Chapter Five

Temporally Forced Models

In this chapter, we consider how seasonally varying parameters act as a forcing mechanism and examine their dynamical consequences. For the most part, we will use measles as a prototypical directly transmitted infectious disease. We demonstrate how such temporally forced models allow us to better capture the observed pattern of recurrent epidemics in contrast to unforced models, which predict oscillations that are damped toward equilibrium (see Chapter 2). We will follow the historical progress of work in this field, because it provides a natural progression from simple models to their more complex and realistic refinements.

5.1. HISTORICAL BACKGROUND

Understanding the mechanisms that generate periodic outbreaks of childhood infectious diseases had been the subject of much debate among Victorian epidemiologists (e.g., Farr 1840; Ransome 1880, 1882; Hamer 1897). In 1880, for example, Arthur Ransome systematically considered numerous "plausible" mechanisms that may generate regular epidemics of measles, whooping cough, and smallpox. Having dismissed factors including meteorological elements (for example, sunspots), isoclinal magnetic lines, the "age-theory" of disease (where only specific ages may be prone to infection), reduced virulence following successive transmission events, Ransome settled on changes in the density of susceptibles as the most likely explanation. He argued that exanthematous diseases wipe out nearly all susceptibles and, as a consequence, "must necessarily wait a number of years before the requisite nearness of susceptible individuals has been again secured." This is essentially a verbal version of the threshold theorem, which, as discussed in Chapter 2, gives rise to damped oscillations. The next important conceptual breakthrough came as a result of classic work by H. E. Soper in 1929. He noticed that in a large population, case report data for measles, which conforms well to the assumptions of the SIR model, show large amplitude recurrent epidemics with very dramatic peaks and troughs. This is in direct contrast to the equilibrium dynamics predicted by simple models, with a steady incidence of disease (Chapter 2). This pattern of pronounced fluctuations in incidence has, since then, been documented for a number of other human infectious diseases such as chickenpox, whooping cough, mumps, and rubella (see, for example, Figure 5.1). When data and models disagree in such a stark manner, there is invariably an important opportunity to re-examine the key assumptions of the model and explore ways in which it can be made more realistic. Focusing on the monthly case reports for measles in Glasgow from 1905–1916, Soper (1929) proceeded to estimate relative transmission rates per month.

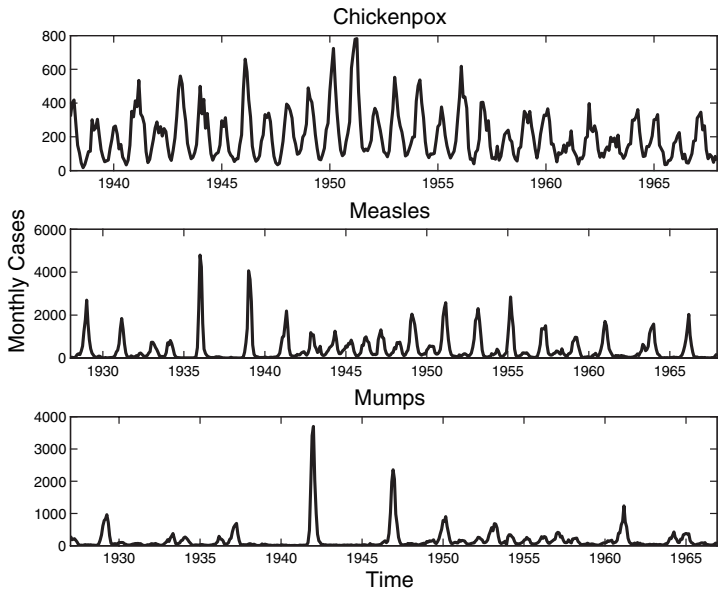

Figure 5.1. Monthly case reports of chickenpox, measles, and mumps in Copenhagen during the twentieth century, demonstrating dramatic patterns of recurrent outbreaks.

His methodology centered on the argument that

$$\frac{\text{Cases this interval}}{\text{Cases last interval}} \sim \frac{\text{Number of susceptibles now}}{\text{Equilibrium number of susceptibles}},$$

which can be expressed as the following equation

$$\left(\frac{C_{t+1}}{C_t}\right)^{\alpha} = k_\theta \frac{X_{t+1}}{X^*}. \tag{5.1}$$

The parameter α relates the realistic infection "time interval" (the sum of the infectious and latent periods) to the time scale of the data. For measles, this is approximately two weeks; therefore for monthly data, we set $\alpha = 1/2$. The term k_θ is "the factor representing the influence of season θ" (Soper 1929). To estimate X^*, Soper followed Hamer's (1906) calculation that the mean number of susceptibles (X^*) is approximately equivalent to 70 weeks' case reports. Then, once we take into account the fact that at the peak of an epidemic, $C_{t+1} \sim C_t$, we have an initial estimate for X_{t+1}, which can be updated by adding the documented births and subtracting the number of cases. All that remains now is to fit the seasonality parameter k_θ.

Soper's findings, averaged over the 12-year period of the data, are presented in Figure 5.2 and clearly demonstrate that estimated transmission was very low in the summer months, and peaked dramatically in the early autumn (October). Soper argued that based on his results, a key missing ingredient in the SIR model proposed originally by Hamer (1906) was seasonal change in "perturbing influences, such as might be brought about by school breakup and reassembling, or other annual recurrences."

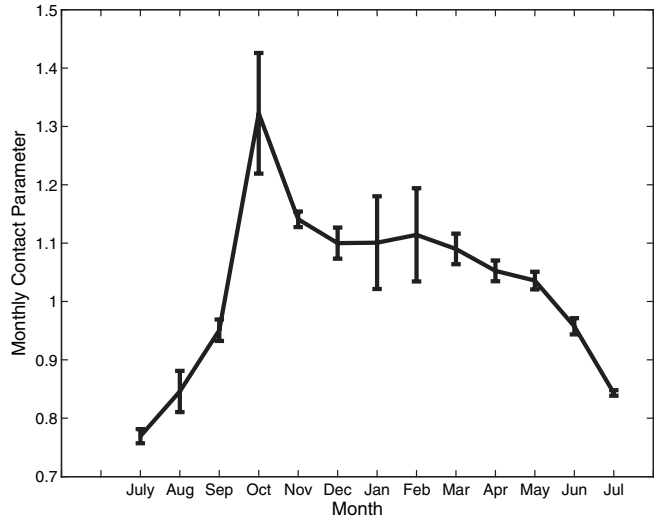

Figure 5.2. Average monthly contact coefficient for measles in Glasgow 1905–1916, as estimated by Soper (1929). The graph clearly highlights the non-constant nature of transmission, with the highest intensity observed during the fall/winter months. Error bars represent standard errors.

The important, though too often ignored, work of Soper was followed by the highly influential studies of London and Yorke (1973) and Yorke and London (1973). These authors were also interested in exploring seasonal influences on transmission, and estimated the mean monthly transmission rates for measles, mumps, and chickenpox in New York City from 1935 to 1972. The key concept in their analysis was based on an earlier empirical observation by Hedrich (1933) that the number of susceptibles has the same value, X_p, at the peak of every outbreak. Mathematically, we can see this is true because at the epidemic peak, the number of infectious individuals, Y, has reached its maximum; therefore, $\frac{dY}{dt} = 0$:

$$\frac{dY}{dt} = 0 \;\Rightarrow\; \beta XY/N - \gamma Y = 0 \;\Rightarrow\; X = X_p = \frac{\gamma N}{\beta}.$$

Hence, at the start of the epidemic year (which they defined to be from the beginning of September to the end of August), the number of susceptibles is X_p plus the cumulative number of reported cases for that year. Then, by using a discrete-time model (see Section 2.7), they were able to explore the pattern of transmission rates that provided model exposures consistent with the observed case reports. Although there are subtle differences in the details of London and Yorke's results compared to those of Soper, they also found a clearly seasonal pattern of transmission for all three diseases, with a peak that coincided with the start of school terms in the autumn and a trough that occurred during the summer months. Since then, more mechanistic approaches for the estimation of transmission rates have been developed, which involve a more detailed "reconstruction" of the number of susceptibles in the population. First Fine and Clarkson (1982) and later Finkenstädt and Grenfell (2000) used case report data for measles in England and Wales, together with information on the population size and birth rates to estimate transmission rates

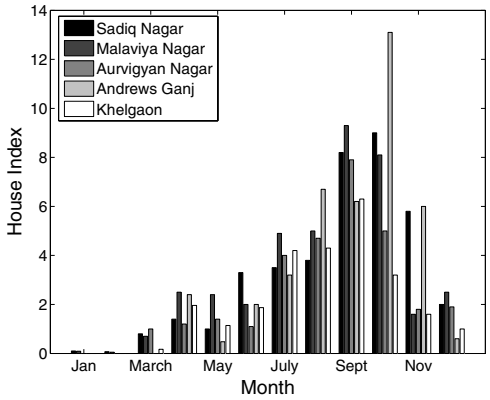

Figure 5.3. House index for *Aedes aegypti* larvae in five different localities in Delhi, India during 1996–1998. The chart clearly demonstrates the seasonal nature of fluctuations in mosquito numbers. Data from Ansari and Razdan (1998).The house index is defined as the percentage of houses infected with mosquito larvae or pupae.

through time. The overall signature detected in these data is consistent with the work of Soper (1929) and London and Yorke (1973).

A range of statistical approaches have revealed that transmission of childhood infections varies seasonally, peaking at the start of the school year and declining significantly in the summer months.

Although epidemiologists have been aware of the importance of seasonal factors in the transmission of childhood diseases, the modeling of such phenomena has been facilitated by the advent of accessible computational power. Analytical methods for dealing with forced models are woefully lacking, and therefore detailed computer integration of forced equations is often the only practical means of understanding or predicting the dynamics.

5.1.1. Seasonality in Other Systems

Changes in transmission rates through time is increasingly recognized as important in a range of infectious diseases (for a review, see Altizer et al. 2006; Grassley and Fraser 2006). For human infectious diseases that are vector transmitted (such as malaria or dengue), seasonality plays an important dynamical role. In these instances, however, the time dependency in transmission is brought about by the biology of the vector population. Hence, models need to capture the fact that mosquito numbers in the tropics, for example, are substantially higher during the rainy season (Figure 5.3). As a result, the forcing required in these models would essentially represent environmental trends, such as precipitation levels through the year (vectored populations are covered in detail in Chapter 4).

In food or waterborne infections, such as cholera, the role played by temporal forcing is more subtle and interesting. There is strong evidence, for example, that the multi-annual dynamics of cholera are interlinked with long-term environmental factors. Studying

historical records of cholera outbreaks in Bangladesh, Pascual et al. (2000) have established a correlation with the El Niño Southern Oscillation. However, cholera data also contain a pronounced annual signature, which is thought to be due to an increase in transmission during the monsoon seasons (Koelle and Pascual 2004).

Interesting seasonal components are also found in wildlife diseases. These most frequently arise from changes in host behavior throughout the year. Increased transmission may result from increased contact arising from flocking behavior (e.g., housefinches; Hosseini et al. 2004), seasonal migration (e.g., Monarch butterflies; Altizer 2002), or congregations during the breeding and molting season (e.g., harbor seals; Swinton et al. 1998).

These well-cited studies have established seasonal changes in the contact rates between susceptible and infectious individuals as an important feature of the dynamics of many infectious diseases. This chapter, reviews the various methods used to model time-dependent transmission in human and animal systems.

5.2. MODELING FORCING IN CHILDHOOD INFECTIOUS DISEASES: MEASLES

The last section reviewed the historical studies of seasonality in case reports of childhood infections. A large body of theoretical work has also examined the dynamical consequences of temporal changes in transmission. These studies started with the work of Soper (1929), Bartlett (1956), and Bailey (1975), who incorporated seasonality in SIR models with the primary aim of establishing the amplitude of variation in contact rates necessary to produce the observed 80% fluctuation in epidemics. (It is often difficult to consider the forcing of childhood infections without considering age structured models (Chapter 3). Throughout this chapter, age structure is ignored for simplicity; however, toward the end we highlight that a true mechanistic description of any childhood disease must take into account the interaction of forcing and age structure.) Bailey (1975) explored a simplified SIR model:

$$\frac{dX}{dt} = \mu N - \beta(t)XYN, \tag{5.2}$$

$$\frac{dY}{dt} = \beta(t)XY/N - \gamma Y. \tag{5.3}$$

As usual, μ is the per capita birth rate and γ is the recovery rate from the infection. These equations ignore the death of susceptible and infectious individuals; it therefore assumes that all individuals contract the infection during their lifetime, which is a reasonable approximation for measles. The transmission rate is a function of time, $\beta(t)$, and was taken by Bailey to be a sinusoid:

$$\beta(t) = \beta_0(1 + \beta_1 \cos(\omega t)). \tag{5.4}$$

The parameter β_0 denotes the baseline or average transmission rate, ω is the period of the forcing, and β_1 is the amplitude of seasonality which is restricted to the unit interval. For this form of forcing, the basic reproductive ratio, R_0, is given by $\frac{\beta_0}{\gamma}$; this value represents a yearly average and at certain times of the year (when $\cos(\omega t) \approx 1$), instantaneous growth rates may be much larger than predicted by this average value. Bailey (1975) proceeded to explore the dynamics of small perturbations to the unforced equilibrium assuming a

small amplitude of seasonality ($\beta_1 \ll 1$). This was achieved by making the substitutions $X = X^*(1+x)$ and $Y = Y^*(1+y)$, which, after omitting some intermediate steps, gives a second order differential equation in the small infectious perturbation y:

$$\frac{d^2y}{dt^2} + \mu R_0 \frac{dy}{dt} + \mu\beta_0 y = -\beta_1 \omega\gamma \sin(\omega t).$$

The particular integral of this equation (see, for example, Strang 1986) gives the period and amplitude of oscillations as driven by the seasonal term. Although the period of the oscillations is the same as the period of the forcing, the amplitude, M, of the oscillations is given by:

$$M = \beta_1\omega\gamma \left\{ \left(\mu\beta_0 - \omega^2\right)^2 + \left(\omega\mu R_0\right)^2 \right\}^{-\frac{1}{2}}. \tag{5.5}$$

Now, making the appropriate substitutions for measles, we set $1/\gamma = 2$, $\mu R_0 \sim 0.014$, and $\omega = \frac{\pi}{26}$ (taking the week as our basic time unit), which gives $M \sim 7.76\beta_1$. The implication of this result is that a 10% variation in the transmission parameter translates into seasonal variations of 78% in case notifications, as envisaged by Soper (1929).

Relatively modest levels of variation in the transmission rate can translate into large amplitude fluctuations in the observed disease incidence.

5.2.1. Dynamical Consequences of Seasonality: Harmonic and Subharmonic Resonance

The first systematic examination of seasonality affecting the dynamical pattern of epidemics was made, as far as we are aware, by Klaus Dietz in his seminal 1976 paper. Dietz carried out a stability analysis of the familiar SIR model:

$$\frac{dX}{dt} = \mu N - \left(\beta(t)\frac{Y}{N} + \mu\right)X, \tag{5.6}$$

$$\frac{dY}{dt} = \beta(t)X\frac{Y}{N} - (\mu+\gamma)Y, \tag{5.7}$$

This is online program 5.1

where $\beta(t) = \beta_0(1 - \beta_1\cos(\omega t))$ (note that he used a minus sign in his formulation in order to ensure that contact rates were at their lowest at the start of the epidemic year). He demonstrated that in the absence of seasonal forcing, the system fluctuated with frequency F (c.f. Chapter 2, Box 2.4), where

$$F^2 = \mu(\gamma+\mu)(R_0 - 1) - \left(\frac{\mu R_0}{2}\right)^2. \tag{5.8}$$

In many realistic situations, $\mu R_0 \ll 1$, hence we may ignore the final term in equation (5.8). Dietz pointed out that for cases in which the natural period of oscillations in the SIR model are approximately the same as that of the seasonal forcing (i.e., $F \approx \omega$), we observe *harmonic resonance*, where model dynamics mimic those of the forcing, although the amplitude of oscillations may be greatly increased.

When the forcing is relatively small, this result can be made more precise. Looking at equation (5.5), the amplification of sinusoidal forcing (M/β_1) is largest whenever the forcing frequency, ω, is close to the natural frequency of oscillations, $F \approx \sqrt{\mu\beta_0}$.

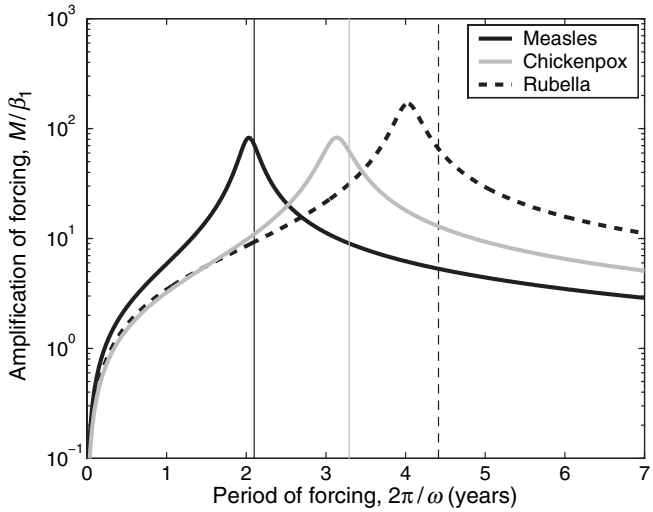

Figure 5.4. The predicted amplification of small amounts of forcing for three childhood infections using equation (5.5). Clear peaks of amplification exist for all three diseases, and these are reasonably close to the natural period of oscillations, $2\pi/F$, shown by the vertical lines. ($\mu = 0.02$ per year. Measles: $1/\gamma = 13$ days, $R_0 = 17$. Chickenpox: $1/\gamma = 20$ days, $R_0 = 11$. Rubella: $1/\gamma = 18$ days, $R_0 = 6$)

Essentially, by forcing the transmission in synchrony with the natural period, we ensure that sequential forcing effects accumulate rather than cancel (Dushoff et al. 2004). By and large, we all have firsthand experience of the effects of resonance in a forced oscillator—on the playground. When pushing (the forcing) a child on a swing (the oscillator), we aim to optimize our effort by timing each push to coincide with the natural period (i.e., the high point) of each oscillation. In Figure 5.4, we show how the amplification of sinusoidal forcing varies with both the forcing frequency and the natural frequency of oscillations (vertical lines) for three childhood diseases.

Forcing is most greatly amplified when the forcing period is close to the natural oscillatory frequency of the unforced dynamics.

For different ratios of ω:F, however, it is possible for forcing to excite *sub-harmonic resonance* that gives rise to oscillations with a longer period than the period of the forcing. This phenomenon can occur whenever the natural period of the oscillations $1/F$ is close to an integer multiple of the period of the forcing $1/\omega$. Subharmonic resonance is dependent on the nonlinearities within the transmission process dynamics and requires substantial levels of forcing. As a result, subharmonic resonance is generally studied numerically (see, for example, Greenman et al. 2004; Choisy et al. 2006). Dietz (1976) showed that for β_1 small, some analytical understanding can be gained by setting $T = \omega/F$, and noting that the subharmonic resonance occurs only when T is an integer. We can rearrange equation (5.8) to obtain the following equation that links the observed period of oscillations (T) to the infectious period of the infection ($1/(\mu + \gamma)$) and its mean age at

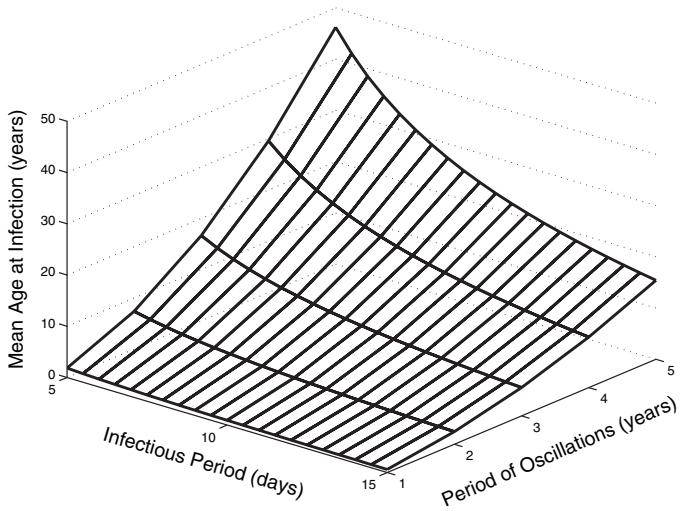

Figure 5.5. The subharmonics of a seasonally forced SIR model, as given by equation (5.9). The surface depicts the relationship between the infectious period, the mean age at infection and the observed period of oscillations. The period of forcing is assumed to be one year ($\omega = 2\pi$).

infection ($A = 1/[\mu(R_0 - 1)]$):

$$A = (\gamma + \mu)\frac{T^2}{\omega^2}. \tag{5.9}$$

For childhood infections, we can assume that the period of forcing is one year ($\omega = 2\pi$) and then examine the predicted period of oscillations as the infectious period and mean age at infection are varied (Figure 5.5). This permits a relationship to be established between the natural period of oscillations resulting from the low-level forcing of the system (T), pathogen transmissibility (in terms of the transmission rate β), and host demography (birth rate μ).

To explore in more general terms how the *amplitude* of seasonality affects dynamics, Dietz (1976) resorted to numerical integration of the underlying equations. In this way, he was able to demonstrate that changes in either R_0 or β_1 can lead to qualitatively different epidemic patterns. For example, when R_0 is large and the level of seasonal forcing is small, the fraction of infecteds shows harmonic oscillations with small-amplitude annual epidemics (Figure 5.6, top left). As β_1 increases to 0.1, we observe subharmonic resonance (as $\omega \approx 2F$) giving rise to biennial dynamics (Figure 5.6, middle left). A further increase in β_1 gives rise to four-year cycles that have a noticeable and pronounced biennial as well as annual component (Figure 5.6, bottom left). However, Dietz (1976) noted that the precise sequence of dynamical transitions that is observed depends on R_0. The second and third columns in Figure 5.6 exhibit qualitatively different dynamics in response to changes in the level of forcing. In these cases, when R_0 is smaller, increases in seasonal amplitude do not influence the period of epidemics that remain annual, but do substantially alter the magnitude of oscillations.

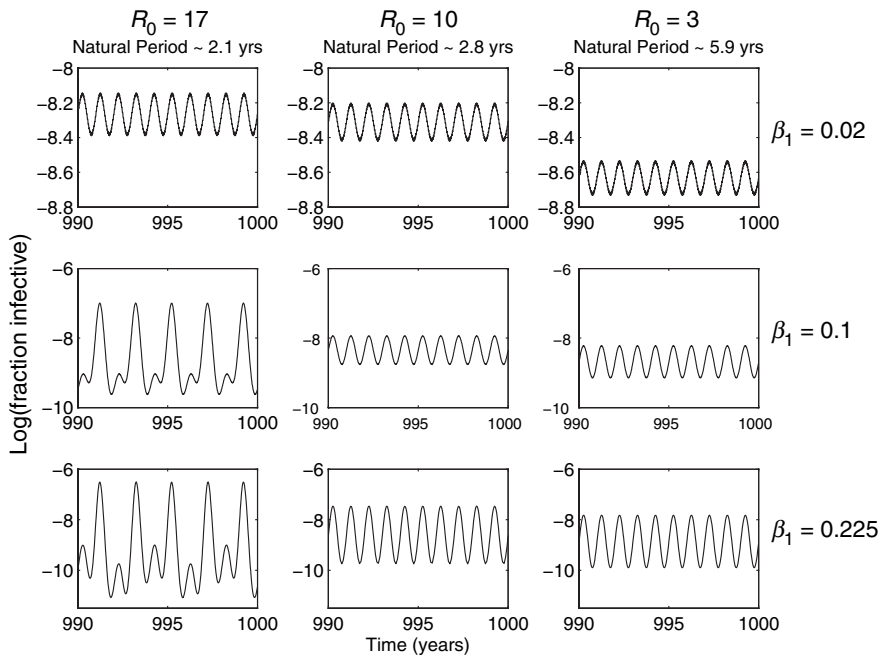

Figure 5.6. The time course of the number of infectives predicted by an $SEIR$ model as the amplitude of seasonal forcing (β_1) and R_0 are varied. The parameters used to generate these panels were $\mu = 0.02$ per year, $1/\sigma = 8$ days, $1/\gamma = 5$ days. All simulations were started with $S(0) = 6 \times 10^{-2}$ and $E(0) = I(0) = 10^{-3}$. The logarithm of the proportion of infectives is plotted for clarity. The low troughs predicted between epidemics could be sustained only by large population sizes.

In the absence of seasonal forcing, the SIR family of models exhibit a stable equilibrium (see Chapter 2). The introduction of time-dependent transmission rates can generate a variety of dynamical patterns—depending on parameter values—ranging from simple annual epidemics to multiennial outbreaks and eventually chaos.

5.2.2. Mechanisms of Multi-Annual Cycles

It is straightforward and intuitive to understand how making the transmission rate oscillate with a period of one year may generate annual epidemics. As described above, this phenomenon is referred to as simple harmonic resonance, whereby the dynamical system (in this case, the $S[E]IR$ model) simply tracks the temporal changes in the forcing (i.e., $F = \omega$). There is also the possibility (as shown in Figure 5.6) of obtaining subharmonic resonance, where oscillatory dynamics have a period that is an integer multiple of the forcing. This phenomenon can be understood by thinking about dynamics within a more ecological framework.

Consider the equation giving the rate of change of infectives in the SIR system (equation (5.7)). As demonstrated by Kermack and McKendrick (1927) and discussed in Chapter 2,

the condition for growth of disease incidence is:

$$\begin{aligned}\frac{dY}{dt} &= \beta X\frac{Y}{N} - (\gamma+\mu)Y > 0,\\ &= Y(\mu+\gamma)(R_0\frac{X}{N} - 1) > 0,\\ \Longrightarrow \frac{X}{N} &> \frac{1}{R_0} \approx \frac{\gamma}{\beta}.\end{aligned}$$

Therefore, the spread of the pathogen can occur only if there is a sufficient fraction of susceptibles in the population, with the critical value determined by the reproductive ratio (R_0). Although this concept is typically thought of in terms of the introduction of a pathogen into a population, it is also informative when thinking about seasonal systems. In such systems we can use the seasonally varying value of $\beta(t)$ to inform whether the current level of susceptibles is sufficient for the number of cases to increase. We demonstrate this concept in Figure 5.7, with the top panel depicting two sinusoidal transmission rates and two lower panels showing the relationship between the fraction of susceptibles (thick lines) and the threshold for infection spread (thin black lines) when epidemics are annual (middle graph) and biennial (bottom graph).

In the case of annual epidemics, there is a straightforward sequence of events. The peak in disease incidence coincides with the point at which the fraction of susceptibles ($\frac{X}{N} = S$) falls below $\gamma/\beta(t)$ (labeled point 1 in Figure 5.7). The fractions of susceptibles and infectives continue to decline until the rate of transmission is less than births, at which point S begins to increase. Once $S > \gamma/\beta(t)$ (labeled point 2 in Figure 5.7) disease incidence rises. This pattern is repeated once transmission has depleted susceptibles below the threshold (labeled point 3 in Figure 5.7). In this instance, the seasonal changes in the transmission rate are intimately associated with the dynamics of susceptibles in driving harmonic resonance.

For a slightly higher amplitude of seasonality, however, the picture changes in important ways (Figure 5.7, bottom graph). As before, the peak in the fraction of infectives coincides with $S = \gamma/\beta(t)$ (point 4). In this instance, however, the peak in the infectives is substantially larger than in the middle graph (due to the greater transmission rate), and as a result the fraction of susceptibles falls to much lower levels than before. Consequently, it takes much longer for births to replenish the susceptibles above the critical threshold (point 5). While at this point $S > \gamma/\beta(t)$, the transmission rate is very near its annual maximum (the threshold is near its minimum), and as a result the susceptibles do not remain above the threshold long enough to produce a large epidemic. It is only when the level of susceptibles exceeds the threshold for a second time (point 6) that a large epidemic begins. The entire process from point 4 to point 7 takes two years, representing subharmonic resonance. What these results show is that as the amplitude of seasonality increases, larger epidemics are generated that lower the level of susceptibles such that recovery to above the threshold takes far longer, resulting in longer period cycles.

5.2.3. Bifurcation Diagrams

How do we visually summarize the dynamics of seasonally forced models without needing to resort to figures containing numerous panels (such as those in Figure 5.6)? This may be achieved by constructing *bifurcation diagrams*, where a bifurcation refers to a qualitative

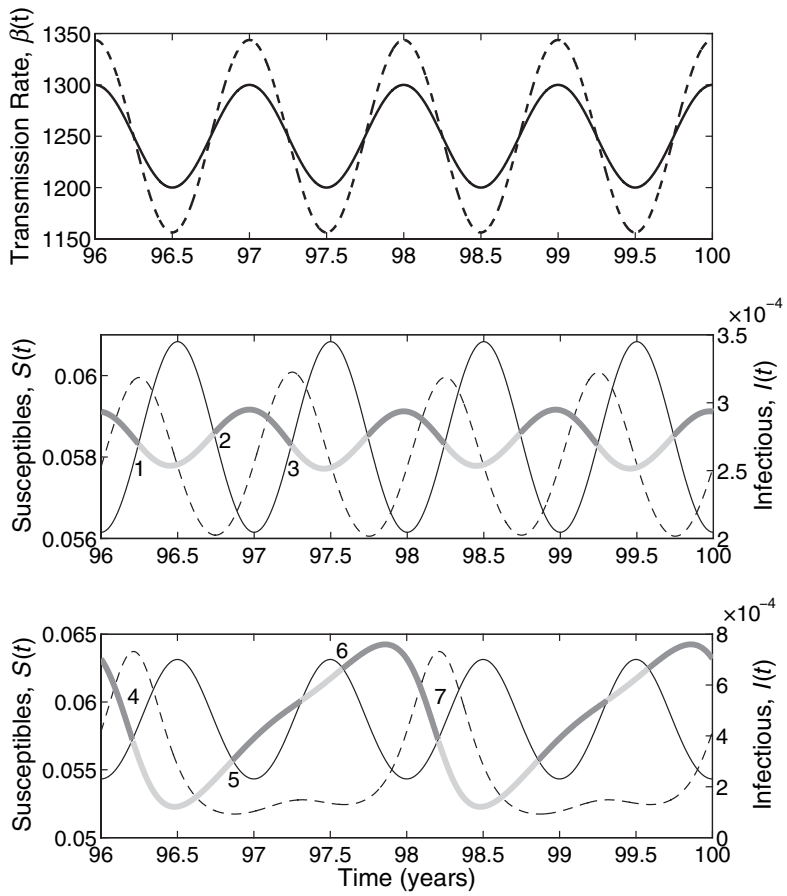

Figure 5.7. The anatomy of seasonally forced epidemics. In the top panel the seasonally varying transmission rate ($\beta(t)$ per year) is plotted, with the solid line representing the weaker forcing used in the middle graph and the dashed line representing the stronger forcing used in the bottom graph. The lower two panels demonstrate outbreak dynamics in annual (middle) and biennial (bottom) regions of parameter space (with $\beta_1 = 0.04$ and $\beta_1 = 0.075$, respectively). In addition to the fraction of susceptibles (thick line) and infectives (dashed line), we have plotted the threshold level of S required for instantaneous epidemic growth (thin line). The level of susceptibles is color coded indicating when it is above or below the threshold. The parameters used to generate these panels were $\mu = 0.02$ per year, $1/\sigma = 8$ days, and $1/\gamma = 5$ days. All simulations were started with $S(0) = 6 \times 10^{-2}$ and $E(0) = I(0) = 10^{-3}$.

change in model dynamics as a *control parameter* is altered. This is most painlessly and commonly achieved by the numerical integration of the equations as the parameter of interest (in this case, β_1) is systematically varied (Figure 5.8). For any specific parameter value, the model is started according to some specified initial conditions and integrated for a "reasonable" period of time, after which it is assumed the dynamics have reached their long-term (or asymptotic) state. Then, some measure of the population is plotted at one

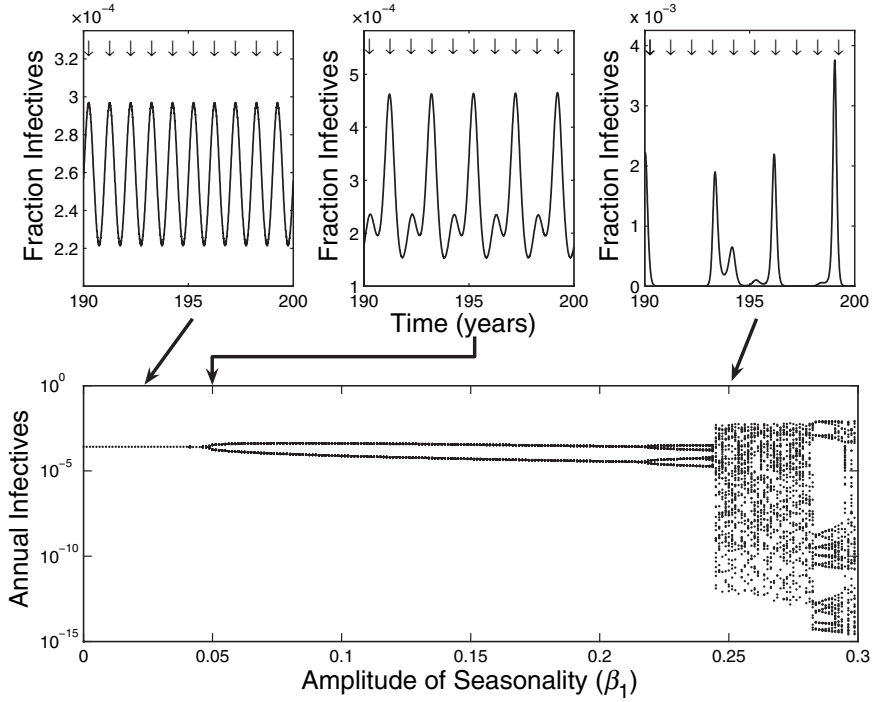

Figure 5.8. Constructing a bifurcation diagram. The top three panels depict time-series data for the $SEIR$ model with different levels of seasonality ($\beta_1 = 0.025, 0.05$, and 0.25, respectively). The arrows at the top of the panels indicate the points when the time series are sampled in order to construct the bifurcation diagram below. The parameters used to generate these panels were $\mu = 0.02$ per year, $\beta_0 = 1250$, $1/\sigma = 8$ days, and $1/\gamma = 5$ days. All simulations were started with $S(0) = 6 \times 10^{-2}$ and $E(0) = I(0) = 10^{-3}$.

particular time-point each year for the subsequent n years; the precise value of n depends on the timescale of the system but for epidemiological systems 50–100 years represent good rule-of-thumb values. Often, we produce a graph with different values of β_1 along the x-axis and the prevalence of infection (at one time each year) on the y-axis (Figure 5.8).

A typical bifurcation diagram for parameter values representative of measles is shown in Figure 5.8. To interpret the figure, we need to consider a value of β_1 and count the corresponding number of points found on the graph. For example, for $\beta = 0.1$ we observe two dots—this informs us that the dynamics are biennial and repeat every two years (see Figure 5.6, middle left, for an example of the dynamics). The bifurcation diagram shows the increasing dynamical complexity as seasonality becomes stronger. For modest levels of forcing, the dynamics mimic those of the forcing function and are rigidly annual. For β_1 greater than approximately 0.0455, the dynamics are biennial, which give way to multiennial and then aperiodic dynamics when the amplitude of seasonality exceeds 0.2. For measles, attempts to fit β_1 from time-series data have provided estimates of around 0.1–0.2 (Keeling and Grenfell 2002), therefore we have focused our attention to a restricted region of parameter space: $\beta_1 \in [0, 0.3]$. For values of β_1 greater than 0.3, the dynamics

are largely chaotic with occasional "windows" (regions of β_1 that give qualitatively similar dynamics) of multiennial cycles.

However, many elements can complicate the formulation of a bifurcation diagram, which is why several different bifurcation diagrams for the $SEIR$ model with measles parameters are given throughout this chapter. For example, although Figure 5.8 uses the amplitude of seasonality, β_1, as the control parameter, many other parameters or combinations of parameters could be used. Additionally, in Figure 5.8 a fixed set of initial conditions were used to generate each point on the bifurcation diagram and, as we show below, different initial conditions can lead to very different bifurcation patterns. Therefore, the richness of the dynamics and the large number of possible scenarios that can be considered means that a single bifurcation diagram can never fully capture the entire range of behavior.

5.2.4. Multiple Attractors and Their Basins

In Chapter 2, the solutions we obtained were globally attracting—as long as we started with some infecteds we eventually settled to the same equilibrium level. However, for seasonally forced models this simple property does not always hold; for some parameter values (generally higher values of β_1 associated with more complex behavior) the qualitative dynamics are sensitive to the initial conditions. Therefore, more than one possible dynamical behavior may exist at each point on the bifurcation diagram. In dynamical systems terminology, there are *multiple stable attractors* and which attractor is observed depends on the initial conditions—we no longer have a single globally attracting solution. Each attractor has an associated *basin of attraction*, such that whenever we start from a specific combination of variables (e.g., S, E, and I) within the basin, we eventually observe the same dynamics after transients. To demonstrate this idea, in the top graph of Figure 5.9 the number of infectives is plotted for $\beta_1 = 0.19$, depicting clear biennial dynamics as predicted by the bifurcation diagram in Figure 5.8. However, the bottom graph is also for $\beta_1 = 0.19$—the only difference between these figures is the initial conditions. The dynamics of the lower graph are qualitatively different, showing a pronounced six-year cycle.

The basin of attraction for the biennial and six-year cycles can be determined by extensive simulations examining different combinations of initial conditions. To achieve this, we fix all parameter values and systematically explore a grid of initial conditions and the resulting long-term dynamics. In Figure 5.10, we plot the results of such an exploration for $\beta_1 = 0.19$. The dark regions represent the basin of attraction for the six-year cycle and the light regions represent the combination of initial conditions for which the biennial dynamics are observed. For some combinations of initial conditions, there is considerable structure (top-right quadrant of Figure 5.10), whereas in other regions there is extreme sensitivity to the initial conditions and very small deviations can flip the long-term dynamics between attractors.

In seasonally forced systems, qualitatively different dynamical patterns can be stable for any specific combination of parameter values. Which attractor is observed depends on whether initial conditions are within the basin of attraction.

Given the possibility of many coexisting attractors for the same parameters, it may be impossible to produce a complete bifurcation diagram that displays and differentiates between the various multiple attractors (see Figure 5.11). This is especially true if some of the attractors have very small basins such that finding the appropriate set of initial

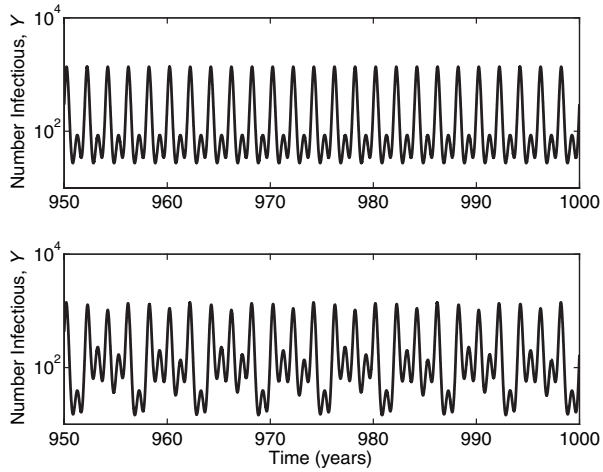

Figure 5.9. Time series from an $SEIR$ model demonstrating the dynamical consequences of using different initial conditions. The two panels share identical levels of seasonality ($\beta_1 = 0.19$), but their initial conditions vary in the number with the exposed class (top graph $W(0) = 10^{-2} \times N$, whereas for the bottom graph $W(0) = 10^{-3} \times N$). All other initial variables are the same between the two graphs ($X(0) = 6 \times 10^{-2} \times N$ and $Y(0) = 10^{-3} \times N$). The parameters used to generate these panels were $\mu = 0.02$ per year, $1/\sigma = 8$ days, $1/\gamma = 5$ days, $R_0 = 17$, and $N = 5 \times 10^6$. Note that the number of infectives is plotted on a logarithmic scale.

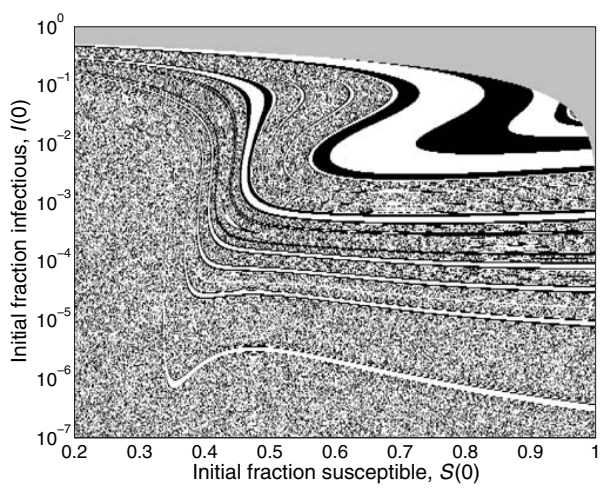

Figure 5.10. The basins of attraction for biennial and six-year dynamics. The black regions represent the combination of the initial fraction of susceptibles ($S(0)$) and infectives ($I(0)$) that give rise to six-year dynamics, and the white regions lead to the biennial attractor. Gray regions show initial conditions where $S(0) + E(0) + I(0) > 1$. Note that we set $E(0) = I(0)\gamma/\sigma$ corresponding to an equal rate of movement through the two classes. A grid of 1,000 susceptible and 460 infectious initial values was used to construct this picture. Model parameters were $\mu = 0.02$ per year, $1/\sigma = 8$ days, $1/\gamma = 5$ days, $\beta_1 = 0.19$, $R_0 = 17$.

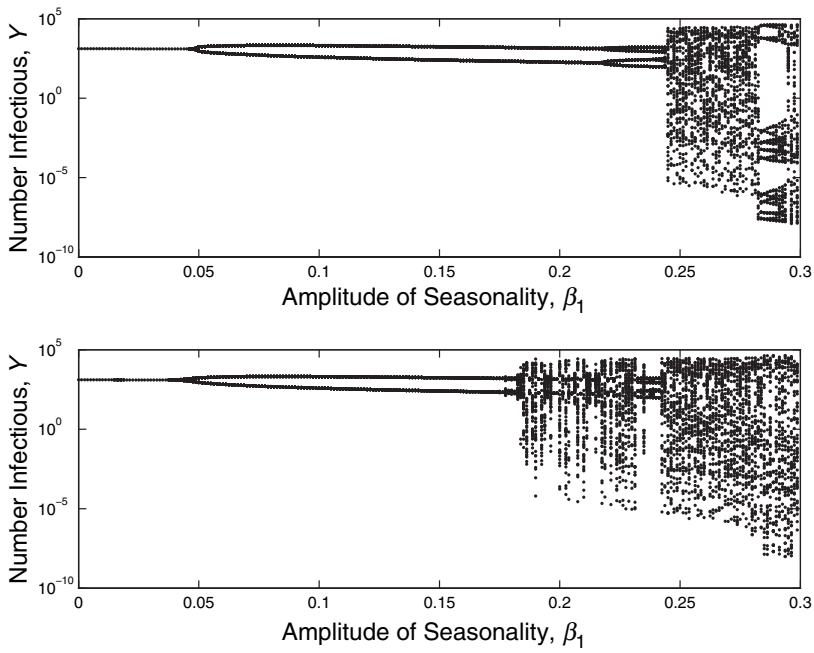

Figure 5.11. Bifurcation diagrams demonstrating the dynamical consequences of using different methods of generating initial conditions. The top graph is constructed using extrapolated initial conditions starting with $\beta_1 = 0$, where the lower graph is plotted starting with $\beta_1 = 0.3$. The parameters used to generate these panels were $\mu = 0.02$ per year, $1/\sigma = 8$ days, $1/\gamma = 5$ days, and $N = 5 \times 10^6$. Note that the number of infectives is again plotted on a logarithmic scale.

conditions is very difficult. In contrast, using a single fixed set of initial conditions can produce a rather disjoint bifurcation diagram because the changing shape of the basins of attraction means that we may suddenly jump between different attractors. An alternative is to use "extrapolated" initial conditions whereby the numbers of susceptibles, exposed, and infectives at the end of one simulation (when $\beta_1 = 0.19$, for example) are used to start the next simulation (for $\beta_1 = 0.2$). This approach ensures that in general our initial conditions start near an attractor, leading to more continuous behavior. Hence, any structural changes in the bifurcation diagram as the control parameter is varied can be confidently attributed to bifurcations (the loss of stability of one attractor giving way to a new stable regime), rather than resulting from the effects of crossing the "basin of attraction."

The use of "extrapolated" initial conditions clearly means that we could obtain different plots depending on whether we start the bifurcation diagram from the right or from the left (Figure 5.11). The top graph shows a bifurcation diagram starting from $\beta_1 = 0$ and increasing β_1 with each increment, whereas the lower graph starts with $\beta_1 = 0.3$ and decreases β_1. The overall qualitative patterns in the two graphs are similar, though they clearly differ in much of the dynamical detail.

As usual in mathematical modeling, there is a trade-off between obtaining a speedy understanding of model dynamics and mathematical rigour. Some of the issues we have

Box 5.1 Bifurcation Methods

The bifurcation diagrams plotted in, for example, Figures 5.8 and 5.11 were obtained numerically, by integrating the model equations and documenting the observed patterns. Ideally, however, we would explore model dynamics using the same kinds of principles employed for unforced models, where we used the Jacobian matrix to evaluate eigenvalues and establish the stability or otherwise of equilibria. In the forced model, we also need to establish whether perturbations made to a known trajectory are likely to die out or grow. This is achieved by studying the stability of the map $P^{(k)}$, which iterates the dynamics forward by k years:

$$P^{(k)} : (X(0), W(0), Y(0)) \mapsto (X(k), W(k), Y(k)).$$

In particular, we are interested in orbits of period k, which corresponds to $P^{(k)}$ having a fixed point such that after waiting for a period of k years we recover the state variables that we started with. These periodic solutions are referred to as *period k fixed points*. To examine their stability, we need to study the fate of perturbations to the fixed point.

Specifically, we define $X_k^*(t)$, $Y_k^*(t)$, $Z_k^*(t)$ as a k-period trajectory and make the substitutions $X(t) = X_k^*(t) + x(t)$, $Y(t) = Y_k^*(t) + y(t)$, and $Z(t) = Z_k^*(t) + z(t)$, then form equations for the dynamics of perturbations $(x(t), y(t), z(t))$ (ignoring terms that are of order $x^2(t)$ or higher). This leads to the following differential equation:

$$\begin{pmatrix} dx/dt \\ dy/dt \\ dz/dt \end{pmatrix} = J_{X_k^*(t), Y_k^*(t), Z_k^*(t)} \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}.$$

Whether the perturbations eventually die out is determined by the dominant eigenvalue $\Lambda^{(k)}$ of the Jacobian matrix J_P of the map $P^{(k)}$ evaluated at its fixed point. (The Jacobian J_P of the map can be formed by integrating the Jacobian $J_{X_k^*(t), Y_k^*(t), Z_k^*(t)}$ over a k-year period). The dominant eigenvalue is called the "floquet multiplier" of the period k solution. A fixed point is stable if and only if it has no multipliers with $\|\Lambda^{(k)}\| \geq 1$. In practical terms, this means that if we are interested in exploring the stability of the annual trajectory, for example, then we need to integrate both the $SEIR$ equations and their Jacobian for one year and then examine the resultant Jacobian's dominant eigenvalue. Note that the initial conditions here are crucial. We need to integrate the model equations starting *on* the annual attractor, whereas the initial conditions for the Jacobian (also referred to as the relational equations) are the identity matrix. It is possible to use relatively simple root-finding schemes (such as the Newton-Raphson method) to work out the period k solutions before we establish the stability.

The main advantage of this dynamical-systems approach is that abrupt changes in the floquet multipliers are very informative about the dynamics we may expect. For example, if at some value of the control parameter the multiplier $\Lambda^{(k)}$ becomes -1, then we know that we expect a period doubling bifurcation, leading to solutions with a period of $2k$. The major disadvantage of this method is that it perhaps requires much more experience with dynamical systems theory. There are, however, at least two well-established freeware programs that can be used to carry out numerical bifurcation analyses: AUTO (Doedel et al. 1998) and Matcont (Dhooge et al. 2003).

come across concerning multiple stable states and the resulting numerical bifurcation diagrams may be overcome by using more mathematically sophisticated methods that allow us to establish the stability of different solutions from basic principles (see Box 5.1). Such approaches, though substantially more technically involved, will allow us to establish *a priori* the range of dynamics we might expect to observe (see, for example, Kuznetsov 1994; Kuznetsov and Piccardi 1994; Seydel 1994).

TABLE 5.1.
Timings of the major school holidays when $Term = -1$; during all other times $Term = +1$. Note that the autumn half-term break is included because this is the only short holiday that has an identifiable signature in the England and Wales data.

Holiday	*Model Days*	*Calendar Dates*
Christmas	356–6	December 21–January 6
Easter	100–115	April 10–25
Summer	200–251	July 19–September 8
Autumn Half Term	300–307	October 27–November 3

5.2.5. Which Forcing Function?

So far, we have explored seasonality by assuming that the transmission rate is time-dependent and specifically is determined by a simple sinusoidal function. The precise shape of seasonality was historically thought to be dynamically unimportant, though this view has changed in recent years. Starting with the influential work of Schenzle (1984), seasonally forced models of childhood infections now more often use a square wave, attempting to capture the aggregation of children in schools. Specifically, the transmission rate is assumed to be high during school terms and low at other times (Bolker and Grenfell 1993; Keeling and Grenfell 1997a; Rohani et al. 1999; Earn et al. 2000; Keeling et al. 2001a). In this manner, equation (5.4) is rewritten as follows:

$$\beta(t) = \beta_0(1 + b_1\, Term(t)), \tag{5.10}$$

where $Term(t)$ is $+1$ during the school term and -1 at other times. We now use the parameter b_1 to represent the amplitude of seasonality: a slightly different notation in order to distinguish between sinusoidal and term time forcing. The historical dates of school holidays in England and Wales are presented in Table 5.1, and the resulting transmission rate throughout the year is plotted in Figure 5.12 (top graph). An important point to note is that if we sum the number of school holidays given in Table 5.1, we obtain 92, leaving 273 days of school. Because there are many more "$+1$" days than "-1" days, adopting the basic equation (5.10) is going to give rise to a mean transmission rate—averaged over the year—that exceeds β_0, with the level of excess depending on b_1. As a result, in order to ensure R_0 is constant irrespective of the precise forcing function used and the amplitude of seasonality, we need to implement a correction to equation (5.10). In general terms, if there are D_+ days of school and D_- holidays, then our forcing function would be:

$$\beta(t) = \frac{\beta_0}{\frac{1}{365}\big((1+b_1)D_+ + (1-b_1)D_-\big)}\Big(1 + b_1\, Term(t)\Big). \tag{5.11}$$

This is online program 5.2

The denominator in equation (5.11) is the mean of the forcing term, division by which ensures $\overline{\beta(t)} = \beta_0$. The changes in the transmission rate resulting from the above correction are depicted in Figure 5.12 (top graph), and the dynamical consequences of this correction are illustrated in the lower two panels. It is evident that dynamics of the basic term-time forced model (equation (5.10)) are more complex than when the corrected forcing is applied, as demonstrated by the period-doubling bifurcation occurring at a smaller amplitude of seasonality ($b_1 \sim 0.0975$ as opposed to $b_1 \sim 0.1285$). Also, without the correction term, quadrennial and higher-period epidemics ensue once $b_1 > 0.5$, whereas

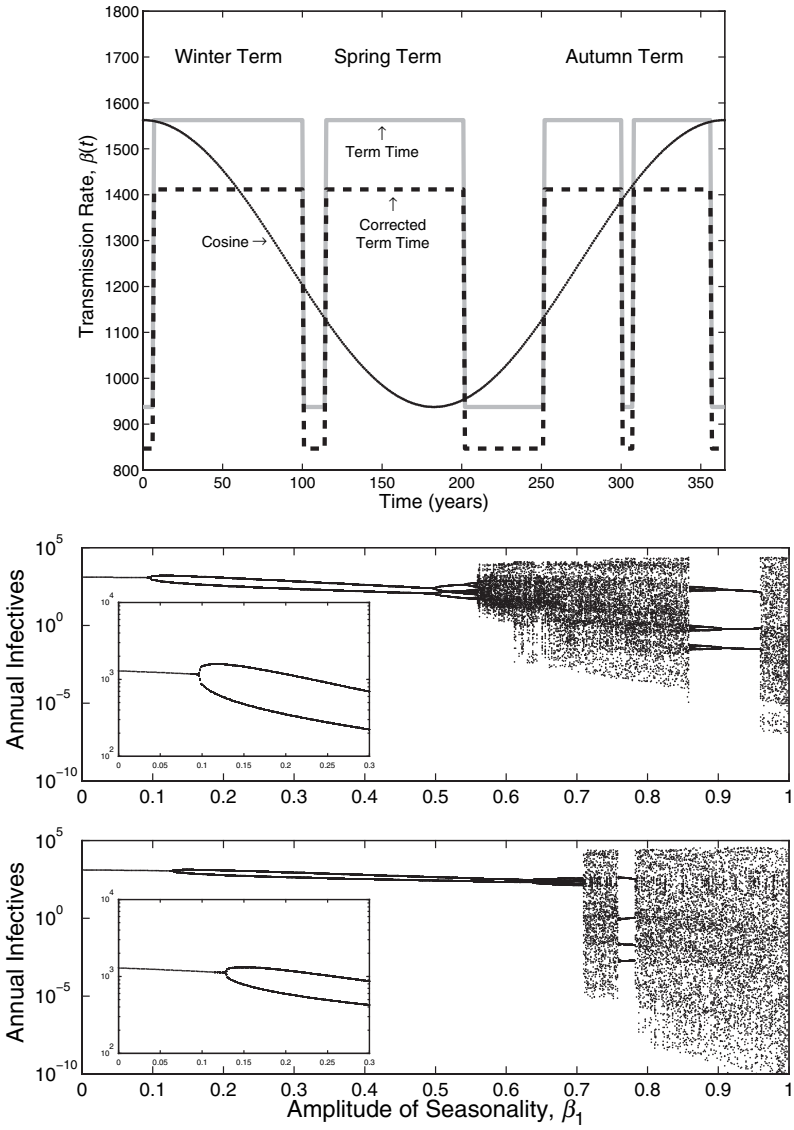

Figure 5.12. The top graph gives the transmission rate plotted through the calendar year for three different forcing functions: the sinusoidal function (solid line), the basic term time (gray line), and the corrected term time (dashed line). The figure was plotted assuming $\beta_0 = 1250$, $\beta_1 = b_1 = 0.25$, and $N = 5 \times 10^6$. The mean transmission rates ($\bar{\beta}$) are 1,250 for the sinusoidal and corrected term time ($R_0 = 17$) compared with 1,384 for the basic term time ($R_0 = 19.35$). The school term dates are given in Table 5.1. The lower two graphs show measles bifurcation dynamics with the basic (middle panel) and corrected term-time forcing function (bottom panel). The insets depict the region of parameter space around the first period-doubling bifurcation.

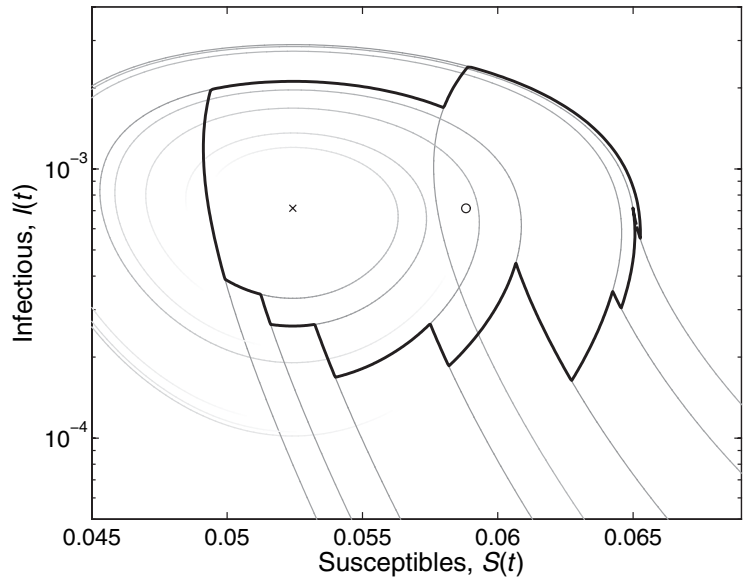

Figure 5.13. The biennial dynamics of a seasonally forced measles-like $SEIR$ model depicted in state space: the time-evolution of susceptibles versus infectious individuals as school terms change through the year. The black line demonstrates the actual biennial cycle whereas the gray lines represent continuation of the trajectory had there not been a switch at those specific points. The cross represents the position of the term-time fixed point (the holiday-time fixed point is off the right-hand side of the graph), whereas the circle represents the fixed-point of the unforced model. The figure was plotted assuming measles parameters: $\beta_0 = 1250$, $1/\sigma = 8$ days, $1/\gamma = 5$ days, $\mu = 0.02$, and $b_1 = 0.25$.

if equation (5.11) is used, complex dynamics are observed only after $b_1 > 0.65$. However, term-time forcing in general produces a much simpler bifurcation picture than sinusoidal forcing (compare Figures 5.11 and 5.12, noting the very different ranges of forcing), with complex/chaotic dynamics only occurring for relatively high values of b_1.

One way to conceptualize the behavior of a disease subject to such binary term-time forcing is as switching between two stable points or spiral sinks (one for term-time when β is high, another for holidays when β is low). Therefore, during term-times β remains constant at the higher value and the trajectories spiral toward the fixed point given by this transmission rate in exactly the same manner as predicted by an unforced model. When holidays start, a new fixed point exists and the trajectories spiral toward that. We therefore view the changing values of β as switching the model between two attracting fixed points. This idea is demonstrated in Figure 5.13 for measles parameter values; by "extending" the orbits (gray lines), the switching between the two spiral-sink attractors is clearly visible. The orbits are traced out counter-clockwise, with an abrupt change in direction every time a switch from term-time to holidays (or vice versa) occurs.

Naturally, we need to establish qualitative and the quantitative consequences of different forcing functions on model dynamics. One way of assessing this is to explore comparable bifurcation diagrams. In Figure 5.14, we present two-dimensional bifurcation figures for

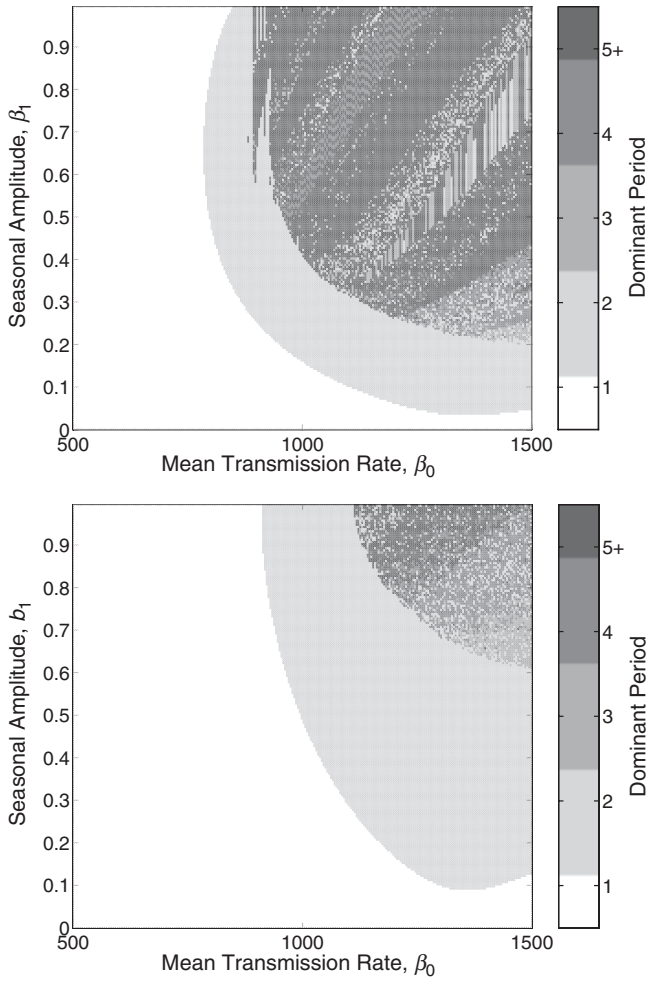

Figure 5.14. Bifurcation diagrams for the $SEIR$ model as the mean transmission rate (per year) and amplitude of seasonality are altered for sinusoidal (top graph) and term-time forcing (bottom graph). The diagrams were constructed using extrapolated initial conditions and as a result are much "cleaner" than equivalent diagrams with fixed initial conditions because these figures do not exhibit multiple stable attractors, especially for small β_1 and b_1. The period plotted is the dominant multiennial period of the Fourier spectrum (Box 5.2). The figure was plotted assuming $1/\sigma = 8$ days, $1/\gamma = 5$ days, and $\mu = 0.02$.

the sinusoidal and corrected term-time forced $SEIR$ model, using extrapolated initial conditions. (Results using a fixed set of initial conditions tend to show more multiennial cycles; extrapolation allows the annual attractor to be tracked through the parameter space.) A number of observations need to be made here. In comparison with the sinusoidally forced model (top graph), term-time forcing (bottom graph) is in some general sense more "stable"

(Bolker and Grenfell 1993). For example, with the mean transmission rate fixed at 1,250 per year, the bifurcation from annual to biennial epidemics occurs at a larger amplitude of seasonality for the term-time forcing ($b_1 \sim 0.1285$ compared with $\beta_1 \sim 0.0455$). In addition, the term-time forced models exhibit biennial epidemics for a far larger region of the parameter space, with irregular (quasiperiodic or chaotic) outbreaks observed only once b_1 exceeds approximately 0.6 (compared to $\beta_1 > 0.2$). Additionally, the sinusoidal-forced model generates very large amplitude dynamics with the proportion of infectives often falling below 10^{-20} when β_1 is large; in contrast, term-time forcing the proportion of infectives in the troughs of epidemics in Figure 5.12 always exceeds 10^{-10}. These figures have been produced using extrapolated initial conditions and as a result do not document the full dynamical complexity of these systems, especially when mean transmission rates are small (explained in detail below). Finally, if we consider the dominant period rather than the dominant multiennial period, then the region of annual behavior is extended further in both models as expected.

The choice of functional form used to represent seasonality in the transmission term can have a substantial qualitative, as well as quantitative, dynamical effect.

The introduction of time-dependence in transmission has introduced a wide array of interesting dynamics to the model. The unforced models discussed in Chapter 2 focused on equilibrium properties, whereas the models here deal with periodic epidemics. The exact period of these oscillations is determined by the characteristics of the disease, such as its mean transmission rate (β_0) and infectious period ($1/\gamma$), as well as host characteristics, such as the per capita birth rate (μ) or the amplitude of seasonality (β_1 or b_1). In order to relate models to case report data from a specific population we therefore need to be able to establish the appropriate amplitude of seasonality, as well as determine the more usual demographic and disease parameters. Unfortunately, a straightforward approach to this problem does not exist, although a number of authors have proposed different methods.

Bolker and Grenfell (1993), for example, studied the dynamics of measles in England and Wales during the 1950s and 1960s, when epidemics were clearly biennial. They constructed the average biennium (gray lines in Figure 5.15) and explored the relative goodness of fit of different models. This kind of exercise allows a clear visual inspection of the comparison between model output and data, but is not statistically rigorous. An alternative approach, adopted by Keeling and Grenfell (2002), has been to attempt to fit, in a rigorous sense, the amplitude of seasonality that is most consistent with the data. Keeling and Grenfell's findings are presented in Table 5.2. The best-fit model with sinusoidal forcing results in a lower error because it more accurately captures the timing of the epidemic peak; in contrast, the best-fit model with term-time forcing generates an epidemic peak that is slightly delayed, even though it captures more of the qualitative properties of the biennial cycle. This discrepancy highlights the need for extra biological detail: A more realistic distribution for the latent and infectious periods or including age structure (see Chapter 3) greatly reduces the error associated with term-time forcing.

For other childhood diseases, such as pertussis and rubella, where there is no systematic epidemic pattern, determining the correct level of seasonality is more complex. One approach is to adapt the time-series methods used for measles (Finkenstädt and Grenfell 2000, Bjørnstad et al. 2002), where the weekly transmission rate is estimated from a statistical viewpoint. Alternatively, we can attempt to match more generic features of

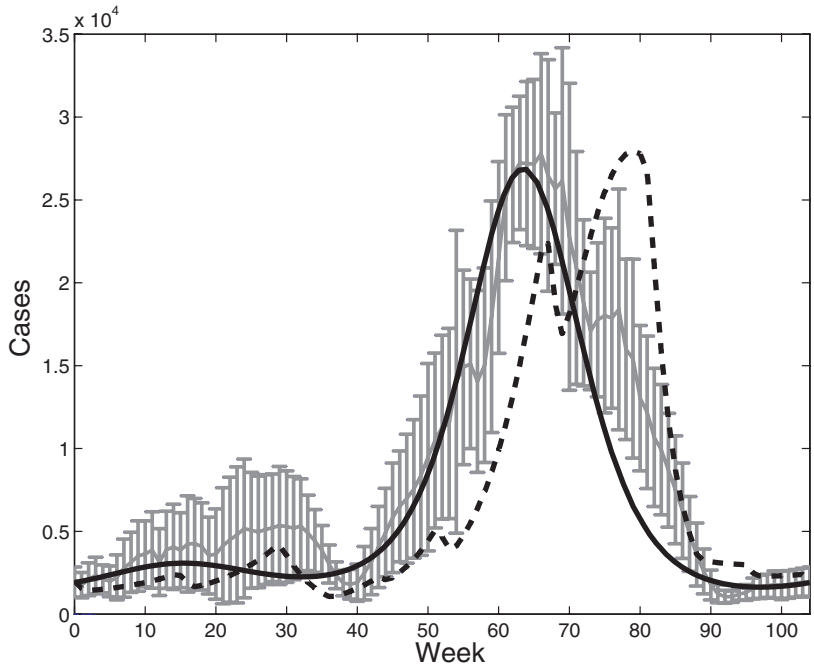

Figure 5.15. Comparing model dynamics and measles data from England and Wales. The gray lines represent the number of weekly cases averaged over the nine biennia from 1950 to 1968, with the error bars representing the standard error. The black solid line is the best fit $SEIR$ model with sinusoidal transmission ($\beta_1 = 0.09$), whereas the dotted line is the fit from the corrected term-time forced model ($b_1 = 0.29$). (The results from the basic (uncorrected) term-time model are very similar.) The figure was plotted assuming $\beta_0 = 1,250$ per year, $1/\sigma = 8$ days, $1/\gamma = 5$ days, $\mu = 0.02$ per year, $\beta_1 = 0.09$, and $b_1 = 0.29$. The estimates for β_1 and b_1 were taken from Keeling and Grenfell 2002.

TABLE 5.2.
Optimum level of seasonality estimated from measles case reports in England and Wales data and the goodness of fit (associated error, E_V) for the $SEIR$ model with the two different seasonal forcing functions (Keeling and Grenfell 2002).

Seasonal Forcing	*Term-time*	*Sinusoidal*
Best Fit	$b_1 = 0.29$	$\beta_1 = 0.11$
Associated Error	$E_V = 1.18$	$E_V = 0.64$

the data, such as the strength of annual, biennial, and multiennial signals in the data (found by taking the Fourier transform). However, if low troughs exist between the major epidemics, a stochastic framework might be required (Chapter 6) that can account for chance extinctions—with stochastic models there is the possibility of obtaining likelihood estimates which provides an alternative means of parameterization.

Box 5.2 Determining Periodicity

Here we outline two basic methods to characterize two distinct but related elements of periodicity. In particular, we look at (i) the period of the deterministic attractor and (ii) the dominant period of the epidemic cycle.

Period of Attractor

If we have an attractor with period n years, then once the dynamics are on the attractor we have the relationship:

$$S(t+n) = S(t) \qquad \text{and} \qquad I(t+n) = I(t) \qquad \forall t,$$

where the time, t, is measured in years. Despite this obvious definition, the computational reality is more far complex due to the time taken for a set of initial conditions to converge close to the attractor. We therefore stipulate that the dynamics are of period n if:

$$\begin{aligned} &|\log(S(t+n)) - \log(S(t))| < \varepsilon \ \text{ and} \\ &|\log(I(t+n)) - \log(I(t))| < \varepsilon, \end{aligned} \qquad \forall t > T_0 \tag{5.12}$$

where the time T_0 allows for convergence and ε is a small numerical tolerance. The difficulty comes in deciding on appropriate convergence times and tolerances—long times and small tolerances are more accurate, but more computationally intensive.

A further complication is that multiples of the true period will also meet criteria (5.12); in particular if the dynamics are period n, then, once close to the attractor:

$$|\log(S(t+2n)) - \log(S(t))| < |\log(S(t+n)) - \log(S(t))| \qquad \forall t > T_0,$$

and similarly for I. We must therefore insist that our dynamics are period n, if n is the smallest value for which criteria (5.12) holds.

Dominant Epidemic Period

The difficulty with the above definition is that it cannot be readily applied to real observational data or the results from stochastic models (Chapter 6). In addition, it is possible for an attractor to be of very high period and yet display annual epidemics that may take many years before they precisely repeat. An alternative is therefore to look for the *dominant* period of the epidemic cycle using Fourier spectra. We define the strength, Q_m, of the m-year cycle to be:

$$Q_m = \frac{1}{T}\left\| \int_{T_0}^{T_0+T} Y(t)\exp\left(\frac{2\pi t}{m} i\right) dt \right\|,$$

$$Q_m = \frac{1}{T}\left[\left(\int_{T_0}^{T_0+T} Y(t)\cos\left(\frac{2\pi t}{m}\right) dt \right)^2 + \left(\int_{T_0}^{T_0+T} Y(t)\sin\left(\frac{2\pi t}{m}\right) dt \right)^2 \right]^{\frac{1}{2}}$$

where T_0 is again a convergence time, and time is measured in years.

For observational data that is collected or collated at discrete time points, the integrals in the above equations are replaced by sums. To make a fair comparison, the length of the time-series used, T, must be a multiple of the period, m, being investigated. We define the dominant period as that which gives rise to the largest Q value.

One difficulty with this Fourier Spectra approach is that many multiennial cycles have a strong annual signature, due to the annual pattern of seasonal forcing. For this reason we often consider the dominant multiennial period ($m \geq 2$) and define the dynamics as annual only if the strength of all other periods is zero or insignificant. A similar result can be achieved by

calculating the Fourier Spectra of annual data Y_t (such as disease incidence at a particular time each year, or the total annual cases).

$$Q_m = \frac{1}{T}\left\| \sum_{t=T_0}^{T_0+T-1} Y_t \exp\left(\frac{2\pi t}{m} i\right)\right\|$$

For a deterministic model that has converged onto the attractor, the largest m for which Q_m is nonzero generally corresponds to the period of the attractor n defined earlier.

When dealing with observational data, or models where the periodicity appears to change, a refinement to the Fourier spectra is possible. *Morlet wavelet analysis* provides a means of determining the dominant period at any given time. At time T, we define the strength $Q_m(T)$ of the m-year cycle to be:

$$Q_m(T) = \left\| \int_{-\infty}^{\infty} \log(I(T+t)+1) \exp\left(\frac{2\pi(T+t)}{m} i\right) \exp\left(-\frac{t^2}{2m^2 V}\right) dt \right\| .$$

In effect this provides a moving average at each time, weighted by a normal distribution. Such methodology has been used to great effect in interpreting the dynamical effects of changing birth rates for measles in England and Wales (Grenfell et al. 2001).

5.2.6. Dynamical Transitions in Seasonally Forced Systems

Now that we have a best-fit estimate for the amplitude of forcing for the average biennial pattern of measles in England and Wales during the 1950s and 1960s, we wish to better understand why case notification data exhibit a more annual pattern of epidemics from 1944–1950 and irregular 2–3-year fluctuations from 1968 onward (Figure 5.16). This kind of dynamical variability has been noticed in other childhood disease incidence data (such as measles in Baltimore or rubella in Copenhagen) and has generated a substantial body of work in the search for explanations. During the 1980s, the general consensus among theoretical epidemiologists was that the epidemics of measles were chaotic, determined by a strange attractor (see, for example, Olsen and Schaffer 1990). In more recent years, however, authors have increasingly focused on more biological explanations of childhood disease epidemics. Finkenstädt et al. (1998), for example, demonstrated strong dependence of measles epidemics on population birth rates with relatively high rates associated with annual patterns, whereas modest birth rates coincided with biennial cycles.

The interaction between factors affecting host population demography and infection dynamics was clarified by Earn et al. (2000). As explained in Chapter 8, Earn et al. pointed out that changes in the recruitment rate of susceptibles—either via systematic trends in the birth rate or resulting from vaccination—can be dynamically expressed as an effective change in the mean transmission rate (β_0) of the disease. For example, by carrying out a simple change of variables, it can be demonstrated that vaccination at level p will induce epidemic patterns identical to those in an unvaccinated population with mean transmission rate $\beta(1-p)$. Similarly, changes in the birth rate μ by a given factor should produce exactly the same dynamical transitions as changing β by the same factor (see Chapter 8). Although a simple observation, this has powerful implications because it allows a single bifurcation diagram to be constructed in order to examine the dynamical consequences of varying host demography.

Such a summary diagram for the term-time forced $SEIR$ model, with epidemiological parameters chosen to correspond to measles, is shown in Figure 5.17. The control

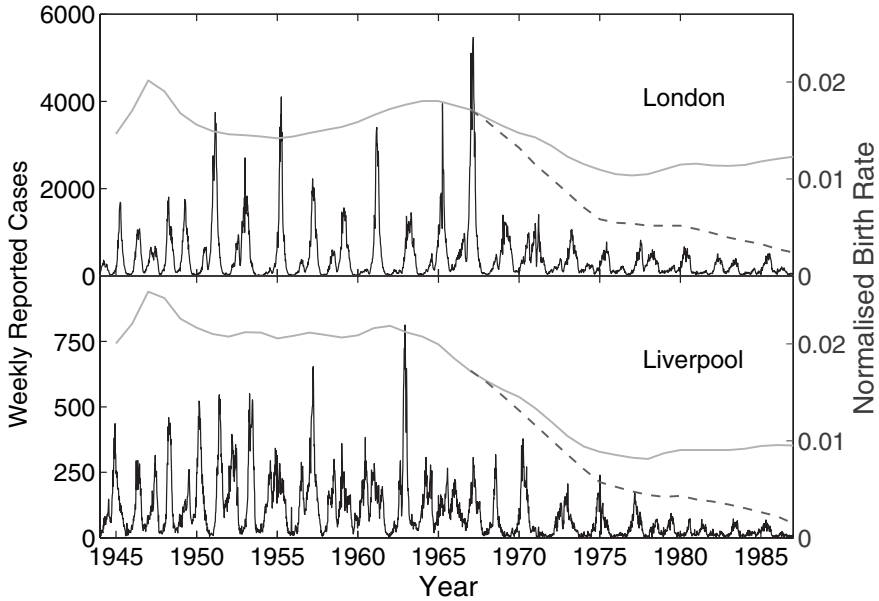

Figure 5.16. Case reports for measles in London and Liverpool from 1944 until 1988. The black line demonstrates weekly reported cases (around 60% of the true cases), with the gray line depicting the per capita birth rate. The darker gray dashed lines demonstrate the effective birth rate, correcting for vaccination that started in 1968 (see Earn et al. 2000).

parameter (x-axis) is the mean transmission rate $\overline{\beta}$ (per year), although this is representative of changes in the vaccination level or birth rate as well. The y-axis shows measles incidence on January 1 of each year, so annual cycles are represented by a single point, biennial cycles by two points, and so on. Different shades correspond to different stable solutions of the model, which attract different sets of initial conditions (basins of attraction). For four values of the mean transmission rate ($\overline{\beta} = 500, 750, 1000, 1750$), basins of attraction of the various coexisting attractors are shown above the bifurcation diagram. Where multiple stable solutions coexist, stochasticity can induce complicated dynamics due to shifts among attractors (Chapter 6). The upper panels of Figure 5.17 show that the basins of coexisting attractors are more intermixed if β is smaller, so we expect the effects of stochasticity to be greater for smaller β (or, equivalently, when the effective β is reduced by vaccination or a decrease in birth rate).

How can we use this diagram to understand measles epidemics in London and Liverpool (Figure 5.16)? The estimated mean transmission rate for this period is $\beta_0 \sim 1{,}240$ per year (Anderson and May 1991), corresponding to a biennial attractor (gray line, Figure 5.17). Before 1950, epidemics were roughly annual; over the same brief period the birth rate was much higher, which greatly increased the effective mean transmission rate, leading to an annual cycle becoming stable (black line, far right of Figure 5.17). After 1968, recruitment rates steadily decreased because of mass vaccination (for example, when vaccine uptake reached 60%, the effective mean transmission rate was reduced to $\beta_0 \sim 750 = 0.6 \times 1240$ per year); this brought the system into the parameter region where there are multiple

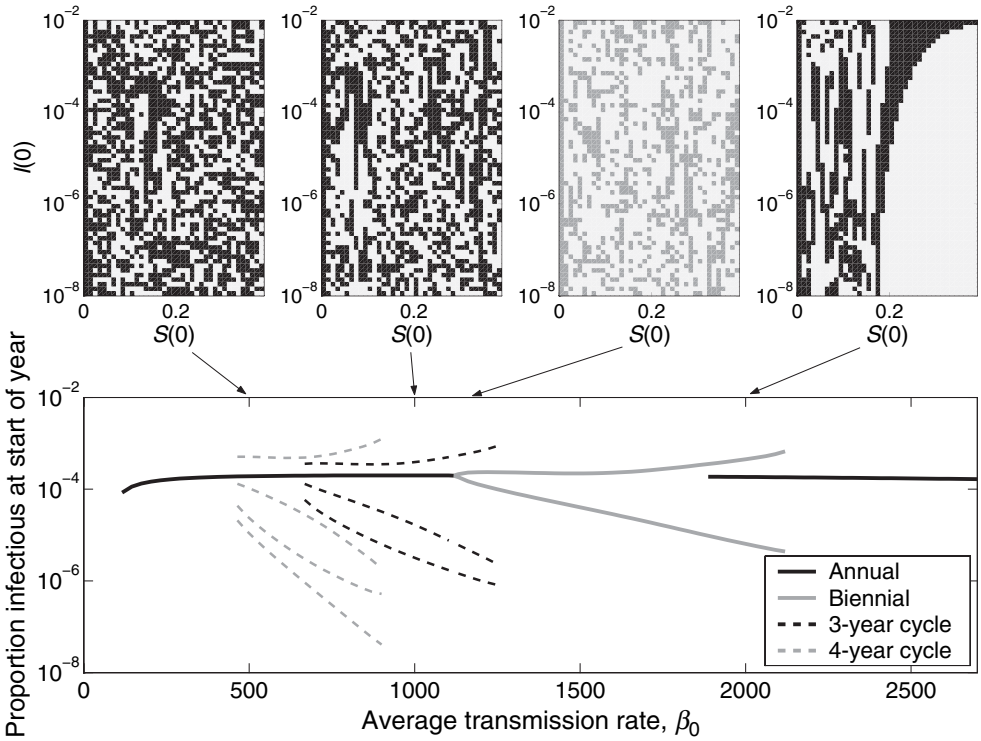

Figure 5.17. Exploring the consequences of changes in the mean transmission rate equivalent to changing the effective host birth rates (μ) for measles-like parameters. In the top panels, initial conditions that lead to annual epidemics are colored black, those that lead to biennial epidemics are dark gray, and those that lead to higher period cycles are light gray. Given the analyses of Keeling and Grenfell (2002), we assumed $b_1 = 0.29$ throughout the dynamic range. We also take a per capita birth rate of $\mu = 0.02$ per year, together with a latent period of $1/\sigma = 8$ days and an infectious period of $1/\gamma = 5$ days.

coexisting attractors with extremely intermixed basins. Stochastic effects then appear to cause frequent random jumps between these attractors (this can be confirmed using simulations—see Chapter 6), providing an explanation for the irregular epidemics in the vaccine era (Rohani et al. 1999). In Liverpool (Figure 5.16, lower graph), the birth rate was much higher than the mean in England and Wales from 1944 to 1968, leading to a higher effective transmission rate. This explains the roughly annual cycle of measles epidemics in this location over the same period. After 1968, the combination of vaccination and a lower birth rate brought Liverpool, like London, into the regime where irregular dynamics are predicted.

The bifurcation diagram in Figure 5.17 is plotted for a particular seasonal amplitude ($b_1 = 0.29$), but the qualitative conclusions of the above discussion are similar for a wide range of amplitudes. For much higher seasonality, the region with many attractors contains chaotic attractors as well. Such high seasonal amplitudes would not change our conclusion that measles dynamics will be irregular in this region. For lower seasonality, many of the

attractor sequences cease to exist or end at higher β_0, but the "ghosts of departed attractors" influence the dynamics for low β_0. This means that often the attractor becomes weakly unstable, such that if the initial conditions are close to this attractor they may take a long time to converge to a stable attractor; this generates extremely long and erratic transient dynamics especially when the unstable attractor is chaotic such that many initial conditions are close by (Rand and Wilson 1991; Earn et al. 2000). Again, this supports the prediction of irregularity, so we would expect the same qualitative dynamical picture to emerge in places that have significantly higher or lower externally imposed seasonality.

5.3. SEASONALITY IN OTHER DISEASES

Although the dynamics of measles provides an ideal test case for our ideas about the effects of seasonality, it is important to extend these concepts to other diseases (with different parameters) or to other forms of seasonal forcing.

5.3.1. Other Childhood Infections

The approach outlined above allows us to understand primarily measles epidemics in large population centers in the modern era. We followed the analysis of Earn et al. (2000) to demonstrate how changes in birth rates and the onset of vaccination can give rise to different epidemic patterns, which may provide a qualitative explanation of observed case notifications in the big cities of the United Kingdom and the United States. However, this argument ignores a potentially very important issue concerning age structure (see Chapter 3). Whether thinking about measles epidemics under extensive vaccination regimes or different infections with smaller R_0 values, we need to consider the fact that a reduction in the basic reproductive ratio gives rise to increasing transmission among older age groups, for which the effects of seasonality may be substantially different. (From Chapter 2, we know that the average age of infection is $A \approx \frac{1}{\mu(R_0-1)}$.)

To understand these effects, we use some straightforward analysis to predict how the amplitude of seasonality might change with R_0. The value of b_1 or β_1 can be approximated by calculating the extent of mixing between susceptible and infectious school children of the same age. Based on the differential equations describing the SIR system, and ignoring heterogeneities in transmission and host vital dynamics, we can calculate (at equilibrium) the proportion of the population of age a that are susceptible by solving:

$$\frac{dX^*}{da} = -\beta(t)X^*(a)\frac{Y_T^*}{N} = -\frac{X^*(a)}{A}, \tag{5.13}$$

where Y_T is the total number of infecteds summing over all ages and A is the average age of infection (Anderson and May 1991). This gives a solution $X^*(a) = \exp(-\mu(R_0-1)a)$. We can now use equation (5.13) to obtain an explicit expression for the number of infectives of age a. Assuming $\frac{dY^*(a)}{dt} \sim 0$, we get $Y^*(a) = \frac{1}{\gamma}\frac{dX^*}{da}$, which yields

$$Y^*(a) = \frac{\mu(R_0-1)}{\gamma}e^{-\mu(R_0-1)a}. \tag{5.14}$$

Now, having established explicit equations describing $X(a)$ and $Y(a)$, we can estimate the relative importance of mixing at school by studying the ratio (ψ) of mixing within

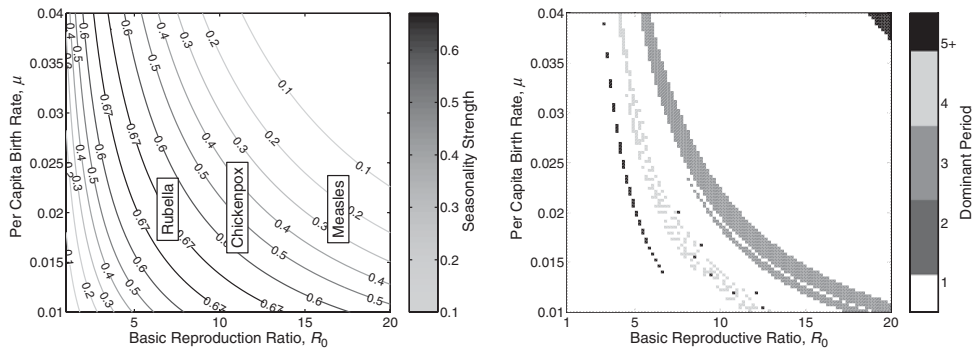

Figure 5.18. Exploring the consequences of changes in infection R_0 and host birth rates (μ) for the strength of seasonality in dynamics. Given the analyses of Keeling and Grenfell (2002), we assumed $b_1 = 0.29$ for measles during the 1950s and 1960s, when per capita birth rates were approximately 0.018 and R_0 is estimated to be 17. This allows us to use equation (5.15) to scale the amplitude of seasonality for other infections, such as chickenpox ($R_0 \sim 11$) and rubella ($R_0 \sim 7$). To produce this figure, we assumed children enter school at age $A_S = 5$ and leave at age $A_L = 16$. The precise shape of the contours (left-hand) giving the highest levels of seasonality varies slightly given different assumptions about A_S and A_L values, though the qualitative picture remains largely unaffected. The right-hand graph shows the dominant epidemic period as determined from the Fourier spectrum (Box 5.2) using an $SEIR$ with the average incubation and infection periods set to one week. All simulations are started at the unforced equilibrium solution.

school compared to random mixing:

$$\psi \propto \frac{\int_{\text{school ages}} X(a)Y(a)da}{\int X(a)da \int Y(a)da},$$

$$\propto \mu R_0\{e^{-2\mu R_0 A_S} - e^{-2\mu R_0 A_L}\}, \tag{5.15}$$

where A_S and A_L represent ages at which children start and leave school. Equation (5.15) allows us to establish the relative role of seasonality as host demographic rates (μ) and the basic reproductive ratio (R_0) vary (Figure 5.18). Intuitively, we would expect that for a given rate of births into the susceptible population, a range of R_0 values ensures that significant transmission occurs in the school-age classes and within this range the effects of seasonality in transmission will be most pronounced. For England and Wales data in the 1950–1960 era, with mean annual per capita births of approximately 0.018 per year, the window of R_0 values resulting in the largest seasonal variation in transmission is from 6–8, consistent with the estimated reproductive ratio for rubella, which is observed to have complex multiennial dynamics (Anderson and May 1982). As R_0 increases, however, the mean age at infection decreases and so does the variation in seasonal transmission. Hence, when attempting to explain the dynamics of childhood infectious diseases, it is crucial not only to obtain accurate estimates for the primary epidemiological parameters (such as transmission rates, the infectious and latent periods), but also to take into account the differential exposure to seasonality (Keeling et al. 2001a). This varying seasonality obviously has dynamic consequences for the infection within a population (Figure 5.18, right-hand graph). Multiennial epidemic cycles can be generated; the correspondence

between the contours in the left-hand graph and the regions of particular periods in the right-hand graph (Figure 5.18) shows that the degree of mixing between school children (ϕ), and therefore the level of seasonality, plays the dominant role in determining the periodicity.

For a specified pattern of school openings and closures, differences in host demography (e.g., the per capita birth rate) and the basic reproductive ratio (R_0) result in different amplitudes of seasonality in disease transmission.

The methodology outlined here is only an approximation to using fully age-structured models (Chapter 3) in which the seasonal variability is naturally included in the mixing between schoolchildren. In addition, changes in the age distribution of infectious and susceptible individuals during the course of an epidemic may cause changes in the strength of seasonality experienced, leading to deviations away from the simple binary pattern.

5.3.2. Seasonality in Wildlife Populations

In wildlife populations, seasonal changes in flocking and social mixing could also generate pulses of high transmission rates for directly transmitted infections (Altizer et al. 2006). Indeed, seasonal changes in social grouping have been demonstrated for a wide range of species in response to variation in food resources and breeding behavior (e.g., Newton-Fisher et al. 2000). Fall and winter flocking behavior and aggregations at bird feeders have been suggested as increasing the transmission and prevalence of Mycoplasma gallisepticum in house finches (Altizer et al. 2004; Hosseini et al 2004). Regular increases in the incidence of rabies in skunks during the winter and spring could be driven by seasonal host crowding (Gremillion-Smith and Woolf 1988). Outbreaks of phocine distemper virus in seals have coincided with the breeding period when animals haul out and aggregate on beaches (Swinton et al. 1998). Seasonal changes in aggressive interactions (male:male) or courtship-related contacts during the breeding season could provide further opportunities for the transmission of directly transmitted pathogens. In many animals that breed in seasonal environments, annual aggregations coincide with long-distance movement events (Dingle 1996), and seasonal migration in insects, birds, and mammals could further drive variation in parasite pressure (Folstad et al. 1991; Loehle 1995). These seasonal changes to the transmission rate for wildlife infections can be dealt with in a similar manner to the seasonal changes experienced for human diseases because their root cause—the aggregation of hosts—is essentially the same.

5.3.2.1. Seasonal Births

An alternative source of seasonality can arise from concentrating host births into a period that is short relative to the full year. This will generate a pulse of hosts that are recruited into the population at approximately the same time each year, thus effectively expanding and contracting the base of susceptible hosts throughout the course of each year (Gremillion-Smith and Woolf 1988; Bolker and Grenfell 1995). Furthermore, juveniles recruited into the population are likely to be immunologically naive and more susceptible to a variety of pathogens. Levels of herd immunity could also decline when a pulse of new juvenile hosts enters the population, leading to greater risks of infection among susceptible adults.

A growing number of empirical studies point to seasonal births, possibly in combination with changes in social behavior, as a factor important to the dynamics of wildlife pathogens, with examples spanning cowpox virus and macroparasites infecting voles and wood mice (Montgomery and Montgomery 1988; Begon et al. 1999) to phocine distemper virus in seals (Swinton et al. 1998). Mathematical models based on the biology of both vertebrate and invertebrate systems further show that the seasonal timing of reproduction can influence the dynamics of host-pathogen systems (e.g., White et al. 1996; Dugaw et al. 2004; Ireland et al. 2004; Bolzoni et al. 2006).

One way to approach modeling seasonal births is simply to make the influx rate into the susceptible population time dependent, such that the susceptible equation becomes:

$$\frac{dX}{dt} = \alpha(t)N - \beta XY - \mu X, \tag{5.16}$$

This is online program 5.3

where $\alpha(t)$ represents the time-dependent per capita birth rate. Here we have assumed transmission is density dependent, in keeping with the standard models of wildlife diseases. Often, it is assumed that $\alpha(t) = \alpha_0(1 + \alpha_1 \cos(2\pi t))$, where as before α_0 is the baseline per capita birth rate and α_1 is the amplitude of seasonality. In this instance, because birth rates are seasonally varying, it is useful to add one extra equation to keep track of the total population size N:

$$\frac{dN}{dt} = \alpha(t)N - \mu N.$$

Using the same bifurcation analysis approach as described in previous sections, it is straightforward to demonstrate that dynamics in such a model are substantially simpler than when seasonality affects the transmission rate (Figure 5.19). For illustration purposes, if we assume model parameters for measles, we find that annual outbreaks dominate unless the variation in the birth rate is extreme. This is clearly in stark contrast with Figure 5.8, where oscillations with period 2 or higher are observed for $\beta_1 > 0.0455$. This would seem to suggest that incorporating environmental variability in births is perhaps less likely to give rise to complex dynamics than similar levels of variability in transmission.

The assumption of a sinusoidally varying birth rate is somewhat naive, we therefore consider an alternative extreme where births occur in an annual pulse. This parallels the work in measles where researchers have generally switched from sinusoidal to term-time seasonality. The per capita birth rate now becomes:

$$\alpha(t) = \begin{cases} \frac{\alpha_0}{T} & \text{if } 0 \leq mod(t, 1) < T \\ 0 & \text{otherwise.} \end{cases}$$

Therefore all births for the year are compressed into a portion T at the start of each year. As such, the dynamics represent the behavior of a large number of species in temperate regions where the breeding season may be relatively short. Additionally, we stipulate that $\alpha_0 = \mu$ such that births and deaths over the year cancel each other.

Figure 5.20 shows the period of attractors that exist for hosts with an annual pulse of births. For simplicity we have let $T \to 0$, such that all births happen at the same instant; however, further numerical studies have shown that the general pattern of periods remains unchanged even when T takes quite large values (such as $T = 0.5$ years). Several interesting features exist in this pattern of coexisting periods. First, low transmission and

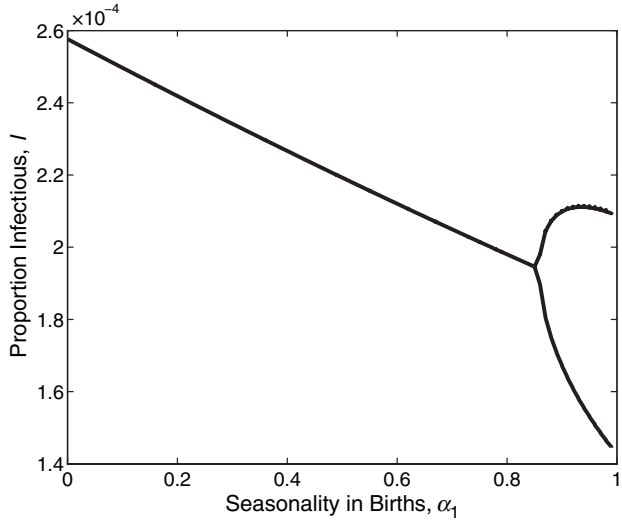

Figure 5.19. The bifurcation diagram showing the period of solutions when the birth rate is seasonally forced: The control-parameter, plotted on the x-axis, is the level of seasonality in the birth rate. Measles-like parameters are chosen so that comparison can be made with previous results: $\beta = 1240$ per year, $b_1 = 0$, $1/\sigma = 8$ days, $1/\gamma = 5$ days, $\alpha_0 = \mu = 0.02$ per year.

long life expectancy, or high transmission and short life expectancy, lead to an annual attractor (pale gray) being globally stable. Short life expectancy and low transmission can give rise to a four-year cycle (dark gray) which is again globally stable because no other attractors coexist for these parameters. Finally, we turn our attention to the bands of period 2 (diagonal hashing) and period 3 (vertical hashing). We observe that the 3-year cycle can coexist with both the annual and biennial attractor; in addition, the appearance and loss of this 3-year cycle as we move through parameter space is independent of the stability of the other periods. In contrast, the biennial cycle at its upper edge (long life expectancy) is formed by the annual cycle undergoing a period-doubling bifurcation, whereas at its lower edge both biennial and annual attractors coexist.

Despite the simplicity of the seasonality within this model, the emergent dynamics are relatively complex. This complexity increases as the infectious period becomes smaller such that longer-period attractors and even chaotic dynamics can be found.

5.3.2.2. Application: Rabbit Hemorrhagic Disease

In order to examine the relative dynamical consequences of different forms of seasonality, we now focus on wildlife diseases more closely and explore model dynamics using parameters estimated for Rabbit Hemorrhagic Disease in the United Kingdom (White et al. 2001, Ireland et al. 2004). In contrast to the above examples of seasonal variation in wildlife populations, we now allow *both* the transmission rate, β, and the per capita birth rate, ν, to be seasonally varying. In addition, following the generally observed behavior of

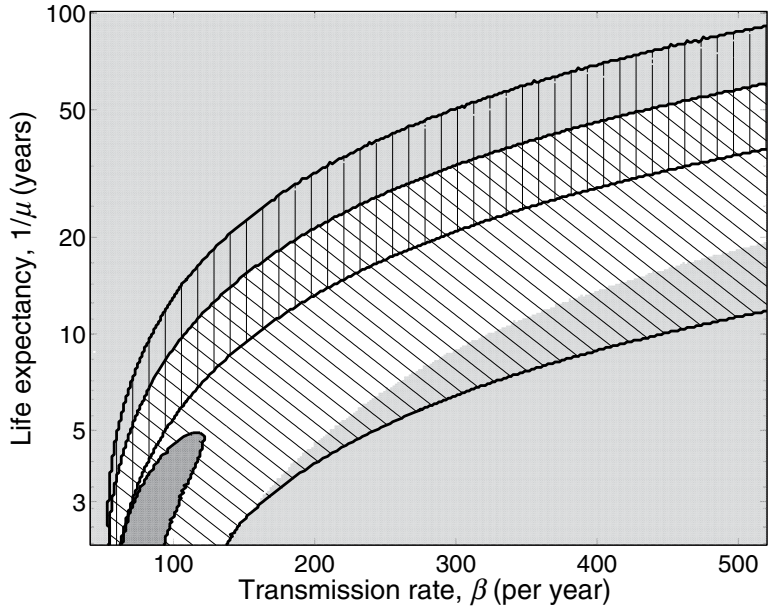

Figure 5.20. For an annual pulse of births, the parameter regimes for which various period attractors exist. Pale gray signifies annual dynamics, diagonal hashing shows the existence of biennial dynamics, vertical hashing represents three-year cycles, and dark gray signifies four-year dynamics. The diagram is computed using $1/\gamma \approx 13$ days, while the duration of the birth pulse T is reduced to zero. The period shown is the exact period of the attractor (Box 5.2), and the dominant period from the Fourier spectrum generally gives similar results with the exception that the region of exact period 4 is dominant by biennial epidemics.

natural populations, we include density-dependent regulation of the hosts (in terms of an increased death rate), which results in rewriting of the SIR system taking the basic host ecology into account:

$$\frac{dX}{dt} = \nu(t)N - \beta XY - \left(\mu + \frac{N}{K}\right)X, \tag{5.17}$$

This is online program 5.4

$$\frac{dY}{dt} = \beta XY - \left(\mu + m + \gamma + \frac{N}{K}\right)Y, \tag{5.18}$$

$$\frac{dZ}{dt} = \gamma Y - \left(\mu + \frac{N}{K}\right)Z, \tag{5.19}$$

$$\frac{dN}{dt} = \left(\nu(t) - \mu - \frac{N}{K}\right)N - \alpha Y. \tag{5.20}$$

Note that $(\nu - \mu)K$ represents the carrying capacity of the habitat, and competition for resources is assumed to affect all categories equally. In the absence of any seasonality, this

system possesses a disease-free and an endemic equilibrium. The key difference between this system of equations (excluding seasonality) and the classic SIR equations is the presence of density-dependent host regulation (represented by the $\frac{N}{K}$ per capita mortality term), which makes this system analytically intractable. For the very high mortality levels associated with this disease ($m \gg \gamma$), the SIR model above approximates SI dynamics with very few rabbits entering the recovered class. Numerical exploration, however, reveals that as expected density dependence has a strongly stabilizing effect on the endemic equilibrium (Ireland et al. 2004).

We can explore the dynamics of incorporating a breeding season by making the birth rate seasonal (such that $\nu(t) = \nu_0(1 + \alpha_1 \cos(2\pi t))$). As shown in Figure 5.21 (top graph, black points), increasing levels of seasonality in the birth rate can generate a cascade of dynamical exotica with long-period multi-ennial cycles and eventually chaos. A qualitatively similar picture emerges when seasonality is incorporated solely into the transmission term, Figure 5.21 (top graph, gray points). An important aspect of these analyses is, however, that the bifurcation from annual to biennial cycles occurs for larger values of β_1 compared to the human versions of the SIR equations (without density dependence). Indeed, the results shown in Figure 5.21 (top graph) are rather sensitive to the choice of ν_0. Higher values of this parameter result in a large influx of susceptibles, leading to annual epidemics irrespective of the amplitude of seasonality.

In some wildlife disease systems, such as conjunctivitis transmitted among house finches, two sources of seasonality arise from a breeding season which forces the birth rate and the winter/fall flocking behavior which forces the transmission rate (Hosseini et al. 2004). It is perhaps surprising that the incorporation of time-dependence in both host birth rates ($\nu(t)$) and transmission ($\beta(t)$) in equations (5.17)–(5.20) does not substantially change the dynamics, Figure 5.21 (bottom graph). Once again, we find that when the amplitude of seasonality is small, purely annual cycles ensue. As the total seasonality exceeds approximately 0.4, dynamics become extremely complex; we find a window of coexisting annual, biennial, and multiennial dynamics is found which gives way to long-period and eventually chaotic epidemics. However, when both breeding and transmission are very strongly seasonal, we once again recover more simple behavior with stable biennial cycles. In fact, if we examine the Fourier spectra (Box 5.2) in some detail we find that annual epidemics occur across the entire parameter regime, although their amplitude is modulated by multiennial cycles. This aanlysis although demonstrating some broad dynamical effects, is clearly both rather brief and somewhat ecologically naive. In reality, we may need to examine carefully the precise timing of the two seasonal mechanisms. For the house finch system, for example, the breeding season is during April to September whereas flock sizes peak during the winter months. To take these factors into account, any detailed study of the system would need to incorporate a fixed phase difference (ψ) between seasonality in breeding and transmission dynamics (i.e., $\beta(t) = \beta_0(1 + \cos(2\pi t + \psi))$ and $\nu(t) = \nu_0(1 + \cos(2\pi t + \psi))$. As shown by Hosseini et al. (2004), the addition of such features generates realistic double epidemics within a calendar year.

5.4. SUMMARY

In this chapter, we have explored epidemics in seasonally varying environments. The source of this environmental variability is varied, ranging from social aggregation of hosts (as seen in measles, phocine distemper, or mycoplasma), pulses of susceptible recruitment

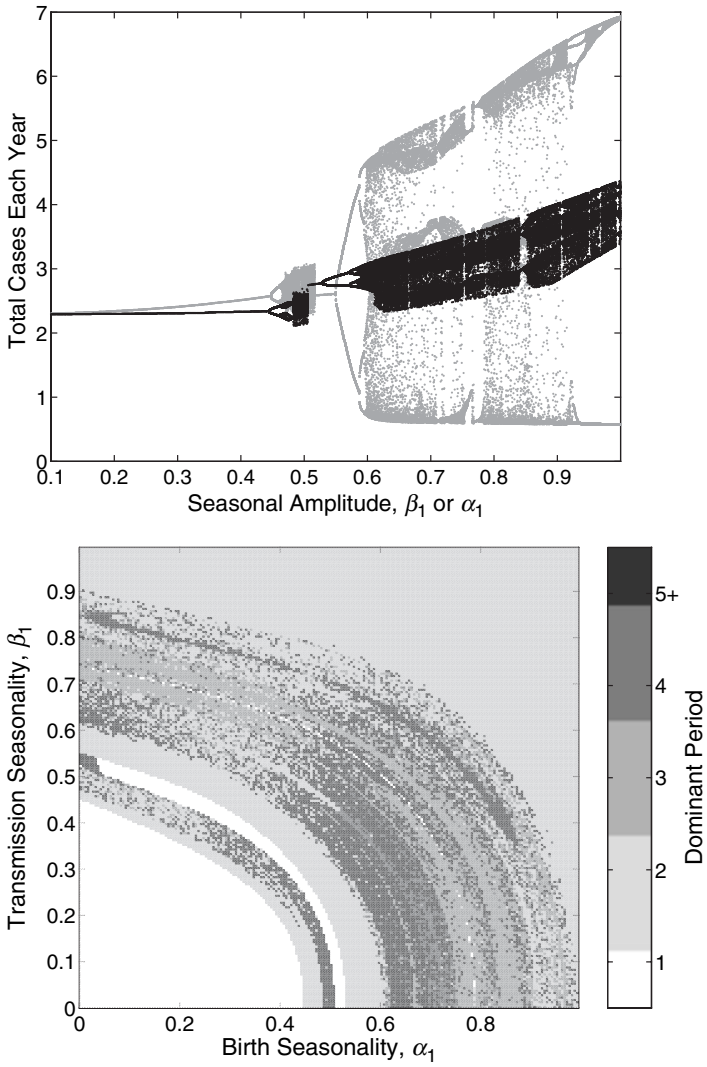

Figure 5.21. The dynamical consequences of seasonality in equations (5.17-5.20). In the top graph seasonality is implemented in either the *per capita* birth rate ($\nu = \nu_0(1+\alpha_1\cos(2\pi t))$, $\beta = \beta_0$; black points) or the transmission rate ($\beta = \beta_0(1+\beta_1\cos(2\pi t))$, $\nu = \nu_0$; gray points). In the bottom graph both disease transmission and population birth rates are assumed seasonal and in phase, although asynchronous forcing produces similar results. The period shown is the dominant multiennial period of the Fourier spectrum (Box 5.2). These bifurcation diagrams are computed using $\gamma = 0.025$ per day, $\mu = 0.01$ per day, $m = 0.475$ per day, $\nu_0 = 0.02$ per day, $\beta_0 = 0.936$ per day and $K = 10000 \Rightarrow N \approx 100$.

following breeding seasons (such as rabbit hemorrhagic disease), and regular disturbances (such as lettuce stuff). We have seen that a number of different dynamical changes can follow the introduction of seasonality. In some cases, we get harmonic oscillations, whereby epidemics simply track the periodic forcing (e.g., low levels of forcing in the measles model). As the amplitude of forcing increases, however, the phenomenon of subharmonic resonance is observed with epidemics that are integer-multiples of the forcing period (Greenman et al. 2004). In other cases, if the period of forcing matches the underlying natural period of the system, we obtain harmonic resonance.

By concentrating largely on the SIR paradigm, we have been able to demonstrate a number of important epidemiological principles:

➤ For childhood infections such as measles, chickenpox, and rubella, it is empirically established that rates of transmission peak at the start of the school year and steadily decline, reaching a trough during the summer months.

➤ Small levels of seasonal forcing in transmission can give rise to harmonic oscillations, with large-amplitude cycles in disease prevalence that have the same period as the forcing function.

➤ Moderate levels of seasonality can result in dynamics that have periods that are integer-multiples of the forcing function period—this phenomenon is known as subharmonic resonance.

➤ The effects of forcing are most pronounced when its period matches the natural period of the system—this is referred to as harmonic resonance.

➤ In systems that experience seasonal forcing, it is possible to observe qualitatively different dynamics (or multiple attractors) for the same combination of parameter values, depending on initial conditions.

➤ Whenever multiple attractors coexist, a full understanding of the system requires the basins of attraction to be determined.

➤ Seasonality may be mathematically implemented using alternative formulations, which qualitatively affect the predicted dynamics.

➤ The strength of seasonality in transmission depends on the mean age at infection, as determined by an infection's basic reproductive ratio, R_0.

➤ Seasonality in births is dynamically less destabilizing, with bifurcations from annual to multiennial dynamics occurring at larger seasonal amplitudes.