Res. Lett. Inf. Math. Sci., 2004, Vol.**6**, pp159-170
Available online at http://iims.massey.ac.nz/research/letters/

Modelling repeated epidemics with general infection kernels

J. Mann & M. Roberts

Institute of Information & Mathematical Sciences
Massey University at Albany, Auckland, New Zealand.

An integral equation approach is taken to explore the characteristics of a general infectious disease in a homogeneous population. It is shown that the final size of the epidemic depends on the basic reproduction ratio for the infection and the initial number of susceptibles. A discrete map for the susceptible population from epidemic generation to epidemic generation is formed to consider the long term behaviour of the disease in a population of constant size.

---

1 Introduction

Consider a population of constant size that we can split into three distinct classes in relation to an infection. Let $S(t)$ be the number of people susceptible to the infection, $I(t)$ be the number infected and $R(t)$ be the number removed from the infection (either through immunity or death). Susceptibles become infected at a rate λ resulting from contact with infectives. Contact here is very loosely defined, as the amount of contact needed to become infected will depend on the infection being modelled. Infectives then become part of the removed compartment at a constant rate γ. As the population size is constant, we know that the change in the population size is zero, that is:

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0 \tag{1}$$

The differential equation model used to describe this system is then:

$$\begin{aligned} \frac{dS}{dt} &= -\beta\chi\frac{SI}{N} \\ \frac{dI}{dt} &= \beta\chi\frac{SI}{N} - \alpha I \\ \frac{dR}{dt} &= \alpha I \end{aligned} \tag{2}$$

Where χ is the rate at which susceptibles contact other members of the population and β is the probability of a susceptible becoming infected given contact with an infected member of the population. We have assumed that the population size is constant, that is: $S + I + R = N$, so it can be seen that one of the above equations is redundant.

A differential equation approach has been used for numerous mathematical models, and there is a large amount of information available for the analysis of such systems. However, with constant contact parameters, the amount of time spend in each compartment by members of the population is exponentially distributed - this does not fit actual data. So we turn to a slightly different way of constructing a model with the use of integral equations.

Using integral equations to model an infection is more intuitive than a differential equations approach, and will match the actual data more closely. However, the down side is, there is not a lot of information published relating to the analysis of such systems.

2 Integral Model in One Dimension

Before we introduce the integral equation method, we must first define some terms and relations that will be used throughout the work.

The probability of a susceptible becoming infected depends on their contact with an infective and the probability of infection given this contact, which depends on the time since the infective was itself infected. If we let $p(\tau)$ be the probability of infection and contact, and $\chi(\tau)$ be the contact rate with an infective, where τ is the time since the infective was initially infected, then:

$$A(\tau) = p(\tau)\chi(\tau) \tag{3}$$

We can then think of $A(\tau)$ as representing the probability of contact and infection with an infective at infection time τ.

To see if an infection will persist within a population, we consider the basic reproduction ratio, represented by R_0, of the epidemic. The basic reproduction ratio is the number of secondary cases that arise from a primary case in a susceptible population (Diekmann & Heesterbeek 2000). So the critical value of R_0 is one. If $R_0 < 1$ then the epidemic will not persist in the population, and the number of infectives will decrease. If $R_0 > 1$ then the epidemic will continue through the population, and the number that have been infected will increase while the number of susceptibles will decrease. We can see that R_0 will depend on the population size, the contact rates and the probability of infection, hence:

$$R_0 = S(0) \int_0^\infty A(\tau) d\tau \tag{4}$$

The incidence of infection $i(t)$ is the number of new cases per unit time. So we

see that it will be equal to the rate of change of the susceptible population (as we have ignored changes in the susceptible population due to other causes).

At time t, the number of new cases of the infection depends on the contacts between susceptibles and infectives - those who were infected themselves before time t. So we have:

$$i(t) = i_0\delta(t) + S(t)\int_0^t A(\tau)i(t-\tau)d\tau \tag{5}$$

where the $i_0\delta(t)$ accounts for the initial introduction of the infection into the population and $\delta(t)$ is the Dirac delta function.

We may also rewrite this in terms of the change in the susceptible population:

$$-\frac{dS(t)}{dt} = i_0\delta(t) - S(t)\int_0^t A(\tau)\frac{dS(t)}{dt}(t-\tau)d\tau \tag{6}$$

The number of infectives can be modelled by an exponential during the initial phases of infection. So we let

$$-\frac{dS}{dt} = i(t) \approx ke^{rt} \tag{7}$$

for some positive constant r, which we call the initial growth rate of the infection. Substituting in equation (6):

$$ke^{rt} = i_0\delta(t) + S(t)k\int_0^t A(\tau)e^{r(t-\tau)}d\tau \tag{8}$$

As we are examining the initial growth of the infection, we do not need to include the initial introduction of the infection into our population; hence we can omit the $i_0\delta$ term. We also set the size of the susceptible population equal to its initial value, $S(t) \equiv S(0)$[1]. So we solve:

$$1 = S(0)\int_0^t A(\tau)e^{-r\tau}d\tau \tag{9}$$

It is shown in Diekmann and Heesterbeek (2000), that there is a unique real r that solves equation (9). Note that equation (9) is similar to our equation for the basic reproduction ratio (equation (4)). The correlation between the two lead to two important facts: $r > 1$ if and only if the basic reproduction ratio is greater than one, and $r < 1$ if and only if the basic reproduction ratio is less than one. That is, we only have initial growth of the infection if we have an epidemic.

We may solve equation (9) to find r in two ways: by using normal integration techniques and by using the method of Laplace transforms. To solve this equation

[1] $S(t) \gg i_0$, and so we let $S(0^+) = S(0^-)$. i_0 will usually be assumed to be equal to one, i.e. there will be one initial case to introduce the infection into the susceptible population.

without using Laplace transforms, we let $t \to \infty$ in the integral, as we want to include all contacts between susceptibles and infectives. The integral then becomes:

$$1 = S(0) \int_0^\infty A(\tau) e^{-r\tau} d\tau \tag{10}$$

Which we may solve either analytically or numerically for r.

To use the method of Laplace Transforms, we start from equation (5) and approximate $S(t) \equiv S(0)$, as we assume that the entire population is initially susceptible to the infection. From this we gain:

$$i(t) = i_0 \delta(t) + S(0) \int_0^t A(\tau) i(t-\tau) d\tau \tag{11}$$

Then taking the Laplace transform, and using the convolution product:

$$\begin{aligned} \bar{i}(s) &= i_0 + S(0)\overline{A}(s)\bar{i}(s) \\ &= \frac{i_0}{1 - S(0)\overline{A}(s)} \end{aligned} \tag{12}$$

(the over bar represents the Laplace transform.) The initial growth rate is taken to be the real part of the value s with the largest real part. Both the Laplace transform method and the normal integration techniques yield the same solution for the initial growth rate.

To calculate the final size of the epidemic, we must calculate how many people were infected, i.e. $S(0) - S(\infty)$. To calculate the value of $S(\infty)$ for a small epidemic (where we assume that the number of susceptibles is constant, $S(t) \equiv S(0)$), we use the method of Laplace transforms on the equation for the incidence of infection (equation (5)). We know that:

$$\int_0^\infty i(t) dt = \lim_{s \to 0} \int_0^\infty i(t) e^{-st} dt = \lim_{s \to 0} \bar{i}(s) \tag{13}$$

and

$$i(t) = -\frac{dS(t)}{dt} \tag{14}$$

Combining the above two equations:

$$-\int_0^\infty \frac{dS(t)}{dt} dt = \lim_{s \to 0} \bar{i}(s) \tag{15}$$

Calculating this integral and substituting in the result from equation (12) gives:

$$S(0) - S(\infty) = \lim_{s \to 0} \frac{i_0}{1 - S(0) \int_0^\infty A(\tau) e^{-s\tau} d\tau} \tag{16}$$

Evaluating the limit and using the definition of the basic reproduction ratio we have the final size equation for a small epidemic:

$$S(0) - S(\infty) = \frac{i_0}{1 - R_0} \tag{17}$$

To calculate the final size equation for large epidemics we cannot assume that the susceptible population remains constant. We let $t > \infty$ in equation (6) and neglect the initial introduction of the infection into our population, hence:

$$\frac{dS(t)}{dt} = S(t) \int_0^\infty A(\tau) \left[\frac{dS(t-\tau)}{dt} \right] d\tau \tag{18}$$

Integrating and rearranging then leads to:

$$\log\left(\frac{S(\infty)}{S(0)}\right) = \left(\frac{S(\infty)}{S(0)} - 1\right) R_0 \tag{19}$$

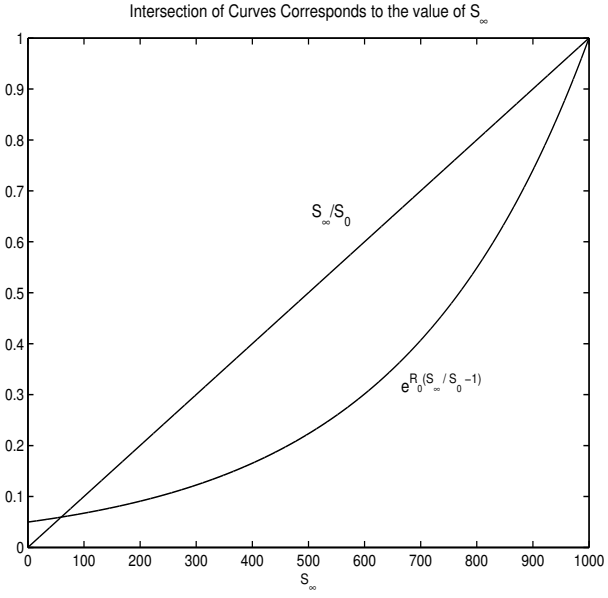

Figure 1: Sample final size curves, intersection of lines gives $S(\infty)$, where $S(0) = 1000$ and $R_0 = 3$

Figure 1 shows the two curves $y = \frac{S(\infty)}{S(0)}$ and $y = e^{(\frac{S(\infty)}{S(0)} - 1)R_0}$ plotted against the same axis. The two curves intercept when $S(0) = S(\infty)$ (corresponding to no epidemic) and at a second point where $S(\infty) < S(0)$ (corresponding to an epidemic). Given the values for $S(0)$ and R_0 we can solve equation (19) for $S(\infty)$ and calculate the final size equation.

3 Repeated Epidemics

When an infective is introduced into a population, an epidemic may or may not occur depending on the number of people in the population who are susceptible to the infection. If an epidemic does occur, the next natural question is, will there be another epidemic in the future?

We can develop a discrete map for the susceptible population from epidemic generation to epidemic generation[2]. We already know that the final size of the epidemic can be calculated when the initial population and basic reproduction ratio are known.

We assume that the entire population is initially susceptible, then let an epidemic occur and calculate the number of susceptibles left after the epidemic ($S(\infty)$). We then let a proportion (θ) of these susceptibles remain in the population, and introduce new susceptibles to keep the population size constant. By calculating the basic reproduction ratio for this new population, we can tell if another epidemic will occur. At the end of each epidemic, we assume that there are no infectives from the previous epidemic present.

An example of this method would be children at a school. Suppose that all the children are susceptible to an infection, and the reproduction ratio for this infection is greater than one - then there will be an epidemic at the school. At the end of the school year, a proportion of the children will leave the school and new students will attend. We then see if another epidemic occurs. A "reshuffling" of susceptibles will occur at the end of every epidemic, and then another epidemic may or may not occur depending on the new reproduction ratio.

Initially we have $S_{0,0} = N$, that is, the susceptible population before the first epidemic is just the entire population. The proportion of susceptibles is then $x_0 = \frac{S_{0,0}}{N} = 1$, and the basic reproduction ratio is $R_0 = N\int_0^\infty A(\tau)d\tau$. If $R_0 < 1$ then there is no epidemic, and the proportion of susceptibles in the next epidemic population will be:

$$x_1 = 1 - \theta + \theta x_0 \tag{20}$$

therefore, we have $x_1 = 1$. However, if $R_0 > 1$ then there will be an epidemic and we must solve the final size equation

$$\log\left(\frac{z}{x_0}\right) = R_0\left(\frac{z}{x_0} - 1\right) \tag{21}$$

where $z = \frac{S_{\infty,0}}{N}$. For this case, we then replace x_0 with z in equation (20) to calculate the susceptible proportion for the next epidemic generation.

We can then iterate the process outlined above for subsequent epidemics. So in

[2]This idea initially came from Andreason (2003), although in his paper a new infection is introduced in each generation.

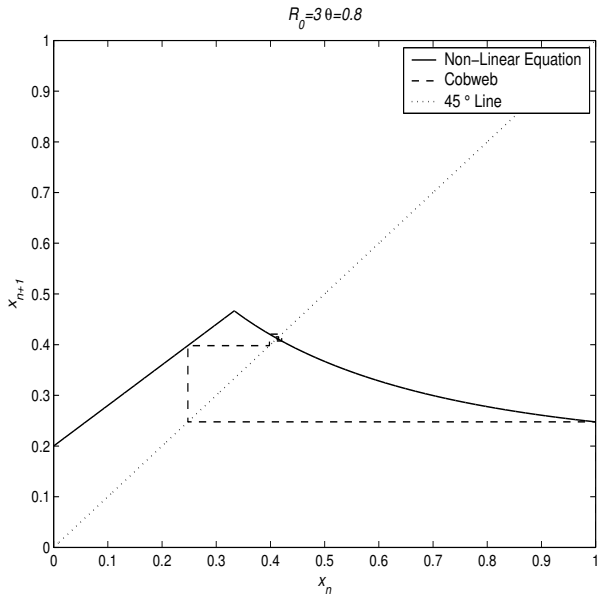

Figure 2: Cobweb plot given: $R_0 =$, $\theta =$ and $x_0 = 1$

general we have:

$$\begin{aligned} x_n &= 1 - \theta + \theta z_{n-1} \\ R_n &= x_n R_0 \end{aligned} \tag{22}$$

If $R_0 > 1$ we have an epidemic, so we solve:

$$\log\left(\frac{z_n}{x_n}\right) = x_n R_0 \left(\frac{z_n}{x_n} - 1\right) \tag{23}$$

for z_n. This has a unique solution[3] $z_n \neq x_n$. If $R_0 < 1$ then this is no epidemic, and we set $z_n = x_n$.

The discrete map is:

$$x_{n+1} = f(x_n) \tag{24}$$

where

$$y = \begin{cases} 1 - \theta + \theta z_n & \text{if } x_n > \frac{1}{R_0}, \text{ i.e. no epidemic} \\ 1 - \theta + \theta x_n & \text{otherwise, i.e. no epidemic} \end{cases}$$

and z_n solves equation (23).

The cobweb plots shown demonstrate the convergence of repeated iteration of the piecewise continuous function given in equation (3), when the entire population is initially susceptible. On the "cobweb" curve, the straight line corresponds to no epidemics, while the decreasing curve corresponds to epidemics. To answer our

[3]See Diekmann and Heesterbeek (2000) for a complete proof

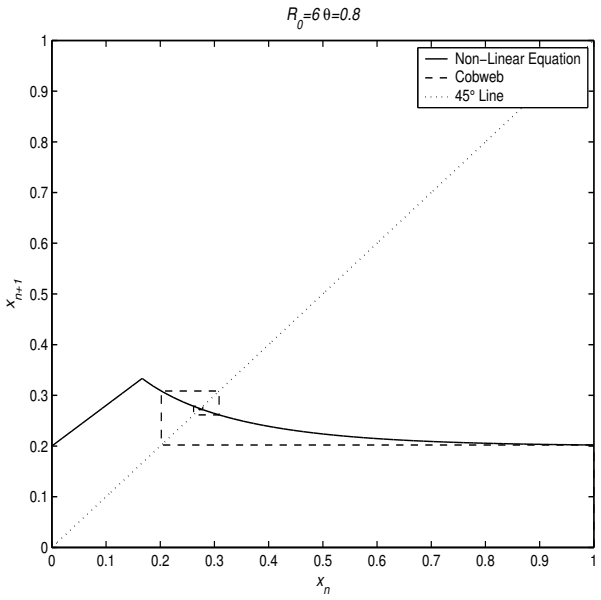

Figure 3: Cobweb plot given: $R_0 =$, $\theta =$ and $x_0 = 1$

original question, we need to calculate the fixed point of the cobweb plot and see if it is stable or unstable.

We know that there is no epidemic when the basic reproduction ratio is less than one for each epidemic generation, that is:

$$R_n = x_n R_0 < 1 \tag{25}$$

So we see that the straight line spans $0 < x_n < \frac{1}{R_0}$.

If we let the straight line part of the cobweb be $g(x) = 1-\theta+\theta x$ and the curved part be the solution to

$$\log\left(\frac{x_{n+1}-(1-\theta)}{\theta x_n}\right) = x_n R_0\left(\frac{x_{n+1}-(1-\theta)}{\theta x_n} - 1\right) \tag{26}$$

defined by the function $x_{n+1} = f(x_n)$, we have:

$$\log\left(\frac{f(x_n)-(1-\theta)}{\theta x_n}\right) = R_0\left(\frac{f(x_n)-(1-\theta)}{\theta} - x_n\right) \tag{27}$$

Note that $g(x)$ and the line $y = x$ only intersect at $x = 0$ and $x = 1$, so a non-zero steady state must lie at some value in the interval $\frac{1}{R_0} < X < 1$.

To find the stability of the fixed point (from the cobweb plots it appears to be stable), we find the derivative of $f(x)$ from equation (27). We can utilise the fact that at the fixed point $f(X) = X$, which gives us:

$$f'(X) = \frac{\theta(1-R_0X)(X-(1-\theta))}{X(\theta - R_0(X-(1-\theta)))} \tag{28}$$

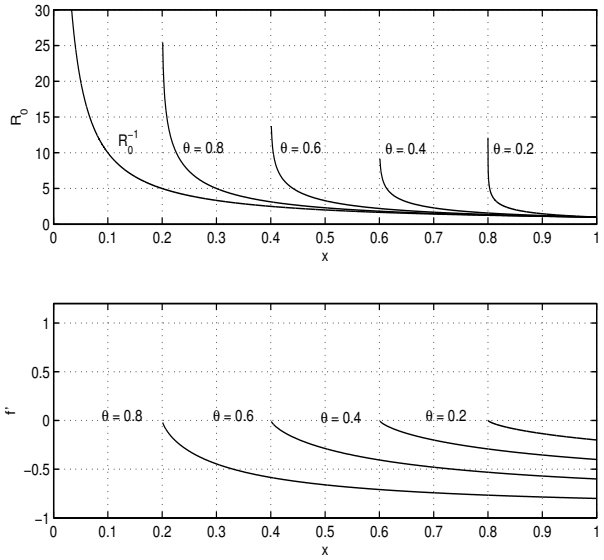

Figure 4: The top plot show the value of R_0 at fixed points X, given θ. The second plot shows the derivative of the nonlinear equation at fixed points X, given θ with corresponding R_0 values seen in the top plot.

We may also use equation (27) to find an implicit relationship between the fixed point X and the basic reproduction ratio:

$$R_0 = \frac{\theta}{(1-\theta)(1-X)} \log\left(\frac{\theta X}{X-(1-\theta)}\right) \tag{29}$$

Recall that R_0 and θ are known and we are solving for the fixed point X. We know that the fixed point lies within the range $(\frac{1}{R_0}, 1)$, using equation (29) we can evaluate the reproduction ratio at every fixed point, and then use this information to calculate equation (28). We have used this method to produce figure (4). This first plot shows the value of R_0 at varying fixed points X and selected values of θ. The line $X = \frac{1}{R_0}$ has been added to emphasize the fact that the fixed point lies in the range $(\frac{1}{R_0}, 1)$. The second plot shows $f'(X)$ at varying fixed points X and different values of θ.

From the numerical examples presented in figures (2 and 3), we would infer that the fixed point is stable. We shall now present an analytical proof to show this.

We can find the values of $f'(X)$ at the end points $X = \frac{1}{R_0}$ and $X = 1$, using asymptotic expansion and simple algebra (respectively), to gain

$$\begin{aligned} f'(\frac{1}{R_0}) &= 0 \\ f'(1) &= -\theta \end{aligned} \tag{30}$$

To show that the fixed point is stable, we must first show $|f'(X)| < 1$ within our range of consideration, $(\frac{1}{R_0}, 1)$, where $R_0 > 1$. It can be shown that $f'(X)$ is

continuous in our range (the functions singularity lies at $R_0 \equiv 1$), and

$$X > 1 - \theta + \frac{\theta}{R_0} \tag{31}$$

We need now only show that $f'(X)$ has no turning point over $X \in (\frac{1}{R_0}, 1)$, i.e. $f'(X)$ has no critical points over $X \in (\frac{1}{R_0}, 1)$

Proof. We have

$$f'(X) = \frac{\theta(1 - R_0 X)(X - (1 - \theta))}{X(\theta - R_0(X - (1 - \theta)))} \tag{32}$$

With some calculation is can be shown that

$$f''(X) = \frac{\theta(1 - \theta)(\theta(1 - R_0) + R_0 X(X - 2) + R_0)}{X^2(\theta(R_0 - 1) + R_0(X - 1))^2} \tag{33}$$

A critical point must satisfy $\frac{df'(X)}{dX} = 0$, that is

$$R_0 X^2 - 2R_0 X + R_0 + \theta(1 - R_0) = 0 \tag{34}$$

Solving the above quadratic for $X < 1$ yields

$$X = 1 - \sqrt{\frac{\theta(R_0 - 1)}{R_0}} \tag{35}$$

We have two possibilities:

If $\frac{\theta(R_0-1)}{R_0} \geq 1$, then from equation (35) X would be negative (or zero), and hence not lie in our region of consideration.

If $\frac{\theta(R_0-1)}{R_0} < 1$ then $X < 1 - \frac{\theta(R_0-1)}{R_0}$ which contradicts equation (31).

Therefore, $f'(X)$ has no critical points over $X \in (\frac{1}{R_0}, 1)$ □

We have shown $f'(1) = -\theta$, $f'(\frac{1}{R_0}) = 0$, the singularity $X = 1-\theta+\frac{R_0}{\theta}$ in equation (28) does not lie in the region we are considering ($\frac{1}{R_0} < X < 1$) and that $f'(X)$ does not have any turning points within this range. So $0 \geq f'(X) \geq -\theta > -1$, thus the fixed point X that satisfies equation (28) is stable.

This means that there is an epidemic each and every year when an infection is introduced into a fully susceptible population, when the basic reproduction ratio is greater than one and the population size remains constant after each epidemic.

4 Discussion

Compartmental models using differential equations have been used to model many infections, and there is a vast amount of information available on the analysis of such systems. However, when constant parameters are used to model the dynamics between the susceptible, infective and recovered populations, the time spent in

each compartment is exponentially distributed among the members of the population. This kind of assumption does not often model the data closely, so an integral equation method has been developed.

We have shown methods to calculate the basic reproduction ratio, the initial growth rate of an infection and the final size equation of an infection. Two methods were presented for calculating the final size of an epidemic bases on the size of the epidemic: for a small epidemic we approximate the susceptible population by the initial number of susceptibles in the population, and for large epidemics we do not assume the susceptible population is constant.

A discrete map was developed to look at the effect of an epidemic on the susceptible population from one epidemic generation to the next, and this was shown to converge to a stable solution. This means that an epidemic occurs within the population each and every year - which is a surprising result consider the similarity of the cobweb map (figure 2 and 3) to the tent map that produces periodic orbits[4]. This result has also been established by Boni (2004) for a the differential equation model.

The integral method for calculating the basic reproduction ratio, the initial growth rate and the final size equation can be extended into two dimensions, where we consider the population to be split into two independent classes. Special cases can also be considered to explore different types of mixing behavior within the total population. To solve the two dimensional final size equation, a greater amount of numerical computation is required than for the simple two dimension model, but if certain mixing patterns are applied the two dimensional case will break down to form a simple one dimensional problem. We can also incorporate the effect of vaccination into the model, by changing the initial susceptible population after each epidemic generation.

The work presented here has been submitted as part of a Masters of Information Science thesis in 2003. This work was supported by a postgraduate scholarship in industrial mathematics, from the New Zealand Institute for Mathematics and its Applications.

References

Anderson, R. M. and May, R. M. 1991. Infectious Diseases of Humans. Oxford: Oxford University Press.

Andreasen, V. 2003. Dynamics of Annual Influenza A Epidemics with Immuno-Selection. Journal of Mathematical Biology, 46, 504-536.

Boni, M., Gog, J., Andreasen, V. and Christiansen F. 2004 Influenza drift and

[4]See Strogatz (1994) for a detailed analysis of the tent map

epidemic size: the race between generating and escaping immunity. Theoretical Population Biology, 65:179-191.

Diekmann, O. and Heesterbeek, J. A. P. 2000. Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation. New York: John Wiley & Sons.

Strogatz, S. H. 1994. Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering. Cambridge, MA: Westview Press.