two.

A rapidly growing set of reports on the COVID-19 infection parameters is becoming available online. Table 1 lists those used in this study (9) (10) (11) (12) (13) (14) (15) (16)

[Table 1 about here]

The model allows the study of an epidemic from beginning to end. All of the model parameters can vary with time to reflect, for example, a change in social distancing, a mutation that alters the pathogen’s infectiousness, or a reduction in morbidity expected from a new therapy available as of some point in the future. Coupled with a modern differential equation solver (8), the model generates accurate results in seconds, which makes it feasible to perform sensitivity analyses with respect to one or more of the parameters used, as shown below in

Figure 3. Models based on local data could be run and the results easily aggregated as needed, in order to increase the granularity of predictions as data availability permits.

Results

Figures 1 and 2 present the results of two runs with identical parameters, except as follows. In Figure 1 there is no attempt to slow down the infection rate with quarantine or social distancing, and R_{eff} is 4.0 throughout. The run in Figure 2 uses a R_{eff} of 4.0 up to 0.2 years (2.4 months) into the epidemic, at which point it goes down to 2.0 until 0.4 years (4.8 months), to simulate the potential effect of a quarantine extending for a bit over two months. A modified reopening with R_{eff} of 3.0 is assumed for all subsequent times. Shading is used to highlight the three time periods. The numbers inside each plot show the maximum value and/or end value of the dependent variable, as relevant. The last two plots in each figure are examples of *phase-space plots*, or *phase portraits* (17) (18); see also (19) and are discussed near the end of the paper.

Each plot in Figure 2 should be compared with the corresponding plot in Figure 1. This shows that, applied at an opportune time, the quarantine succeeds in slowing down the infection (“flattening the curve”). The maximum values in the Exposed, Infected and daily death rate categories, all decrease by better than 50%. The impact on cumulative deaths and Recovered is less marked, as the epidemic is slowed down but not stopped (since R_{eff} remains greater than 1.)

Figure 3 shows the variation in the maximum daily infection rate with R_{eff} and with the time at which this maximum is estimated to occur. These are the curves in red. Because the infection probability increases with R_{eff} , the larger maxima occur at correspondingly shorter times. The curves in blue show similar results for the estimated maximum daily death rate, or new deaths per day. As can be seen from Figure 3 and Table 2, the effect of a change in R_{eff} on these maximum daily rates is quite large. This type of information could be used to assess whether a

relatively early weakening of quarantine and social distancing in a given region, leading to a substantial increase in R_{eff} , would result in an acceptable number of additional casualties.

[Table 2 about here]

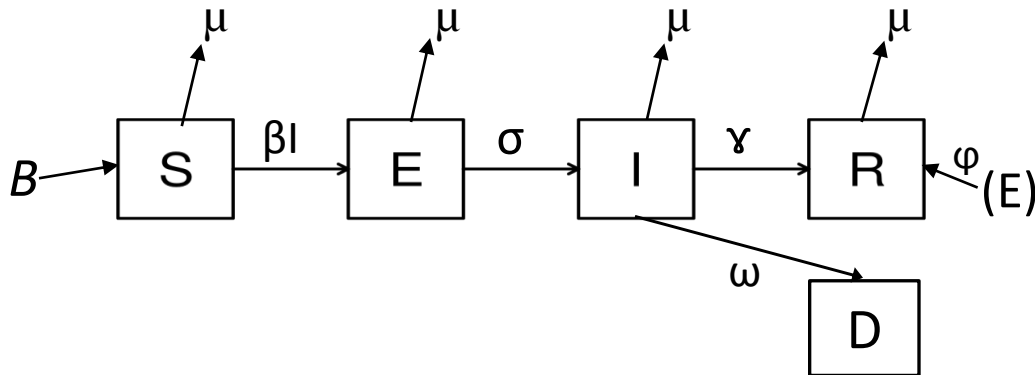
Figure 4 depicts the results of the run at $R_{\text{eff}} = 4.0$ when integrated out to 100 years. The use of logarithmic time as covariate allows the plot to cover a long time period and still show the early part of the epidemic in detail. Beginning at time near 30 years, *the model predicts a series of increasingly less severe resurgences of the epidemic*. Periodic and seasonal epidemic recurrences are well known and have been studied using SIR-class models, both autonomous and demographically forced (17) (18). The timing of the COVID-19 recurrences observed here varies with the birth and death rates used, B and μ . Why? Every year a certain proportion of the population dies for reasons unrelated to COVID-19, and a similar number of people are born, all of which into the Susceptible population — certainly true after maternal immunity, if any, disappears. Assume the disease is not eradicated, the pathogen has not lost its virulence, and no effective vaccine has been found. In time the Susceptible group becomes large enough for the virus to propagate again, albeit at a slower pace since some of the hosts it would encounter would be immune. Eventually the proportion of Susceptible and Infected hosts approach constant (equilibrium) values, at which point the epidemic becomes endemic. This is shown in the bottom right-hand plot of Figure 4, a phase portrait, where time increases as the curve is traced in a counterclockwise direction. Over the years the epidemic behaves like a damped oscillator, whose dampening is related to the variation in “herd immunity” over time. Endemic equilibrium is reached in approximately 100 years. Doubling B and μ would decrease the time to equilibrium by about half.

Discussion

A SEIRD epidemiological model with time-dependent parameters was presented, able to follow an epidemic from inception to eradication or endemic equilibrium, and applied to COVID-19 data. By varying the value of the pathogen's effective reproduction ratio, the model was used to assess the impact of quarantine and social distancing on the number of infections and deaths, on their daily changes ("new infections per day", "new deaths per day") and on the maxima in these daily rates expected during the epidemic. The effect of changing R_{eff} is substantial and ought to inform policy decisions around resource allocation to hospitals, appropriate mitigation strategies and their duration, and economic tradeoffs. Since all parameters can vary with time, the model is also able to quantify the effect of a change in the pathogen's infectiousness or morbidity as the virus mutates, or the expected effects of a new therapy or vaccine arriving at some future date. Finally, the long-term potential endemic end of COVID-19 absent eradication is discussed which, the model suggests, might involve times of the order of 100 years.

Appendix

SEIRD model diagram and its differential equations



$$\frac{dS}{dt} = B - (\beta I + \mu)S$$

$$\frac{dE}{dt} = \beta IS - (\sigma + \mu + \varphi)E$$

$$\frac{dI}{dt} = \sigma E - (\gamma + \mu + \omega)I$$

$$\frac{dR}{dt} = \gamma I + \varphi E - \mu R$$

$$\frac{dD}{dt} = \omega I$$

The basic reproduction number, R_0 , is the average or expected number of secondary cases one typical case would produce in a completely susceptible population (20) (21). The precise relationship between the transmission rate, β , and the pathogen's R_0 is model dependent. For the SEIRD model described, it is

$$R_0 = \frac{B\sigma\beta}{\mu(\gamma + \mu + \omega)(\sigma + \mu + \varphi)}$$

173 R_0 has the following interpretation: it is the product of the production rates of E and I per unit
174 contact, weighted by B/μ (22).

175 In this paper we define R_{eff} as $R_{\text{eff}} = (1-f) R_0$ where f is the fraction of the
176 Susceptible population that has been “removed” from the susceptible pool at a given time *though*
177 *interventions* such as quarantine, social distancing, or immunity due to a vaccine. (The model
178 already accounts for immunity in the Recovered group.) Other definitions of R_{eff} may differ. All
179 of the model’s parameters can be arbitrary (known) functions of time.

180

References

1. ECDC. COVID-19 situation update worldwide, as of 2 July 2020 [Internet]. Available from: <https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases>
2. Jackson J, Weiss M, Schwarzenberg A, Nelson R. Global Economic Effects of COVID-19. Congr Res Serv [Internet]. 2020;(20):78. Available from: <https://crsreports.congress.gov>
3. World Bank. Updated estimates of the impact of COVID-19 on global poverty [Internet]. Washington, DC; 2020. Available from: <https://blogs.worldbank.org/opendata/updated-estimates-impact-covid-19-global-poverty>
4. CDC and ESPR. COVID-19 Pandemic Planning Scenarios. Centers for Disease Control and Prevention. 2020.
5. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna: R Foundation for Statistical Computing; 2020. Available from: <https://www.r-project.org/>
6. Soetaert K, Petzoldt T, Setzer RW. Solving Differential Equations in R. R J. 2010;2(2):5–15.
7. King AA, Bolker B, Drake J, Rohani P, Smith D. Integrating ordinary differential equations in R with contributions from. 2012;1–8.
8. Hindmarsh AC, Brown PN, Grant KE, Lee SL, Serban R, Shumaker DE, et al. SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers. ACM Trans Math Softw. 2005;31(3):363–96.
9. Aronson JK, Brassey J, Mahtani KR. “When will it be over?”: An introduction to viral reproduction numbers, R0 and Re - CEBM. Cent Evidence-Based Med [Internet]. 2020; Available from: <https://www.cebm.net/covid-19/when-will-it-be-over-an-introduction-to-viral-reproduction-numbers-r0-and-re/>

10. Sanche S, Lin YT, Xu C, Romero-Severson E, Hengartner N, Ke R. High Contagiousness and Rapid Spread of Severe Accute Respiratory Syndrome Coronavirus 2. *Emerg Infect Dis*. 2020;26(7).
11. Zhang S, Diao MY, Yu W, Pei L, Lin Z, Chen D. Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. *Int J Infect Dis* [Internet]. 2020;93:201–4. Available from: <https://doi.org/10.1016/j.ijid.2020.02.033>
12. Wang H, Wang Z, Dong Y, Chang R, Xu C, Yu X, et al. Phase-adjusted estimation of the number of Coronavirus Disease 2019 cases in Wuhan, China. *Cell Discov* [Internet]. 2020;6(1):4–11. Available from: <http://dx.doi.org/10.1038/s41421-020-0148-0>
13. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis*. 2020;20(6):669–77.
14. Basu A. Estimating The Infection Fatality Rate Among Symptomatic COVID-19 Cases In The United States. *Health Aff (Millwood)*. 2020;1–6.
15. Nature. How deadly is the coronavirus? Scientists are close to an answer. *Nature Online News*. 2020 Jun;
16. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. *Int J Infect Dis* [Internet]. 2020;93:284–6. Available from: <https://doi.org/10.1016/j.ijid.2020.02.060>
17. Earn DJ. Mathematical Epidemiology. In: Bauer F, van der Driessche P, Wu J, editors. Heidelberg: Springer International Publishing; 2008.
18. Greer M, Saha R, Gogliettino A, Yu C, Zollo-Venecek K. Emergence of oscillations in a simple epidemic model with demographic data. *R Soc Open Sci*. 2020;7(1).

19. Torres BY, Oliveira JHM, Thomas Tate A, Rath P, Cumnock K, Schneider DS. Tracking Resilience to Infections by Mapping Disease Space. PLoS Biol [Internet]. 2016;14(4):1–19. Available from: <http://dx.doi.org/10.1371/journal.pbio.1002436>
20. Delamater PL, Street EJ, Leslie TF, Yang YT, Jacobsen KH. Complexity of the basic reproduction number (R_0). Emerg Infect Dis. 2019;25(1):1–4.
21. Ridenhour B, Kowalik JM, Shay DK. Unraveling R_0 : Considerations for public health applications. Am J Public Health. 2018;108(2):S445–54.
22. Jones JH. Notes on R naught [Internet]. Class Notes. 2007. Available from: papers2://publication/uuid/B09D6611-5FA2-4743-9F25-70030551AE4E

Address for correspondence: Sergio Bienstock, 221 Lake Rd, Morristown, NJ 07960-2727; email: sbie52@aol.com

Table 1. Table title: **COVID-19 model parameters**

<i>Name</i>	<i>Symbol</i>	<i>Value</i>	<i>Comments</i>
Basic reproduction number	R_0	4.0	The range 1-14 was examined.
Infection rate if exposed	σ	0.8/4.2	Shown as probability/average period in days.
Recovery rate if exposed	φ	0.2/11.5	Same as above.
Death rate if Infected (IFR)	ω	0.0066/16.5	Same as above.
Recovery rate if infected	Υ	0.9934/11.5	Same as above.
Death rate unrelated to COVID-19	μ	0.02/365	Precise value immaterial in first years of infection.
Gross birth rate	B	0.02/365	Same as above.

240

241 Table 2. Table two title: **Maximum daily infection and death rates vs R_{eff}**

R_{eff}	$\max(dI/dt)$	$1000 \max(dD/dt)$
2.0	0.00189	0.03947
3.0	0.00634	0.07640
4.0	0.01128	0.10137
6.0	0.02050	0.13152
8.0	0.02835	0.14849
10.0	0.03504	0.15912

242

243 Figure 1. Label: No community mitigation assumed; $R_{\text{eff}} = 4.0$ at all times

244 Figure 2. Label: R_{eff} values differ in the three shaded regions to simulate quarantine and

245 reopening

246 Figure 3. Label: Variation of maximum daily infection and death rates with R_{eff} , and with time to

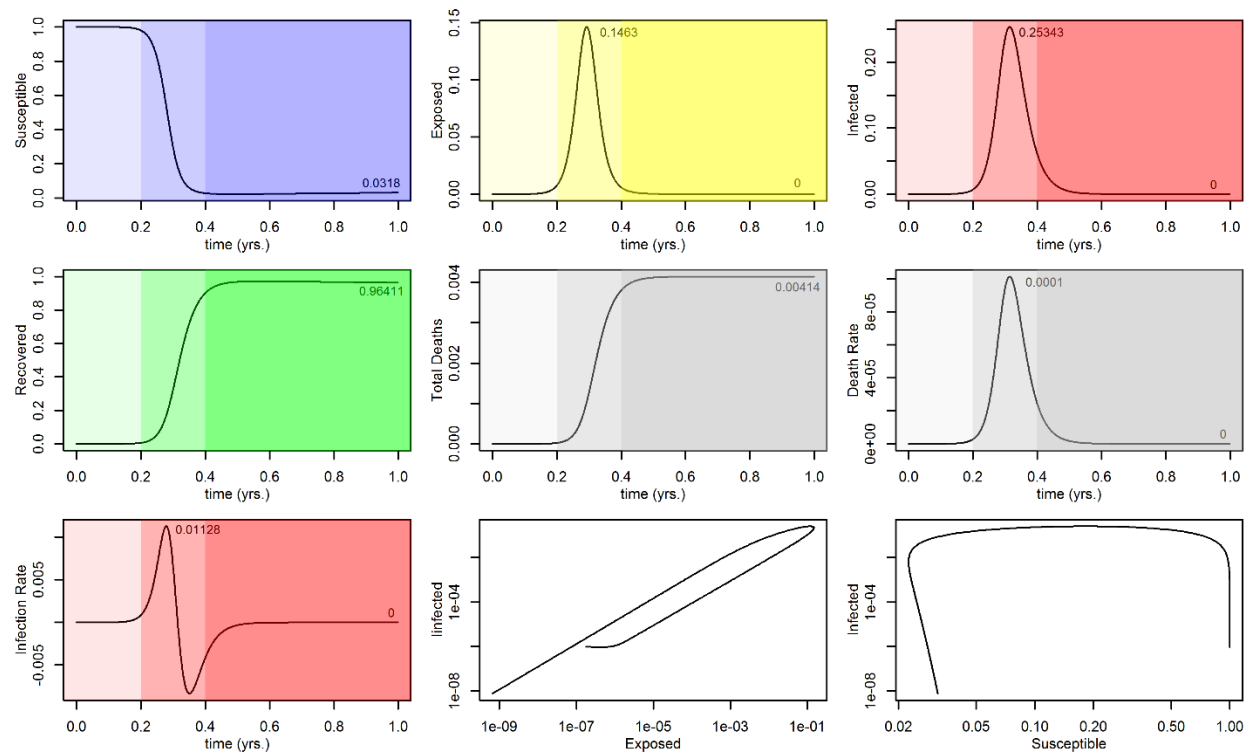
247 maximum

248 Figure 4. Label: Covid-19 epidemic modeled with $R_{\text{eff}} = 4.0$ from inception to endemic

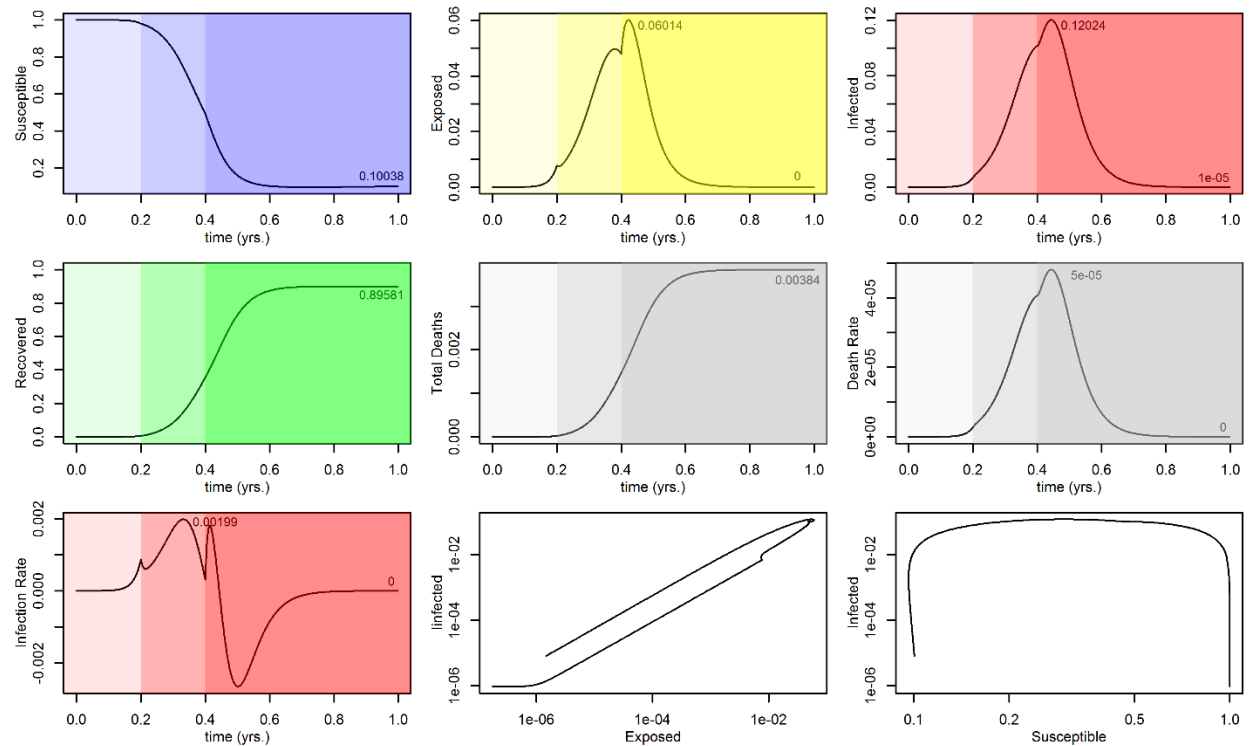
249 equilibrium

250

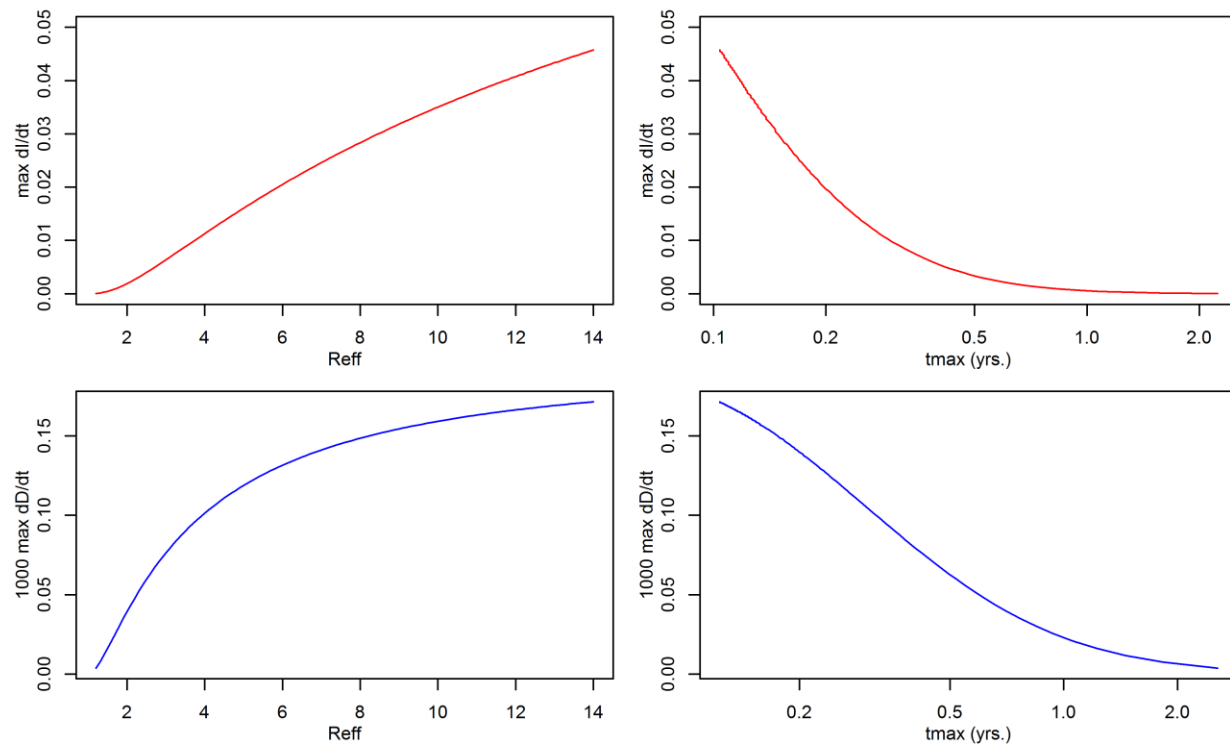
251 **Figure 1**



252
253
254 **Figure 2**
255

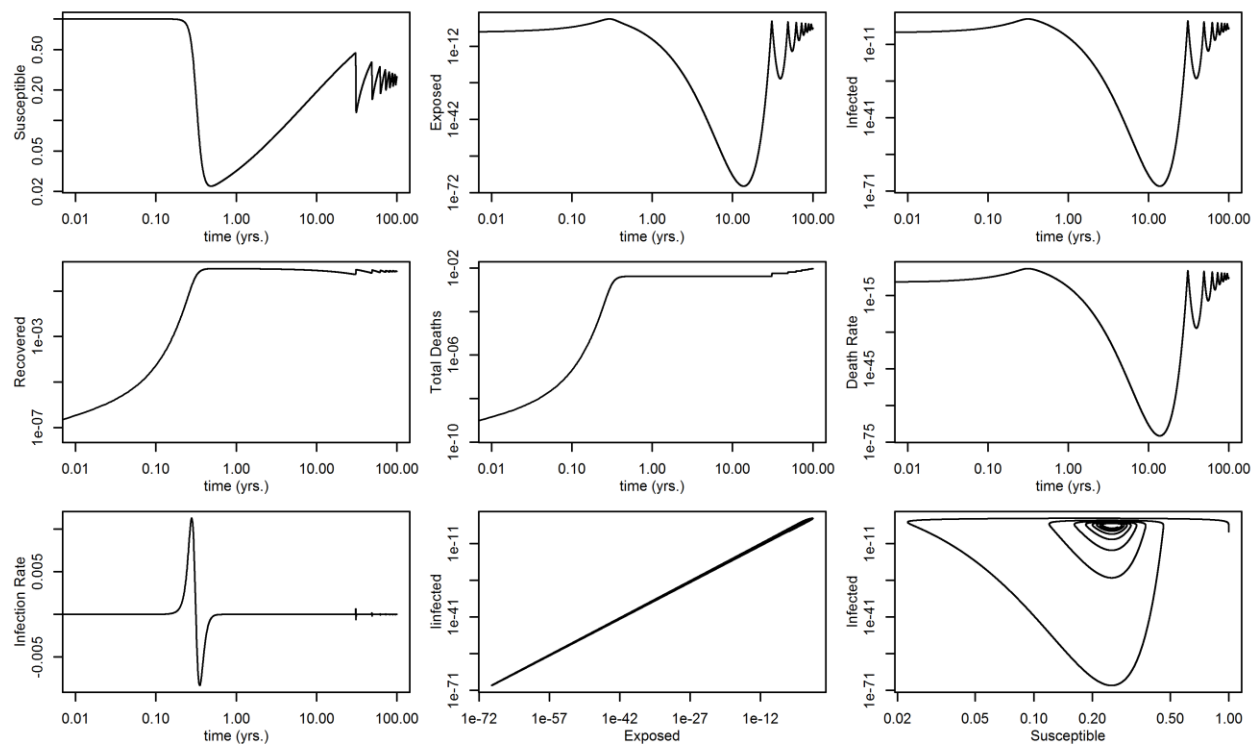


257 **Figure 3**



258

259 **Figure 4**



260

261