



## Spatial Heterogeneity in Epidemic Models

ALUN L. LLOYD<sup>†</sup> AND ROBERT M. MAY

*University of Oxford, Department of Zoology, South Parks Road, Oxford OX1 3PS, U.K.*

*(Received on 30 March 1995, Accepted in revised form on 15 September 1995)*

Spatial heterogeneity is believed to play an important role in the persistence and dynamics of epidemics of childhood diseases because asynchrony between populations within different regions allows global persistence, even if the disease dies out locally. A simple multi-patch (metapopulation) model for spatial heterogeneity in epidemics is analysed and we examine conditions under which patches become synchronized. We show that the patches in non-seasonal deterministic models often oscillate in phase for all but the weakest between patch coupling. Synchronization is also seen for stochastic models, although slightly stronger coupling is needed to overcome the random effects. We demonstrate that the inclusion of seasonal forcing in deterministic models can lead to the maintenance of phase differences between patches. Complex dynamic behaviour is observed in the seasonally forced spatial model, along with the coexistence of many different behaviours. Compared to the non-spatial model, chaotic solutions are observed for weaker seasonal forcing; these solutions have a more realistic minimum number of infectives.

© 1996 Academic Press Limited

### 1. Introduction

Much of the theoretical discussion of the dynamics of epidemics of childhood diseases has centred on the susceptible, exposed, infective and recovered (SEIR) compartmental model (Anderson & May, 1991 and references therein). The long-term dynamics of the basic model are simple: either the disease dies out or a stable equilibrium is reached in which the disease is endemic. A threshold condition determines which of these two fixed points is stable. If the value of the basic reproductive ratio ( $R_0$ , as defined below) is greater than one, the system settles down in the endemic state. In this case, the equilibrium is approached via damped oscillations.

Many childhood diseases exhibit recurrent epidemics, however, often with annual or biennial cycles. Such oscillations are sustained in the model if a stochastic formulation of the SEIR equations is used, as the random effects prevent the system from settling into the stable endemic equilibrium (Bartlett, 1957; Bartlett, 1960). In the deterministic framework,

oscillations can be sustained if the contact rate is allowed to vary seasonally (London & Yorke, 1973; Dietz, 1976). When the deterministic model is seasonally forced, a wide range of complex dynamic behaviour is seen, including chaos and coexisting cycles of different periods (Schwartz, 1985; Olsen & Schaffer, 1990; Engbert & Drepper, 1994). Much of this complex behaviour is only seen, however, when the system is strongly forced, and many authors argue that the degree of seasonal forcing required is unrealistic (Pool, 1989). There are other deficiencies in the model's behaviour, such as the frequency of fade-outs (episodes where there are no new cases of the disease) and the often unrealistically small number of infective individuals during minima, for instance as low as  $10^{-10}$  of population size (Bolker & Grenfell, 1993). Such a low number of infectives is clearly at variance with the recurrent nature of the epidemics.

Various kinds of heterogeneities have been proposed as answers to these problems. Age structure has been most widely studied (Anderson & May, 1984, 1991; Schenzle, 1984). It clearly plays an important role as children within the same schools spend a lot of time together during the term, which

<sup>†</sup> Author to whom correspondence should be addressed.  
E-mail: ALUN.LLOYD@ZOO.OX.AC.UK

causes an increase in the probability of disease transmission between individuals within these age classes. The opening and closing of schools is the main cause of the seasonal forcing discussed above. Age structured models can reproduce the observed disease incidences fairly well and also lead to more realistic estimates of the average age,  $A$ , at which individuals acquire infection (Schenzle, 1984). Their dynamics tends to be less complex than those of strongly seasonally forced models (Bolker & Grenfell, 1993). This has led advocates of the importance of complex dynamics in epidemics to claim that the (more complex) age structured models do not reproduce the observed dynamics as well as their strongly forced seasonal models (Tidd *et al.*, 1993; see also Grenfell *et al.*, 1994a, in which the basis of the analysis of Tidd *et al.* is questioned).

It has been suggested that spatial heterogeneity may address many of the deficiencies of epidemic models. At the simplest level these heterogeneities are included by adding an immigration term, where infective individuals enter the system at a constant rate (Olsen *et al.*, 1988; Engbert & Drepper, 1994). This clearly allows the persistence of the disease because if it dies out in one region then the arrival of an infective from elsewhere can trigger another epidemic. Indeed, the arrival of new infectives has been demonstrated as being important in the outbreaks of measles observed in Iceland, a small island community (Cliff *et al.*, 1993). A constant immigration term has a mildly stabilizing effect on the dynamics, and tends to increase the minimum number of infective individuals observed in the models (Bolker & Grenfell, 1995).

A more sophisticated way of introducing spatial effects into the model is to divide the population into  $n$  subpopulations and allow infective individuals in one patch to infect susceptible individuals in another. The equilibrium behaviour of such models has been studied widely (Lajmanovich & Yorke, 1976; Hethcote, 1978; Nold, 1980; Post *et al.*, 1983; Hethcote & Thieme, 1985; Hethcote & Van Ark, 1987), particularly with regard to the effects of spatial heterogeneity on the design of immunization programmes (Anderson & May, 1984; May & Anderson, 1984). Simulation studies have been presented (Murray & Cliff, 1975), and it has been shown that spatial heterogeneity can reduce the occurrence of fade-outs in epidemic models (Bolker & Grenfell, 1995; Grenfell *et al.*, 1995). Some attention has been directed towards understanding the dynamics of spatial models (Schwartz, 1992; Bolker & Grenfell, 1995; Grenfell *et al.*, 1995). If spatial effects are important for the persistence of the disease it is crucial

to examine phase differences between oscillations in different patches in the model. If all the patches become synchronized, then we have essentially recovered the homogeneous system, and spatial effects (at least within the model framework) will not be able to explain persistence. As part of his seminal work on epidemic models, Bartlett (1956) considered a two patch spatial model for which he demonstrated that both patches oscillate in phase as the system approaches its endemic equilibrium, although no estimate was given of the speed at which synchronization occurs. Bartlett anticipated that similar behaviour would be seen for many patch models, and we show here that this is the case.

This paper is organized as follows. Section 2 describes the general mathematical model. Section 3 examines the equilibrium properties of the model, and, using a linear approximation, considers the way in which the system approaches the endemic equilibrium. In Section 3.1 we examine a symmetric special case, and show that rapid synchronization of the oscillations in different patches occurs for all but the weakest between patch coupling. Section 4 presents deterministic and stochastic simulations of the model, which illustrate the analytic results. The effects of including seasonality in the deterministic model are examined in Section 5.

## 2. The Model

The basic SEIR model can be written as a set of three coupled nonlinear ordinary differential equations:

$$\frac{dS}{dt} = \mu N - \mu S - \lambda S \quad (1)$$

$$\frac{dE}{dt} = \lambda S - (\mu + \sigma)E \quad (2)$$

$$\frac{dI}{dt} = \sigma E - (\mu + \gamma)I \quad (3)$$

$$\lambda = \beta I. \quad (4)$$

Here  $S$ ,  $E$ ,  $I$  and  $R$  represent the numbers of susceptible, exposed (but not yet infectious), infectious, and recovered individuals. It is assumed that the number of births balances the number of deaths, so that the total population size,  $N = S + E + I + R$ , is constant. The average life expectancy,  $L$ , is  $1/\mu$ , the average latent period of the disease is  $1/\sigma$ , and the average infectious period  $1/\gamma$ . The net infection rate per susceptible,  $\lambda$ , is known as the force of infection. The constant  $\beta$  involved in calculating this rate is a measure of the rate at which each infective makes

effective contacts (those leading to transmission of the disease) with each susceptible. This form of the model assumes perfect mixing of the population; it allows no heterogeneity.

In order to incorporate spatial effects we divide the population into  $n$  subpopulations. The force of infection in patch  $i$  is given by

$$\lambda_i = \sum_{j=1}^n \beta_{ij} I_j. \quad (5)$$

We assume the epidemiological system cannot be decomposed into two or more non-interacting sub-systems. In this case the population is said to be connected (Lajmanovich & Yorke, 1976). This generalized model can be written as

$$\frac{dS_i}{dt} = \mu N_i - \mu S_i - \lambda_i S_i \quad (6)$$

$$\frac{dE_i}{dt} = \lambda_i S_i - (\mu + \sigma) E_i \quad (7)$$

$$\frac{dI_i}{dt} = \sigma E_i - (\mu + \gamma) I_i. \quad (8)$$

We assume that the total population of each patch remains constant, that is the number of births exactly balances the number of deaths and that individuals do not move permanently from one patch to another:  $N_i = S_i + E_i + I_i + R_i$ .

If the number of patches is set to equal one, then the system of equations reduces to the standard homogeneous SEIR equations. If the latent period tends to zero, corresponding to  $\sigma \rightarrow \infty$ , there will not be any individuals in the exposed classes, and the model reduces to a susceptible, infective and recovered (SIR) model. The analysis in the SIR case is much simpler than that for the full SEIR system, and in much of what follows we shall only examine this limiting case. We do not include age structure in our model as this increases both the number of variables and parameters within it, making the effects of spatial heterogeneity less transparent.

### 3. Equilibrium and stability analyses

The equilibrium values of  $S$ ,  $E$ ,  $I$  and hence  $R$  in each patch are given by

$$S_i^* = \frac{\mu N_i}{\mu + \lambda_i^*} \quad (9)$$

$$E_i^* = \frac{\lambda_i^*}{(\mu + \sigma)} \frac{\mu N_i}{(\mu + \lambda_i^*)} \quad (10)$$

$$I_i^* = \frac{\sigma}{(\mu + \gamma)} \frac{\lambda_i^*}{(\mu + \sigma)} \frac{\mu N_i}{(\mu + \lambda_i^*)}. \quad (11)$$

Then, substituting into eqn (5) for  $\lambda_i^*$ , we get

$$\lambda_i^* = \frac{\sigma}{(\mu + \gamma)} \frac{\mu}{(\mu + \sigma)} \sum_{j=1}^n \beta_{ij} \frac{\lambda_j^* N_j}{(\mu + \lambda_j^*)}. \quad (12)$$

In principle, eqn (12) can be solved for the  $n$  variables  $\lambda_i^*$  and then everything else is known. Alternatively, the equilibrium can be found by eliminating  $S_i^*$  and  $E_i^*$  and solving the following equations

$$\mu I_i^* = \left\{ \frac{\sigma \mu N_i}{(\mu + \sigma)(\mu + \gamma)} - I_i^* \right\} \sum_{j=1}^n \beta_{ij} I_j^*. \quad (13)$$

Just as in the homogenous case, there is a threshold condition which determines whether the trivial fixed point (with  $S_i^* = N_i$  and  $E_i^* = I_i^* = 0$ ) is stable or whether the endemic equilibrium exists and is stable. Above the threshold, Lajmanovich & Yorke (1976) proved that the equations for  $I_i^*$  have a unique solution with  $0 < I_i^* < N_i$ , which means there is a unique endemic equilibrium. The threshold condition is that the maximum of the real parts of the eigenvalues of the matrix  $\mathbf{T}$  is greater than one, where

$$T_{ij} = \frac{\beta_{ij} N_i \sigma}{(\mu + \sigma)(\mu + \gamma)}. \quad (14)$$

We are interested in how the system approaches the endemic equilibrium; we linearize the system, writing  $S_i = S_i^* + s_i$ ,  $E_i = E_i^* + e_i$ ,  $I_i = I_i^* + k_i$ ,  $\lambda_i = \lambda_i^* + l_i$ , where the perturbations are small (and time dependent) and second and higher order terms are ignored. Linearization leads to a set of linear first order differential equations, whose solution is a linear combination of exponentials. (If, as is seen below, the system has an eigenvalue which is repeated  $m$  times, the solution takes the form  $p(t)\exp(\Lambda t)$ , where  $p(t)$  is a polynomial of degree  $(m - 1)$ . Notice it is still the exponential term which dominates growth or decay.) Hence each variable is given time dependence  $\exp(\Lambda t)$ :  $s_i = \tilde{s}_i \exp(\Lambda t)$ , and so on. The time derivatives of these quantities are simply given by multiplying by  $\Lambda$ , and this gives the following set of linear equations (where the tildes have been dropped for convenience)

$$\Lambda s_i = -\mu s_i - \lambda_i^* s_i - l_i S_i^* \quad (15)$$

$$\Lambda e_i = \lambda_i^* s_i + l_i S_i^* - (\mu + \sigma) e_i \quad (16)$$

$$\Lambda k_i = \sigma e_i - (\mu + \gamma) k_i \quad (17)$$

$$\Lambda l_i = \sum_j \beta_{ij} \dot{k}_j = \sum_j \beta_{ij} (\sigma e_j - (\mu + \gamma) k_j). \quad (18)$$

Algebraic manipulation gives the following equation for the  $s_i$

$$(\Lambda + \mu + \lambda_i^*)s_i - \frac{\sigma(\Lambda + \mu)}{(\Lambda + \mu + \gamma)(\Lambda + \mu + \sigma)} S_i^* \sum_j \beta_{ij} s_j = 0. \quad (19)$$

Writing

$$A_{ij} = (\Lambda + \mu + \lambda_i^*)\delta_{ij} - \frac{\sigma(\Lambda + \mu)}{(\Lambda + \mu + \gamma)(\Lambda + \mu + \sigma)} S_i^* \beta_{ij}, \quad (20)$$

the equation for the  $s_i$  can be rewritten in matrix form

$$\mathbf{A}\mathbf{s} = \mathbf{0}, \quad (21)$$

where  $\mathbf{s}$  is a vector whose  $i$ -th component is  $s_i$ . Hence we need

$$\det \mathbf{A} = 0 \quad (22)$$

for there to be a non-trivial solution. This polynomial of degree  $3n$  determines the values taken by  $\Lambda$ .

In certain special cases, analytic expressions for  $S_i^*$ ,  $E_i^*$ ,  $I_i^*$  and  $\lambda_i^*$  can be obtained and the nature of the stability of the equilibrium can be examined analytically. One such case occurs when each patch is of the same size ( $N_i = N$ ) and the  $\beta_{ij}$  are such that  $S_i^* = S^*$ ,  $E_i^* = E^*$ ,  $I_i^* = I^*$  and  $\lambda_i^* = \lambda^*$ , independently of  $i$ . The quadratic for  $\lambda^*$  can be solved to give

$$\lambda^* = \frac{\mu\sigma N}{(\mu + \gamma)(\mu + \sigma)} \sum_{j=1}^n \beta_{ij} - \mu. \quad (23)$$

Once  $\lambda^*$  is known, then  $S^*$ ,  $E^*$  and  $I^*$  can be calculated. Notice that the expression for  $\lambda^*$  requires the sum  $\sum_{j=1}^n \beta_{ij}$  to be independent of  $i$ . The matrix  $\mathbf{T}$ , defined by eqn (14), is such that the sum of all entries in each row is the same. It therefore has an eigenvector which has a one in every entry, and the corresponding eigenvalue is

$$R_0 = \frac{\sigma N}{(\mu + \sigma)(\mu + \gamma)} \sum_{j=1}^n \beta_{ij}. \quad (24)$$

If  $R_0 > 1$  then the endemic equilibrium exists and is stable. The equilibrium values are

$$S^* = N/R_0 \quad (25)$$

$$E^* = \frac{\mu N}{\mu + \sigma} (1 - 1/R_0) \quad (26)$$

$$I^* = \frac{\mu\sigma}{(\mu + \gamma)(\mu + \sigma)} N(1 - 1/R_0) \quad (27)$$

$$\lambda^* = \mu(R_0 - 1). \quad (28)$$

In this situation, the matrix  $\mathbf{A}$ , defined by equation (20), can be written as follows

$$A_{ij} = \frac{-\sigma(\Lambda + \mu)S^*}{(\Lambda + \mu + \gamma)(\Lambda + \mu + \sigma)} \times \left\{ \beta_{ij} - (\Lambda + \mu + \lambda_i^*) \frac{(\Lambda + \mu + \gamma)(\Lambda + \mu + \sigma)}{\sigma(\Lambda + \mu)S^*} \delta_{ij} \right\}. \quad (29)$$

Hence, in eqn (22),  $\det \mathbf{A}$  is zero if  $\det \mathbf{B}$  is zero, where

$$B_{ij} = \beta_{ij} - \Gamma \delta_{ij} \quad (30)$$

and

$$\Gamma = (\Lambda + \mu + \lambda_i^*) \frac{(\Lambda + \mu + \gamma)(\Lambda + \mu + \sigma)}{\sigma(\Lambda + \mu)S^*}. \quad (31)$$

The values taken by  $\Gamma$  are simply the eigenvalues of the  $\beta$  matrix. More detailed information about how the system tends to the equilibrium can be obtained by finding these eigenvalues and then solving the following cubic equation for  $\Lambda$ .

$$\Gamma = (\Lambda + \mu R_0) \frac{(\Lambda + \mu + \gamma)(\Lambda + \mu + \sigma)}{\sigma(\Lambda + \mu)N} R_0 \quad (32)$$

In the limit  $\sigma \rightarrow \infty$ , when the SEIR equations reduce to the SIR equations, eqn (32) reduces to the following quadratic

$$\Gamma = (\Lambda + \mu R_0) \frac{(\Lambda + \mu + \gamma)}{(\Lambda + \mu)N} R_0, \quad (33)$$

where the definition of  $R_0$  has been modified to

$$R_0 = \frac{N}{(\mu + \gamma)} \sum_{j=1}^n \beta_{ij}. \quad (34)$$

In childhood diseases the timescale on which births occur is much longer than the timescale on which disease processes occur. In terms of the model parameters, this is because  $\sigma$  and  $\gamma \gg \mu$  and  $\mu R_0$ . Anderson & May (1991) show that under these conditions a cubic of the form (32) has one fast decaying root, which is approximately  $-(\gamma + \sigma)$ , and two other roots given by a modified form of the quadratic (33). This approximation could be used here, but in order to keep the analysis simple we consider the quadratic obtained in the SIR limit (i.e. when  $\sigma \rightarrow \infty$ ).

### 3.1. A SYMMETRIC SPECIAL CASE

One example of a  $\beta$  matrix which naturally leads to solutions of the form discussed above arises from a symmetric situation in which the contact rate is the

same within each patch, and another (usually different and smaller) rate between each pair of distinct patches. The  $\beta$  matrix takes the form

$$\beta_{ij} = \begin{cases} \beta & \text{if } i = j \\ \epsilon\beta & \text{otherwise} \end{cases}, \quad (35)$$

with  $0 \leq \epsilon \leq 1$ .

Notice that the symmetry of the matrix implies that the sum  $\sum_{j=1}^n \beta_{ij}$  is independent of  $i$ . The basic reproductive ratio is given by

$$R_0 = \frac{N}{(\mu + \gamma)} \beta(n\epsilon + 1 - \epsilon). \quad (36)$$

Notice that  $R_0$  depends on the number of patches and the strength of the coupling. In many cases, it is this overall value of  $R_0$  that can be estimated from available epidemiological data and so in some models the value of  $\beta$  is scaled according to the number of patches and coupling strength (Grenfell, 1992; Grenfell *et al.*, 1995). We choose not to do this here and it makes little difference to the analysis or in the numerical simulations discussed later.

The  $\beta$  matrix defined by eqn (35) has eigenvalues given by

$$\Gamma = \beta(1 - \epsilon),$$

repeated  $(n - 1)$  times, and

$$\Gamma = \beta(n\epsilon + 1 - \epsilon). \quad (37)$$

If the eigenvectors of the  $\beta$  matrix are examined, it can be seen that the single eigenvalue corresponds to an in-phase mode, where the numbers of individuals in each class in all patches oscillate in phase. This corresponds to the homogeneous situation. The repeated eigenvalues correspond to internal modes of the system.

For the single eigenvalue, the quadratic (33) for  $\Lambda$  simplifies (using the definition of  $R_0$ ) to

$$(\Lambda + \mu R_0)(\Lambda + \mu + \gamma) - (\mu + \gamma)(\Lambda + \mu) = 0. \quad (38)$$

This gives the standard equation for the stability of the endemic equilibrium of the homogeneous SIR equations

$$\Lambda^2 + \mu R_0 \Lambda + \mu(\mu + \gamma)(R_0 - 1) = 0. \quad (39)$$

As discussed in Dietz (1976) and Anderson & May (1991), we write  $\mu R_0 = 1/A$  ( $A$  is the average age at first infection) and  $1/\tau = \mu + \gamma \approx \gamma$  ( $\tau$  is the average duration of infectiousness). Then, as  $A \gg \tau$ , we approximately have

$$\Lambda^2 + \frac{1}{A} \Lambda + \frac{1}{A\tau} \approx 0. \quad (40)$$

This quadratic has the approximate solutions

$$\Lambda \approx -\frac{1}{2A} \pm \frac{i}{\sqrt{A\tau}} \quad (41)$$

which represents a weakly damped oscillation towards the equilibrium, with the period of the oscillation much shorter than the damping time.

For the repeated eigenvalue, the quadratic becomes

$$(\Lambda + \mu R_0)(\Lambda + \mu + \gamma)(n\epsilon + 1 - \epsilon) - (1 - \epsilon)(\Lambda + \mu)(\mu + \gamma) = 0. \quad (42)$$

Using the definitions of  $A$  and  $\tau$  given above, this gives

$$\Lambda^2 + \left[ \frac{1}{A} + \frac{n\epsilon}{\tau(1 - \epsilon + n\epsilon)} \right] \Lambda + \frac{1}{A\tau} \left[ 1 - \frac{\mu A(1 - \epsilon)}{n\epsilon + 1 - \epsilon} \right] = 0. \quad (43)$$

Because  $A \gg \tau$  and  $\mu A = A/L \ll 1$ , the eigenvalues are given to an excellent approximation by

$$\Lambda^2 + \left[ \frac{1}{A} + \frac{n\epsilon}{\tau(1 - \epsilon + n\epsilon)} \right] \Lambda + \frac{1}{A\tau} = 0. \quad (44)$$

We write

$$\frac{n\epsilon}{1 - \epsilon + n\epsilon} = \frac{1}{\alpha}, \quad (45)$$

which gives  $\alpha = 1 - 1/n + 1/(n\epsilon)$ . Unless  $\alpha$  is very large, the term in square brackets in equation (44) is of the order of  $1/\tau$ , and the  $1/A$  term may be neglected. If  $\alpha < \sqrt{A/(4\tau)}$  then the resulting quadratic has two real roots given by

$$\Lambda = \frac{-1}{2\tau\alpha} \pm \frac{1}{2\tau\alpha} \sqrt{1 - \frac{4\alpha^2\tau}{A}}. \quad (46)$$

Because  $\sqrt{1 - x} \leq 1 - x/2$ , the roots satisfy

$$\Lambda \leq -\frac{\alpha}{A}. \quad (47)$$

We have  $\alpha \geq 1$  (because  $\epsilon \leq 1$ ) and so the internal modes decay on a faster timescale than the overall mode, which decays on the timescale of  $2A$ .

If  $\alpha \geq \sqrt{A/(4\tau)}$ , then the quadratic has roots with real parts given by

$$\text{Re}\Lambda = \frac{-1}{2\tau\alpha}. \quad (48)$$

Therefore, unless  $\alpha$  is very large, the internal modes are strongly damped, decaying on the fast timescale

of  $\tau$ , compared with the overall mode which decays on the timescale of  $A$ .

The above argument relies on  $\alpha$  not being too large. The internal modes can persist for longer if

$$\alpha = 1 - \frac{1}{n} + \frac{1}{n\epsilon} \gtrsim \frac{A}{\tau}. \quad (49)$$

Because  $A \gg \tau$ , this condition can be rewritten as

$$\epsilon \lesssim \left(\frac{\tau}{A}\right) \frac{1}{n}. \quad (50)$$

In short, we see that—for all but the smallest values of the coupling—the oscillations quickly become phase locked.

Such an analysis can be repeated for other symmetric choices of the couplings between patches, such as one where the patches are taken to lie in a linear array, and each patch interacts just with its nearest two neighbours, or more generally in a way which diminishes systematically with increasing patch separation distance in the linear array. Phase locking can result in these cases, although there can be special circumstances under which the oscillations can drift apart.

This result is based on a linear stability analysis of the equilibrium and therefore only holds when the system is close enough to the equilibrium for the linearized system of equations to provide a good description of the dynamics. If the equilibrium is globally attracting, which is believed to be the case, then this will occur once enough time has passed for the system to be approaching the equilibrium. Before this time, nonlinear effects may be able to increase phase differences between oscillations in different patches.

#### 4. Simulation Results

The consequences of the above analytic results can clearly be seen in simulations of coupled SEIR (or SIR) models. For simplicity, we consider a two patch model with the  $\beta$  matrix as given by (35). We use the following set of parameter values:  $n = 2$ ,  $N_1 = N_2 = 10^6$ ,  $\mu = 0.02 \text{ year}^{-1}$ ,  $\gamma = 73.0 \text{ year}^{-1}$ ,  $\sigma = 45.6 \text{ year}^{-1}$  and  $\beta = 0.0010107 \text{ year}^{-1} \text{ infective}^{-1}$ . These values were chosen because they have previously been used in simulations of measles epidemics (Olsen & Schaffer, 1990; Grenfell, 1992; Bolker & Grenfell, 1993). The average latent period ( $1/\sigma$ ) is 8 days and the infectious period ( $\tau$ ) 5 days. For the SIR simulations, we let  $\sigma$  tend to infinity, corresponding to a disease with no latent period and an average infectious period of 5 days.

For the SIR case, the above expression for the basic reproductive ratio gives  $R_0$  equal to about 13.8, when  $\epsilon$  is zero. This gives  $A$  to be about 3.6 years and hence rapid phase locking occurs when  $\epsilon$  is larger than about 0.002. Figure 1(a) shows a simulation with the coupling strength set equal to 0.01; the initial conditions were chosen to be reasonably close to the endemic equilibrium, but such that the oscillations start in anti-phase. Phase locking is also seen to occur for smaller couplings, but occurs more slowly [Fig. 1(b), where  $\epsilon = 0.001$ ]. Similar behaviour is seen in the SEIR case (results not shown).

Stochastic formulations of SEIR equations can exhibit behaviour different from their deterministic counterparts, as discussed earlier. The consequences of stochasticity were studied using the standard Monte Carlo simulation technique (Olsen & Schaffer, 1990; Grenfell, 1992). As in the spatially homogeneous stochastic case, the random effects cloud the deterministic result. For small values of the coupling, which lead to phase locking in the deterministic system, the oscillations drift in and out of phase [Fig. 1(c)]. This should be expected because the coupling between patches is very weak and is easily swamped by the random effects and, compared with the deterministic case, the oscillations in the stochastic system are much less regular. The relative importance of stochasticity depends on the population size, and so this effect is most visible in simulations with small populations. For stronger coupling, however, the patches do become synchronized, even though the oscillations are fairly irregular [Fig. 1(d)]. Notice that, even in this case where the populations are oscillating in phase, the stochastic effects cause the two populations to follow each other less closely than they would in the deterministic case.

When the two patches are of unequal size we can solve the equations for the equilibrium and its stability numerically. An important question which arises is the dependence of  $\beta$  (or equivalently  $R_0$ ) on the size of the population in each patch. For large community sizes,  $R_0$  is only weakly dependent on community size (Anderson, 1982). This implies that the coefficient  $\beta_{ij}$  should be inversely proportional to  $N_i$  (Hethcote & Van Ark, 1987). Notice that if  $R_0$  were strongly dependent on the size of the population then the period of oscillation in each uncoupled patch would depend strongly on its population size.

We let  $N_1$  equal  $10^6$  and  $N_2$  equal  $2 \times 10^6$ . The  $\beta_{ij}$  coefficients are appropriately scaled to these population sizes, and  $\epsilon$  is taken to be 0.02. The numerically calculated values of  $\Lambda$  are approximately  $-119$ ,  $-117$ ,  $-0.11 \pm 2.7i$  and  $-0.80 \pm 2.6i$ . The first two values correspond to the approximate roots

$-(\gamma + \sigma)$  as discussed earlier. The remaining pairs of values correspond to the damped oscillations of the in-phase and internal models. The latter die away more quickly than the former and so, after a short time, the two patches oscillate in phase [Fig. 1(e)].

### 5. Maintaining Phase Differences: Inclusion of Seasonality

We have seen that coupling can often lead to synchronization. The addition of spatial degrees of freedom cannot, therefore, always explain the

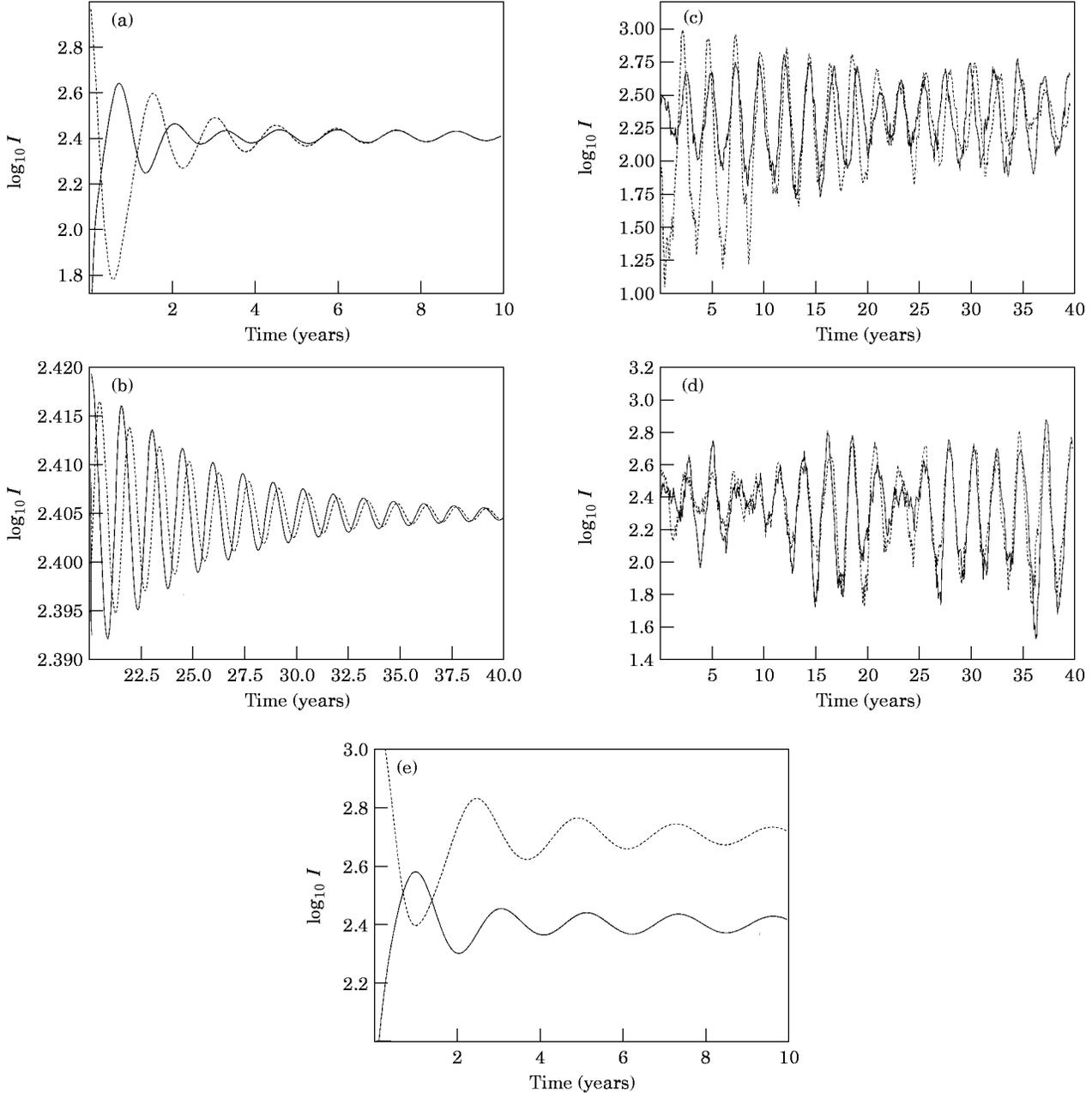


FIG. 1. Simulation of two-patch SIR model with coupling parameter  $\epsilon$  equal to (a) 0.01, (b) 0.001. The two curves show the numbers of infectives in the two patches, plotted on a logarithmic scale. Both simulations are started with the same initial conditions. Note the different scales used in (b), reflecting the longer time taken for synchronization with the weaker coupling. Simulation of two patch stochastic SEIR model with a coupling parameter  $\epsilon$  equal to (c) 0.002, (d) 0.02. With the weaker coupling in (c) the patches drift in and out of phase, which can be seen, for instance, between 15 and 20 years, and 25 and 35 years. (e) Simulation of two unequally sized SEIR patches,  $N_2 = 2N_1 = 2 \times 10^6$ , with coupling  $\epsilon = 0.02$ . In all cases, other parameter values are as defined in the text.

deficiencies of the homogeneous model. If seasonal forcing is included in the deterministic models then phase differences can be maintained. In SEIR models, a moderate level of seasonal forcing leads to a solution which oscillates with a period which is an integral number of years. The inclusion of seasonality also imposes a definite phase on oscillatory solutions; the minima occur at definite times of the year. In unforced models the phase depends on the initial conditions; the governing equations do not contain time explicitly and hence there is nothing which makes minima occur at certain times of the year.

We now consider the case in which we couple a pair of patches whose seasonal forcing causes them to undergo oscillations with a two year period. We use the phenomenological form of forcing, in which the contact rates vary sinusoidally

$$\beta_{ij} = \beta_{ij}^0 (1 + \beta_{ij}^1 \cos 2\pi t). \quad (51)$$

$\beta_{ii}^1$  is taken to be 0.2. We assume that there is no seasonality in the coupling between patches,  $\beta_{ij}^1 = 0$  for  $i \neq j$ . This assumption corresponds to the situation where the patches are taken to be collections of school catchment areas and so the increased contact rate between children during term time is a within-patch effect.

For a single uncoupled patch, this seasonal forcing produces epidemics which have large maxima every second year, with smaller peaks in the intervening years. The seasonality forces the maxima to occur at definite times of the year and so there are only two possible phases that can be seen; the large maxima can occur either in odd or in even years. This phase is determined by the  $S$ ,  $E$  and  $I$  values at time  $t = 0$ . We cannot easily depict the three dimensional plot showing the phase as a function of the three initial values, but a satisfactory compromise is a two dimensional plot showing phase as a function of the initial  $S$  and  $I$  values. The initial  $E$  is taken so that the ratio of infective to exposed individuals is  $\sigma/\gamma$  (Schwartz, 1985). (This  $E$  value is chosen since, after a very short time, the numbers of infective and exposed individuals are approximately linearly related, with their ratio being  $\sigma/\gamma$  to first order. This is another consequence of the presence of multiple timescales in the model, as discussed earlier.) Figure 2 is a plot showing how the phase depends on these initial values, with each  $(S, I)$  pair on a 200 by 200 grid coloured according to the outcome starting from these initial conditions.

Away from the fixed points, dependence of the behaviour on the initial value depends sensitively on initial conditions, and the figure shows self-similar features. Long transients are often seen for these

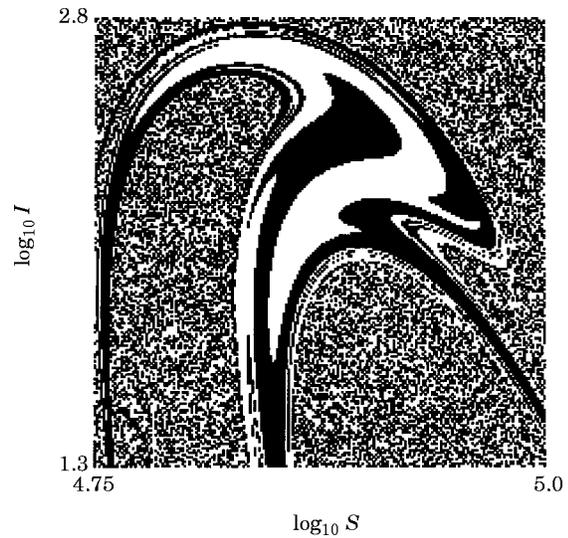


FIG. 2. Plot showing dependence of the phase of biennial oscillations on initial conditions. A white point is plotted if the corresponding  $(S, I)$  pair leads to large epidemics in odd years, a black point if large epidemics occur in even years. The initial value of  $E$  is chosen so that  $I/E = \sigma/\gamma$ , as described in the text.

initial conditions, often lasting many hundreds of years in the simulations. Such behaviour is associated with the existence of a strange repelling set on which trajectories behave chaotically for a while, but from which they eventually escape. The properties of such a set in the SEIR equations, and the implications of its existence for epidemic models is considered by Rand & Wilson (1991).

When the two patches are ‘‘coupled’’ with the trivial coupling  $\epsilon = 0$ , two different biennial behaviours are seen. Either both patches oscillate in phase, or there is a one year phase difference between them. The in-phase oscillations occur if both patches have large epidemics in the same years. The basins of attraction of the ‘‘coupled’’ system can be considered as being a product of the basins of the uncoupled system, but because of the high dimension of phase space we cannot produce a satisfactory plot of them.

In-phase and out of phase biennial oscillations are still seen as the coupling is strengthened. Larger values of the coupling favour in-phase oscillations, as one would expect as with large coupling the model approaches that of a single homogeneously mixed patch. These results echo those seen in another spatial metapopulation model, the coupled logistic map, where strong coupling causes patches to become synchronized (Lloyd, 1995). The basin of attraction for the in-phase solution grows, and that of the out of phase shrinks as the coupling is increased. As the coupled logistic map is only two dimensional, plots showing these changing basins of attraction can be produced (Lloyd, 1995).

Other solutions become more commonly observed as the coupling is strengthened. For instance, when  $\epsilon = 10^{-3}$ , a chaotic solution is seen where both patches exhibit variable amplitude oscillations (Fig. 3). For one patch, these oscillations are superimposed on a basic biennial cycle, and for the other they are superimposed on a basic triennial cycle. In order to confirm the chaotic nature of the system, the (exponential) rate at which nearby trajectories diverge was calculated using standard methods (Eckmann & Ruelle, 1985; Wolf *et al.*, 1985). This quantity, known as the largest Lyapunov exponent of the system, was found to be  $0.093 \pm 0.005 \text{ year}^{-1}$ ; its positive value demonstrates that the system shows sensitivity to initial conditions and hence can be chaotic for these parameter values.

The minimum number of infectives in either patch is about one individual (about  $10^{-6}$  of population size); this occurs in the patch undergoing the larger amplitude period three oscillations. The minimum in the period two patch is about 50 individuals. We notice that in the spatial system chaotic solutions are possible for much weaker seasonal forcing and that the minimum number of infectives in any patch is much larger than in the chaotic non-spatial model. Similar chaotic solutions have previously been observed in spatial measles models (Schwartz, 1992), although they were seen in the context of coupling a large city undergoing a high period oscillation to a small city undergoing annual oscillations. In the SEIR equations it is quite common to observe a high period, large amplitude solution coexisting with the low period oscillation (Schwartz, 1985). Despite trying a large number of initial conditions (over 5000), we did not find such a high period solution for our set of parameter values.

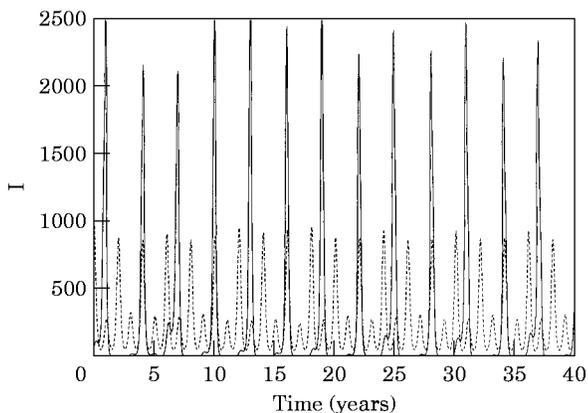


FIG. 3. Numbers of infectives in the two patches for the chaotic solution seen when two patches undergoing biennial oscillations are coupled with  $\epsilon = 10^{-3}$  (note the linear scale). The simulations were run for 500 years before results were recorded in order to run off any transient behaviour.

As we cannot make a complete exploration of the possible initial conditions, we can never be sure that further solutions do not exist: an attractor may easily be missed if it has a small basin of attraction. One technique for following the evolution of an attractor as a parameter is changed is to alter the parameter in small steps, using a point on the previous attractor as the initial condition for the system after the parameter is altered. This method is not so useful in this case because long transients are often seen before the system settles onto the new attractor, even if the parameter value is changed by an extremely small amount. The one-dimensional bifurcation diagram, which summarizes the behaviour of a model as a single parameter is altered (see, for example, Bolker & Grenfell, 1993), is difficult to interpret here as it is difficult to disentangle points which lie on coexisting attractors.

## 6. Discussion

Clearly, the general behaviour of coupled seasonally forced patch models is quite complex, and we do not address the general behaviour here. Behaviour with multiple attractors is common, and so the qualitative nature of the dynamics may depend on initial conditions as well as on the values taken by the various parameters. This will become more pronounced as the number of patches is increased, or if individual patches exhibit higher period or more complex dynamics. If stochasticity, such as the demographic stochasticity discussed earlier, is included in the model then it may be possible for the random effects to move the system from one attractor into another, for instance from biennial into triennial oscillations. Furthermore, some solutions may have small basins of attraction or exist only for small sets of parameter values, and so are easily missed in a numerical study. We have also assumed, in common with many epidemiological modellers, that demographic parameters such as the birth and death rates are constant. The effects of demographic changes can be observed, for instance, in the England and Wales measles incidence time series where the effects of the “baby-boom” following the war are clearly seen (Grenfell *et al.*, 1994b).

The data for measles incidence in England and Wales is available on a city-by-city scale, allowing the simulation results to be compared with a real epidemiological system. In the post war years, before the introduction of mass vaccination, the epidemics in various cities generally occurred at biennial intervals, and were mainly in phase (Grenfell *et al.*, 1994a). During other periods, out of phase biennial

oscillations have been seen, for instance between London and Birmingham in the period between the world wars (Bryan Grenfell, personal communication). With the introduction of mass vaccination during the sixties, which reduced the size of epidemics, a change in pattern was seen, with a reduced coherence between oscillations in different cities. A similar change was observed in the United States (Cliff *et al.*, 1992).

A simple modification to the SEIR model allows the effects of vaccination to be studied (Anderson & May, 1991). It is assumed that a fraction  $p$  of infants are vaccinated at birth, and move immediately into the recovered class. The term  $\mu N$  representing the birth of new susceptibles is replaced by  $\mu N(1 - p)$ . Vaccination reduces the value of  $R_0$ , leading to an increase in the average age at infection ( $A$ ). Equation (50) suggests that this change would increase the tendency for patches to become synchronized, which is the opposite of what is seen in the observed time series. Deterministic and stochastic simulations of the unforced system echo the analytic result, showing an increase in coherence after vaccination. The analytic result is obtained from a linear approximation to a model which itself neglects age structure and seasonality. In addition, vaccination reduces the size of the population which can be involved in the epidemic process, increasing the effects of demographic stochasticity. Clearly, some or all of these factors need to be retained in order to obtain realistic simulations. Models including all of these effects do show a decreased coherence between patches after the introduction of vaccination (Grenfell & Bolker, manuscript in preparation), and preliminary investigations of our own show that decreased coherence is seen in seasonally forced, non-age structured stochastic models.

Grenfell *et al.* (1995) show that stronger seasonal forcing leads to increasing synchrony between patches, and a higher proportion of fade-outs. In seasonally forced spatial models, there is a complex interaction between the within-patch dynamics and the two forms of forcing; seasonal and spatial coupling, with high levels of either kind of forcing often leading to synchrony. As found in other metapopulation models, strong spatial coupling effectively reduces the system to a single patch. In contrast with other metapopulation models (for example Allen *et al.*, 1993), Grenfell *et al.* (1995) show that coupling chaotic patches (i.e. those with strong seasonal forcing) does not lead to greater asynchrony between patches.

Other forms of heterogeneity, such as age structure or genetic heterogeneity, can be incorporated into the

basic SEIR framework using a methodology fairly similar to that described here (Anderson & May, 1991). The analysis laid out above can be altered in a straightforward way to cover these situations. We would not be surprised to see subpopulations in other unforced models to become synchronized in a similar way.

The study of spatially extended systems poses many interesting challenges throughout ecology. One lesson which should be learnt from this simple model is that it may be essential to incorporate many biological factors, such as seasonality, demographic or environmental stochasticity, as well as other heterogeneities such as age structure, into the model. One of the most exciting avenues for future work is the study of the disease incidence time series which contain spatial data, such as the England and Wales series mentioned above. These give important information regarding the spatial processes involved in epidemics (Cliff *et al.*, 1993), and their use should help in the construction of more realistic spatial models.

We wish to thank Bryan Grenfell for discussions and comments on the manuscript. A.L.L. is supported by the Wellcome Trust (Biomathematical Scholarship, grant number 036143/Z/92/Z) and R.M.M. by the Royal Society.

## REFERENCES

- ALLEN, J. C., SCHAFFER, W. M. & ROSKO, D. (1993). Chaos reduces species extinction by amplifying local population noise. *Nature, Lond.* **364**, 229–232.
- ANDERSON, R. M. (1982). Transmission dynamics and control of infectious disease agents. In: *Population Biology of Infectious Diseases* (Anderson, R. M & May, R. M., eds) pp. 149–176. Berlin: Springer-Verlag.
- ANDERSON, R. M. & MAY, R. M. (1984). Spatial, temporal and genetic heterogeneity in host populations and the design of immunization programmes. *IMA J. Math. Appl. Med. Biol.* **1**, 233–266.
- ANDERSON, R. M. & MAY, R. M. (1991). *Infectious Diseases of Humans: Dynamics and Control*. Oxford: Oxford University Press.
- BARTLETT, M. S. (1956). Deterministic and stochastic models for recurrent epidemics. In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, vol 4 (Neyman, J. ed.) pp. 81–109. Berkeley: University of California Press.
- BARTLETT, M. S. (1957). Measles periodicity and community size. *J. R. Stat. Soc. A* **120**, 48–70.
- BARTLETT, M. S. (1960). The critical community size for measles in the United States. *J. R. Stat. Soc. A* **123**, 37–44.
- BOLKER, B. M. & GRENFELL, B. T. (1993). Chaos and biological complexity in measles dynamics. *Proc. R. Soc. Lond.* **B251**, 75–81.
- BOLKER, B. M. & GRENFELL, B. T. (1995). Space, persistence, and dynamics of measles epidemics. *Phil. Trans. R. Soc. Lond.* **B348**, 309–320.
- CLIFF, A. D., HAGGETT, P., STROUP, D. F. & CHENEY, E. (1992). The changing geographical coherence of measles morbidity in the United States, 1962–88. *Stat. Med.* **11**, 1409–1424.

- CLIFF, A., HAGGETT, P. & SMALLMAN-RAYNOR, M. (1993). *Measles: An Historical Geography of a Major Human Viral Disease*. Oxford: Blackwell.
- DIETZ, K. (1976). The incidence of infectious diseases under the influence of seasonal fluctuations. *Lect. notes Biomath.* **11**, 1–15.
- ECKMANN, J.-P. & RUELE, D. (1985). Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **57**, 617–656.
- ENGBERT, R. & DREPPER, F. R. (1994). Chance and chaos in population biology—models of recurrent epidemics and food chain dynamics. *Chaos, Solitons and Fractals* **4**, 1147–1169.
- GRENFELL, B. T. (1992). Chance and chaos in measles dynamics. *J. R. Stat. Soc.* **B54**, 383–398.
- GRENFELL, B. T., KLECZKOWSKI, A., ELLNER, S. P. & BOLKER, B. M. (1994a). Measles as a case study in nonlinear forecasting and chaos. *Phil. Trans. R. Soc. Lond.* **A348**, 515–530.
- GRENFELL, B. T., BOLKER, B. M. & KLECZKOWSKI, A. (1994b). Seasonality, demography, and the dynamics of measles in developed countries. In: *Epidemic Models: Their Structure and Relation to Data* (Mollison, D., ed.) pp. 248–268. Cambridge: Cambridge University Press.
- GRENFELL, B. T., BOLKER, B. M. & KLECZKOWSKI, A. (1995). Seasonality and extinction in chaotic metapopulations. *Proc. R. Soc. Lond.* **B259**, 97–103.
- HETHCOTE, H. W. (1978). An immunization model for a heterogeneous population. *Theor. Popul. Biol.* **14**, 338–349.
- HETHCOTE, H. W. & THIEME, H. R. (1985). Stability of the endemic equilibrium in epidemic models with subpopulations. *Math. Biosci.* **75**, 205–227.
- HETHCOTE, H. W. & VAN ARK, J. W. (1987). Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation, and immunization programs. *Math. Biosci.* **84**, 85–118.
- LAJMANOVICH, A. & YORKE, J. A. (1976). A deterministic model for gonorrhoea in a nonhomogeneous population. *Math. Biosci.* **28**, 221–236.
- LLOYD, A. L. (1995). The coupled logistic map: a simple model for the effects of spatial heterogeneity on population dynamics. *J. theor. Biol.* **173**, 217–230.
- LONDON, W. P. & YORKE, J. A. (1973). Recurrent outbreaks of measles, chickenpox and mumps. I. Seasonal variation in contact rates. *Am. J. Epidem.* **98**, 453–468.
- MAY, R. M. & ANDERSON, R. M. (1984). Spatial heterogeneity and the design of immunization programmes. *Math. Biosci.* **72**, 83–111.
- MURRAY, G. D. & CLIFF, A. D. (1975). A stochastic model for measles epidemics in a multi-region setting. *Trans. Inst. Brit. Geog.* **2**, 158–174.
- NOLD, A. (1980). Heterogeneity in disease-transmission modeling. *Math. Biosci.* **52**, 227–240.
- OLSEN, L. F. & SCHAFFER, W. M. (1990). Chaos versus noisy periodicity: alternative hypotheses for childhood epidemics. *Science* **249**, 499–504.
- OLSEN, L. F., TRUTY, G. L. & SCHAFFER, W. M. (1988). Oscillations and chaos in epidemics: a nonlinear dynamic study of six childhood diseases in Copenhagen, Denmark. *Theor. Popul. Biol.* **33**, 344–370.
- POOL, R. (1989). Is it chaos, or is it just noise? *Science* **243**, 25–28.
- POST, W. M., DEANGELIS, D. L. & TRAVIS, C. C. (1983). Endemic disease in environments with spatially heterogeneous host populations. *Math. Biosci.* **63**, 289–302.
- RAND, D. A. & WILSON, H. B. (1991). Chaotic stochasticity: a ubiquitous source of unpredictability in dynamics. *Proc. R. Soc. Lond.* **B246**, 179–184.
- SCHENZLE, D. (1984). An age-structured model of pre- and post-vaccination measles transmission. *IMA J. Math. Appl. Med. Biol.* **1**, 169–191.
- SCHWARTZ, I. B. (1985). Multiple stable recurrent outbreaks and predictability in seasonally forced nonlinear epidemic models. *J. Math. Biol.* **21**, 347–361.
- SCHWARTZ, I. B. (1992). Small amplitude, long period outbreaks in seasonally driven epidemics. *J. Math. Biol.* **30**, 473–491.
- TIDD, C. W., OLSEN, L. F. & SCHAFFER, W. M. (1993). The case for chaos in childhood epidemics. II Predicting historical epidemics from mathematical models. *Proc. R. Soc. Lond.* **B254**, 257–273.
- WOLF, A., SWIFT, J. B., SWINNEY, H. L. & VASTANO, J. A. (1985). Determining Lyapunov exponents from a time series. *Physica D* **16**, 285–317.