



Multiplicative Moments and Measures of Persistence in Ecology

MATT J. KEELING*

Zoology Department, Cambridge University, Downing Street, Cambridge, CB2, 3EJ, U.K.

(Received on 9 June 1999, Accepted in revised form on 30 March 2000)

Ecologists and epidemiologists have begun focusing on demographic stochasticity and spatial heterogeneity as important biological factors. With high-powered computers simulation of such systems is a common modelling technique; however we lack a detailed understanding of the processes involved. Moment closure approximations provide a simple method which can be used to capture the main features of a wide variety of stochastic models and to gain a more intuitive understanding. In this paper we give an alternative variation based on multiplicative moments which is equivalent to taking a novel third-order cumulant approximation. The differential equations for these multiplicative moments are far more robust than their additive counterparts. We use this technique to consider the behaviour and persistence of finite metapopulations for two common ecological systems.

© 2000 Academic Press

Introduction

For many years, ecological and epidemiological theory has been dominated by deterministic equations and the fixed points or other forms of attractor that are predicted. More recently it has been realized that, due to the individual nature of the elements involved and the in-built stochasticity of the interactions, no biological system conforms to this deterministic ideal (Bartlett, 1956; Durrett & Levin, 1994; Bolker *et al.*, 1997). Once the essential stochasticity of a system has been acknowledged, the persistence of populations becomes an important consideration. For deterministic systems, it is common for the absence of a population to be an unstable fixed point—so from any non-zero state all populations persist. However, for stochastic integer-based populations, chance events may cause the number of individuals to hit zero, after which recovery can

only come from an external source. A population is said to persist if such extinction events are rare.

Spatial heterogeneity has been considered as a major factor in the greater persistence of natural systems compared to simple models (Hassell & May, 1974; Hassell *et al.*, 1991; Grenfell *et al.*, 1995; Keeling, 1997). One of the simplest and therefore most common forms of incorporating space into mathematical models is the use of metapopulation or patch models (Gilpin & Hanski, 1991; Hanski & Gilpin, 1997; Grenfell & Harwood, 1997). These frequently take the form of local interactions within each sub-population together with coupling between the sub-populations (Keeling, 1999). Most theoretical studies have assumed a classical Levins-type metapopulation (Levins, 1969) where internal dynamics are ignored and each sub-population is considered to be in one of two states, empty or occupied (Gilpin & Hanski, 1991; Hanski & Gilpin, 1997; ter Braak *et al.*, 1998). Here we consider full stochastic metapopulations where

* E-mail: matt@zoo.cam.ac.uk

there are an integer number of interacting individuals within each subpopulation (cf. Grenfell & Bolker, 1998; Lande, 1998; Swinton *et al.*, 1998), and it is these stochastic interactions that lead to localized extinctions.

The standard view is that stochasticity increases the likelihood for a population to go extinct, but that spatial heterogeneity promotes persistence. In this paper we categorize extinctions into two distinct classes, local and global. Local extinctions are situations where an organism has died out within a small localized habitat (a subpopulation), whereas global extinctions refer to the complete loss of the organism (e.g. the species is absent from the entire metapopulation). In conservation terms, a localized extinction may be worrying but is reversible, whereas with global extinctions the species is lost forever. When considering a disease, local extinctions at the community scale are common place, and it is global eradication that is the goal of vaccination programs. These two forms of extinction will be shown to behave very differently under changes in the model parameters.

Although large-scale computer simulations are often extremely useful, they are usually specific to a given problem; here we seek analytical results from which we can develop a generic and intuitive understanding. To model the full dynamics of a stochastic system, we in general need an infinite set of ODEs (the master equations) which describe the probability distribution for every possible populations size. However, the dynamics of many ecological systems is often low-dimensional (Ellner & Turchin, 1995) and therefore the master equations contain vast amounts of redundant information. An alternative is to change basis, and consider the dynamics on the space of moments rather than on the space of probability distributions. For many biological systems, all the important information is contained within the first few moments, with higher-order moments either negligible or tending to fixed measurable quantities (Isham, 1995; Bolker & Pacala, 1997).

The usual moment closure assumption is to ignore third-order cumulants (by setting them to zero); this should be expected to perform better than standard deterministic models which ignore second order cumulants and concentrate exclus-

ively on the mean. However, this assumption can frequently lead to unrealistic (negative or complex) population sizes or unbounded growth. Here we introduce the concept of multiplicative moments (cf. Keeling *et al.*, 1997), where the natural third-order approximation leads to more biologically acceptable behaviour.

Moment closure techniques have generally been used to consider either the average behaviour of a single stochastic population or infinitely many coupled populations. This paper examines the behaviour of a *finite* sized metapopulation, so that equations need to be developed for the moments at both the subpopulation and metapopulation levels. Looking exclusively at the equilibrium dynamics, and comparing theoretical results with stochastic simulations, we formulate measures for the local and global persistence of a population.

A Simple Stochastic Population

The logistic model is one of the most simple ecological models and one of the most widely used,

$$\frac{dx}{dt} = bx - dx^2. \quad (1)$$

It has a fixed point at $x = 0$, which is unstable, and a non-trivial stable fixed point at $x = b/d$ (cf. carrying capacity).

When moving from a deterministic to a stochastic framework, even the simple logistic model has numerous interpretations. Here we insist on a constant birth rate per individual ($b + a$), but a density-dependent death rate ($dx + a$). Naturally in the deterministic formulation the a terms in the birth and death rates cancel, so the deterministic equations are unaffected by the value of a . However, in the stochastic model, although the expected rate of change is independent of a , having a large a means that on average more events occur per unit time. Therefore, this extra a component increases the amount of stochasticity experienced.

Assume that many realizations of the stochastic model are performed; although we cannot predict the outcome of any one realization, the ensemble average over infinitely many realizations should behave deterministically. The

master equation for the distribution of population sizes is

$$\begin{aligned} \frac{dD_x}{dt} &= (b+a)(x-1)D_{x-1} + (dx+d+a)(x+1)D_{x+1} \\ &- (b+a+dx+a)x D_x \quad \forall x \geq 0, \\ D_x &= 0 \quad \forall x < 0. \end{aligned}$$

Let $\langle \cdot \rangle$ represent the ensemble average of any given quantity, e.g.

$$\langle f(x) \rangle = \sum_x D_x f(x).$$

We can now find differential equations for the behaviour of the mean population level, X ,

$$X = \langle x \rangle, \tag{2}$$

$$\begin{aligned} \frac{dX}{dt} &= \left\langle \sum_{\text{events}} \text{rate of event} \right. \\ &\quad \left. \times \text{change to } x \text{ due to event} \right\rangle \\ &= \langle (a+b)x \times (+1) + (ax+dx^2) \times (-1) \rangle \\ &= b\langle x \rangle - d\langle x^2 \rangle \\ &= bX - dX^2 - dV. \end{aligned} \tag{3}$$

Notice that the mean of x^2 ($\langle x^2 \rangle$) has been decomposed into the square of the mean ($\langle x \rangle^2$) plus the variance (V). Occasionally, our understanding of the biology means that we can approximate the variance in terms of known quantities or from available data. However, in general, we must treat the variance as another unknown parameter, and formulate differential equations for its evolution:

$$V = \langle (x - X)^2 \rangle = \langle x^2 \rangle - X^2, \tag{4}$$

$$\begin{aligned} \frac{dV}{dt} &= \langle (a+b)x(2x+1) \\ &\quad + (ax+dx^2)(-2x+1) \rangle - 2X \frac{dX}{dt} \end{aligned}$$

$$\begin{aligned} &= 2(a+b)\langle x^2 \rangle + (a+b)\langle x \rangle - 2d\langle x^3 \rangle \\ &\quad - 2a\langle x^2 \rangle + d\langle x^2 \rangle + a\langle x \rangle \\ &\quad - 2X(bX - dX^2 - dV) \\ &= b(2V + X) \\ &\quad - d(4XV + 2T - X^2 - V) \\ &\quad + 2aX, \end{aligned} \tag{5}$$

where T is the third-order cumulant of x . Here the standard moment-closure approximation is to ignore T , which is equivalent to assuming a Gaussian distribution (Whittle, 1957). However, as discussed below, setting third-order moments to zero is fraught with difficulties and can lead to negative average densities. Instead, by writing products in terms of multiplicative rather than additive moments (cf. Keeling *et al.*, 1997), this difficulty can be removed. As discussed below this is equivalent to assuming a log-normal distribution.

Writing multiplicative moments with hats, for the mean of x^2 we find,

$$\langle x^2 \rangle = X^2 + V = X^2 \hat{V} \Rightarrow \hat{V} = 1 + \frac{V}{X^2},$$

then we shall take

$$\langle x^3 \rangle = X^3 + 3XV + T = X^3 \hat{V}^3 \hat{T}.$$

The simplest assumption about the multiplicative third-order cumulant (\hat{T}) is that it is unity; this leads to

$$X^3 + 3XV + T = X^3 \left(1 + \frac{V}{X^2} \right)^3,$$

$$T = 3 \frac{V^2}{X} + \frac{V^3}{X^3}. \tag{6}$$

The standard assumption, that the higher-order additive cumulants are all zero, leads to a Gaussian distribution—although this is not precise if we wish the distribution to only be defined on the positive integers. It will now be shown that the

log-normal distribution has the same moments as our multiplicative assumption, and therefore the two distributions are equivalent. Assuming that all higher-order multiplicative cumulants are one,

$$\Rightarrow \langle x^n \rangle = X^n \hat{V}^{n(n-1)/2}. \quad (7)$$

Calculating moments for the log-normal distribution we have

$$\begin{aligned} \langle x^n \rangle &= \int_0^\infty x^n \sqrt{\frac{\lambda}{\pi}} \frac{e^{-\lambda(\ln(x/x))^2}}{x} dx \\ &= \sqrt{\frac{\lambda}{\pi}} \int_{-\infty}^\infty \frac{e^{ny} e^{-\lambda y^2}}{\alpha^n} dy \\ &= \left(\frac{1}{\alpha}\right)^n (e^{1/4\lambda})^{n^2}. \end{aligned} \quad (8)$$

Therefore, matching terms in eqns (7) and (8), we find that our multiplicative assumption is equivalent to a log-normal of the form,

$$D(x) = \frac{1}{x} \exp\left(\frac{1}{2\ln(\hat{V})} \left[\ln\left(\frac{\sqrt{\hat{V}}x}{X}\right)\right]^2\right)$$

which has mean X and variance $X^2(\hat{V} - 1)$ as required.

Although we could substitute eqn (6) into the variance eqn (5), it will be mathematically simpler to reformulate all equations in terms of multiplicative moments and ignore \hat{T} (by setting it equal to one). Hence, after some algebra, our new equations for the mean and multiplicative variance are

$$\frac{dX}{dt} = bX - dX^2 \hat{V}, \quad (9)$$

$$\begin{aligned} X^2 \frac{d\hat{V}}{dt} &= bX + 2aX - 2dX^3 \hat{V}^2 (\hat{V} - 1) \\ &\quad + dX^2 \hat{V}. \end{aligned} \quad (10)$$

This has a stable fixed point for all values of a , b and d ,

$$X^* = \frac{b^3}{d(ad + bd + b^2)}, \quad \hat{V}^* = \frac{ad + bd + b^2}{b^2}$$

$$\Rightarrow V^* = \frac{b^4(a + b)}{d(ad + bd + b^2)^2}. \quad (11)$$

Had the additive moments been used, together with the closure assumption that $T = 0$, then there is an unnatural bifurcation and the mean is zero for a range of parameters (Appendix A). Obviously, when b/d is small the system should frequently fail to persist, but the mean should still be real and positive. Therefore, even at this simple level, there are strong advantages to using multiplicative moments.

For simplicity, we shall set $d = 1$ by rescaling time. From eqn (11) it is found that if $a = 0$, the population is always under-dispersed, that is the variance is less than the mean. However, observations of real populations would indicate that over-dispersal is the rule, which implies that $a > b/(1 + b)$.

It is clear that for a single isolated population the rate of extinction is expected to be an increasing function of a and a decreasing function of b ; persistence is maximized by large populations and little stochasticity. However, in order to consider the role of spatial heterogeneity, we now want to use our moment closure technique to consider finitely many coupled populations.

FINITE METAPOPULATION AND PERSISTENCE

Suppose that there exists n distinct subpopulations, with $x_i \in \mathbb{N}$ individuals in sub-population i . Although we wish the subpopulations to obey the same logistic model as before, we will insist that the birth and death rates now have two components; one based on the dynamics within a subpopulation and one influenced by the coupling to the average over all the subpopulations, \bar{x} . Each subpopulation will therefore be taken to have the following underlying deterministic equations:

$$\begin{aligned} \frac{dx_i}{dt} &= b([1 - \sigma_b]x_i + \sigma_b \bar{x}) \\ &\quad - dx_i([1 - \sigma_d]x_i + \sigma_d \bar{x}) \end{aligned} \quad (12)$$

where $\bar{x} = \frac{1}{n} \sum_i x_i$.

More precisely this corresponds to a situation where a proportion σ_b of the births are into

a randomly chosen population, and that the density dependent death rate is influenced to some degree (σ_d) by the global density. This deterministic model predicts that all subpopulations asymptote to a single fixed point, $x_i^* = b/d \forall i$. Therefore, in the deterministic model any heterogeneity is soon lost.

The master equation for the metapopulation now has to consider the distribution across all n sub-populations—therefore finding its solution, either analytically or numerically is increasingly unlikely. The low dimensionality of moment closure models is an appealing method for understanding this type of problem. As before we shall wish to consider the mean and multiplicative variance at the subpopulation level for the stochastic system with birth and death rates given by eqn (12), and with a set to zero for simplicity. We notice that for any metapopulation system, the average for one subpopulation over many realizations, is the same as the average over all sub-populations and many realizations,

$$X = \langle x_i \rangle = \langle \bar{x} \rangle$$

this dramatically simplifies the equations. The evolution equation for the mean density is now

$$\begin{aligned} \frac{dX}{dt} &= \langle b[1 - \sigma_b]x_i + b\sigma_b\bar{x} \\ &\quad - d[1 - \sigma_d]x_i^2 - d\sigma_dx_i\bar{x} \rangle \\ &= bX - dX^2\hat{V} + d\sigma_dX^2(\hat{V} - \hat{C}) \end{aligned} \quad (13)$$

where \hat{V} is the multiplicative variance within a subpopulation (as before) and we introduce \hat{C} , the multiplicative covariance between a subpopulation (x_i) and the mean metapopulation level (\bar{x}). However, \hat{C} can also be expressed as the multiplicative variance in the mean metapopulation level, and this formulation is an easier quantity to work with

$$X^2\hat{C} = \langle x_i\bar{x} \rangle = \langle \bar{x}^2 \rangle.$$

The calculation of other ensemble averages is performed by considering all possible pairwise combinations of the elements and multiplying by

the appropriate moment. The following two averages will be required below:

$$\langle x_i^2 \bar{x} \rangle = X^3 \hat{C}^2 \hat{V},$$

$$\langle x_i \bar{x}^2 \rangle = X^3 \hat{C} \hat{V}^2.$$

Using the birth and death rates, it is now possible to construct an equation for the variance at the subpopulation level,

$$\hat{V} = \frac{\langle x_i^2 \rangle}{\langle x_i \rangle^2}, \quad (14)$$

$$\begin{aligned} X^2 \frac{d\hat{V}}{dt} &= \langle (b[1 - \sigma_b]x_i + b\sigma_b\bar{x})(2x_i + 1) \\ &\quad + (d[1 - \sigma_d]x_i^2 + d\sigma_dx_i\bar{x})(-2x_i + 1) \rangle \\ &\quad - 2X\hat{V} \frac{dX}{dt}, \\ &= bX - 2b\sigma_bX^2(\hat{V} - \hat{C}) - d(1 - \sigma_d)X^2\hat{V} \\ &\quad \times [2X\hat{V}(\hat{V} - 1) - 1] - d\sigma_dX^2\hat{C} \\ &\quad \times (2X\hat{V}(\hat{C} - 1) - 1]. \end{aligned} \quad (15)$$

Similarly, for the covariance term,

$$\hat{C} = \frac{\langle \bar{x}^2 \rangle}{\langle \bar{x} \rangle^2}, \quad (16)$$

$$\begin{aligned} X^2 \frac{d\hat{C}}{dt} &= \left\langle n(b[1 - \sigma_b]x_i + b\sigma_b\bar{x}) \left(\frac{2\bar{x}}{n} + \frac{1}{n^2} \right) \right. \\ &\quad \left. - n(d[1 - \sigma_d]x_i^2 + d\sigma_dx_i\bar{x}) \left(\frac{2\bar{x}}{n} - \frac{1}{n^2} \right) \right\rangle \\ &\quad - 2X\hat{C} \frac{dX}{dt} \\ &= -2dX^3\hat{C}(\hat{C} - 1)([1 - \sigma_d]\hat{V} + \sigma_d\hat{C}) \\ &\quad + \frac{bX}{n} + \frac{dX^2}{n} ([1 - \sigma_d]\hat{V} + \sigma_d\hat{C}). \end{aligned} \quad (17)$$

Here, terms in n arise because a single event causes the average over all sites \bar{x} to change by $1/n$.

It should be noticed from eqn (17) that $\hat{C} \rightarrow 1$ in the limits as $n \rightarrow \infty$. This is to be expected, as when the number of subpopulations becomes large, the behaviour of the mean will become independent of the value of any individual population and the additive variance at the metapopulation level will tend to zero. At the opposite extreme, when $n = 1$ as there is only one subpopulation the local variance is the same as the global variance and $C \equiv V$.

Figure 1(a) gives the mean density of individuals in a stochastic metapopulation ($d = 1$, $a = 0$, $n = 10$); the model was run for 200 time steps and results averaged over the last 100. For low birth rate, b , and low coupling σ , extinctions dominate the system; whereas elsewhere the mean density increases strongly with the birth rate b , but there is also a slight increase with the amount of coupling $\sigma (= \sigma_b = \sigma_d)$. Figure 1(b) shows the extinction rate for this metapopulation; this was calculated by measuring the average time to extinction starting all populations at the deterministic fixed point. This result agrees with our understanding of graph (a), showing that persistence increases with both the coupling σ and the birth rate b .

Figure 1(c) gives the mean density of individuals as predicted by the moment closure eqns (13), (15) and (17). We find very good agreement between the stochastic metapopulation and the moment equations, except that the moment equations persist for all parameters—this deterministic formulation for the mean is unable to account for stochastic extinctions. However, we can capture the likelihood of extinctions by considering the behaviour of the variance and covariance.

Given a fixed shape for the distribution of population sizes, there is expected to be a greater risk of extinction when the standard error (the standard deviation relative to the mean) is large. In particular, it is supposed that the probability of extinction is an increasing function of the variance divided by the mean squared. Therefore, define the local extinction measure L_E to be

$$L_E = \frac{\langle x_i^2 \rangle}{\langle x_i \rangle^2} = \frac{V}{X^2} = \hat{V} - 1;$$

similarly the probability that all the sites are extinct can be quantified by the global extinction measure G_E ,

$$G_E = \frac{\langle \bar{x}^2 \rangle}{\langle \bar{x} \rangle^2} = \frac{C}{X^2} = \hat{C} - 1.$$

Throughout this work, it is assumed that the theoretically predicted extinction measures (L_E and G_E) are monotonically increasing functions of the true extinction rates, and hence allows us to assess the qualitative effects of parameter changes on the likelihood of extinction. To confirm this assumption, we shall return our consideration to the dynamics of a single patch. Setting $a = 0$ and $d = 1$, consider the equilibrium distribution of population size, D_n , conditional on the population having not gone extinct,

$$D_n = \frac{D_1 b^{n-1}}{n!n} \quad \text{where } D_1 \text{ is such that } \sum_{n=1}^{\infty} D_n = 1.$$

This formulation comes from solving the detailed balance equations, such that the transition from D_{n+1} to D_n are balanced by the transitions from D_n to D_{n+1} . For this conditional distribution the exact extinction rate can be seen to be

$$\begin{aligned} E_{exact} &= \text{death rate} \times \mathbb{P}(\text{pop size} = 1) \\ &= d \times D_1 = D_1. \end{aligned}$$

We are now in a position to compare the extinction measure from the moment equations, L_E , with the exact extinction rate E_{exact} [Fig. 1(d)]. It is clear that the extinction measure from the moment equations is a monotonic function of the true extinction rate as claimed. In fact,

$$E_{exact} \approx \frac{1}{L_E^2} \exp\left(-\frac{1}{L_E}\right)$$

is a good approximation whenever the exact extinction rate is less than 0.5.

Figure 1(e) and (f) shows the local and global extinction measures, L_E and G_E , respectively; these results compare well with the stochastic system [Fig. 1(a) and (b)]. It is only when the

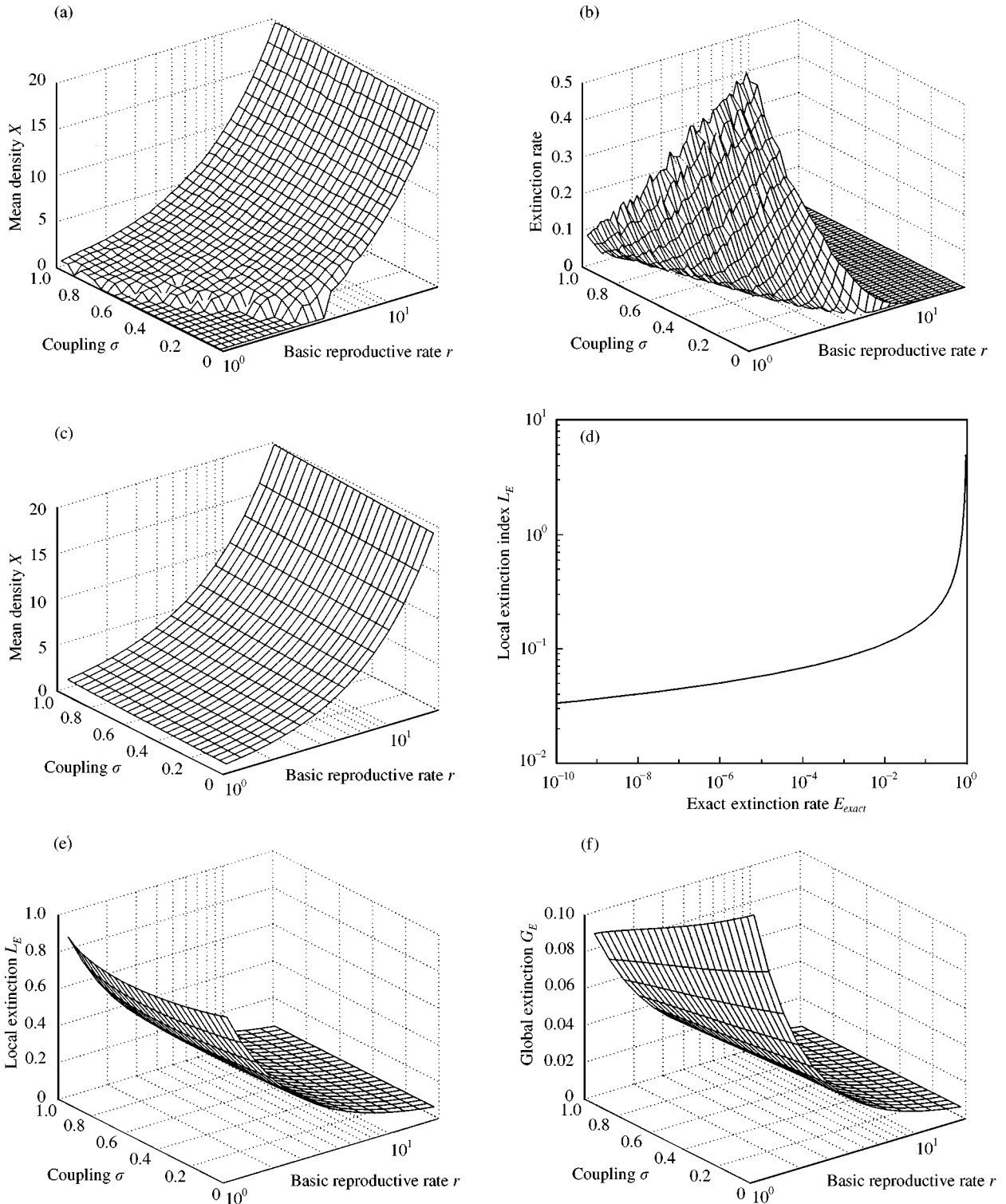


FIG. 1. Results from the single-species metapopulation model ($d = 1, n = 10$). Graphs (a) and (b) are from the stochastic (Monte Carlo) model. Graph (a) gives the mean density of individuals in a stochastic simulation, averaged over 100 time steps. Graph (b) shows the average rate of global extinctions; starting at the deterministic fixed point, the average time to global extinction is measured—this is then used to determine the rate. Graphs (c), (e) and (f) show the mean density of individuals and the measures of local and global extinction, respectively, from the moment closure eqns (13), (15) and (17). Finally, graph (d) considers the single population model and shows theoretical predictions of the extinction measure (L_E) against the exact extinction rate (E_{exact}) for a range of birth rates.

system's behaviour is dominated by extinctions that the approximations fail. Therefore, it appears that for this simple one-species model, the moment equations can capture the main qualitative features.

Consider what this has told us about the ecology of such populations. As expected the rate of global extinctions is an increasing function of a and d , but a decreasing function of b . Hence, when the deterministic carrying capacity is high and the level of stochasticity is low we expect the population to persist. However, the behaviour with respect to the coupling, σ , goes against current ecological perception (Hassell & May, 1974; Grenfell & Harwood, 1997) although it conforms to theories of conservation (Hanski, 1982; Fahrig & Merriam, 1985; Burkey, 1989). For this single-species model, both the analytic moment equations and the stochastic metapopulation predict the lowest rates of extinction when the coupling is largest. Therefore, this system persists better in one large patch—spatial heterogeneity actually increases the risk of extinction.

The surprising behaviour with changes in coupling can be explained by the nonlinearities present in the system. First note that the average birth rate is independent of the distribution of individuals between the sub-populations, as it is linear. However, due to its density-dependent nature, the death rate is minimized when the distribution is such that each individual experiences the same mean value. Therefore, any form of spatial heterogeneity merely increases the average death rate without providing any benefits.

From these results I believed that the vast majority of such single component models will persist better as the amount of coupling increases. For a given birth rate $B(x)$ per individual and death rate $D(x)$ per individual, the following would appear to be sufficient (although not necessary) conditions for persistence to be maximized when the coupling is large,

(1) Wider distributions of population sizes have lower average reproductive rates. This means that there are strong density dependent effects so that large populations have large negative growth rate.

(2) The number of events and therefore the amount of stochasticity also increases for wider distributions.

Both these conditions can be made mathematically precise,

$$\frac{d^2}{dx^2} (B - D) < 0 \quad \text{and} \quad \frac{d^2}{dx^2} (B + D) \geq 0 \quad \forall x.$$

Therefore, there is a very wide class of single-species models where large amounts of coupling is good for global persistence.

An Epidemic Model

So far only a single species, which has very simple dynamics, has been considered. The moment closure technique will now be extended to examine multi-species systems, which have slightly more complex behaviour. Spatial heterogeneity that arises from the limited coupling between communities is believed to be important in the persistence of diseases and other natural enemy systems (Hassell & Pacala, 1990; Hassell *et al.*, 1991; Bolker & Grenfell, 1995; Lloyd & May, 1996; Grenfell & Harwood, 1997). Attention is now focused on a simplified two-species model of disease dynamics, to see if the moment closure method can again capture the extinction behaviour.

Suppose the standard SIR (Susceptible–Infectious–Recovered) equations (Anderson & May, 1992) for the state of subpopulation i can be simplified to

$$\begin{aligned} \dot{S}_i &= B - \beta S_i ([1 - \sigma] I_i + \sigma I), \\ \dot{I}_i &= \beta S_i ([1 - \sigma] I_i + \sigma \bar{I}) - g I_i + \varepsilon. \end{aligned} \quad (18)$$

In this formulation, death only occurs in the recovered class, and therefore eqn (18) provides a reasonable model of many childhood diseases. This slightly unusual form was taken so as to minimize the number of events that had to be considered—thus simplifying the algebra. The coupling σ is only present in the infection terms, and as such represents new cases in a given subpopulation caused by infectious individuals in others. The ε term, which mimics imports from outside the metapopulation, is included to prevent permanent extinctions.

As before, differential equations can be developed for the ensemble means (S and I) of the

stochastic system:

$$\dot{S} = B - \beta SI(\xi - \sigma[\xi - \psi]), \quad (19)$$

$$\dot{I} = \beta SI(\xi - \sigma[\xi - \psi]) - gI + \varepsilon, \quad (20)$$

where ξ is the multiplicative covariance between infectious and susceptible individuals within a subpopulation,

$$SI\xi = \langle S_i I_i \rangle$$

and ψ is the multiplicative covariance between the mean density of infectious and susceptible individuals in the entire metapopulation. As shown in the previous section, covariances such as ψ can be written in several forms,

$$SI\psi = \langle \bar{S}\bar{I} \rangle = \langle S_i \bar{I} \rangle = \langle \bar{S} I_i \rangle.$$

Consideration of eqns (19) and (20) shows that there exists a unique non-trivial fixed points for the average number of infectious individuals, namely $I^* = (B + \varepsilon)/g$; thus the average number of cases is unaffected by either the stochasticity or the spatial heterogeneity of the system. This is of great benefit when approximating the persistence from the local and global extinction measures, as changes in the coupling σ will only affect the variance and not the mean.

Differential equations can now be found for the two covariances, ξ and ψ , as shown in Appendix B. This brings in another four new parameters, \hat{V}_S and \hat{V}_I , the variance in susceptibles and infectious within a sub-population, and \hat{C}_S and \hat{C}_I , the covariance in susceptibles and infectious between a sub-population and the metapopulation average—as well as various third order cumulants.

It is for this two-species system that the use of multiplicative, rather than additive, moments is of the greatest benefit. Figure 2(a) gives the additive covariance between susceptible and infectious individuals within a sub-population ($\langle S_i I_i \rangle - \langle S_i \rangle \langle I_i \rangle = SI(\xi - 1)$). It is due to these large negative values that the multiplicative form of the correlations is necessary. If the addi-

tive moments had been used (and third-order cumulants ignored) then when $\langle SI \rangle = 0$, so that the covariance is large and negative, then the approximation of $\langle S^2 I \rangle$ or $\langle SI^2 \rangle$ is rarely zero and frequently negative, whereas the multiplicative approximation to these terms is always zero. Hence, the multiplicative moments appear to have a clear advantage whenever large negative correlations arise. It is found that differential equations that use additive moments frequently “blow up” due to this behaviour.

Figure 2(b) and (c) shows the local and global extinction measures, respectively, for the infection,

$$L_E = \hat{V}_I - 1, \quad G_E = \hat{C}_I - 1.$$

These measures compare well to the results of stochastic metapopulation simulations—Fig. 2(d) and (e). The rate of local extinction is a monotonically decreasing function of the coupling—this is as expected, as when the coupling is large an extinction will only occur if infection is rare in all subpopulations. However, the global extinction measure, G_E , and the rate of extinction from the stochastic simulations both demonstrate a minimum for intermediate levels of coupling.

Although there is clearly good qualitative agreement between the general trends predicted by the moment equations and the simulations, we consistently find that the moment equations underestimate the extinction rate when the coupling is low. This discrepancy is primarily due to the intense negative correlations that develop and the long periods for each subpopulation when the disease is extinct. These two conditions mean that the disease would be better characterized by long extinctions with erratic epidemics, breaking our basic assumption that the epidemic can be described as a continuous process.

Finally, Fig. 2(f) shows the probability distribution for the number of extinct subpopulations (solid line) and this is compared to the standard metapopulation assumption that extinctions happen at random (dashed line). Clearly, large range extinctions are much more common than a simple prediction from the local dynamics would suggest—this is because the coupling between subpopulations leads to the formation of positive correlations.

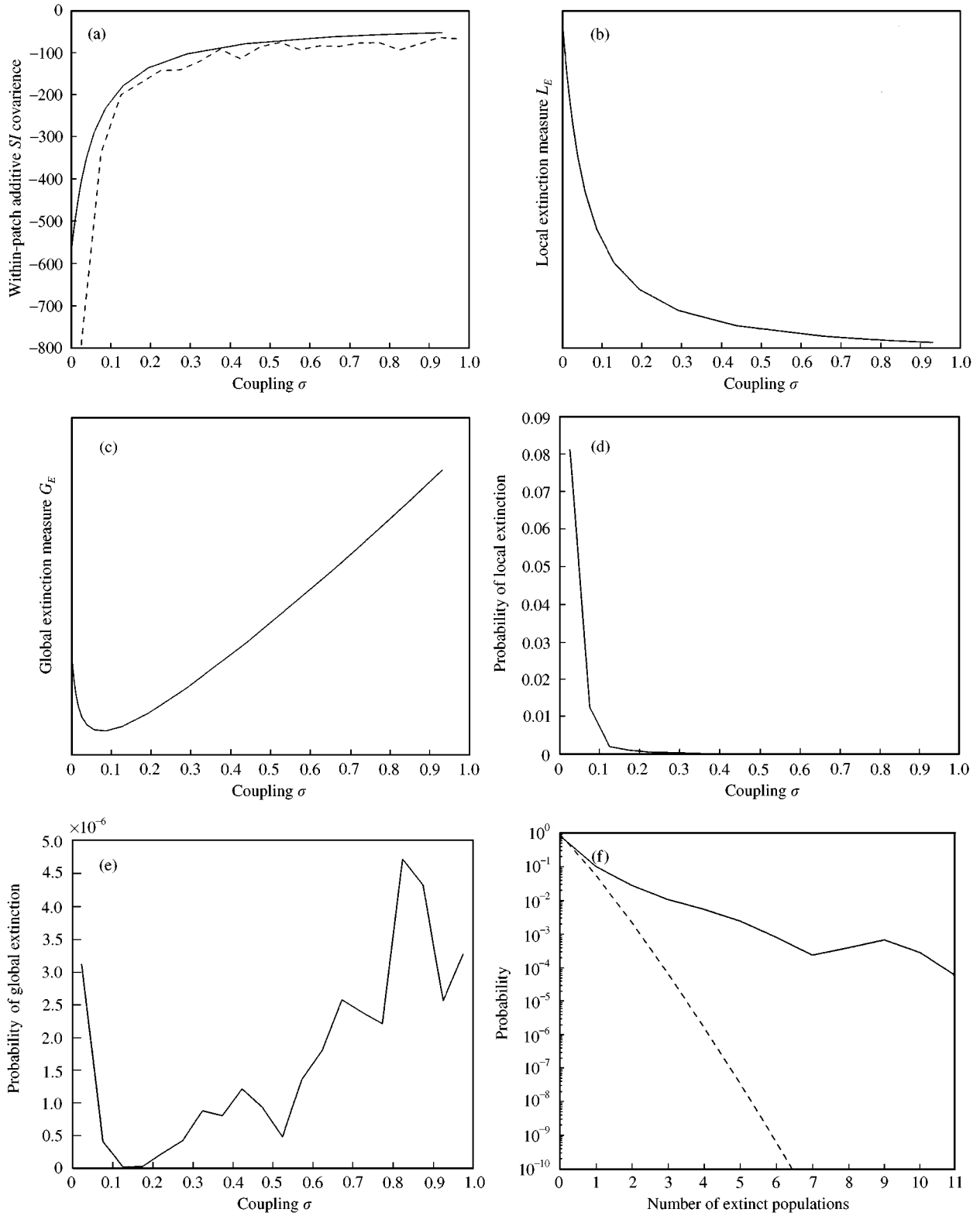


FIG. 2. (Caption opposite)

Discussion

This paper has introduced two novel concepts which makes the estimation of extinction rates for a variety of systems possible.

The use of multiplicative, rather than additive, moment closure assumptions has greatly improved the robustness of the equations. Despite the long history of using additive moments, it is often far easier to think about multiplicative cumulants as these are dimensionless and possess a more natural scaling behaviour. For continuous populations, the standard additive assumption leads to a Gaussian distribution, whereas the multiplicative assumption introduced here leads to a log-normal distribution. Therefore, the multiplicative model has a far fatter tail than any exponential or Gaussian distribution, but has the strong advantage of only being defined for positive values. Which assumption is the most valid will depend greatly on the problem being examined; although it is believed that, for a wide class of multi-species models, the multiplicative assumptions are more robust as unnatural negative densities are prevented.

The second advancement introduced is to consider a *finite* metapopulation. Much of the work using moment closure models has either considered a single patch in isolation (Renshaw, 1991; Isham, 1995) or an infinite system (Bolker & Pacala, 1997; Keeling *et al.*, 1999). With an infinite number of subpopulations the global mean is expected to remain constant, whereas for finitely many subpopulations the global mean will experience some stochasticity (although less than each individual subpopulation). It is the variability at this level that leads to stochastic global extinctions. Therefore, despite the added complexity, the benefits of being able to consider global extinctions and the effects coupling are compelling.

In many single-species models, strong density-dependent effects will mean that global persist-

ence is maximized by large amounts of coupling. In these models, heterogeneities reduce the overall reproductive rate of the species so there is no advantage to be gained by limiting the coupling. In such cases, the species persists best in one large well-mixed population. For the SIR disease model (and presumably for predatory-prey systems as well), spatial heterogeneity has considerable benefits for persistence allowing the susceptible (prey) population the chance to recover. Therefore, for these two-species models there is a tradeoff between spatial heterogeneity and rescue effects, which in turn means that persistence is maximized for intermediate levels of coupling.

Throughout this work only global coupling has been considered. In almost all real ecological systems the influence of an organism is expected to decrease with distance, so that many effects are localized. Although models with local interactions are far more complex and can display far richer dynamics, the same qualitative results can be expected to hold as persistence will frequently be a tradeoff between spatial heterogeneity and rescue events. Many spatial models consider the dynamics of individuals on a homogeneous plane (Keeling, 1999a), however, for humans and many other social animals a metapopulation is a good model as distinct subpopulations do exist.

Although this type of moment closure model has been shown to provide a good description of both single-species logistic growth and two-species natural enemy systems much more work needs to be done to assess just how generic this technique can be. For particular systems, a different functional form for L_E and G_E may be more appropriate. However, I believe that for systems where extinctions are rare and rescue effects common, this technique will be applicable to a wide class of ecological and epidemiological problems.

FIG. 2. Results from the infectious disease metapopulation model ($B = 2$, $\beta = 3 \times 10^{-5}$, $g = 0.077$, $\varepsilon = 4 \times 10^{-3}$, $n = 20$); these parameters correspond closely to those of measles, with a metapopulation size close to the critical community size. Graph (a) gives the additive covariance between infectious and susceptible individuals for both the moment closure equations (—) and the stochastic simulation (---). Graphs (b) and (c) show the measures of local and global extinction, respectively, from the moment closure equations. Graphs (d), (e) and (f) are from the full stochastic simulations. Graph (d) is the probability of a sub-population being extinct, while graph e is the probability that the whole metapopulation is extinct. Graph (f) shows the probability distribution for the number of extinct patches (—) and compares it to the standard metapopulation assumption of random independent extinctions (---) ($\sigma = 0.1$).

I would like to thank Bryan Grenfell, Pej Rohani and the two anonymous referees for their helpful comments on this manuscript, special thanks also go to Stephen Cornell for recognizing the log-normal distribution. This research was supported by the Royal Society.

REFERENCES

- ANDERSON, R. M. & MAY, R. M. (1992). *Infectious Diseases of Humans*. Oxford: Oxford University Press.
- BARTLETT, M. S. (1956). Deterministic and Stochastic Models for recurrent epidemics. *Proc. 3rd Berkley Symp. on Math. Stats. and Prob.*, Vol. 4, 81–108.
- BOLKER, B. M., DEUTSCHMAN, D. H., HARTVIGSEN, G. & SMITH, D. L. (1997). Individual-based modelling—What is the difference? *Trends Ecol. Evol.* **12**, 111.
- BOLKER, B. M. & GRENFELL, B. T. (1995). Space, persistence and dynamics of measles epidemics. *Phil. Trans. Roy. Soc. Lond. B* **348**, 309–320.
- BOLKER, B. & PACALA, S. W. (1997). Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theor. Pop. Biol.* **52**, 179–197.
- TER BRAAK, C. J. F., HANSKI, I. & VERBOOM, J. (1998). The incidence function approach to modeling of metapopulation dynamics. In: *Modeling Spatiotemporal Dynamics in Ecology*. Bascombe, J. & Sole, R. V., eds pp.167–188. Berlin: Springer.
- BURKEY, T. V. (1989). Extinction in nature reserves: the effect of fragmentation and the importance of migration between reserve fragments. *Oikos* **55**, 75–81.
- DURRETT, R. & LEVIN, S. A. (1994). Stochastic spatial models: a user's guide to ecological applications. *Phil. Trans. Roy. Soc. Lond. B* **343**, 329–350.
- ELLNER, S. & TURCHIN, P. (1995). Chaos in a noisy world new methods and evidence from time series analysis. *Am. Nat.* **145**, 343–375.
- FAHRIG, L. & MERRIAM, G. (1985). Habitat patch connectivity and population survival. *Ecology* **66**, 1762–1768.
- GILPIN, M. E. & HANSKI, I. A. (eds) (1991). *Metapopulation Dynamics: Empirical and Theoretical Investigations*. New York: Academic Press.
- GRENFELL, B. T. & BOLKER, B. M. (1998). Cities and villages: infection hierarchies in a measles metapopulation. *Ecol. Lett.* **1**, 63–70.
- GRENFELL, B. T., BOLKER, B. M. & KLECZKOWSKI, A. (1995). Seasonality and extinction in chaotic metapopulations. *Proc. Roy. Soc. Lond. B* **259**, 97–103.
- GRENFELL, B. & HARWOOD, J. (1997). (Meta)population dynamics of infectious diseases. *TREE* **12**, 395–399.
- HANSKI, I. A. (1982). Dynamics of regional distribution: the core and satellite hypothesis. *Oikos* **38**, 210–221.
- HANSKI, I. A. & GILPIN, M. E. (eds) (1997). *Metapopulation Biology. Ecology, Genetics and Evolution*. New York: Academic Press.
- HASSELL, M. P., COMINS, H. & MAY, R. M. (1991). Spatial structure and chaos in insect population dynamics. *Nature* **353**, 255–258.
- HASSELL, M. P. & MAY, R. M. (1974). Aggregation in predators and insect parasites and its effects of stability. *J. Anim. Ecol.* **43**, 567–594.
- HASSELL, M. P. & PACALA, S. W. (1990). Heterogeneity and the dynamics of host–parasitoid interactions. *Phil. Trans. Roy. Soc. Lond. B* **330**, 203–220.
- ISHAM, V. (1995). Stochastic models of host-macroparasite interaction. *Ann. Appl. Probab.* **5**, 720–740.
- KEELING, M. J. (1997). Modelling the persistence of measles. *Trends MicroBiol.* **5**, 513–518.
- KEELING, M. J. (1999). Spatial models of interacting populations. In: *Advanced Ecological Theory* (McGlade, J., ed.) London: Blackwell Scientific Publishing.
- KEELING, M. J., RAND, D. A. & MORRIS, A. J. (1997). Correlation models for childhood diseases. *Proc. Roy. Soc. B* **264**, 1149–1156.
- LANDE, R., ENGEN, S. & SAETHER, B. E. (1998). Extinction times in finite metapopulation models with stochastic local dynamics. *Oikos* **83**, 383–389.
- LEVINS, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Ent. Soc. Am.* **5**, 237–240.
- LLOYD, A. L. & MAY, R. M. (1996). Spatial heterogeneity in epidemic models. *J. theor. Biol.* **179**, 1–11.
- RENSHAW, E. (1991). *Modelling Biological Populations in Space and Time*. Cambridge: University Press.
- SWINTON, J., HARWOOD, J., GRENFELL, B. T. & GILLIGAN, C. A. (1998). Persistence thresholds for phocine distemper virus infection in harbour seal *Phoca vitulina* metapopulations. *J. Anim. Ecol.* **67**, 54–68.
- WHITTLE, P. (1957). On the use of the normal approximation in the treatment of stochastic processes. *J. Roy. Stat. Soc. B* **19**, 268–281.

APPENDIX A

Using the standard *additive* moments, together with the assumption that the third-order cumulant T is zero, leads to the following fixed point solution:

$$X^* = \frac{3b \pm \sqrt{b^2 - 8bd - 8ad}}{4d} \quad \text{or} \quad X^* = 0.$$

We therefore find that when the birth rate is small relative to the death rate ($b^2 < 8(b+a)d$), the only solution is $X^* = 0$. No such bifurcation are observed in the full stochastic model, and would seem to be biologically unreasonable.

APPENDIX B

$$\begin{aligned} SI\dot{\xi} &= \langle BI_i + \beta S_i([1 - \sigma]I_i + \sigma\bar{I}) (S_i - I_i - 1) \\ &\quad - gS_iI_i + \varepsilon S_i \rangle - SI\dot{\xi} - IS\dot{\xi} \\ &= -\beta I(\xi - 1) + \beta SI\xi[1 - \sigma](S\hat{V}_S - 1) \\ &\quad - I\xi(\hat{V}_I - 1) - 1 \\ &\quad + \beta SI\psi\sigma(S\psi\hat{V}_S - S\hat{\xi} - I\hat{\xi}(\hat{C}_I - 1) - 1) \\ &\quad - \varepsilon S(\xi - 1), \end{aligned} \tag{B.1}$$

$$\begin{aligned}
 SI\dot{\psi} &= \left\langle B\bar{I} + \beta S_i([1 - \sigma]I_i + \sigma\bar{I}) \left(\bar{S} - \bar{I} - \frac{1}{n} \right) \right. \\
 &\quad \left. - gS_i\bar{I} + \varepsilon S_i \right\rangle - SI\dot{\psi} - I\dot{S}\psi \\
 &= BI(1 - \psi) + \beta SI(\xi[1 - \sigma] + \psi\sigma) \left(S\psi(\hat{C}_S - 1) \right. \\
 &\quad \left. - I\psi(\hat{C}_I - 1) - \frac{1}{n} \right) \\
 &\quad - \varepsilon S(\psi - 1). \tag{B.2}
 \end{aligned}$$

This brings in another four new parameters—the definitions and evolution equations for these variances and covariances are

$$\begin{aligned}
 I^2\hat{V}_I &= \langle I_i^2 \rangle \\
 I^2\hat{V}_I &= \langle \beta S_i([1 - \sigma]I_i + \sigma\bar{I})(2I_i + 1) \\
 &\quad - gI_i(2I_i - 1) + \varepsilon(2I_i + 1) \rangle - 2I\dot{I}\hat{V}_I \\
 &= \beta SI\xi[1 - \sigma](2I\hat{V}_I(\xi - 1) + 1) + \beta SI\psi\sigma(2I\xi\hat{C}_I \\
 &\quad - 2I\hat{V}_I + 1) + g - \varepsilon(2I(\hat{V}_I - 1) - 1), \tag{B.3}
 \end{aligned}$$

$$\begin{aligned}
 S^2\hat{V}_S &= \langle S_i^2 \rangle \\
 S^2\hat{V}_S &= \langle B(2S_i + 1) - \beta S_i([1 - \sigma]I_i + \sigma\bar{I}) \\
 &\quad (2S_i - 1) \rangle - 2S\dot{S}\hat{V}_S \\
 &= -B(2S(\hat{V}_S - 1) - 1) - \beta SI\xi[1 - \sigma]
 \end{aligned}$$

$$\begin{aligned}
 &\times (2S\xi\hat{V}_I - 2S\hat{V}_S - 1) \\
 &\quad - \beta SI\psi\sigma(2S\psi\hat{V}_I - 2S\hat{V}_S - 1), \tag{B.4}
 \end{aligned}$$

$$\begin{aligned}
 I^2\hat{C}_I &= \langle \bar{I}^2 \rangle \\
 I^2\hat{C}_I &= \left\langle \beta S_i([1 - \sigma]I_i + \sigma\bar{I}) \left(2\bar{I} + \frac{1}{n} \right) \right. \\
 &\quad \left. - g\bar{I} \left(2\bar{I} - \frac{1}{n} \right) + \varepsilon \left(2\bar{I} + \frac{1}{n} \right) \right\rangle - 2I\dot{I}\hat{C}_I \\
 &= \beta SI(\xi[1 - \sigma] + \psi\sigma) \left(2I\hat{C}_I(\psi - 1) + \frac{1}{n} \right) \\
 &\quad + \frac{gI}{n} - \varepsilon \left(2I(\hat{C}_I - 1) - \frac{1}{n} \right), \tag{B.5}
 \end{aligned}$$

$$\begin{aligned}
 S^2\hat{C}_S &= \langle \bar{S}^2 \rangle \\
 S^2\hat{C}_S &= \left\langle B \left(2\bar{S} + \frac{1}{n} \right) - \beta S_i([1 - \sigma]I_i + \sigma\bar{I}) \right. \\
 &\quad \left. \left(2\bar{S} - \frac{1}{n} \right) \right\rangle - 2S\dot{S}\hat{C}_S \\
 &= -B \left(2S(\hat{C}_S - 1) - \frac{1}{n} \right) - \beta SI(\xi[1 - \sigma] \\
 &\quad + \psi\sigma) \left(2S\hat{C}_S(\psi - 1) - \frac{1}{n} \right). \tag{B.6}
 \end{aligned}$$