

M.A. Lewis · S. Pacala

## **Modeling and analysis of stochastic invasion processes\***

Received: 11 August 1997 / Revised version: 22 September 1998 /  
Published online: 4 October 2000 – © Springer-Verlag 2000

**Abstract.** In this paper we derive spatially explicit equations to describe a stochastic invasion process. Parents are assumed to produce a random number of offspring which then disperse according to a spatial redistribution kernel. Equations for population moments, such as expected density and covariance averaged over an ensemble of identical stochastic processes, take the form of deterministic integro-difference equations. These equations describe the spatial spread of population moments as the invasion progresses. We use the second order moments to analyse two basic properties of the invasion. The first property is ‘permanence of form’ in the correlation structure of the wave. Analysis of the asymptotic form of the invasion wave shows that either (i) the covariance in the leading edge of the wave of invasion asymptotically achieves a permanence of form with a characteristic structure described by an unchanging spatial correlation function, or (ii) the leading edge of the wave has no asymptotic permanence of form with the length scales of spatial correlations continually increasing over time. Which of these two outcomes pertains is governed by a single statistic,  $\phi$  which depends upon the shape of the dispersal kernel and the net reproductive number. The second property of the invasion is its patchy structure. Patchiness, defined in terms of spatial correlations on separate short (within patch) and long (between patch) spatial scales, is linked to the dispersal kernel. Analysis shows how a leptokurtic dispersal kernel gives rise to patchiness in spread of a population.

---

### **1. Introduction**

The classical mathematical model for a single-species biological invasion is a traveling wave of reproducing, dispersing organisms spreading into virgin territory [17, 12, 33, 26, 23, 1, 35]. This wave describes the expected density of organisms as a function of space and time. The analysis of such a model typically yields a traveling wave speed which describes the rate of spread and a corresponding

---

M.A. Lewis: Department of Mathematics, University of Utah, 150 South 1400 East, JWB 233, Salt Lake City, UT 84112, USA. e-mail: mlew@math.utah.edu

S. Pacala: Department of Ecology and Evolutionary Biology, Princeton University Princeton, NJ 08544-1003, USA

\*Research supported in part by a research fellowship from the Alfred P. Sloan Foundation and by the National Science Foundation under grant no. DMS-9457816.

**Key words:** Stochastic – Invasion – Covariance – Dispersal – Integro – Difference equation – Patchiness

monotone traveling wave profile which describes variation in the expected density of organisms.

Actual biological invasions are far more complex than these mathematical models indicate. Not only are there spatial and temporal variations in factors affecting spread, but the observed density of organisms rarely can be described by a simple expanding wave front (Figure 1a); rather, it is typical to observe a series of invaded patches which spread, coalesce and spawn new patches [27]. This can be seen clearly for species as diverse as cheat grass (*Bromus tectorum*), and house finch (*Carpodacus mexicanus*) (Figure 1b and 1c).

Even though the classical invasion models have often been successful in predicting spread rates, as measured by  $\sqrt{\text{area invaded}/\text{time}}$  [1], they give no indication of the possibly patchy structure of the wave. Such structure can play an important role in the ability to make field observations. For example, a single observer with limited observation range would miss the 'leading edge' of the invasion entirely unless one of the patches happened to be close by. Furthermore, not all attempts to predict spread rates have been successful, with failure typically attributed to difficult-to-measure very long-distance dispersal events [1].



**Fig. 1a.** (a) Invasion map for the Japanese beetle *Popilla japonica* in the eastern United States. Contours show the location of the invasion front at successive dates. (b) Invasion map for cheat grass (drooping brome) *Bromus tectorum* in North America. Black regions indicate area invaded. (c) Invasion map for the house finch (*Carpodacus mexicanus*) in eastern North America. Shaded regions indicate area invaded. (Based on [15].)

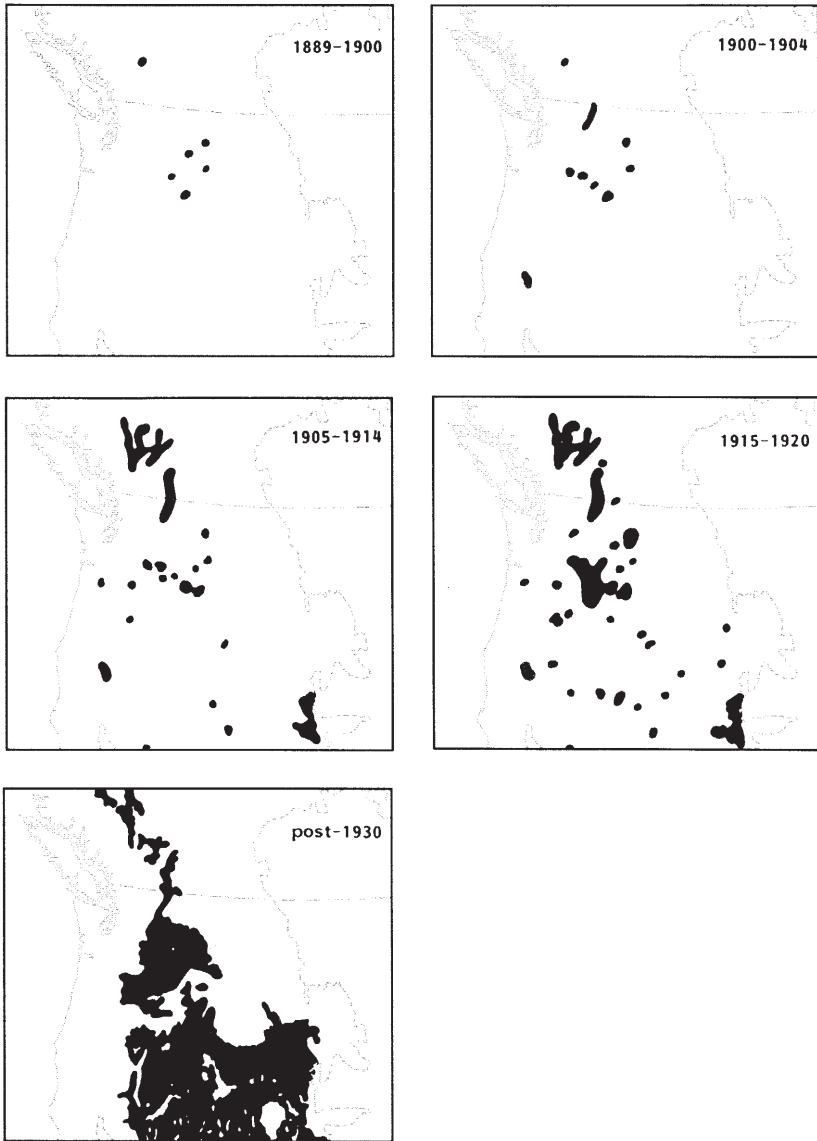


Fig. 1b.

Stochastic factors play a key role here. In a spatially homogeneous environment, stochastic movements of individuals affect the location of the patches. This can be seen by modeling the behavior of individuals by a branching process where individuals have given probabilities of reproducing and dying per unit time and redistribute spatially. Using Monte Carlo simulations of the branching process we show in Figure 2 that long distance dispersal over multiple space scales yields

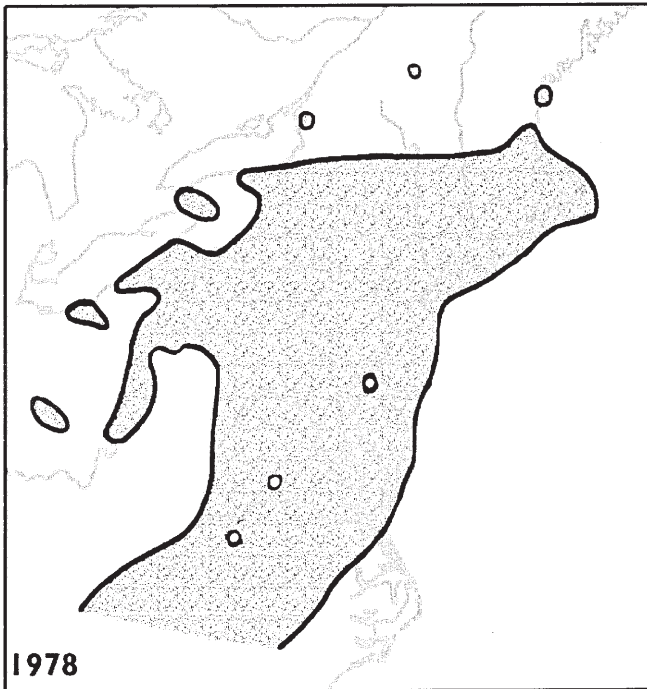
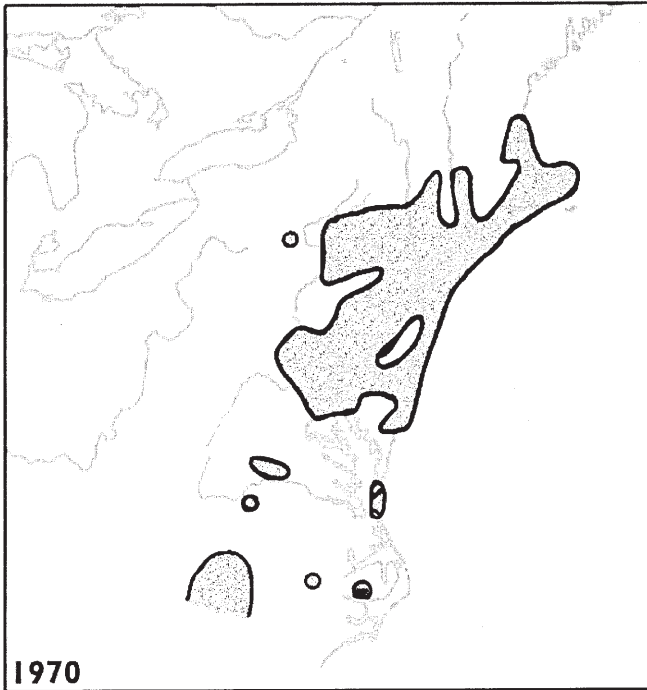
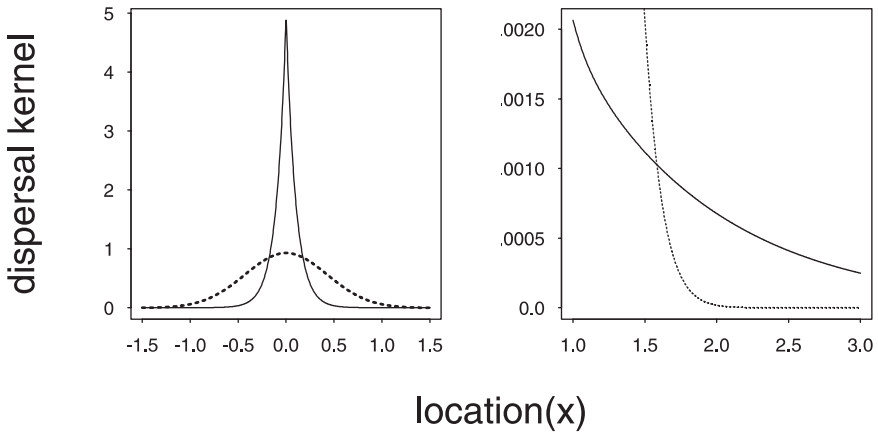


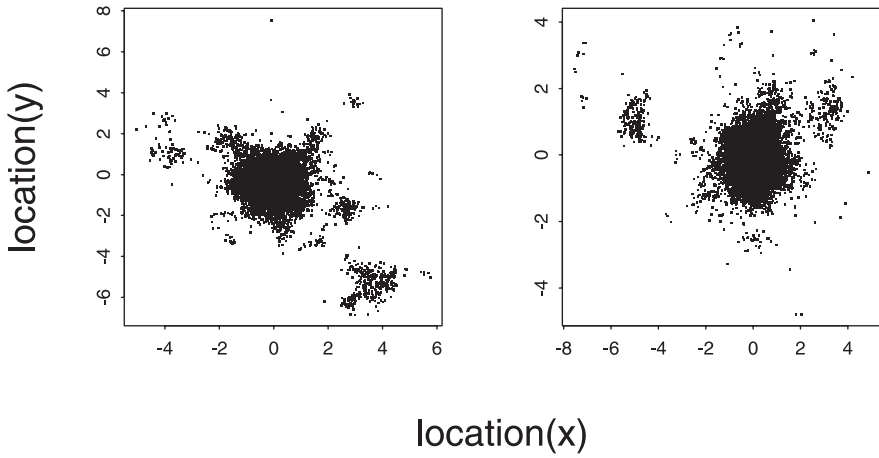
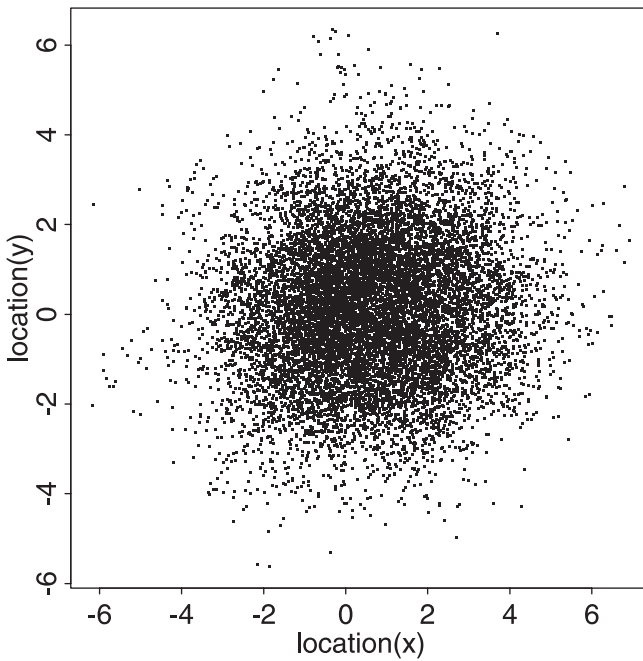
Fig. 1c.



**Fig. 2a.** When dispersal events occur on markedly different spatial scales, populations spread in a patchy manner. Monte–Carlo simulation of reproducing dispersing individuals. Each time step: (i) individuals have a Poisson number of offspring, with mean of 1.2, (ii) offspring disperse with random distances drawn from the dispersal kernel. Angles are drawn from a uniform distribution, and (iii) the parent individual then dies. Initially 10 individuals were released at  $x = 0$ . The distribution of individuals is shown after 40 time steps. (a) Composite Laplace kernel (113) with  $\alpha_1 = 10$ ,  $\alpha_2 = 1$  and  $p = 0.99$  (solid line) and Normal kernel (11) with  $D = 0.092$ . Notice that the Laplace kernel is leptokurtic: there are more very short and very long dispersal distances than found in a comparable Normal kernel. (b) Two Monte–Carlo simulation using the composite Laplace kernel. Here the two relevant spatial scales are: mean dispersal distance given by a Laplace kernel with  $\alpha_1 = 10$  (0.1 units), and mean dispersal distance given by a Laplace kernel with  $\alpha_2 = 1$  (1.0 units). The only difference between the two simulations is the random number seed used. (c) Monte–Carlo simulation using the Normal kernel. Here there is a single relevant spatial scale: mean dispersal distance given by a Normal kernel with  $D = 0.092$  (0.429 units). When dispersal events occur on markedly different spatial scales, there is increased stochastic variation about the theoretically predicted ARS. Using simulations as described above, but in one spatial dimension, the location of the individual dispersing farthest in the positive  $x$ -direction is plotted as a function of time. (d) Two simulations with parameters and kernel as described in (b) above. (e) Two simulations with parameters and kernel as described in (c) above. In each of (d) and (e) the dashed line has slope 0.2529, which is the theoretical ARS given by (107). The large vertical jumps in the graphs in (d) indicate single individuals making long-distance jumps thereby pushing forward the location of the farthest dispersing individual.

‘patchy’ spread, with locations of the patches varying from simulation to simulation. Furthermore, the spatial extent of the spread, as measured by the distance from the origin of the farthest dispersing individual, is not constant from simulation to simulation. While it is clear that the underlying dispersal process, as described by a dispersal kernel, plays a role in determining the average spread rate of individuals, variability in the rate of spread is also governed by this kernel (Figures 2d and 2e).

Stochastic models of population spread have been widely studied in the context of the spread of an infection. An introduction to the early work in this area can be found in a review by Mollison [26]. We will not attempt comprehensive review in this paper. The idea that equations for ‘second-order densities’, such as spatial covariance, would yield information about the spread of an invading population

**Fig. 2b.****Fig. 2c.**

was first pursued by Bartlett [3,4], and later by Daniels [9] and others [6]. Later approaches to nonlinear spatial epidemic models include particles interacting on a lattice [10] and coupling methods [2].

The simplest stochastic model involves only density-independent birth and dispersal. In this case equations for the successive moments can be written down

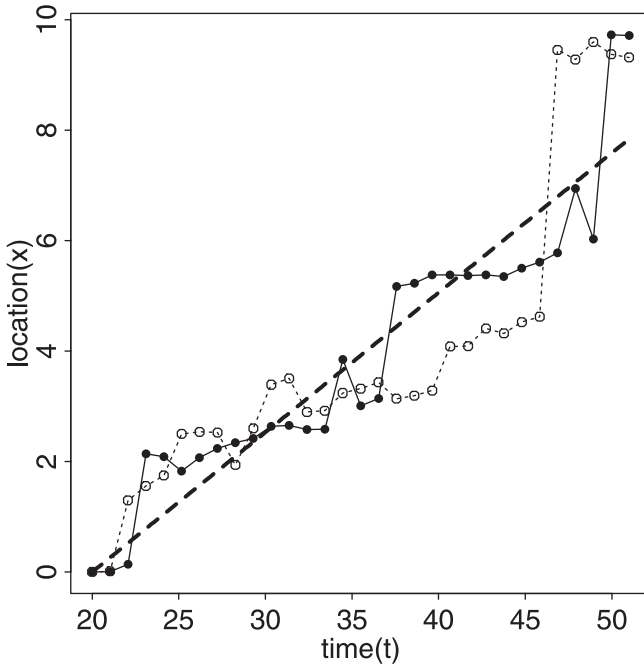


Fig. 2d.

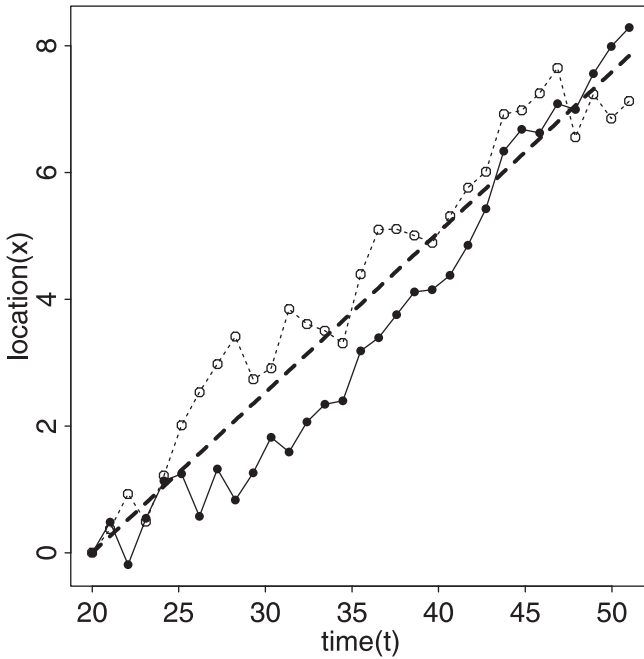


Fig. 2e.

explicitly (see [9] for a linear integro-differential equation formulation). Using transform methods and a saddle point approximation for inverting the transformed solution, Daniels [9] analysed the asymptotic velocity of the moments and derived approximate formulae for their asymptotic form.

Although population processes typically involve density-dependent interactions, simulations show that the phenomena of ‘patchy’ spread is evident even in the absence of such effects (Figure 2). In an attempt to quantitatively analyse the patchiness of a spreading population we will revisit linear models, but in doing so will derive a rather different system of equations than those studied by earlier researchers. By analyzing the covariance wave structure we will show how the shape of the dispersal kernel determines the asymptotic behavior of this wave and also the patchy structure of the invasion. Finally we will formulate a model for the spatial moments which includes density-dependence.

## 2. The dispersal kernel

Dispersal patterns have been measured for many organisms, either as seed shadows [40], plant disease dispersal gradients [13] or mark-recapture data [34]. In spite of variability in these patterns, there is an strong preponderance of leptokurtic data. Here there are more very short and more very long dispersal distances than found in a Normal kernel with comparable mean and variance.

To analyse the kurtosis of the dispersal data we define the spatial redistribution kernel by

$$\kappa(z, x)\Delta z\Delta x = \text{Probability of dispersing from } (z, z + \Delta z) \text{ to } (x, x + \Delta x). \quad (1)$$

In the case where dispersal is simply a function of distance, so that  $\kappa(z, x) = k(z - x)$ , moments of the kernel on an infinite domain can be defined as

$$k^{(n)} = \int_{-\infty}^{\infty} x^n k(x) dx, \quad (2)$$

provided they exist. The kurtosis of the kernel,

$$B_2 = \frac{k^{(4)}}{k^{(2)2}}, \quad (3)$$

is a measure of the disparity of spatial scales for the dispersal process. Leptokurtic kernels have  $B_2 > 3$ .

To understand how leptokurtic kernels can arise naturally from biological dispersal, we observe that the assumption of random movement, modeled by diffusion, a constant settling rate per unit time, and a sufficiently long time for dispersers to settle gives rise to a leptokurtic kernel described by a Laplace (decaying exponential) dispersal function in one spatial dimension, or a modified Bessel function in two spatial dimensions [8, 39, 21]. The effect of other movement assumptions on the shape of the dispersal kernel can be found in Neubert et al. [29] or Van Kirk [16].



### 3. Derivation of equations

In this section we derive equations for the expected density of individuals and the spatial covariance of individuals when the individuals undergo birth, death and dispersal. Basic definitions of these statistics and related density functions are given in Appendix A.

The population of individuals is assumed to undergo discrete-time reproduction. Immediately after reproduction the parent dies (or is simply considered an offspring) and the offspring redistribute according to the dispersal kernel (1). It is assumed that the individuals survive the redistribution process and thus

$$\int_{\Omega} \kappa(z, x) dx = 1, \tag{4}$$

where  $\Omega$  is the spatial domain of interest. The subscript  $t$  ( $t = 0, 1, 2, \dots$ ) is now used to denote the time step.

In Appendix B we show that if  $\mathcal{R}$  is a random variable representing each individual's reproductive output with expected value  $E[\mathcal{R}] = R$ , then equations governing the first two moments of the population are

$$p_{t+1}(x) = \int_{\Omega} R \kappa(z, x) p_t(z) dz \tag{5}$$

and

$$c_{t+1}(x, y) = \int_{\Omega} \{R(R - 1) + \text{Var}[\mathcal{R}]\} p_t(z) \kappa(z, x) \kappa(z, y) dz + \int_{\substack{\Omega \\ z_2 \neq z_1}} R^2 c_t(z_1, z_2) \kappa(z_1, x) \kappa(z_2, y) dz_1 dz_2, \tag{6}$$

where  $p_t(x)$  is the expected density of individuals and  $c_t(x, y)$  is the spatial covariance density.

The first term on the right hand side of (6) describes the spatial covariance due to two individuals having an identical parent, and the second term describes the spatial covariance due to two individuals from different but spatially correlated parents. The spatial covariance density  $c_t(x, y)$  is defined only for  $x \neq y$  and an extension of the definition to cover the point  $x = y$  introduces a singularity describing self-interactions (see equation (92)). Thus we are careful to exclude such points from domain of integration in the second term of equation (6).

The birth rate variance  $\text{Var}[\mathcal{R}]$  causes a increase in the contribution to spatial covariance that arises from individuals producing spatially correlated offspring. The case with a Poisson number of offspring yields  $R(R - 1) + \text{Var}[\mathcal{R}] = R^2$ .

Note that  $R \geq 0$  means that the term  $R(R - 1) + \text{Var}[\mathcal{R}]$  in (6) is always non-negative. When  $R < 1$  the smallest value arises when  $\mathcal{R}$  is a Bernoulli random variable with probability  $R$  of having a single offspring and probability  $1 - R$  of having no offspring. In this case  $\text{Var}[\mathcal{R}] = R(1 - R)$  and thus  $R(R - 1) + \text{Var}[\mathcal{R}] = 0$ . A consequence is that, for non-negative initial data, equations (5) and (6) remain non-negative for all time (see also [14]).

With (5) and (6) given as equations for the first two spatial moments, we integrate (5) and (6) with respect to space over the entire solution domain  $\Omega$ . This yields equations for the expected total number of individuals  $m_t = E(n_t(\Omega))$  and the variance in this total  $v_t$ ,

$$m_{t+1} = Rm_t \quad (7)$$

$$v_{t+1} = R^2 v_t + \text{Var}[\mathcal{R}] m_t, \quad (8)$$

with initial values

$$m_0 = \int_{\Omega} p_0(x) dx \quad (9)$$

$$v_0 = \int_{\Omega} \int_{\Omega} \tilde{c}_0(x, y) dx dy. \quad (10)$$

Here  $p_0(x)$  and  $c_0(x, y)$  describe the expected density and spatial covariance of the initial individuals over the entire ensemble of stochastically identical processes, and  $\tilde{c}_0(x, y)$  is the spatial covariance according to the extended definition given in Appendix A, equation (92). Using equation (92) from Appendix A and equations (7)–(8) we observe that the case  $c_0(x, y) \equiv 0$  is consistent with a Poisson number of initial individuals ( $v_0 = m_0$ ) which are randomly distributed over a spatial interval.

In summary, the assumptions of no density-dependent interactions between individuals, a random number of offspring per generation, and long-distance dispersal of individuals via a redistribution kernel lead to linear dynamical equations for the moments of the spatial distribution of individuals. Equations for the first two moments are given in (5) (expected density) and (6) (spatial covariance). Note the solution to (5) is used in (6), but the lack of density-dependent interactions between individuals leads to a closed system of equations—higher-order moments do not appear in equations for the mean (5) or spatial covariance (6). This contrasts with the situation shown in Section 9, where density-dependent interactions are included.

#### 4. Connection with continuous-time models

It is instructive to connect the discrete-time equations (5), (6) with Poisson birth dynamics with perhaps more familiar continuous time models. Specifically, if we choose the solution domain  $\Omega$  to be the real line and the redistribution kernel (1) to be the fundamental solution to the diffusion equation evaluated at unit time,

$$k(x) = \frac{1}{\sqrt{4D\pi}} \exp\left(-\frac{x^2}{4D}\right) = N(0, 2D), \quad (11)$$

then (5), (6) are transformed to

$$p_{t+1}(x) = \int_{-\infty}^{\infty} \frac{R}{\sqrt{4D\pi}} \exp\left(-\frac{(x-z)^2}{4D}\right) p_t(z) dz \quad (12)$$

and

$$c_{t+1}(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{R^2}{4D\pi} \exp\left(-\frac{(x - z_1)^2 + (y - z_2)^2}{4D}\right) \tilde{c}_t(z_1, z_2) dz_1 dz_2, \tag{13}$$

where  $\tilde{c}_t(z_1, z_2)$  is spatial covariance density, extended to cover the line  $x = y$  as defined in Appendix A, equation (92). Defining  $\psi(x, t)$  and  $\gamma(x, y, t)$  to be continuous-time analogs for the expected density and extended spatial covariance density, we observe that for  $t = 0, 1, 2, \dots$   $p_t(x) = \psi(x, t)$  and  $\tilde{c}_t(x, y) = \tilde{\gamma}(x, y, t)$  if  $\psi(x, t)$  satisfies the linearized Fisher equation

$$\frac{\partial \psi}{\partial t} = D \frac{\partial^2 \psi}{\partial x^2} + r \psi, \tag{14}$$

$$\psi(x, 0) = p_0(x) \tag{15}$$

and  $\gamma(x, y, t)$  satisfies

$$\frac{\partial \tilde{\gamma}}{\partial t} = D \left( \frac{\partial^2 \tilde{\gamma}}{\partial x^2} + \frac{\partial^2 \tilde{\gamma}}{\partial y^2} \right) + 2r \tilde{\gamma} + \delta(x - y) \frac{\partial \psi}{\partial t}, \tag{16}$$

$$\tilde{\gamma}(x, y, 0) = c_0(x, y) + \delta(x - y) p_0(x), \tag{17}$$

where  $r = \log(R)$  is the continuous-time birth rate.

To derive a continuous time spatial contact process with long-distance jumps we replace (11) with

$$k(x) = \frac{1 - \Delta t \mu}{R} \delta(x) + \frac{R - 1 + \Delta t \mu}{R} k_1(x), \tag{18}$$

and choose  $R$  in terms of the fecundity rate  $f$  and mortality rate  $\mu$  by  $R = \exp(\Delta t(f - \mu))$ . As  $\Delta t \rightarrow 0$ , the first two moments are given by (5), (6) as

$$\frac{\partial \psi}{\partial t} = -\mu \psi + \int_{\Omega} f k_1(z - x) \psi(z, t) dz \tag{19}$$

and

$$\begin{aligned} \frac{\partial \gamma}{\partial t} = & -2\mu \gamma + f (\psi(x, t) k_1(x - y) + \psi(y, t) k_1(y - x)) \\ & + \int_{\substack{\Omega \\ z \neq y}} f \gamma(z, y, t) k_1(z - x) dz + \int_{\substack{\Omega \\ z \neq x}} f \gamma(z, x, t) k_1(z - y) dz. \end{aligned} \tag{20}$$

Here it is assumed that parents do not move, but produce offspring at rate  $f$  and die at rate  $\mu$ . Immediately after birth, the offspring disperse instantaneously according to the redistribution kernel  $k_1$ . For the case  $\mu = 0$  (20) can be rewritten in a form identical to equation (5.8) of [9], once the extended definition of the spatial covariance (92) from Appendix A is used.

Finally, the case with two very different dispersal length scales can be thought of as a mixture of linear Fisher dynamics, describing local reproduction and neighborhood diffusion (14), and of jumps describing the long-distance dispersal of colonizing offspring (19). Here, as  $\Delta t \rightarrow 0$ , the kernel

$$k(x) = \frac{1 - \Delta t \lambda / R}{\sqrt{4D\Delta t\pi}} \exp\left(-\frac{x^2}{4D\Delta t}\right) + \frac{\lambda \Delta t}{R} k_1(x), \tag{21}$$

yields

$$\frac{\partial \psi}{\partial t} = D \frac{\partial^2 \psi}{\partial x^2} + r\psi + \lambda \int_{\Omega} k_1(z - x) \psi(z, t) dz, \tag{22}$$

and

$$\begin{aligned} \frac{\partial \tilde{\gamma}}{\partial t} = D \left( \frac{\partial^2 \tilde{\gamma}}{\partial x^2} + \frac{\partial^2 \tilde{\gamma}}{\partial y^2} \right) + 2r\tilde{\gamma} + \lambda \int_{\Omega} \tilde{\gamma}(z, y) k_1(z - x) dz \\ + \lambda \int_{\Omega} \tilde{\gamma}(z, x) k_1(z - y) dz + \delta(x - y) \frac{\partial \psi}{\partial t}, \end{aligned} \tag{23}$$

where  $r + \lambda = \log(R) / \Delta t$  is the continuous-time birth rate and  $\lambda$  is the colonization rate. Although the equations differ from those analysed in Shigesada et al. [32], they model the same phenomenon, referred to in [32] as ‘stratified diffusion’.

**5. How the redistribution kernel determines invasion speed**

It is possible to analyse the speed of invasion by choosing a threshold value of detection  $p_c$  for the expected density in (5). The location of this threshold is  $x_t$  and the invasion speed is defined to be the speed at which  $x_t$  moves. While a typical invasion scenario with compact initial data would give rise to two such points  $x_t$ , one moving to the right and the other to the left, without loss of generality we simply concentrate on the rightward moving wave (Figure 3).

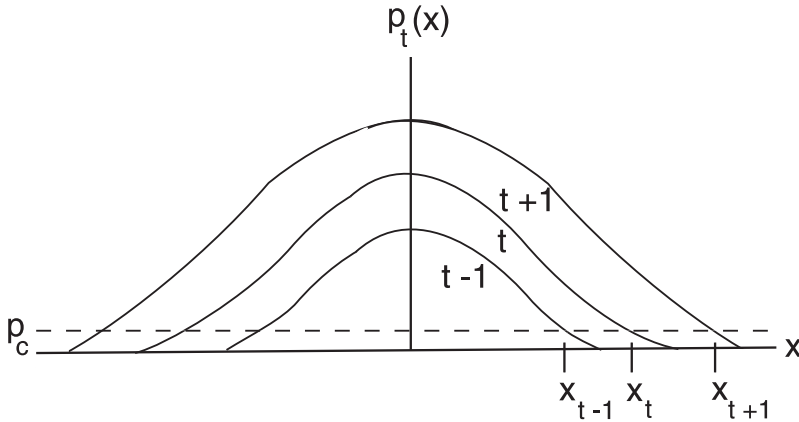
The rate of invasion of expected density in biological populations has been well studied, for both linear models and the related nonlinear deterministic models. Much of the early analysis for contact distributions with continuous time dynamics can be found in Mollison’s paper on the subject [22], and later results for integrodifference models can be found in the papers of Weinberger [37, 38]. A recent review of the subject is given in Kot et al. [18].

Kot et al. [18] describe how three different possible outcomes pertain, depending upon the shape of the redistribution kernel  $k(x)$ . (i)  $k(x)$  has a moment generating function (i.e. has exponentially bounded tails). Such kernels include the Normal distribution (11) and the leptokurtic Laplace kernel

$$k(x) = \frac{\alpha}{2} \exp(-\alpha|x|), \quad \alpha > 0. \tag{24}$$

Here the invasion speed is asymptotically constant provided the initial data are compact (Appendix C). (ii)  $k(x)$  has finite moments of all orders, but no moment generating function. Such ‘fat-tailed’ kernels include, for example,

$$k(x) = \frac{\alpha^2}{4} \exp(-\alpha\sqrt{|x|}), \quad \alpha > 0. \tag{25}$$



**Fig. 3.** Location of the threshold  $x_t$  as a function of time.

Here the spatial extent of spread  $x_t$  typically has polynomial growth in time and thus the invasion speed is asymptotically unbounded. For example, (5) and (25) and a point release with the expected number of individuals  $m_0$  yields a spatial extent

$$x_t = \frac{1}{\alpha^2} \left[ t \log(R) + \log \left( \frac{\alpha^2 m_0}{4 p_c} \right) \right]^2 \tag{26}$$

[18]. (iii)  $k(x)$  has moments that are infinite. These are extremely ‘fat-tailed’ kernels, and include, for example, the Cauchy distribution

$$k(x) = \frac{1}{\pi} \frac{\beta}{\beta^2 + x^2}, \quad \beta > 0. \tag{27}$$

Here the spatial extent of spread  $x_t$  can have exponential growth. For example (5), (27) yields a spatial extent

$$x_t = \sqrt{\frac{\beta t m_0 R^t}{\pi p_c} - \beta^2 t^2} \tag{28}$$

[18].

In summary, the rate of spread of expected density of individuals into the new environment may be linear (exponentially bounded kernel), polynomial (kernel with finite moments of all orders), or exponential (kernel without finite moments of all orders) in time. However, this spread rate gives no information about the actual spatial patterning arising from spatial correlations between individuals. We now analyse these correlations by deducing the qualitative large-time behavior of the covariance equation. It is most convenient to start this analysis in frequency space.

## 6. Analysis in frequency space yields a condition for permanence of form

We consider the solution to (5) and (6) when the spatial redistribution kernel is simply a function of distance, so that  $\kappa(x, y) = k(x - y)$ . Defining the Fourier transform and its inverse as

$$\hat{f}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \exp(i\omega x) dx, \quad f(x) = \int_{-\infty}^{\infty} \hat{f}(\omega) \exp(-i\omega x) d\omega, \quad (29)$$

the transformed versions of (5) and (6) are

$$\hat{p}_{t+1}(\omega) = 2\pi R \hat{k}(\omega) \hat{p}_t(\omega) \quad (30)$$

and

$$\begin{aligned} \hat{c}_{t+1}(\omega_1, \omega_2) &= 2\pi (R(R-1) + \text{Var}[\mathcal{A}]) \hat{k}(\omega_1) \hat{k}(\omega_2) \hat{p}_t(\omega_1 + \omega_2) \\ &\quad + 4\pi^2 R^2 \hat{k}(\omega_1) \hat{k}(\omega_2) \hat{c}_t(\omega_1, \omega_2), \end{aligned} \quad (31)$$

with solutions

$$\hat{p}_t(\omega) = \left(2\pi R \hat{k}(\omega)\right)^t \hat{p}_0(\omega) \quad (32)$$

and

$$\begin{aligned} \hat{c}_t(\omega_1, \omega_2) &= \left(4\pi^2 R^2 \hat{k}(\omega_1) \hat{k}(\omega_2)\right)^t \hat{c}_0(\omega_1, \omega_2) \\ &\quad + \sum_{s=0}^{t-1} \left(4\pi^2 R^2 \hat{k}(\omega_1) \hat{k}(\omega_2)\right)^{t-1-s} 2\pi (R(R-1)) \\ &\quad + \text{Var}[\mathcal{A}] \hat{k}(\omega_1) \hat{k}(\omega_2) \hat{p}_s(\omega_1 + \omega_2) \end{aligned} \quad (33)$$

respectively. Each element summed in (33) gives the contribution to the transformed spatial covariance from a time  $s$  units earlier.

If the initial spatial covariance is zero and the number of offspring is Poisson distributed (33) simplifies to

$$\hat{c}_t(\omega_1, \omega_2) = \frac{1}{2\pi} \sum_{s=0}^{t-1} \left(4\pi^2 R^2 \hat{k}(\omega_1) \hat{k}(\omega_2)\right)^{t-s} \hat{p}_s(\omega_1 + \omega_2) \quad (34)$$

Using the results from the previous section, observe that these initial data constrain  $p_t(x)$  and  $c_t(x, y)$  to be non-negative for all time.

To understand the contribution to the transformed spatial covariance from each time step we consider the case where a Poisson number of individuals ( $n_0$ ) is released from a point source at  $x = 0$ . Here

$$\hat{p}_t(\omega) = \frac{m_0 \left(2\pi R \hat{k}(\omega)\right)^t}{2\pi}. \quad (35)$$

where  $m_0 = E(n_0)$ . The transformed spatial covariance (34) can be expressed solely in terms of the transformed expected densities at previous time steps

$$\hat{c}_t(\omega_1, \omega_2) = \frac{2\pi}{m_0^2} \sum_{s=0}^{t-1} \hat{p}_{t-s}(\omega_1) \hat{p}_{t-s}(\omega_2) \hat{p}_s(\omega_1 + \omega_2) \tag{36}$$

and thus we calculate the spatial covariance directly in terms of expected densities as

$$c_t(x, y) = \frac{1}{m_0^2} \sum_{s=0}^{t-1} \int_{-\infty}^{\infty} p_{t-s}(x - z) p_{t-s}(y - z) p_s(z) dz. \tag{37}$$

The  $s^{\text{th}}$  term in this sum is the contribution from time step  $s$  to the spatial covariance at time  $t$  (Figure 4). For the case with a Normal kernel (11), (37) can be calculated explicitly as

$$c_t(x, y) = \frac{1}{m_0^2} \sum_{s=0}^{t-1} \frac{R^{2(t-s)}}{2} \frac{\exp\left(-\frac{\left(\frac{x-y}{\sqrt{2}}\right)^2}{4D(t-s)}\right)}{\sqrt{\pi 4D(t-s)}} \frac{\exp\left(-\frac{\left(\frac{x+y}{\sqrt{2}}\right)^2}{4D(t-s)+2Ds}\right)}{\sqrt{\pi(4D(t-s)+2Ds)}}. \tag{38}$$

To analyse the form of the spatial covariance for large  $t$  we can rewrite (35)–(36) in terms of a geometric sum

$$\hat{c}_t(\omega_1, \omega_2) = \frac{\hat{p}_t(\omega_1) \hat{p}_t(\omega_2)}{m_0} \sum_{s=0}^{t-1} \phi^s(\omega_1, \omega_2), \tag{39}$$

provided  $\hat{k}(\omega_1)\hat{k}(\omega_2) \neq 0$ , or

$$\hat{c}_t(\omega_1, \omega_2) = \frac{\hat{p}_t(\omega_1 + \omega_2)}{2\pi} \sum_{s=1}^t \frac{1}{\phi^s(\omega_1, \omega_2)}, \tag{40}$$

provided  $\hat{k}(\omega_1 + \omega_2) \neq 0$ . Here for any pair of wave numbers  $(\omega_1, \omega_2)$

$$\phi(\omega_1, \omega_2) = \left( \frac{\hat{k}(\omega_1 + \omega_2)}{2\pi R \hat{k}(\omega_1) \hat{k}(\omega_2)} \right) \tag{41}$$

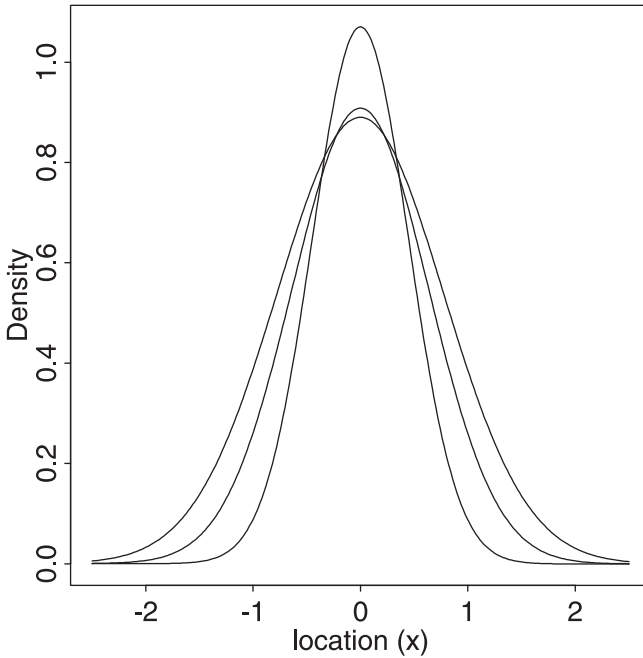
is determined uniquely by  $\hat{k}(\omega)$ , the transform of the redistribution  $k(x)$ , and  $R$ , the basic reproductive number.

Thus from (39) we observe that for any pair of wave numbers,  $\omega_1$  and  $\omega_2$ ,  $|\phi(\omega_1, \omega_2)| < 1$  implies that as  $t \rightarrow \infty$

$$\hat{c}_t(\omega_1, \omega_2) \rightarrow \frac{\hat{p}_t(\omega_1) \hat{p}_t(\omega_2)}{m_0(1 - \phi(\omega_1, \omega_2))} \tag{42}$$

and so

$$\hat{c}_{t+1}(\omega_1, \omega_2) \rightarrow 4\pi^2 R^2 \hat{k}(\omega_1) \hat{k}(\omega_2) \hat{c}_t(\omega_1, \omega_2). \tag{43}$$



**Fig. 4a.** Spread in expected density and covariance decomposition for the first three time steps of an invasion. A single individual is released at  $x = 0$  when  $t = 0$ . Thus the expected density at  $t = 0$  is  $p_0(x) = \delta(x)$ . (a) The expected density is shown at each of the three time steps  $t = 1$  (inside)  $t = 2$  (middle) and  $t = 3$  (outside). Calculations solved (5) using Fast Fourier Transforms on a size 256 grid. (b) The expected densities from (a) are used to calculate the contribution to the spatial covariance (37) from each of the time steps for  $s = 0$  (top left)  $s = 1$  (top right) and  $s = 2$  (bottom left) and finally the covariance at  $t = 3$  as the sum of the three terms (bottom right). Growth is Poisson, given by  $R = 1.2$  and  $\sigma_R^2 = 1.2$ , and the dispersal kernel is Normal (11) with  $D = 0.1$ .

On the other hand when  $|\phi(\omega_1, \omega_2)| > 1$  we observe from (40) that as  $t \rightarrow \infty$

$$\hat{c}_t(\omega_1, \omega_2) \rightarrow \frac{\hat{p}_t(\omega_1 + \omega_2)}{2\pi(\phi(\omega_1, \omega_2) - 1)} \tag{44}$$

and so

$$\hat{c}_{t+1}(\omega_1, \omega_2) \rightarrow \hat{k}(\omega_1 + \omega_2)\hat{c}_t(\omega_1, \omega_2). \tag{45}$$

The dependence of large-time asymptotic behavior of  $\hat{c}_{t+1}(\omega_1, \omega_2)$  on  $\phi(\omega_1, \omega_2)$  (41) suggests a decomposition for the Fourier transformed covariance:

$$\hat{c}_t(\omega_1, \omega_2) = \hat{c}_t^-(\omega_1, \omega_2) + \hat{c}_t^+(\omega_1, \omega_2), \tag{46}$$

where

$$\hat{c}_t^-(\omega_1, \omega_2) = \begin{cases} \hat{c}_t(\omega_1, \omega_2) & \text{if } |\phi(\omega_1, \omega_2)| < 1 \\ 0 & \text{otherwise} \end{cases}, \tag{47}$$

and

$$\hat{c}_t^+(\omega_1, \omega_2) = \begin{cases} \hat{c}_t(\omega_1, \omega_2) & \text{if } |\phi(\omega_1, \omega_2)| > 1 \\ 0 & \text{otherwise} \end{cases}. \tag{48}$$



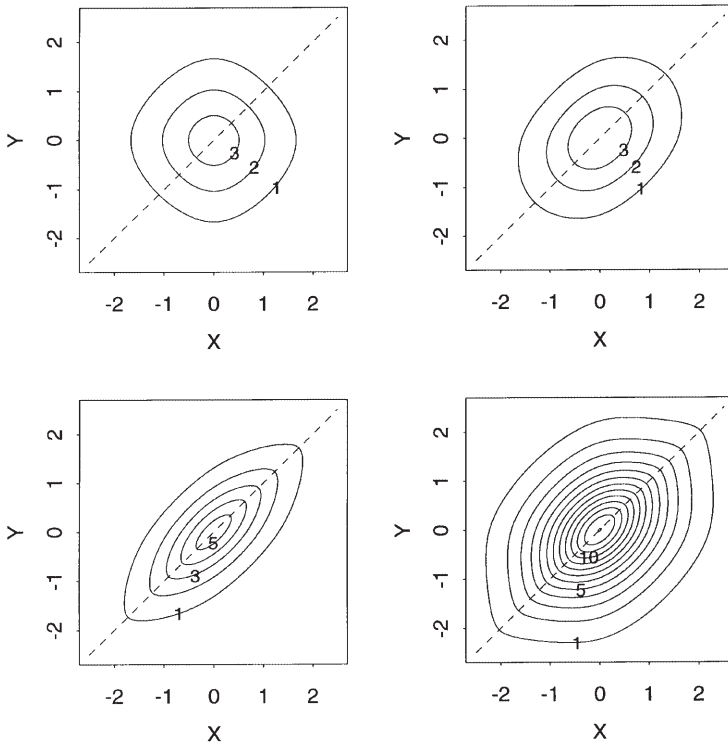


Fig. 4b.

Using these definitions we observe that  $\hat{c}_t^-(\omega_1, \omega_2)$  satisfies (43) for all pairs of wave numbers  $(\omega_1, \omega_2)$ , and  $\hat{c}_t^+(\omega_1, \omega_2)$  satisfies (45) for all pairs of wave numbers  $(\omega_1, \omega_2)$ . Taking inverse transforms of  $\hat{c}_t^-(\omega_1, \omega_2)$  and  $\hat{c}_t^+(\omega_1, \omega_2)$  and using (46) yields

$$c_t(x, y) = c_t^-(x, y) + c_t^+(x, y), \tag{49}$$

where

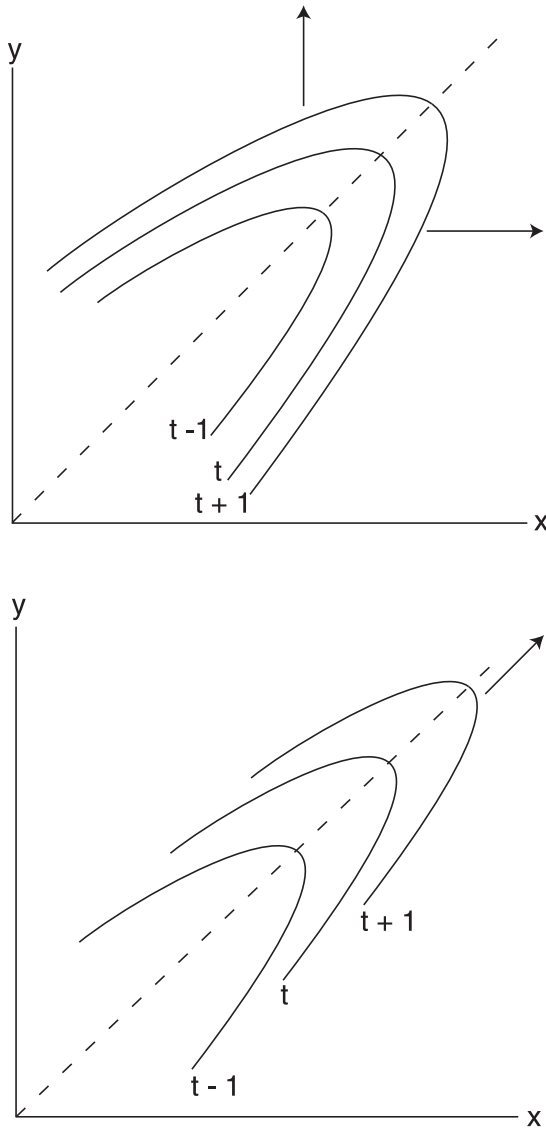
$$c_{t+1}^-(x, y) \rightarrow \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R^2 c_t^-(z_1, z_2) k(z_1 - x) k(z_2 - y) dz_1 dz_2, \tag{50}$$

$z_2 \neq z_1$

and

$$c_{t+1}^+(x, y) \rightarrow \int_{-\infty}^{\infty} R c_t^+(x - z, y - z) k(z) dz. \tag{51}$$

The first term in the covariance decomposition (49)  $c_t^-(x, y)$  has asymptotic behavior governed by (50). Here the value at time  $t + 1$   $c_{t+1}^-(x, y)$  is a scaled weighted spatial average of the value at time  $t$   $c_t^-(x, y)$ . Over time the  $c_t^-(x, y)$  term spreads symmetrically in both  $x$ - and  $y$ -directions according to the spatial weighting function  $k(z_1 - x)k(z_2 - y)$  (Figure 5a).



**Fig. 5.** (a) Covariance isoclines for  $c_t^-(x, y)$  spread in both the  $x$ - and  $y$ -directions. (b) Covariance isoclines for  $c_t^+(x, y)$  spread only in both the  $(x + y)/2$ -direction.

The second term in the covariance decomposition (49)  $c_t^+(x, y)$  has asymptotic behavior governed by (51). Once again the value at time  $t + 1$   $c_{t+1}^+(x, y)$  is a scaled weighted spatial average of the covariance at time  $t$   $c_t^+(x, y)$ . However, over time  $c_t^+(x, y)$  now spreads only with respect to a single variable,  $(x + y)/2$ , representing the mean location. To see this more clearly, one can rewrite equation (51) in a form similar to (50):

$$c_{t+1}^+(x, y) \rightarrow \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R c_t^+(z_1, z_2) k((x+y)/2 - (z_1+z_2)/2) \delta(z_1 - z_2) dz_1 dz_2. \tag{52}$$

If we shift to a local reference frame moving with the wave in the direction  $x = y$  we observe no asymptotic spread of spatial covariance in the transverse direction  $x = -y$ . In other words, level sets, defined by  $c_t(x, y) = c_0 > 0$ , simply move in the direction  $x = y$ . While the speed with which the level set moves can depend upon  $c_0$ , the local shape of the set does not change as it moves (Figure 5b). Symmetry of  $c_t^+(x, y)$  about the  $x = y$  line means that this *wave of permanent form* describes spatial covariance in terms of the spatial lag  $|x - y|$ . This permanent form describes the characteristic spatial correlations between individuals that, after a long time, remain unchanged in the leading edge of the wave. We refer to these unchanging spatial correlations in the leading edge of the wave as *permanence of form*. The permanence of form occurs for the range of wave numbers where  $|\phi| > 1$  and thus  $c_t^+(\omega_1, \omega_2)$  is nonzero.

Looking back to equations (37), (39), (40), (41) we observe that this permanence of form arises from the later terms in (37) dominating the summation when analysed in Fourier space. In other words, the permanence of form is the result of spatial covariances between relatives with a recent common ancestor dominating the covariance structure of the wave.

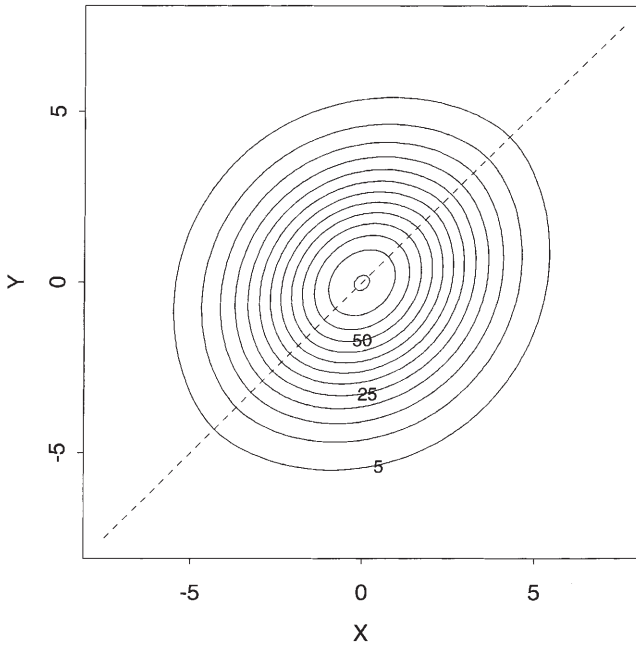
By way of contrast the covariance wave has no permanence of form over the range of wave numbers where  $|\phi| < 1$ . This portion of the covariance is denoted by  $c_t^-(x, y)$ . Although  $c_t^-(x, y)$  does describe spatial correlations, these change continually over time.

Appendix D shows that  $\phi = 1/R$  for delta function and Cauchy kernels. Normal, composite Normal and composite Laplace kernels have their  $\phi$  functions calculated and shown graphically in Figure 9.

Examples of difference in the spread patterns in (50) and (52) are illustrated with a Normal kernel with  $R = 1.2$  ( $|\phi| < 1$  for all  $\omega_1, \omega_2$ ) and an exponentially declining kernel with  $R = 0.5$  ( $|\phi| > 1$  for all  $\omega_1, \omega_2$ ) (Appendix D) (Figure 6). The covariance  $c_t(x, y)$  can be explicitly calculated for a Normal kernel (38). Here the early terms in the series dominate; the  $s$ th term in the series is the product of two Normal distributions with standard deviations of  $2D(t - s)$  and  $D(2t - s)$ . Numerical solution of covariance for the exponentially declining kernel with  $R = 0.5$  ( $|\phi| > 1$ ) yields a wave moving in the direction  $x = y$  while retaining a constant profile in the direction  $x = -y$ .

When permanence of form is seen only over a range of high wave numbers the unchanging covariance structure is most relevant at the level of interactions between individuals at short spatial scales (compare Figures 7a and 7b), but may not be apparent from casual observation of the global spread of the invasion.

In Appendix E we derive an asymptotic relationship between  $\phi$  (41) and kurtosis (3) for a symmetric kernel. We demonstrate that kernels that are sufficiently leptokurtic (i.e., that have sufficiently large values of (3)) can give rise to  $|\phi| > 1$  for wavelengths that are long relative to the standard deviation of the kernel. In



**Fig. 6a.** (a) Asymptotic form of the spatial covariance for a case where  $|\phi(\omega_1, \omega_2)| < 1$  for all wave numbers. Here the dispersal kernel is Normal (11) with  $D = 0.1$  (as in Figure 4), growth is given by  $R = 1.2$  and  $\sigma_R^2 = 1.49$ , the initial data for the covariance is zero everywhere, and the initial data for the expected density is as described in Figure 8a. Isoclines, shown at time  $t = 30$ , range from 5 to 65. (b) Asymptotic form of the spatial covariance for a case where  $|\phi(\omega_1, \omega_2)| > 1$  for all wave numbers. As discussed in Appendix D, at low wave numbers this inequality requires  $R < 1$ . Here the dispersal kernel is composite Laplace, as described in Figure 2a, growth is given by  $R = 0.5$  and  $\sigma_R^2 = 0.375$ , and initial data is as described above in (a). Isoclines, shown at time  $t = 30$  range from  $1 \times 10^{-10}$  to  $5 \times 10^{-10}$ . Calculations in parts (a) and (b) solved (5), (6) using Fast Fourier Transforms on a size 256 by 256 grid.

other words, leptokurtic kernels can give rise to permanence of form on length scales that are many times larger than the average dispersal distance.

In summary, the spatial covariance can be broken into two terms (49), the first of which ( $c_t^-(x, y)$ ) spreads symmetrically in both  $x$ - and  $y$ -directions and describes correlations arising from distant ancestors sharing spatial locations, and the second of which ( $c_t^+(x, y)$ ) spreads only in the direction  $x = y$ , has permanence of form, and describes spatial correlations from sharing close ancestors (Figure 7). The Fourier spectrum of the first term is nonzero on the set where  $|\phi(\omega_1, \omega_2)| < 1$  and the Fourier spectrum of the second term is nonzero on the set where  $|\phi(\omega_1, \omega_2)| > 1$ . In turn the function  $\phi(\omega_1, \omega_2)$  (41), determining the qualitative behavior of modes with given wave numbers  $(\omega_1, \omega_2)$ , involves two ingredients (41): the shape of the redistribution kernel  $k(x)$ , as reflected by its Fourier transform  $\hat{k}(\omega)$ , and the

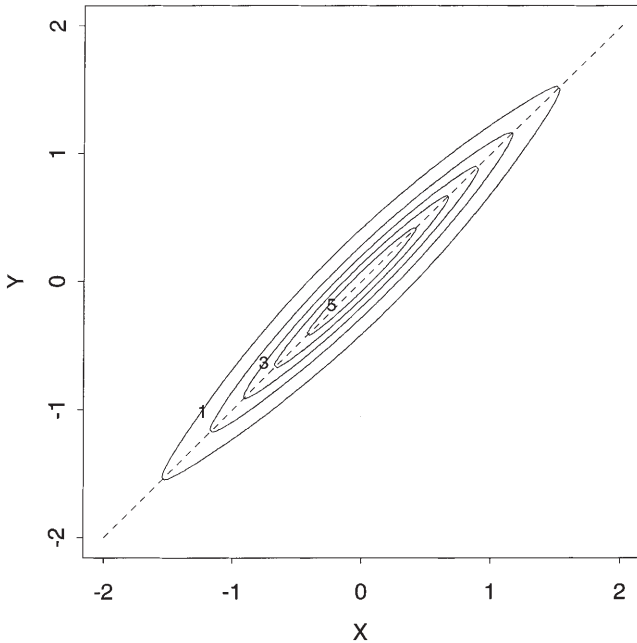


Fig. 6b.

basic reproductive number of the population  $R$ . We now take a closer look at the relationship between covariance form of the wave and the function  $\phi$ .

**7. Approximations for the covariance wave form in terms of  $\phi$**

While typical values for  $\phi$  may vary above and below 1, depending upon wave numbers, analyses of the limiting cases  $\phi \rightarrow 0$  and  $\phi \rightarrow \infty$  provide insight as to the limiting behaviors of the spatial covariance. In practice, these limits would require either a very large or very small value for the net reproductive number  $R$  (41).

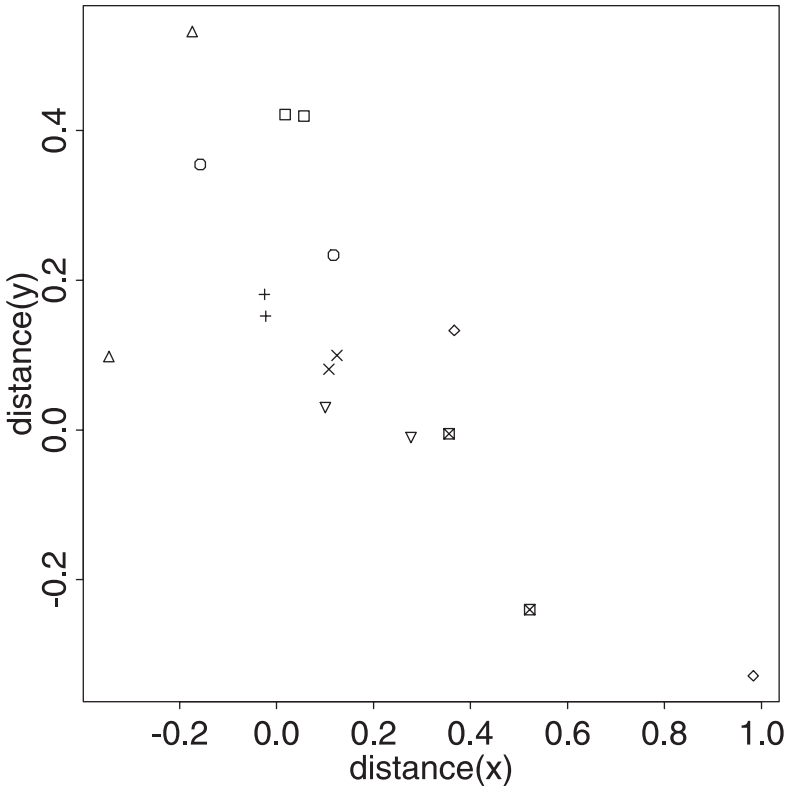
Observe that the limiting case  $\phi \rightarrow 0$  in (43) yields

$$c_{t+1}(x, y) \rightarrow \frac{p_{t+1}(x)p_{t+1}(y)}{n_0} \tag{53}$$

whereas the limiting case  $\phi \rightarrow \infty$  in (45) yields

$$c_{t+1}(x, y) \rightarrow \int_{-\infty}^{\infty} R^2 p_t(z) k(z-x) k(z-y) dz. \tag{54}$$

The former limiting case (53) has no permanent form whatsoever and spatial covariance determined solely by the spatial expansion of the expected density at time  $s = 0$  for  $t + 1$  time steps (37). In other words, the spatial covariance at time  $t$  simply arises from the initial correlations of individuals, which were distributed



**Fig. 7a.** Monte–Carlo simulations show the locations of siblings at time step  $t = 5$ . Simulations are as described in Figure 2, except (i) one individual is released at the origin at time  $t = 0$ , and (ii) each parent has precisely two offspring per generation ( $R = 2$ ,  $\text{Var}[\mathcal{R}] = 0$ ). Thus after five generations there are 32 individuals. Symbols show the locations of pairs of siblings. (a) The kernel is composite Laplace (113) with  $\alpha_1 = 10$ ,  $\alpha_2 = 1$  and  $p = 0.99$ , as in Figure 2a. The function  $\phi(\omega_1, \omega_2)$ , as shown in Figure 9c, but scaled by a factor describing the ratios of the growth rates  $1.2/2 = 0.6$ , indicates permanence of form at high wave numbers, including  $\omega_1, \omega_2 \geq 40$  and thus at short wave lengths, including  $\lambda_1 = 2\pi/\omega_1, \lambda_2 = 2\pi/\omega_2 \leq 0.08$ . On these scales spatial covariance has permanence of form. Thus asymptotically spatial correlations between closely related individuals predominate as described by the later terms in (37). Note three close pairs of siblings separated by such short distances (box, cross and times). (b) The kernel is a composite Normal kernel (111) with  $D_1 = 0.02$ ,  $D_2 = 2$  and  $p = 0.99$ . The function  $\phi(\omega_1, \omega_2) < 1$  for all wave numbers. The covariance thus has no permanence of form on any spatial scale and thus asymptotically correlations between distantly related individuals predominate, as described by the later terms in (37). Note that there are fewer closely paired siblings than shown in (a).

as a point source (35). The latter limiting case (54) has permanence of form and spatial covariance determined solely by the spatial expansion of expected density at time  $s = t$  for one time step (37). In other words, the spatial covariance simply arises from siblings with the same parent. Cases for less extreme values of  $\phi$  are shown in Figure 8.

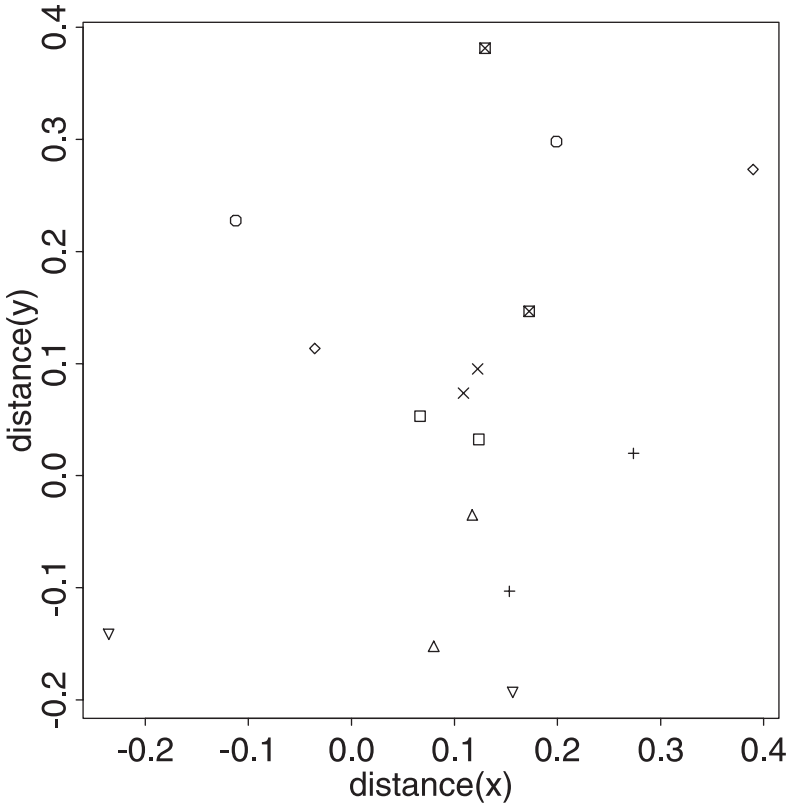


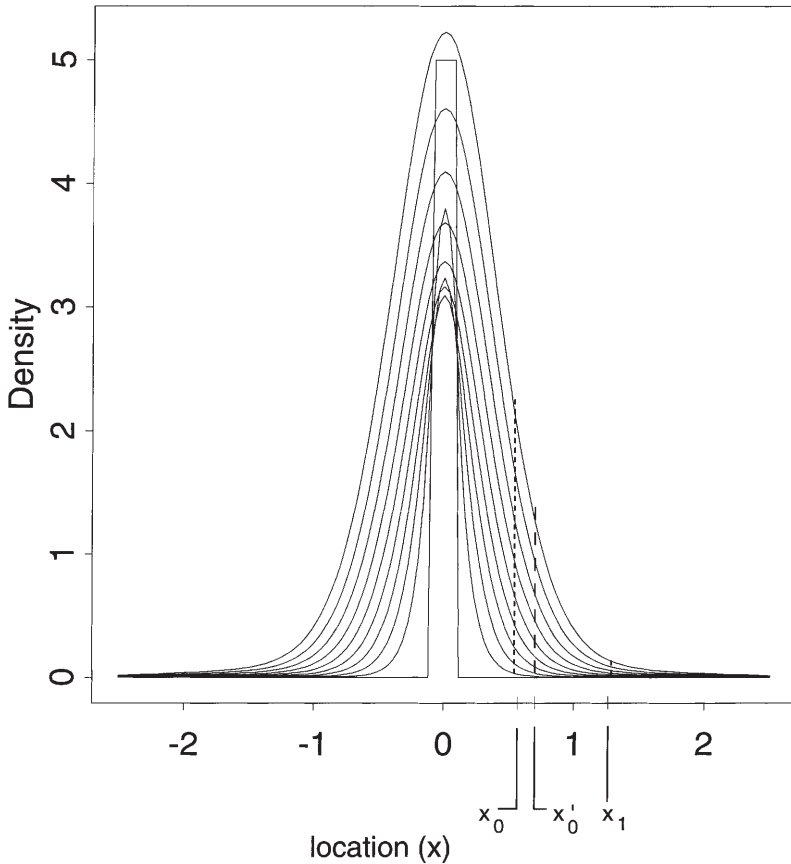
Fig. 7b.

The correlation density function about the leading edge of the wave in  $p_t(x)$  is dramatically different for the cases  $\phi \rightarrow 0$  and  $\phi \rightarrow \infty$ . The case  $\phi \rightarrow 0$  (53) yields a correlation density function (90)  $\rho_t(x, y) = \sqrt{p_t(x)p_t(y)} \rightarrow 0$  at the leading edge of the wave.

For exponentially bounded dispersal kernels  $k(x)$ , it is possible to heuristically analyse the case  $\phi \rightarrow \infty$  (54). Substitution of (108) into (54), evaluation of the integral and application of (90) and (107) yields the correlation density about the leading edge of the wave as

$$\rho_t(x, y) = \rho_t(x - y) = \frac{R}{M(s)} \int_{-\infty}^{\infty} k(\xi + (y - x)/2)k(\xi + (x - y)/2) \exp(-s\xi) d\xi. \tag{55}$$

Thus the correlation density is a positive function that simply depends upon spatial lag, and not upon the location in the wave front. This can be seen in some numerical examples, even when  $|\phi| < 1$  for some wave numbers (Figure 8c). Note that the correlation density function for short spatial scales is given in terms of moment generating functions by  $\lim_{x \rightarrow y} \rho_t(x - y) = RM_{k^2}(s)/M_k(s)$ .



**Fig. 8a.** Spread governed by a composite Laplace kernel as described in Figure 2a with growth given by  $R = 1.2$  and  $\sigma_R^2 = 1.49$ . (a) Expected density, shown from  $t = 0$  (rectangular initial conditions) to  $t = 10$  (outer) (b) Spatial covariance shown at  $t = 10$  has isoclines ranging from  $2 \times 10^1$  to  $24 \times 10^1$ . (c) Spatial correlation function shown at time  $t = 10$  has isoclines ranging from 5 to 40. At time  $t = 10$  the location  $x'_0$  with probability  $q = 0.5$  of having an individual to the right of it, given by  $\int_{x'_0}^{2.5} p_{10}(\xi) d\xi$  is  $x'_0 = 0.703$  (see (a)). The corresponding integrated covariance is  $\int_{x'_0}^{2.5} \int_{x'_0}^{2.5} c_{10}(x, y) dx dy = 4.99$  (see (b)). Thus (92), (93) give the variance as  $4.99 + 0.5 = 5.49$ , and a variance to mean ratio of  $5.49/0.5 = 10.98$ , indicating a high degree of variability in the location of individuals in the leading edge of the wave. (See also Figure 2d.) At time  $t = 10$  the expected number of individuals between  $x_0 = 0.625$  and  $x_1 = 1.25$  is  $0.593$  (see (a)). The corresponding integrated covariance is  $5.30$  (see (b)). Thus the mean crowding in the region  $0.625 \leq x \leq 1.25$  is  $C = 0.593 + 5.30/0.593 = 9.53$  (61). In other words, given that there is an individual in this region, the expected number of other individuals in this region is  $9.53$ .



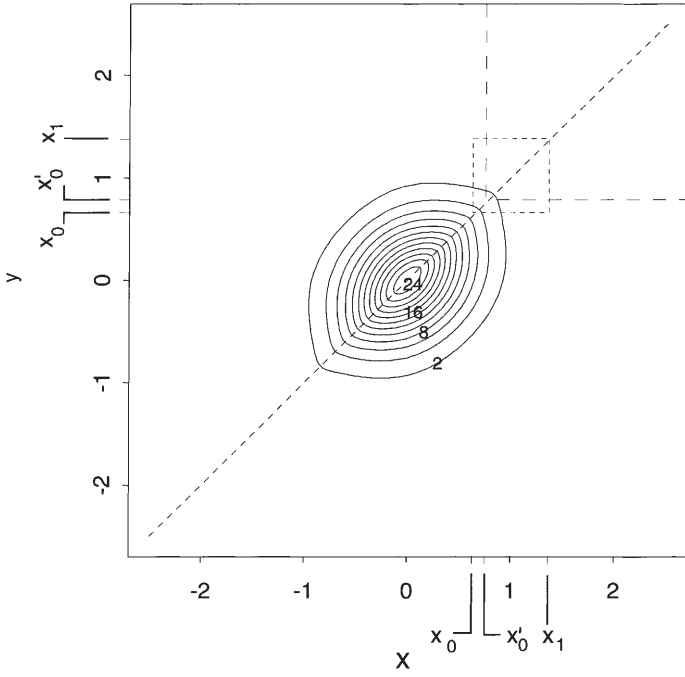


Fig. 8b.

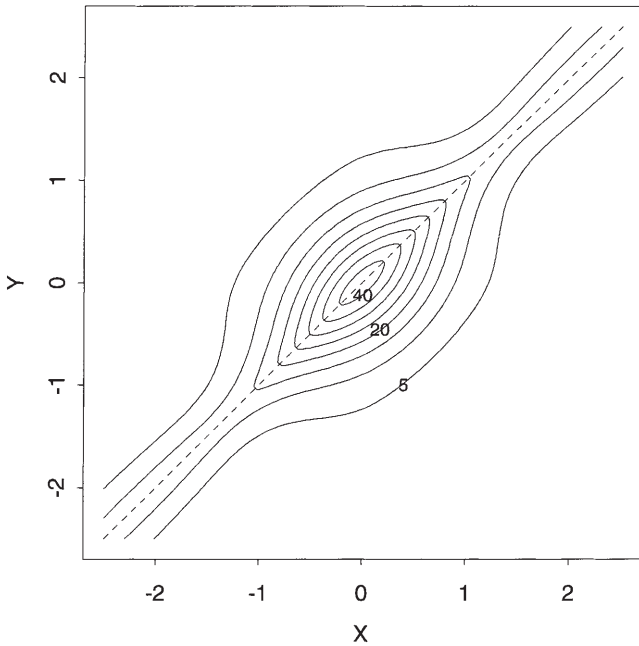


Fig. 8c.

The variance over an interval in the leading edge of the wave can also be calculated from (92)–(93) for the two limiting cases. As  $\phi \rightarrow 0$  the variance over the interval  $(x_0, x_1)$  is given by (53), (108) as

$$V(x_0, x_1) = \left( \frac{p_t(x_0) - p_t(x_1)}{s} \right) \left( 1 + \frac{p_t(x_0) - p_t(x_1)}{s} \right). \tag{56}$$

As  $\phi \rightarrow \infty$  the variance over the interval  $(x_0, x_1)$  is given by (54), (108) as

$$V(x_0, x_1) = \left( \frac{p_t(x_0) - p_t(x_1)}{s} \right) \left( 1 + \frac{R}{M(s)} \int_{-\infty}^{\infty} H^2(u) \exp(-su) du \right), \tag{57}$$

where  $H$  is the cumulative density function for  $k$

$$H(x) = \int_{-\infty}^x k(\xi) d\xi. \tag{58}$$

Equations (56)–(58) can be used to understand variability in the location of the farthest dispersing individual about the theoretically predicted value (Figure 2) as a function of  $\phi$ . If we choose  $x_1 = \infty$  and choose  $x_0$  so that there is probability  $q < 1$  of an individual to the right of  $x_0$  then

$$\int_{x_0}^{\infty} p_t(\xi) d\xi = \frac{p_t(x_0)}{s} = q. \tag{59}$$

This means that the variance to mean ratio over the interval  $(x_0, \infty)$  is  $1 + q$  for the case  $\phi \rightarrow 0$  (56) but is a quantity independent of  $q$

$$1 + \frac{R}{sM(s)} \int_{-\infty}^{\infty} H^2(u) \exp(-su) du, \tag{60}$$

for the case  $\phi \rightarrow \infty$  (57). This indicates that, when the likelihood of an individual dispersing past a given point  $x_0$  is low, the variability in the number dispersing past that point is higher for  $\phi \rightarrow \infty$  than for  $\phi \rightarrow 0$ . Another way of interpreting this result is to say that the precision with which the location of the farthest dispersing individual can be determined is lower for  $\phi \rightarrow \infty$  than for  $\phi \rightarrow 0$ . Calculations for mean and variance in the tails of the expectation and covariance waves are shown for a numerical example in Figures 8a and 8b.

Alternatively, (56)–(58) can be used to calculate a ‘mean crowding index’ in the leading edge of the wave. This is based on the idea of Lloyd [20] that the appropriate index for mean crowding over a region is the expected number of other individuals in the region seen by a given individual or, using the terminology of this paper

$$C = \frac{\int_{x_0}^{x_1} \int_{x_0}^{x_1} p_t^{(2)}(x, y) dx dy}{\int_{x_0}^{x_1} p_t(x) dx} = \int_{x_0}^{x_1} p_t(x) dx + \frac{\int_{x_0}^{x_1} \int_{x_0}^{x_1} c_t(x, y) dx dy}{\int_{x_0}^{x_1} p_t(x) dx} \tag{61}$$

A sample calculation for  $C$  is made in the caption for Figure 8. As  $\phi \rightarrow 0$  the mean crowding index (61), calculated using (89), (92), (93) and (56) is

$$C = \frac{2(p_t(x_0) - p_t(x_1))}{s}. \tag{62}$$

As  $\phi \rightarrow \infty$  the mean crowding index (61), calculated using (89), (92), (93) and (57) is

$$C = \frac{2(p_t(x_0) - p_t(x_1))}{s} + \frac{R}{sM(s)} \int_{-\infty}^{\infty} H^2(u) \exp(-su) du. \tag{63}$$

Note that near the front of the wave ( $x_0$  large) the mean crowding goes to zero for the  $\phi \rightarrow 0$  case (62), but remains a positive constant for the  $\phi \rightarrow \infty$  case (63). This means that, for the  $\phi \rightarrow \infty$  case, even when the mean density goes to zero, the mean crowding remains positive.

In summary, the cases  $\phi \rightarrow 0$  and  $\phi \rightarrow \infty$  allow us to analyse limiting behaviors of the spatial covariance equation. In the case  $\phi \rightarrow 0$  (no permanence of form and earliest term in (37) dominates) (i) the correlation density approaches zero in the leading edge of the wave (ii) the mean to variance ratio in the leading edge of the wave approaches  $1 + q$  where  $q$  is the expected number of individuals in the leading edge (56) (iii) the mean crowding index (62) is simply a function of the expected density and approaches zero as the expected density approaches zero. In the case  $\phi \rightarrow \infty$  (permanence of form and latest term in (37) dominates) heuristic arguments are used to show that (i) correlation density approaches a function which is independent of spatial location and is simply a function of spatial lag (55) (ii) the mean to variance ratio in the leading edge of the wave approaches a quantity independent of the expected number of individuals in the leading edge  $q$  (60) (iii) the mean crowding index (63) includes a term which is independent of the expected density and thus does not approach zero as the expected density approaches zero.

We have analysed the form of the invasion in terms of the quantity  $\phi$  which we understand to be a statistic governing the permanence of form of the covariance wave. We now address the issue of ‘patchiness’ directly by relating it to the kurtosis of the redistribution kernel.

### 8. Patchiness and its relationship to the redistribution kernel

Patchiness in the invasion process is a qualitative phenomenon, but is typified by disparity in the spatial scales for correlations between individual locations; short-scale correlations describe relative locations of individuals within a patch, while long-scale correlations describe the relative locations of patches. In this section we restrict ourselves to kernels with finite integer moments, thus excluding, for example, the Cauchy kernel (27) which has ‘very fat’ tails. Defining  $\xi = (x + y)/\sqrt{2}$  as the coordinate describing mean location and  $\eta = (x - y)/\sqrt{2}$  as the coordinate describing spatial lag between locations, and using kurtosis (3) as a description of disparity of spatial scales, a measure of patchiness about a point  $\xi = (x + y)/\sqrt{2}$  can be defined as

$$P_t = \frac{c_t^{(4)}}{c_t^{(2)2}} \tag{64}$$

where

$$c_t^{(n)} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \eta^n c_t(\xi, \eta) d\xi d\eta}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c_t(\xi, \eta) d\xi d\eta}, \tag{65}$$

is the  $n$ th moment of the normalized covariance density function (91). We say that the overall spread is patchy if  $P_t$  is leptokurtic, in other words, if  $P_t > 3$  (64).

We now show that leptokurtic kernels (3) yield overall patchy spread. Our approach is to use the explicit representation for covariance in terms of expected density (37).

In preparation for calculating (64) we first calculate the moments of  $p_t(x)$  in terms of the kernel moments (2) for the case where there is initially a point release of a single individual at time  $t = 0$ :

$$\begin{aligned} p_{t+1}^{(n)} &= \int_{-\infty}^{\infty} x^n p_{t+1}(x) dx \\ &= R \int_{-\infty}^{\infty} p_t(z) \left( \int_{-\infty}^{\infty} (z - u)^n k(u) du \right) dz \end{aligned} \tag{66}$$

Dependence of the dispersal kernel on distance, and thus its symmetry, lead to odd moments of the kernel equaling zero. The case  $n = 0$  yields

$$p_{t+1}^{(0)} = R p_t^{(0)}, \tag{67}$$

$n = 2$  thus yields

$$p_{t+1}^{(2)} = R \left( p_t^{(2)} + k^{(2)} p_t^{(0)} \right), \tag{68}$$

and the case  $n = 4$  yields

$$p_{t+1}^{(4)} = R \left( p_t^{(4)} + 6k^{(2)} p_t^{(2)} + R^t k^{(4)} \right). \tag{69}$$

The point release of a single individual implies that  $p_0^{(0)} = 1$  and  $p_0^{(n)} = 0$  for  $n > 0$  and equation (67) implies that  $p_t^{(0)} = R^t$ . Thus the second (68) and fourth (69) moments are calculated to be

$$p_t^{(2)} = t R^t k^{(2)} \tag{70}$$

and

$$p_t^{(4)} = R^t \left( 3t(t - 1)k^{(2)2} + t k^{(4)} \right), \tag{71}$$

respectively. The kurtosis of the normalized expected density

$$\frac{p_t^{(0)} p_t^{(4)}}{p_t^{(2)2}} = 3 + \frac{1}{t} (B_2 - 3) \tag{72}$$

thus asymptotically approaches that of a Normal distribution.

Appendix F calculates a lower bound for the normalized kurtosis in the  $\eta$ -direction of each term in (37) as

$$3 + \frac{1}{2(t - s)} (B_2 - 3), \quad s = 0 \dots t - 1. \tag{73}$$

Thus a leptokurtic kernel gives leptokurtic contributions to the covariance for each term in (37). Finally Appendix F concludes that a convex combination of leptokurtic terms leads to a leptokurtic sum. Thus when the kernel is leptokurtic, the kurtosis of the normalized sum of terms in (65), (37) exceeds 3 and the overall spread is patchy (64).

Note that the lower bound (73) suggests the most leptokurtic contribution comes from the latest term in the sum ( $s = t - 1$ ) and the least leptokurtic contribution from the earliest term in the sum ( $s = 0$ ). In other words later reproduction and dispersal events appear to give the largest contribution to covariance kurtosis (64) and thus to patchiness. Our earlier analysis of permanence of form (Section 6), showing that later terms in (37) dominate when  $|\phi| > 1$ , suggests that very patchy spread requires both a leptokurtic kernel and  $|\phi| > 1$  for a substantial range of wave numbers. (See also Appendix E.)

### 9. Density-dependent formulation of stochastic model

Finally, we formulate the density-dependent versions of (5) and (6). We assume monotonic growth whose deterministic analog is given by the logistic growth function

$$u_{t+1} = Ru_t(1 - \alpha u_t), \tag{74}$$

where  $u_t$  is the density of individuals at time step  $t$ . To formulate the stochastic model, we assume (i) density-dependent effects control fecundity before dispersal, as opposed to survival after dispersal (ii) only individuals within radius  $r$  exert a density-dependent effect.

Here the expected density of individuals at  $z_1$ , conditional upon there being an individual at  $z$  is given by

$$\frac{p_t^{(2)}(z, z_1)}{p_t(z)}, \tag{75}$$

For ease of notation we define the *local* expected joint density as

$$p_t^{(2)}(z, z) = \lim_{z_1 \rightarrow z} p_t^{(2)}(z, z_1). \tag{76}$$

Averaging the expected density of neighbors over radius  $r$  yields

$$\frac{1}{2r} \int_{z-r}^{z+r} \frac{p_t^{(2)}(z, z_1)}{p_t(z)} dz_1 = \frac{p_t^{(2)}(z, z)}{p_t(z)} + \mathcal{O}(r). \tag{77}$$

To simplify the derivation we assume that interactions occur over a sufficiently small neighborhood that the  $\mathcal{O}(r)$  terms are not significant.

The equation for expected density is

$$p_{t+1}(x) = \int_{-\infty}^{\infty} R \kappa(z, x) \left( p_t(z) - \alpha p_t^{(2)}(z, z) \right) dz, \tag{78}$$

and the equation for expected joint density is

$$\begin{aligned}
 p_{t+1}^{(2)}(x, y) = & \int_{-\infty}^{\infty} \{R(R-1) + \text{Var}[\mathcal{R}]\} \kappa(z, x)\kappa(z, y) \left(p_t(z) - \alpha p_t^{(2)}(z, z)\right) dz \\
 & + \int_{-\infty}^{\infty} \int_{\substack{-\infty \\ z_2 \neq z_1}}^{\infty} R^2 \kappa(z_1, x) \kappa(z_2, y) \left(p_t^{(2)}(z_1, z_2) - \alpha p_t^{(3)}(z_1, z_2, z_2)\right) \\
 & - \alpha p_t^{(3)}(z_1, z_1, z_2) \alpha^2 p_t^{(4)}(z_1, z_1, z_2, z_2) \Big) dz_1 dz_2. \tag{79}
 \end{aligned}$$

Here the local expected higher order densities  $p_t^{(3)}(z_1, z_2, z_2)$ ,  $p_t^{(3)}(z_1, z_2, z_2)$  and  $p_t^{(4)}(z_1, z_2, z_2, z_2)$ , are defined in a manner similar to the local expected joint density (76). A more detailed derivation and analysis of this system is given in a recent paper by one of the authors [19].

Notice that even though this system describes density-dependent interactions, it remains linear. However, it is not closed — the equation for  $p_t^{(2)}(x, y)$  depends on higher order moments. The logistic form of this model means that strong density-dependent terms could possibly drive the moments negative. Bolker and Pacala [7] analyse this possibility and show that the density-dependent terms do not drive the moments negative for a wide range of parameter values.

## 10. Conclusion

In this paper there are only three model ingredients: the expected number of offspring  $R$ , the variance in that number  $\text{Var}[\mathcal{R}]$ , and the shape of the spatial redistribution kernel. When there are discrete, non-overlapping generations these three ingredients lead to integro-difference equations for the spatial moments of the population of individuals (5), (6). Under varying assumptions about the reproduction time steps and spatial scales of dispersal, these integro-difference models have equivalent formulations as partial differential equations (14), (16), integro-differential equations (19), (20), or hybrid models (22), (23). The integro-difference models can be modified to include density-dependent population regulation (Section 9). However, the cost of including these terms is the inability to ‘close’ the system of equations: lower order moments depend upon higher order moments, and vice-versa. This system cannot be decoupled. Thus a full characterization of the density-dependent interactions requires knowing all the moments of the population. Approximate ‘moment closure’ methods to decouple equations for lower-order moments from higher-order moments are given in [19].

A classical invasion problem is the asymptotic rate of spread of a population. For the integro-difference formulation (5), the rate of spread of the expected density of individuals depends crucially upon the tails of the kernel, the fatter the tails the faster the spread. Fat-tailed or very fat-tailed kernels give spread rates that are infinite asymptotically in time. This is derived and discussed in detail by [18] for both the linear case (such as (5)) and the case where there is nonlinear population density regulation.

The spatial covariance is governed by equation (6). The growth rate of the covariance depends not only upon the expected number of offspring but also upon the

variance in that number. The higher the variance, the higher the likelihood of the formation of a patch arising from spatially correlated siblings.

The quantity  $\phi$  determines asymptotic permanence of form in the covariance wave. For wave numbers with  $|\phi| > 1$  the covariance is dominated by spatially correlated individuals at points  $x$  and  $y$  with a more recent common ancestor. The structure of the leading edge of the covariance wave approaches a characteristic permanent form as the wave spreads in a single direction  $x = y$  (Figure 5b). If  $|\phi| < 1$  the covariance is dominated by spatially correlated individuals at points  $x$  and  $y$  with a more distant common ancestor. The average length scale of correlations in the leading edge of the wave grow as the covariance wave expands in both  $x$ - and  $y$ -directions (Figure 5a). The quantity  $\phi$  (41) can be loosely related to the kurtosis  $B_2$  of the dispersal kernel (3). This is done in Appendix E. The larger the growth rate  $R$ , the smaller the magnitude of  $\phi$ , and thus the lower the likelihood of the covariance wave achieving permanence of form.

For a given fixed kernel and random variable describing reproduction, the quantity  $\phi$  is frequency dependent. In other words  $\phi$  is a function of the wave numbers  $\omega_1$  and  $\omega_2$  at locations  $x$  and  $y$ . This suggests a mathematically convenient splitting of the covariance into a portion with wave numbers giving permanence of form, and the remainder which has wave numbers not giving permanence of form. When the permanence of form is seen only over high wave numbers (short spatial scales) the unchanging covariance structure will impact the short-range interactions between individuals, but may not be apparent from casual observation of the global spread of the invasion.

Limiting cases for  $\phi$  can be used to calculate variability in location of the farthest dispersing individual: when  $\phi \rightarrow 0$  the variability is low and when  $\phi \rightarrow \infty$  the variability is high. Ultimately this means that for large  $\phi$  it is possible to have a model that faithfully describes the underlying invasion process, but does not give a precise estimate for spread. I.e., the variability in spread rates between realizations within an ensemble is large.

Patchiness of the invasion can be understood as a disparity in length scales of correlations: high correlation at short and long space scales, as opposed to intermediate length scales. The patchiness can be analysed in terms of the profile of the covariance wave as measured in the  $\eta = (x - y)/\sqrt{2}$ -direction. Here a relevant statistic is the kurtosis of the covariance wave in the  $\eta$ -direction after averaging in the perpendicular  $\xi = (x + y)/\sqrt{2}$ -direction. This kurtosis can be related to kurtosis of the dispersal kernel, demonstrating that leptokurtic kernels give rise to patchy spread. A more precise definition of patchiness would specify the location in the wave front. I.e., averaging in the  $\xi$ -direction would be replaced by holding  $\xi$  constant, and an argument could be made for using the spatial correlation function  $\rho_t$  instead of the spatial covariance function  $c_t$  in the calculation. These are directions for further work.

Even though the model analyzed here is linear, the possibility that for  $\phi \rightarrow \infty$  crowding approaches a constant nonzero level, even when the expected density approaches zero, means that for kernels with large  $\phi$  covariance terms dominate interactions at low densities. Such behavior would, for example, be expected in focal plant epidemics [27] such as potato blight [21], where very local patches of

infestation or hot spots are separated spatially from other hot spots and are only connected through rare dispersal events in which a spore escapes the canopy. Here density dependent effects arising through the second order terms (see (78), (79)) may be evident even when the expected density (averaged over the ensemble) is low. In this case the deterministic nonlinear formulation will fail to reflect the biology (but see [21] for an approximate way of dealing with such multiple dispersal mechanisms).

What little is known about stochastic spatio-temporal models with population regulation through density dependence suggests that this more realistic case may be different from linear stochastic models in several respects. Spread rates are lower [19,22,24], the qualitative conditions for finite velocity may differ [22,24] and the models can exhibit different types of spatial patchiness [11,30] Furthermore, the exponential growth exhibited by linear models allows the variability in population size to be higher and the variability of the position of the farthest dispersing individual may be lower than expected for a density-dependent population [25].

This paper has focused upon the role of stochasticity in determining the correlation structure and patchiness of an invasion process. In doing so we have assumed that the environment is spatially homogeneous — an assumption which is likely to be violated in most biological invasions. For example the westward spread of the house finch population (Figure 1b) was faster up river valleys which provided abundant resources [28]. Likewise our assumption of the absence of density-dependent population regulation is unlikely to be defensible for most populations. While spatial heterogeneity, density-dependence and other factors undoubtedly play a role in invasion processes, we justify their exclusion on the grounds that we have chosen the simplest possible model that incorporates the stochastic aspects which are the focus of this paper.

*Acknowledgements.* Equations (5) and (6) were derived jointly with Mike Neubert. This work was initiated during a visit to Simon Levin's lab at Princeton. Thanks to him for his encouragement and insight. Thanks to J. A. J. Metz for letting us know about the early history of work on covariance models. Thanks to Fred Adler, George Hurtt, Davar Koshnevisan, Rich McLaughlin, Hans Othmer, members of the Pacala and Levin labs and two anonymous referees for ideas and comments. Lastly, thanks to the JMB editor Odo Diekmann for numerous helpful comments on the manuscript.

## A. Spatially distributed probability densities

In this appendix we derive formulae relating the first two spatially distributed probability densities (expected density and expected joint density) to the expected number of individuals in a region and the expected number of pairwise interactions between individuals. Using these probability densities we define the spatial covariance and the correlation density function.

Neighborhoods of area  $\epsilon$ ,  $0 < \epsilon \ll 1$ , about the points  $x$  and  $y$  are given by  $\Omega_x^\epsilon$  and  $\Omega_y^\epsilon$  respectively. It is assumed that no more than one individual can occupy any given point in space. Mathematically we assume that for  $\epsilon$  sufficiently small there is no more than a single individual in  $\Omega_x^\epsilon$  for each  $x$  in the entire domain of interest



$\Omega$ . Thus, at any point in time, for  $\epsilon$  sufficiently small the number of individuals in any  $\Omega_x^\epsilon$  is a Bernoulli random variable with mean

$$P_x^\epsilon = E(n(\Omega_x^\epsilon)) = \Pr \{n(\Omega_x^\epsilon) = 1\}, \tag{80}$$

variance

$$\sigma_x^{\epsilon 2} = P_x^\epsilon(1 - P_x^\epsilon), \tag{81}$$

mean product

$$P_{xy}^\epsilon = E(n(\Omega_x^\epsilon)n(\Omega_y^\epsilon)) = \Pr \{n(\Omega_x^\epsilon) = 1 \cap n(\Omega_y^\epsilon) = 1\}, \tag{82}$$

and spatial correlation

$$R_{xy}^\epsilon = \frac{P_{xy}^\epsilon - P_x^\epsilon P_y^\epsilon}{\sigma_x^\epsilon \sigma_y^\epsilon}. \tag{83}$$

For regions  $\Omega_1$  and  $\Omega_2$  contained in  $\Omega$  we define  $n(\Omega_1)$  as the number of individuals in  $\Omega_1$ , and  $n(\Omega_2)$  as the number of individuals in  $\Omega_2$ . Taking expectations over an ensemble of stochastically identical processes yields

$$E(n(\Omega_1)) = \int_{\Omega_1} p(x) dx \tag{84}$$

$$E(n(\Omega_1)n(\Omega_2)) = \int_{\Omega_2} \int_{\substack{\Omega_1 \\ y \neq x}} p^{(2)}(x, y) dx dy + \int_{\Omega_1 \cap \Omega_2} p(x) dx, \tag{85}$$

where  $p(x)$  is the expected density of individuals and  $p^{(2)}(x, y)$  is the expected joint density of individuals, defined for  $x \neq y$  [31]. The first term on the right hand side of (85) describes the contribution from pairs of distinct individuals, one individual lying in  $\Omega_1$  and the other in  $\Omega_2$  (valid for  $x \neq y$ ), and the second term describes the contribution from single individuals lying in the overlap between  $\Omega_1$  and  $\Omega_2$ .

To relate our definitions of  $p(x)$  and  $p^{(2)}(x, y)$  to the probabilities of finding single individuals in arbitrarily small disks we observe that for  $\epsilon$  sufficiently small and disjoint  $\Omega_x^\epsilon$  and  $\Omega_y^\epsilon$ , equations (80) and (84) give us

$$P_x^\epsilon = E(n(\Omega_x^\epsilon)) = \int_{\Omega_x^\epsilon} p(\xi) d\xi, \tag{86}$$

and equations (82) and (85) give us

$$P_{xy}^\epsilon = E(n(\Omega_x^\epsilon)n(\Omega_y^\epsilon)) = \int_{\Omega_y^\epsilon} \int_{\Omega_x^\epsilon} p^{(2)}(\xi, \eta) d\xi d\eta \tag{87}$$

and, as  $\epsilon \rightarrow 0$ , we have

$$p(x) = \lim_{\epsilon \rightarrow 0} \frac{P_x^\epsilon}{\epsilon}, \quad p^{(2)}(x, y) = \lim_{\epsilon \rightarrow 0} \frac{P_{xy}^\epsilon}{\epsilon^2}, \quad x \neq y. \tag{88}$$

We use definition (88) when deriving equations for the spatially distributed moments  $p(x)$  and  $p^{(2)}(x, y)$  from first principles.

We define a spatial covariance density function as

$$c(x, y) = p^{(2)}(x, y) - p(x)p(y), \quad x \neq y \tag{89}$$

[5]. The covariance can be considered as a scaled measure of the correlation between locations of individuals. To see this we use (81), (83) and (88) to define a correlation density function

$$\rho(x, y) = \lim_{\epsilon \rightarrow 0} \frac{R_{xy}^\epsilon}{\epsilon} = \frac{c(x, y)}{\sqrt{p(x)p(y)}}. \tag{90}$$

Note that, unlike a correlation coefficient,  $\rho$  is not constrained to lie between  $\pm 1$ . If the locations of individuals are uncorrelated then the likelihood of observing an individual in the neighborhood of  $x$  is independent of the likelihood of observing an individual in the neighborhood of  $y$  and thus  $c(x, y) = p^{(2)}(x, y) - p(x)p(y) \equiv 0$ . Positive correlation yields  $c(x, y) > 0$  and negative correlation yields  $c(x, y) < 0$ .

The expected number of individuals in a space interval  $\Omega_1$  can be calculated using (84), and the covariance in the numbers of individuals over two intervals  $\Omega_1$  and  $\Omega_2$  can be calculated from (84)–(85), (89) as

$$\begin{aligned} C(\Omega_1, \Omega_2) &= E(n(\Omega_1)n(\Omega_2)) - E(n(\Omega_1))E(n(\Omega_2)) \\ &= \int_{\Omega_2} \int_{\substack{\Omega_1 \\ y \neq x}} c(x, y) dx dy + \int_{\Omega_1 \cap \Omega_2} p(x) dx. \end{aligned} \tag{91}$$

If  $\Omega_1 = \Omega_2$  then (91) yields the variance in the number of individuals. By extending the definition of (89) to include the line  $x = y$  by

$$\tilde{c}(x, y) = \begin{cases} c(x, y) & x \neq y \\ \delta(x - y)p(x) & x = y \end{cases}, \tag{92}$$

the variance in the number of individuals in  $\Omega_1$  is given by

$$V(\Omega_1) = \int_{\Omega_1} \int_{\Omega_1} \tilde{c}(x, y) dx dy. \tag{93}$$

An intuitive reason can be given for the appearance of the delta function in the extended definition of the spatial covariance density function (92): if the disks  $\Omega_x^\epsilon$  and  $\Omega_y^\epsilon$  overlap as  $\epsilon \rightarrow 0$  then  $P_{xy}^\epsilon = P_x^\epsilon$  and  $p^{(2)}(x, y) \rightarrow p(x)/\epsilon$  in (88).

**B. Stochastic integro-difference model**

*B.1. Deterministic demography*

We start by assuming that each individual has precisely  $R$  (positive integer) offspring per unit time. We denote birth of an individual in  $\Omega$  and dispersal to  $\Omega_x^\epsilon$  in time step  $t$  as  $b_t(\Omega \rightarrow \Omega_x^\epsilon)$ . Thus

$$\Pr\{n_{t+1}(\Omega_x^\epsilon) = 1\} = \Pr\{b_t(\Omega \rightarrow \Omega_x^\epsilon) = 1\} \tag{94}$$

is rewritten as

$$\int_{\Omega_x^\epsilon} p_{t+1}(\xi) d\xi = \int_{\Omega_x^\epsilon} \int_{\Omega} R \kappa(z, \xi) p_t(z) dz d\xi \tag{95}$$

As  $\epsilon \rightarrow 0$ ,

$$p_{t+1}(x) = \int_{\Omega} R \kappa(z, x) p_t(z) dz \tag{96}$$

The joint probability function is governed by

$$\begin{aligned} & \Pr\{n_{t+1}(\Omega_x^\epsilon) = 1 \cap n_{t+1}(\Omega_y^\epsilon) = 1\} = \\ & \Pr\{b_t(\Omega \rightarrow \Omega_x^\epsilon) = 1 \cap b_t(\Omega \rightarrow \Omega_y^\epsilon) = 1\} \end{aligned} \tag{97}$$

An individual at point  $z$  which produces  $R \geq 0$  indistinguishable offspring has  $R(R - 1)$  ways to distribute exactly one offspring to each of  $\Omega_x^\epsilon$  and  $\Omega_y^\epsilon$ , whereas two individuals, one at point  $z_1$  and one at point  $z_2$  each of which produces  $R$  offspring have  $R^2$  ways to distribute exactly one offspring from  $z_1$  to  $\Omega_x^\epsilon$  and one offspring from  $z_2$  to  $\Omega_y^\epsilon$ . Thus

$$\begin{aligned} & \int_{\Omega_x^\epsilon} \int_{\Omega_y^\epsilon} p_{t+1}^{(2)}(\xi, \eta) d\eta d\xi \\ &= \int_{\Omega_x^\epsilon} \int_{\Omega_y^\epsilon} \int_{\Omega} R(R - 1) \kappa(z, \xi) \kappa(z, \eta) p_t(z) dz d\eta d\xi \\ &+ \int_{\Omega_x^\epsilon} \int_{\Omega_y^\epsilon} \int_{\substack{\Omega \\ z_2 \neq z_1}} R^2 \kappa(z_1, \xi) \kappa(z_2, \eta) p_t^{(2)}(z_1, z_2) dz_1 dz_2 d\eta d\xi \end{aligned} \tag{98}$$

So as  $\epsilon \rightarrow 0$ ,

$$\begin{aligned} p_{t+1}^{(2)}(x, y) &= \int_{\Omega} R(R - 1) \kappa(z, x) \kappa(z, y) p_t(z) dz \\ &+ \int_{\substack{\Omega \\ z_2 \neq z_1}} R^2 \kappa(z_1, x) \kappa(z_2, y) p_t^{(2)}(z_1, z_2) dz_1 dz_2 \end{aligned} \tag{99}$$

Note that

$$p_{t+1}(x)p_{t+1}(y) = \int_{\Omega} \int_{\Omega} R^2 \kappa(z_1, x) \kappa(z_2, y) p_t(z_1) p_t(z_2) dz_1 dz_2 \tag{100}$$

Using the definition of spatial covariance (89) we observe that

$$\begin{aligned} c_{t+1}(x, y) &= \int_{\Omega} R(R - 1) \kappa(z, x) \kappa(z, y) p_t(z) dz \\ &+ \int_{\substack{\Omega \\ z_2 \neq z_1}} R^2 \kappa(z_1, x) \kappa(z_2, y) c_t(z_1, z_2) dz_1 dz_2 \end{aligned} \tag{101}$$

*B.2. Stochastic demography*

Now we assume that  $\mathcal{R}$  is a random variable representing each individual's reproductive output. We assume that  $\mathcal{R}$  has a given probability density function  $\zeta$ , which does not depend upon spatial location, such that

$$\Pr\{\mathcal{R} = r\} = \zeta(r). \tag{102}$$

Application of the equation describing the probability of having an individual in a disk of radius  $\epsilon$  (94) yields

$$\int_{\Omega_x^\epsilon} p_{t+1}(\xi) d\xi = \int_{\Omega_x^\epsilon} \int_{\Omega} \sum_{r=0}^{\infty} r \zeta(r) \kappa(z, \xi) p_t(z) dz d\xi. \tag{103}$$

Denoting the expected value of  $\mathcal{R}$  as

$$E[\mathcal{R}] = R \geq 0, \tag{104}$$

we again obtain equation (5) for the expected density of individuals. The joint probability function, governed by (97) satisfies

$$\begin{aligned} & \int_{\Omega_x^\epsilon} \int_{\Omega_y^\epsilon} p_{t+1}^{(2)}(\xi, \eta) d\eta d\xi \\ &= \int_{\Omega_x^\epsilon} \int_{\Omega_y^\epsilon} \int_{\Omega} \sum_{r=0}^{\infty} r(r-1) \zeta(r) p_t(z) \kappa(z, \xi) \kappa(z, \eta) dz d\eta d\xi \\ &+ \int_{\Omega_x^\epsilon} \int_{\Omega_y^\epsilon} \int_{\Omega} \int_{z_2 \neq z_1} \left[ \sum_{r=0}^{\infty} r \zeta(r) \right]^2 p_t(z_1, z_2) \kappa(z_1, \xi) \kappa(z_2, \eta) dz_1 dz_2 d\eta d\xi \end{aligned} \tag{105}$$

Rewriting the sums in (105) in terms of the mean and variance for  $\mathcal{R}$  and employing (100) shows stochastic demography modifies the covariance density equation (101) by adding a variance term which simply increases the contribution that expected density makes to the spatial covariance:

$$\begin{aligned} c_{t+1}(x, y) &= \int_{\Omega} \{R(R-1) + \text{Var}[\mathcal{R}]\} p_t(z) \kappa(z, x) \kappa(z, y) dz \\ &+ \int_{\Omega} \int_{z_2 \neq z_1} R^2 c_t(z_1, z_2) \kappa(z_1, x) \kappa(z_2, y) dz_1 dz_2. \end{aligned} \tag{106}$$

**C. Invasion speed for exponentially bounded kernels**

For an exponentially bounded dispersal kernel  $k(x)$ , and compact initial data, the invasion speed  $c$  at which  $x_t$  moves is determined by the double root with respect to  $s$  to

$$\exp(sc) = R \int_{-\infty}^{\infty} k(u) \exp(su) du = RM_k(s), \tag{107}$$

where  $M_k(s)$  is the moment generating function for  $k$ . Weinberger [37, 38], proved that  $c$  is, in fact, the minimum traveling wave speed for versions of (5) with nonlinear dynamics. A heuristic approach to understanding (107) comes from assuming that the solution for the leading edge of the linear system (5) decays exponentially in space [18, 22],

$$p_t(x) = A \exp(-sx) \tag{108}$$

and moves to the right with velocity  $c > 0$  so that  $p_{t+1}(x) = p_t(x - c)$ . Substitution into (5) yields a dispersion relation between the value of the speed  $c$  and the steepness of the wave  $s$  (107). The minimum value of  $c$  yielding a non-negative solution (108) is found when the value of  $s$  in (107) switches from complex to real and thus (107) has a double root with respect to  $s$ .

**D. Calculation of  $\phi$  for certain kernels**

In general  $|\phi|$  may vary above and below 1, depending upon the relative contributions of higher wave numbers  $\omega_1 + \omega_2$  and lower wave numbers  $\omega_1$  and  $\omega_2$  in the kernel.

Cases where  $\phi$  is independent of wave number are found when the kernel is a Cauchy distribution (27) which has Fourier Transform

$$\hat{k}(\omega) = \frac{\exp(-\beta|\omega|)}{2\pi} \tag{109}$$

or a delta distribution, which has Fourier Transform  $\hat{k}(\omega) = 1/2\pi$ . For each of these (41) yields  $\phi = 1/R$ .

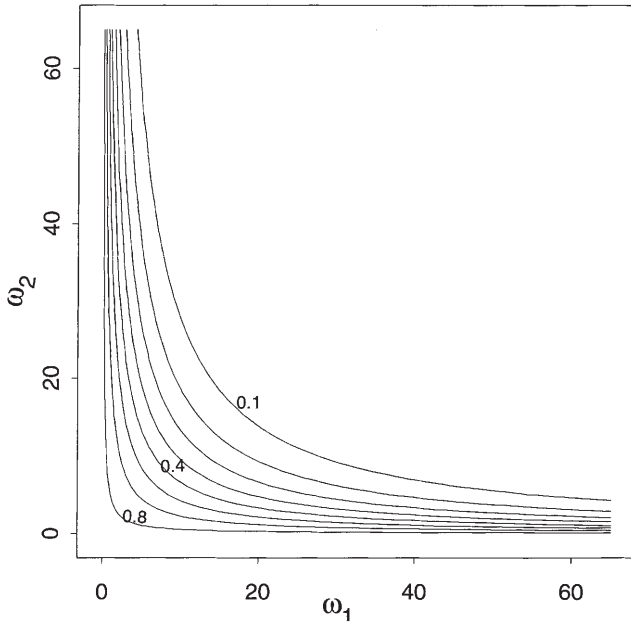
Equations (29) and (41) yield  $\phi(0, \omega_2) = \phi(\omega_1, 0) = 1/R$  and thus a growth rate exceeding unity implies that  $|\phi| < 1$  for small wave numbers.

When the growth rate  $R$  exceeds unity, the Normal kernel (11) with transform

$$\hat{k}(\omega) = \frac{\exp(-D\omega^2)}{2\pi} \tag{110}$$

yields  $|\phi| < 1$  for all  $\omega_1$  and  $\omega_2$  (Figure 9a), whereas a linear composition of Normal kernels

$$k(x) = \frac{p}{\sqrt{4D_1\pi}} \exp\left(-\frac{x^2}{4D_1}\right) + \frac{1-p}{\sqrt{4D_2\pi}} \exp\left(-\frac{x^2}{4D_2}\right), \quad D_1, D_2 > 0, \quad 0 \leq p \leq 1 \tag{111}$$



**Fig. 9a.** Plots of  $\phi(\omega_1, \omega_2)$  for (a) Normal kernel (11) with  $D = 0.01$  and  $R = 1.2$  (b) composite Normal kernel (111) with  $D_1 = 0.01, D_2 = 0.002$  and  $p = 0.99$  and  $R = 1.2$  (c) composite Laplace kernel (113) with  $\alpha_1 = 10, \alpha_2 = 1$  and  $p = 0.99$  and  $R = 1.2$ . Shading shows regions in wave number space where  $\phi > 1$ .

can lead to  $|\phi| > 1$  for a bounded region in  $\omega_1 - \omega_2$  space (Figure 9b). By way of contrast, the Laplace kernel (24) with transform

$$\hat{k}(\omega) = \frac{\alpha^2}{2\pi(\alpha^2 + \omega^2)} \tag{112}$$

has largest  $\phi$  for the highest wave numbers. A linear combination of Laplace kernels leads to

$$k(x) = p \frac{\alpha_1}{2} \exp(-\alpha_1|x|) + (1 - p) \frac{\alpha_2}{2} \exp(-\alpha_2|x|), \quad \alpha_1, \alpha_2 > 0, \quad 0 \leq p \leq 1, \tag{113}$$

which lead  $|\phi| > 1$  for sufficiently large  $\omega_1$  and  $\omega_2$  (Figure 9c).

However, given a small enough  $R < 1$  a kernel such as (113) will yield  $|\phi| > 1$  for all wave numbers (Figure 6).

**E. Relating  $\phi$  to kernel kurtosis**

In this appendix we analyse the asymptotic relationship between  $\phi$  (41) and kurtosis (3) for a symmetric kernel. To simplify matters we consider  $\phi(\omega_1, \omega_2)$  for  $R = 1$  and  $\omega_1 = \omega_2 = \omega$ . These simplifications imply that, on average, individuals simply replace themselves through reproduction and we restrict ourselves to studying correlations with identical frequency components at different locations.

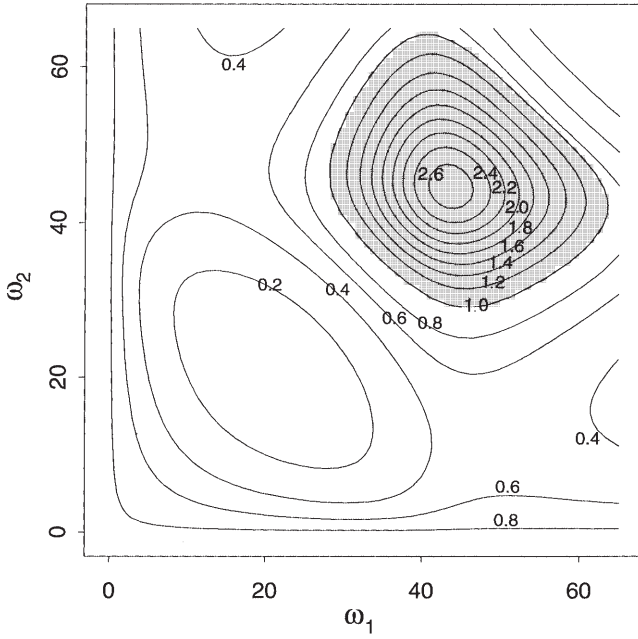


Fig. 9b.

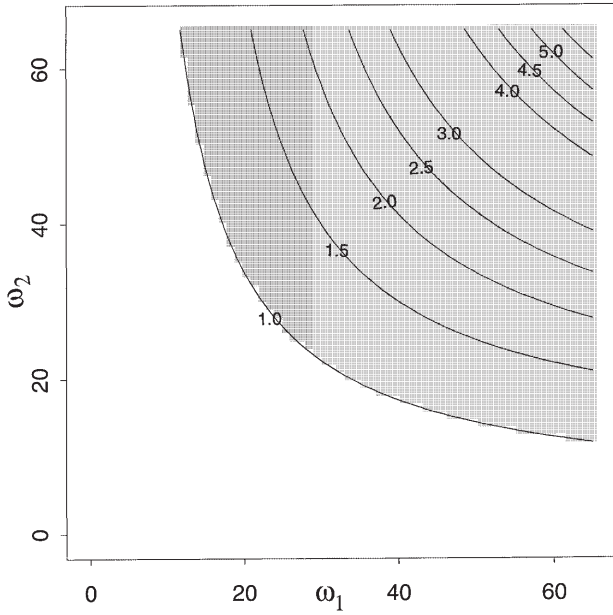


Fig. 9c.

Under the assumption  $R = 1$ , the constraint  $\phi(\omega, \omega) > 1$  can be rewritten as

$$\frac{\hat{k}(2\omega)}{2\pi} - \hat{k}^2(\omega) > 0. \tag{114}$$

We normalize the wavenumber  $\omega$  by the standard deviation of the kernel,  $k^{(2)}$  (2) to give  $\tilde{\omega} = \omega\sqrt{k^{(2)}}$ . Using a series expansion for the exponential in the Fourier transform (29) equation (114) is rewritten as a polynomial inequality

$$\begin{aligned} & \left(\frac{1}{2\pi}\right) \left(-\tilde{\omega}^2 + \frac{1}{4} \left(\frac{7}{12} B_2 - 1\right) \tilde{\omega}^4 + \frac{1}{24} \left(B_2 - \frac{31}{15} B_3\right) \tilde{\omega}^6 \right. \\ & \left. + \frac{1}{720} \left(\frac{254}{56} B_4 - B_3 - \frac{15}{12} B_2^2\right) \tilde{\omega}^8 + \text{h.o.t.}\right) > 0, \end{aligned} \tag{115}$$

where

$$B_n = \frac{k^{(2n)}}{k^{(2)^n}}. \tag{116}$$

Coefficients of the higher order terms are polynomials in the  $B_n$ s. When  $\tilde{\omega}$  is small (i.e., when wavelengths are long relative to the standard deviation of the kernel  $k(x)$ ), the earlier terms in the polynomial dominate, provided the  $B_n$ s are finite. In particular, note that a kurtosis  $B_2 > 12/7$  will ensure that the second term in the expansion is positive and a sufficiently large kurtosis and sufficiently small wavenumber  $\tilde{\omega}$  will ensure that the inequality (115) is satisfied.

**F. Calculation of patchiness as determined by kernel kurtosis**

Defining  $\xi = (x + y)/\sqrt{2}$  as the coordinate describing mean location and  $\eta = (x - y)/\sqrt{2}$  as the coordinate describing spatial lag between locations, we use (37) and (72) to calculate the normalized kurtosis of the two-dimensional redistribution kernel  $p_{t-s}(x)p_{t-s}(y)$  in the  $\eta$ -direction as

$$\frac{p_{t-s}^{(0)2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{x-y}{\sqrt{2}}\right)^4 p_{t-s}(x)p_{t-s}(y) dx dy}{\left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{x-y}{\sqrt{2}}\right)^2 p_{t-s}(x)p_{t-s}(y) dx dy\right)^2} = 3 + \frac{1}{2(t-s)} (B_2 - 3). \tag{117}$$

The inequality

$$\left(\int_{-\infty}^{\infty} a(z)b(z) dz\right)^2 \leq \int_{-\infty}^{\infty} a(z)b^2(z) dz, \tag{118}$$

where

$$\int_{-\infty}^{\infty} a(z) dz = 1, \quad a(z) \geq 0,$$

and  $a(z)$  and  $b(z)$  are square integrable can be derived directly from Schwartz's inequality for infinite integrals [36]. Using an inequality of this form with  $a(z) =$



$p_s(z)/p_s^{(0)}$  and integrating leads to a lower bound for the normalized kurtosis of each term in (37)

$$\begin{aligned} & \frac{p_{t-s}^{(0)2} p_s^{(0)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{x-y}{\sqrt{2}}\right)^4 \int_{-\infty}^{\infty} p_s(z) p_{t-s}(x-z) p_{t-s}(y-z) dz dx dy}{\left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{x-y}{\sqrt{2}}\right)^2 \int_{-\infty}^{\infty} p_s(z) p_{t-s}(x-z) p_{t-s}(y-z) dz dx dy\right)^2} \\ & \geq \frac{p_{t-s}^{(0)2} p_s^{(0)} \int_{-\infty}^{\infty} p_s(z) \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{x-y}{\sqrt{2}}\right)^4 p_{t-s}(x-z) p_{t-s}(y-z) dx dy\right) dz}{\int_{-\infty}^{\infty} p_s(z) \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{x-y}{\sqrt{2}}\right)^2 p_{t-s}(x-z) p_{t-s}(y-z) dx dy\right)^2 dz} \\ & = 3 + \frac{1}{2(t-s)} (B_2 - 3). \end{aligned} \tag{119}$$

Finally, application of the discrete form of (118),

$$\left(\sum_i a_i b_i\right)^2 \leq \sum_i a_i b_i^2, \tag{120}$$

where

$$\sum_i a_i = 1, \quad a_i \geq 0,$$

permits us to conclude that a convex combination of leptokurtic kernels gives a leptokurtic sum and thus the kurtosis of the normalized sum of terms in (65), with  $c_t(x, y)$  given by (37) exceeds 3, and the overall spread is patchy.

**References**

1. Andow, D.A., Kareiva, P.M., Levin, S.A., Okubo, A.: Spread of invading organisms. *Landscape Ecology*, **4**, 177–188 (1990)
2. Ball, F.: Coupling methods in epidemic theory. In: Mollison, D. editor, *Epidemic Models: Their Structure and Relation to Data*, pages 34–52. Cambridge University Press (1995)
3. Bartlett, M.S.: Deterministic and stochastic models for recurrent epidemics. *Proc. 6th Berkely Symp. on Math. Statist. and Prob.*, **4**, 81–109 (1956)
4. Bartlett, M.S.: *Stochastic Population Models in Ecology and Epidemiology*. Methuen, London (1960)
5. Bartlett, M.S.: *An Introduction to Stochastic Processes with Special Reference to Methods and Applications*. Cambridge University Press, Cambridge (1966)
6. Biggins, J.D.: The asymptotic shape of the branching random walk. *Adv. Appl. Prob.*, **10(1)**, 62–84 (1977)
7. Bolker, B., Pacala, S.: Using moment equations to understand stochastically driven pattern formation in ecological systems. *Theor. Pop. Biol.*, (in press) (1997)
8. Broadbent, S.R., Kendall, D.G.: The random walk of *Trichostrongylus Retortaeformis*. *Biometrics*, pages 460–466 (1953)
9. Daniels, H.E.: The advancing wave in a spatial birth process. *J. Appl. Prob.*, **14(4)**, 689–701 (1977)

10. Durrett, R.: Spatial epidemic models. In Mollison, D. editor, *Epidemic Models: Their Structure and Relation to Data*, pages 187–201. Cambridge University Press (1995)
11. Durrett, R., Levin, S.A.: The importance of being discrete (and spatial). *J. Theor. Biol.*, **46**, 363–394 (1994)
12. Fisher, R.A.: The wave of advance of advantageous genes. *Ann. Eugen. London.*, **37**, 355–369 (1937)
13. Fitt, B.D.L., Gregory, P.H., Todd, A.D., McCarney, H.A., MacDonald, O.C.: Spore dispersal and plant disease gradient; a comparison between two empirical models. *Journal of Phytopathology*, **118**, 227–242 (1987)
14. Harris, T.: A correlation inequality for Markov processes in partially ordered state spaces. *Ann. Prob.*, **5**, 451–454 (1977)
15. Hengeveld, R.: *Dynamics of Biological Invasions*. Chapman and Hall, London (1989)
16. Van Kirk, R.W.: *Integrodifference Models of Biological Growth and Dispersal*. PhD thesis, University of Utah, Department of Mathematics (1995)
17. Kolmogorov, A., Petrovsky, P., Piscounov, N.I.: Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application a un problème biologique. *Bull. Moscow Univ. Math. Mech*, **1(6)**, 1–26 (1937)
18. Kot, M., Lewis, M.A., van den Driessche, P.: Dispersal data and the spread of invading organisms. *Ecology* (in press) (1996)
19. Lewis, M.A.: Spread rate for a nonlinear stochastic invasion. *J. Math. Biol.*, **41** (same issue), 430–454 (2000) (DOI: 10.1007/s002850000022)
20. Monte Lloyd: Mean crowding. *Journal of Animal Ecology*, **30**, 1–30 (1967)
21. Metz, J.A.J., van den Bosch, F.: Velocities of epidemic spread. In Mollison, D. editor, *Epidemic Models: Their Structure and Relation to Data*, pages 150–186. Cambridge University Press (1995)
22. Mollison, D.: The rate of spatial propagation of simple epidemics. *Proc. 6th Berkely Symp. on Math. Statist. and Prob.*, **3**, 579–614 (1972)
23. Mollison, D.: Modeling biological invasions: chance, explanation, prediction. *Phil. Trans. Roy. Soc. Lond. B*, **B14**, 173–158 (1986)
24. Mollison, D.: Dependence of epidemic and population velocities on basic parameters. *Mathematical Biosciences*, **107**, 255–287 (1991)
25. Mollison, D., Daniels, H.: The deterministic simple epidemic unmasked. *Mathematical Biosciences*, **117**, 147–153 (1993)
26. Denis Mollison.: Spatial contact models for ecological and epidemic spread. *J. R. Statist. Soc. B*, **39(3)**, 283–326 (1977)
27. Michael Moody, E., Richard Mack, N.: Controlling the spread of plant invasions: the importance of nascent foci. *Journal of Applied Ecology*, **25**, 1009–1021 (1988)
28. Paul Munding, C., Sylvia Hope.: Expansion of the winter range of the house finch: 1947–79. *American Birds*, **36(4)**, 347–353 (1982)
29. Neubert, M., Kot, M., Lewis, M.A.: Dispersal and pattern formation in a discrete-time predator-prey model. *Theor. Pop. Biol.*, **48**, 7–43 (1995)
30. Rand, D.A., Keeling, M., Wilson, H.B.: Invasion, stability and evolution to criticality in spatially extended, artificial host-pathogen ecologies. *Proc. Roy. Soc. Lond. B*, **255**, 55–63 (1995)
31. Ripley, B.D.: *Spatial Statistics*. Wiley, New York (1981)
32. Shigesada, N., Kawasaki, K., Takeda, Y.: Modeling stratified diffusion in biological invasions. *American Naturalist*, **146**, 229–251 (1995)
33. Skellam, J.G.: Random dispersal in theoretical populations. *Biometrika*, **38**, 196–218 (1951)
34. Taylor, R.A.J.: The relationship between density and distance of dispersing insects. *Ecological Entomology*, **3**, 63–70 (1978)

35. van den Bosch, F., Metz, J.A.J., Diekmann, O.: The velocity of spatial population expansion. *J. Math. Biol.*, **28**, 529–565 (1990)
36. Weinberger, H.F.: *A First Course in Partial Differential Equations with Complex Variables and Transform Methods*. Blaisdell Publishing, New York (1965)
37. Weinberger, H.F.: Asymptotic behavior of a model in population genetics. In Chadam, J.M. editor, *Lecture Notes in Mathematics No. 648: Nonlinear Partial Differential Equations and Applications*, pages 47–96. *Proceedings Indiana 1967–1977* (1978)
38. Weinberger, H.F.: Long-time behavior of a class of biological models. *SIAM J. Appl. Math.*, **13(3)**, 353–396 (1982)
39. Williams, E.J.: The distribution of larvae of randomly moving insects. *Australian Journal of Biological Sciences*, **14**, 598–604 (1961)
40. Willson, M.F.: Dispersal mode, seed shadows and colonization patterns. *Vegetatio*, **107/108**, 261–280 (1993)