Available online at www.sciencedirect.com
SCIENCE DIRECT®
Theoretical Population Biology 67 (2005) 241–256
Theoretical Population Biology
www.elsevier.com/locate/ytpbi

Families of discrete kernels for modeling dispersal

Peter Chesson[a,*], Charlotte T. Lee[b,1]

[a]Section of Evolution and Ecology, Division of Biological Sciences, University of California, Davis, CA 95616, USA
[b]Department of Environmental Science and Policy, and Graduate Group in Ecology, University of California, Davis, CA 95616, USA

Received 15 June 2004
Available online 1 April 2005

Abstract

Integer lattices are important theoretical landscapes for studying the consequences of dispersal and spatial population structure, but convenient dispersal kernels able to represent important features of dispersal in nature have been lacking for lattices. Because leptokurtic (centrally peaked and long-tailed) kernels are common in nature and have important effects in models, of particular interest are families of dispersal kernels in which the degree of leptokurtosis can be varied parametrically. Here we develop families of kernels on integer lattices with several important properties. The degree of leptokurtosis can be varied parametrically from near 0 (the Gaussian value) to infinity. These kernels are all asymptotically radially symmetric. (Exact radial symmetry is impossible on lattices except in one dimension.) They have separate parameters for shape and scale, and their lower order moments and Fourier transforms are given by simple formulae. In most cases, the kernel families that we develop are closed under convolution so that multiple steps of a kernel remain within the same family. Included in these families are kernels with asymptotic power function tails, which have provided good fits to some observations from nature. These kernel families are constructed by randomizing convolutions of stepping-stone kernels and have interpretations in terms of population heterogeneity and heterogeneous physical processes.

Keywords: Dispersal kernel; Leptokurtosis; Discrete lattice; Negative binomial; Sichel distribution; Stable distribution

1. Introduction

Spatially explicit models are of increasing importance in population biology, especially in ecology where such models are relatively recent (Dieckmann et al., 2000). Key questions concern the patterning of organisms in space (Levin, 1992), the relationship of the patterning of organisms to patterns in the environment (Roughgarden, 1978), and the rate and pattern of spread of a species or allele across a landscape (Kinezaki et al., 2003; Lewis and Pacala, 2000). How organisms become patterned in space is of intrinsic interest (Klausmeier, 1999; Levin, 1992), but such patterns may also be used to draw conclusions about underlying processes. For example, spatial patterns may help distinguish different mechanisms of species coexistence (Bolker et al., 2003), or indicate dispersal distances (Ouborg et al., 1999). Spatial patterns potentially affect other processes. They may change the nature and outcomes of species interactions (Kareiva and Wennergren, 1995), potentially promoting coexistence of competitors (Bolker and Pacala, 1999; Hassell et al., 1994; Murrell and Law, 2003; Snyder and Chesson, 2003) or stabilizing host–parasitoid and predator–prey relationships (Briggs and Hoopes, 2004; Comins et al., 1992; De Roos et al., 1998).

Spatially explicit models inevitably require the use of functions called kernels, which describe dispersal in space (Snyder and Chesson, 2003) or represent interactions between individuals as functions of their distance apart (Bolker and Pacala, 1999; Snyder and Chesson, 2004). Our concern here is with dispersal kernels. In discrete time, a dispersal kernel defines for each spatial

*Corresponding author. Fax: +1 530 752 1449.
E-mail address: PLChesson@UCDavis.edu (P. Chesson).
[1]Present address: Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA.

0040-5809/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.tpb.2004.12.002

location the probability distribution of places dispersed from that location in one unit of time. The variance of this distribution defines the spatial scale of dispersal, but also important is kurtosis, which reflects the shape of the distribution. In nature, leptokurtosis is common, that is, dispersal kernels are often observed to have a sharp peak at the point of origin and a long tail. They are thus far from the Gaussian (normal) distributions often used in modeling. Of most importance, leptokurtosis has been shown to greatly increase the rate of spread of an invading organism or allele, and has been hypothesized to explain the faster than expected rates of spread sometimes found in nature (Cain et al., 1998; Kot et al., 1996; Lewis and Pacala, 2000). Moreover, recent modeling studies show that dispersal kurtosis may have important repercussions for the dynamics of spatial host–parasitoid interactions (Wilson et al., 1999) and disease (Brown and Bolker, 2004). Finally, leptokurtic models are of value in estimating dispersal characteristics from field data, giving greater precision when kurtosis is appropriately modeled (Clark et al., 1999). It is only recently, however, that suitable families of dispersal kernels, allowing broad ranges of kurtosis, have been in use. Hence, investigations of the full realistic range of kurtosis, including the extremes sometimes observed, are just beginning. This article facilitates this endeavor by providing new families of dispersal kernels for discrete space and time that allow exploration of the effects of kurtosis ranging from the Gaussian value to infinity.

In spatially explicit models, space can be represented as discrete or continuous, but integer lattices in one or two dimensions have advantages for many problems (Snyder and Chesson, 2003; Thomson and Ellner, 2003). However, models of dispersal on integer lattices are not well developed. The earliest integer-lattice models use stepping-stone dispersal: in one unit of time, only nearest neighbors of a lattice point are accessible (Kimura and Weiss, 1964; Malécot, 1969). Such models are useful for qualitative assessment of the effects of localized dispersal (Barton et al., 2002). Quantitative effects, and especially questions about the shape of the dispersal kernel, as discussed above, demand more sophisticated treatments. However, there has been very little development in the statistical literature of suitable probability distributions on integer lattices. Indeed, there is a need for parametric families of probability distributions on integer lattices in which the degree of kurtosis is a parameter so that the effects of leptokurtic dispersal can be studied in models.

Most discrete probability distributions of concern to statisticians are restricted to the nonnegative integers (Johnson et al., 1992). As a consequence, dispersal is often modeled by discretizing distributions on continuous Euclidean space for use on a lattice (e.g. Higgins and Richardson, 1999; Ibrahim et al., 1996). However, theory for the original continuous distributions does not apply to discretizations. Indeed, such features as convolutions, moments, and the relationships between them, transfer at best approximately to discretizations, and may be especially misleading for cases where the median dispersal distance is only a few lattice points. Similar difficulties arise with the common approach of using a probability distribution for distance dispersed to define the probabilities of dispersing to multidimensional lattice points regardless of direction (e.g. Levin and Kerster, 1975; Rousset, 2000).

We develop here a class of integer lattice distributions with a special focus on their applications to modeling dispersal. These distributions are designed to be simulated readily, with properties that are easy to define and control. We provide several families of such distributions defining dispersal kernels in any number of dimensions, although serious applications in population biology rarely go beyond two. These families are defined by time randomizations of convolutions of stepping stone kernels. A counterpart of this technique was recently applied by Yamamura (2002) to create a class of leptokurtic distributions for modeling dispersal in continuous space. For the most part, the families that we derive preserve convolutions, and so multiple dispersal steps of a kernel remain in the same family. The moments of these kernels have an elegant simplicity of interpretation, and their Fourier transforms have compact forms. Calculation of the probabilities for these distributions, i.e. calculation of the kernel itself, is often more complex, but robust numerical techniques are available in general. We build these distributions from stepping-stone distributions as the basic elements. In general, however, they have infinite tails and their properties vary from discrete approximation of multivariate normality to strong leptokurtosis suitable for representing rare long-distance dispersal.

To facilitate understanding, a list of notation is provided as Table 1.

2. Foundations

Given a probability mass function $K(\mathbf{x}) = P(\mathbf{X} = \mathbf{x})$, for some random variable $\mathbf{X} = (X_1, X_2, \ldots, X_d)$ on the d-dimensional integer lattice $(Z^d, \mathbf{x} \in Z^d)$, a dispersal kernel can be defined as the function of two variables $K(\mathbf{y} - \mathbf{x})$. This function gives the probability of dispersing in one unit of time from lattice point $\mathbf{x}$ to lattice point $\mathbf{y}$. Such kernels are translationally invariant because the dispersal probabilities depend only on the displacement $\mathbf{y} - \mathbf{x}$, not separately on the point of origin $\mathbf{x}$. Because of the direct relationship between $K(\mathbf{x})$ and the kernel $K(\mathbf{y} - \mathbf{x})$ derived from it, we refer to $K(\mathbf{x})$ as the kernel. Kernels are commonly chosen with further symmetry properties. In continuous space, radially

symmetric kernels are used so that all directions of movement are equally likely. On an integer lattice, however, exact radial symmetry is impossible because the lattice itself has restricted symmetry: only multiples of right angle rotations of the lattice are invariant. However, one can seek forms of asymptotic invariance where the probability distribution of the dispersal angle approaches uniformity as the distance dispersed approaches infinity.

In continuous space, a multivariate Gaussian distribution with components $X_1, X_2, \ldots, X_d$ independent and identically distributed is radially symmetric, and is the only radially symmetric distribution with independent components (Fang et al., 1990; Feller, 1971). Thus, except for the Gaussian, radially symmetric kernels cannot be obtained from statistically independent component distributions. However, symmetric kernels can be manufactured as averages (commonly called mixtures) of Gaussian kernels with different variances (Ibrahim et al., 1996; Yamamura, 2002). Such kernels have higher kurtosis than Gaussian kernels (Nichols and Hewitt, 1994), a useful property as most dispersal kernels observed in nature are strongly leptokurtic. We use an analogous technique to obtain dispersal kernels on integer lattices with the property of long-range symmetry.

The well-known normal approximation to the binomial distribution suggests the binomial as an appropriate approximation to the Gaussian on the integers, provided the n parameter of the binomial is large enough. We use the binomial as our starting point. In d dimensions, the separate coordinates can be independent binomials, just as independent Gaussians as coordinates yield d-dimensional Gaussian kernels. For use as a kernel, the binomial needs to be centered at zero; and for it to be symmetric on the integers, the usual parameter p of the binomial must equal $\frac{1}{2}$. Thus, if X is a random variable representing dispersal from zero on the integers, the binomial that we use is defined by the formula

$$B_{n,1}(x) = P(X = x) = \binom{n}{\frac{n+x}{2}} \left(\frac{1}{2}\right)^n,$$
$$x = -n, \quad -n+2, \ldots, n-2, n. \tag{1}$$

This kernel describes n steps of the stepping-stone kernel, the kernel allowing movement one unit to the right or left, each with probability $\frac{1}{2}$. In terms of random variables, this means that X is a sum of n independent stepping-stone steps:

$$X = \sum_{i=1}^{n} S_i, \tag{2}$$

where the random variables S_i are the stepping-stone steps. These steps are independent and identically distributed binary random variables with

$$P(S_i = -1) = P(S_i = 1) = \tfrac{1}{2}. \tag{3}$$

Thus, the kernel $B_{n,1}(x)$, which we think of as defining one step on the temporal scale of interest, can be decomposed into n steps on some finer temporal scale.

For d dimensions, we define the kernel

$$B_{n,d}(\mathbf{x}) = \prod_{i=1}^{d} B_{n,1}(x_i). \tag{4}$$

The random vector $\mathbf{X}$ representing one step in d dimensions has coordinates $X_1, X_2, \ldots, X_d$, where the X_i are independent with distribution $B_{n,1}$. Formula (2) still applies but with S_i replaced by the vector $\mathbf{S}_i$ having the components $S_{1i}, S_{2i}, \ldots, S_{di}$, the independent stepping-stone steps for each dimension.

These binomial kernels have some peculiarities, which mean that they are not themselves satisfactory models. Note first that from formula (1) that $n + X_i$ must be even, i.e., every other integer is inaccessible in a single step with these kernels. Second, they have finite ranges ($|X_i| \leqslant n$) though an unbounded range for dispersal is most natural. Third, these kernels have lower kurtosis than the Gaussian, although not by much. For a random variable centered on zero, kurtosis is measured as

$$\kappa = \frac{E[X_i^4]}{(E[X_i^2])^2} - 3, \tag{5}$$

which is the fourth cumulant (Johnson et al., 1992) rescaled to unit variance. A Gaussian has $\kappa = 0$, but for these binomials $\kappa = -2/n$ (Johnson et al., 1992).

The first two of these problems with binomials can be eliminated and the third ameliorated first by forming the following Poisson kernel:

$$P_{\mu,d}(\mathbf{x}) = \sum_{n=0}^{\infty} \frac{e^{-\mu}\mu^n}{n!} B_{n,d}(\mathbf{x}), \tag{6}$$

which is a mixture of binomial kernels over the Poisson distribution with mean μ. In terms of random variables, a step $\mathbf{X}$ now has the representation

$$\mathbf{X} = \sum_{i=1}^{N} \mathbf{S}_i, \tag{7}$$

where each $\mathbf{S}_i$ remains a d-dimensional stepping-stone step, but the number of these steps is a Poisson random variable N with mean μ. In one dimension, this kernel is positive on the entire integer lattice, with maximum at zero and, as we shall see later, $\kappa = 1/E[N] (= 1/\mu)$. It is thus slightly leptokurtic. In one dimension, this Poisson kernel has an explicit representation in terms of modified Bessel functions

$$P_{\mu,1}(x) = e^{-\mu} I_x(\mu) \tag{8}$$

(Feller, 1971, p. 59).

Table 1
Notation

α	Function of Sichel parameters equal to $\sqrt{\omega^2+2\omega\xi}$. Also used for stable distribution shape parameter
$B_{n,d}(\mathbf{x})$	Binomial kernel, Eq. (4)
$\cos(\mathbf{t})$	Product of cosines: $\prod_{l=1}^{d}\cos(t_l)$, Eq. (16)
γ	Sichel shape parameter $\in(-\infty,\infty)$
d	Dimension of the space
$f_U(u)$	Probability density function of U
k	Convolution parameter
κ	Kurtosis, Eq. (5)
$K(\mathbf{x})$	Dispersal kernel specifying the probability of moving from $\mathbf{0}$ to $\mathbf{x}$ in one step
$\tilde{K}(\mathbf{t})$	Fourier transform (characteristic function) of K, with argument $\mathbf{t}$, Eq. (13)
K_v	Modified Bessel function of the third kind of degree v
μ	Common mean of N and U, and kernel scale parameter
N	Random variable giving the number stepping stone steps making up $\mathbf{X}$ or $\mathbf{Z}$
$P_{\mu,d}(\mathbf{x})$	Poisson kernel, Eq. (6)
$\pi_N(\theta)$	Probability generating function for N, Eq. (18)
$\mathbf{S}_i$	One d-dimensional stepping-stone step on the diamond lattice
$\mathbf{S}_i^*$	One d-dimensional stepping-stone step on the rectangular lattice
U	Random variable specifying the conditional mean of N, Eq. (12)
$U_{(k,\mu)}$	Generic parametric family of U variables where k is the number of convolutions of base-level random variable $U_{(1,1)}$. μ is a scale parameter equal to the mean when it exists
$\varphi^2_{(1)}$	Squared coefficient of variation of $U_{(1,1)}$
φ^2_N	Squared coefficient of variation of N
$\mathbf{X}$	Random vector corresponding to one step of the general kernel K, on a diamond lattice
ξ	Scale parameter for Sichel and stable Poisson mixture kernels
$\psi_U(\lambda)$	Laplace transform of U, Eq. (20)
$\psi_{(1)}$	Laplace transform for $U_{(1,1)}$
$\mathbf{Z}$	Random vector corresponding to one step of the kernel $K_{\mathbf{Z}}$ on the rectangular lattice
ω	Sichel shape parameter $\in(0,\infty)$

In more than one dimension, this kernel is not positive on the entire integer lattice because each coordinate of $\mathbf{X}$ is odd or even depending on whether N is odd or even, and so no realization of $\mathbf{X}$ can have a mixture of odd and even coordinates. In two dimensions (the only multidimensional case of practical importance), this means that half the integer lattice is missed, viz all those points where one coordinate is even, and one odd. The lattice represented by values that $\mathbf{X}$ can take consists of integer combinations of the basis vectors $(1,1)'$ and $(-1,1)'$. Thus, it is a diamond lattice obtained from the integer lattice by rotation by 45° and dilation so that the minimum spacing of lattice points is $\sqrt{2}$. This diamond lattice may be just as convenient for modeling purposes as the usual integer lattice because it is just a different representation of the same object. Alternatively, $\mathbf{X}$ can be rotated and contracted to give a random vector $\mathbf{Z}$ defining a positive kernel on the entire integer lattice by the transformation

$$\mathbf{Z}=\begin{pmatrix}\frac{1}{2}[X_1+X_2]\\ \frac{1}{2}[X_2-X_1]\end{pmatrix}. \tag{9}$$

This random vector also has the representation

$$\mathbf{Z}=\sum_{i=1}^{N}\mathbf{S}_i^*, \tag{10}$$

where $\mathbf{S}_i^*$ is a two-dimensional stepping-stone step defined as a random choice from the vectors $(1,0)',(0,1)',(-1,0)'$, and $(0,-1)'$. In contrast, $\mathbf{S}_i$ is a random choice from $(1,1)',(-1,1)',(-1,-1)'$, and $(1,-1)'$. The very simplest case of $\mathbf{Z}$, with $N\equiv 1$, and hence one stepping-stone step, is nearest neighbor dispersal in two dimensions.

The kernel defined by $\mathbf{Z}$ is related to that defined by $\mathbf{X}$ according to the relationship

$$K_{\mathbf{Z}}\begin{pmatrix}z_1\\ z_2\end{pmatrix}=K\begin{pmatrix}z_1-z_2\\ z_1+z_2\end{pmatrix}. \tag{11}$$

Thus, results derived for $\mathbf{X}$ are easily transformed into results for $\mathbf{Z}$. Theory development using $\mathbf{X}$ is simpler than using $\mathbf{Z}$ because the coordinates of $\mathbf{X}$ are conditionally independent given N, while those of $\mathbf{Z}$ are not. Numerical computation of the kernel itself is also more simply approached using $\mathbf{X}$. Hence, the development here is in terms of $\mathbf{X}$ stating results for $\mathbf{Z}$ that follow directly from this development.

So far, we have defined only the binomial and Poisson kernels, but a variety of useful kernels is available by making other choices for the distribution of N. The Poisson kernel, being very similar to the Gaussian in most properties, is at one end of realistic distributions for modeling dispersal. In many instances, the Gaussian is the standard of comparison against which the effects of leptokurtosis may be assessed. Greater degrees of leptokurtosis arise from distributions of N that have higher variance for a given value of the mean than the Poisson. A convenient family of distributions with this property is the Poisson mixture distributions, i.e. those distributions for which

$$P(N=n)=E\left[\frac{e^{-U}U^n}{n!}\right], \tag{12}$$

where U is some nonnegative random variable. Naturally, when U has a continuous distribution with a probability density function f_U, the expectation (12) is given by an integral—see Table 2. Poisson mixture distributions include important classes of discrete probability distributions on the nonnegative integers such as the negative binomial, the Sichel distributions (as discussed below), and many others (Johnson et al., 1992). Before presenting these distributions, however,

Table 2
Parametric families for the distribution of N

Distribution of N	Probability mass function, $p_n = P(N = n)$	Probability generating function, $\pi_N(\theta)$	Squared coefficient of variation φ_N^2
Poisson	$\frac{\mu^n e^{-\mu}}{n!}$	$e^{\mu(\theta-1)}$	$\frac{1}{\mu}$
Negative binomial	$\binom{k+n-1}{k-1}\left(\frac{\mu}{k}\right)^n\left(1+\frac{\mu}{k}\right)^{-n-k}$	$\left[1+\frac{\mu}{k}(1-\theta)\right]^{-k}$	$\frac{1}{k}+\frac{1}{\mu}$
Sichel	$\frac{\xi^n(\omega/\alpha)^{\gamma+n}K_{\gamma+n}(\alpha)}{n!K_\gamma(\omega)}$, where $\alpha=\sqrt{\omega^2+2\omega\xi}$	$\frac{K_\gamma(\omega\sqrt{1+2\xi(1-\theta)/\omega})}{K_\gamma(\omega)(\sqrt{1+2\xi(1-\theta)/\omega})^\gamma}$	$\frac{1}{\mu}\left[1+\frac{2\xi(\gamma+1)}{\omega}\right]-1+\left(\frac{\xi}{\mu}\right)^2$
Infinitely divisible Poisson mixture	$\int_0^\infty f_U(u)\frac{u^n e^{-u}}{n!}\,du$	$\left\{\psi_{(1)}\left(\frac{\mu}{k}[1-\theta]\right)\right\}^k$	$\frac{\varphi_{(1)}^2}{k}+\frac{1}{\mu}$
Stable Poisson mixture	$\int_0^\infty f_U(u)\frac{u^n e^{-u}}{n!}\,du$	$\exp\{-k^{1-\alpha}\xi^\alpha(1-\theta)^\alpha\}$	∞

we derive the general form of the Fourier transform of these kernels, showing how it is related to the transforms of the distributions of N and U used to build these kernels.

3. Fourier transforms

For analytical modeling purposes, one of the most useful features of a kernel is its Fourier transform. The Fourier transform is unique to a kernel and can be used to calculate the kernel, as discussed below. However, the Fourier transform is more useful than the kernel itself when modeling population dynamics that are linear (Lande, 1991), or approximately linear (Lande et al., 1999; Roughgarden, 1977, 1978; Snyder and Chesson, 2003). For studies of invasions, the rate of spread can be understood from the Fourier transform (expressed as the moment generating function), even though population dynamics are nonlinear locally in space (Clark et al., 2001). The Fourier transform of a kernel K is defined slightly differently by different authors, but we choose the following definition

$$\tilde{K}(\mathbf{t}) = \sum_{\mathbf{x}\in Z^d} K(\mathbf{x})e^{i\mathbf{t}\cdot\mathbf{x}}, \tag{13}$$

where the argument $\mathbf{t}$ is the vector $(t_1, t_2, \ldots, t_d)'$, the dot $(\cdot)$ means inner product, and i is $\sqrt{-1}$. In the probability literature, this Fourier transform is called the characteristic function of the probability distribution K, or the random variable $\mathbf{X}$, and can be written as the expected value

$$\tilde{K}(\mathbf{t}) = E[e^{i\mathbf{t}\cdot\mathbf{X}}], \tag{14}$$

(Feller, 1971). Fourier transforms of kernels in the families introduced here are especially easy to obtain. We begin with the fact that the characteristic function of a stepping-stone step in one dimension is, by (14),

$$E[e^{it_jS_{jl}}] = \tfrac{1}{2}(e^{it_j} + e^{-it_j}) = \cos(t_j). \tag{15}$$

Independence of the components implies that the d-dimensional characteristic function of a d-dimensional stepping-stone step is

$$E[e^{i\mathbf{t}\cdot\mathbf{S}_j}] = \prod_{l=1}^{d}\cos(t_l) \stackrel{\text{def}}{=} \cos(\mathbf{t}), \tag{16}$$

where the product of cosines in the center defines the cosine of a d-dimensional argument on the right.

As the binomial kernel is the n-fold convolution of the stepping-stone kernel, it follows that its Fourier transform is $[\cos(\mathbf{t})]^n$, according the standard properties of Fourier transforms. To obtain the Fourier transform of a kernel in a family defined by replacing n with the random variable N, we use the conditional expectation formula:

$$\begin{aligned}\tilde{K}(\mathbf{t}) &= E\{E[e^{i\mathbf{t}\cdot\mathbf{X}}|N]\}\\ &= E[\cos(\mathbf{t})^N]\\ &= \pi_N(\cos(\mathbf{t})).\end{aligned} \tag{17}$$

The function π_N in this last step is the probability generating function of N, which is defined in general as

$$\pi_N(\theta) = E[\theta^N] = \sum_{n=0}^{\infty}\theta^n P(N=n). \tag{18}$$

The probability generating function is readily available for most commonly used discrete distributions (e.g. Johnson et al., 1992, and Table 2 above). The power series representation (18) shows that the probability generating function determines the probability distribution. For the Poisson distribution, the probability generating function is $\exp(\mu(\theta-1))$ (Johnson et al., 1992), and so the characteristic function for the Poisson kernel is

$$\tilde{P}_{\mu,d}(\mathbf{t}) = e^{\mu(\cos(\mathbf{t})-1)}. \tag{19}$$

To obtain the probability generating function of N with a Poisson mixture distribution, we need the Laplace transform of U, which is defined as

$$\psi_U(\lambda) = E[e^{-\lambda U}] = \int_0^\infty f_U(u)e^{-\lambda u}\,du. \tag{20}$$

The Laplace transform is the most common transform used for analysis of nonnegative continuous random variables and is readily available for common distributions (Johnson et al., 1994). Examples are given in Table 2. To get the probability generating function of N from the Laplace transform of U, we again use conditional expectations to obtain

$$\begin{aligned}\pi_N(\theta) &= E\{E[\theta^N|U]\}\\ &= E\{e^{U(\theta-1)}\}\\ &= \psi_U(1-\theta),\end{aligned} \tag{21}$$

where the second step uses the fact that N is Poisson conditional on U with mean U.

We now use formula (21) to see that the general formula for the Fourier transform of a Poisson mixture kernel is

$$\tilde{K}(\mathbf{t}) = \psi_U(1-\cos(\mathbf{t})). \tag{22}$$

Thus, we can go easily from the Laplace transform of a nonnegative random variable U, with desirable properties, to the Fourier transform of the kernel generated by it. In making this connection, we have used three different sorts of transform. These transforms are all variations on the characteristic function. Substituting θ for $\exp(it)$ in the characteristic function gives the probability generating function, and substituting $-\lambda$ for it gives the Laplace transform. These are the most common transforms in use for each of the different sorts of distribution involved, viz, distributions taking both positive and negative values, nonnegative integer values, and positive continuous values.

For two-dimensional kernels defined by $\mathbf{Z}$, the representation (9) implies that $\cos(\mathbf{t})$ (i.e. $\cos(t_1)\cos(t_2)$) is replaced by $\frac{1}{2}(\cos(t_1)+\cos(t_2))$ in each of the formulae above. In particular,

$$\tilde{K}_{\mathbf{Z}}(\mathbf{t}) = \pi_N(\tfrac{1}{2}[\cos(t_1)+\cos(t_2)]). \tag{23}$$

For Poisson kernels, the exponential form of the probability generating function π_N means that this characteristic function is the product of the characteristic functions for each coordinate considered separately. Thus, the coordinates of $\mathbf{Z}$ are statistically independent in the Poisson case. Moreover, the distributions of the coordinates are Poisson kernels as are the coordinate distributions of $\mathbf{X}$. This fact means that an alternative development of this kernel theory could be based on conditional independence of the coordinates of $\mathbf{Z}$ given U. However, the kernel in terms of $\mathbf{X}$ is much more simply obtained numerically because the formulae involve series rather than integrals.

The moment generating function, $M(\mathbf{t}) = E[\exp(\mathbf{t}\cdot\mathbf{X})]$, of a kernel also appears in the literature on long distance dispersal for some applications (Caswell et al., 2003; Kot et al., 1996). This function is obtained by substituting $-i\mathbf{t}$ for $\mathbf{t}$ in the characteristic function. Note, however, that although a strongly leptokurtic kernel need not have a moment generating function (i.e. $M(\mathbf{t})$ may be infinite except for $\mathbf{t}=\mathbf{0}$), all kernels have characteristic functions (Fourier transforms).

4. Distributions for N and U

We have just seen how a discrete distribution on the nonnegative integers for N, or alternatively a positive continuous distribution for U, yields a dispersal kernel. Our task now is to choose N and U distributions that confer desirable properties on the kernel. Because n convolutions of a kernel represent n steps of that kernel, interpretability of the family is aided if convolutions of a kernel remain in the family. We also seek ways of changing the spatial scale of dispersal without changing the basic shape of the kernel. Finally, we seek the ability to vary the shape of the kernel, especially the nature of the tail, to be able to compare different degrees and different sorts of leptokurtosis. We begin our development by inquiring about how to create families that preserve convolution.

The convolution of a kernel with itself is defined as

$$K^{*2}(\mathbf{x}) = \sum_{\mathbf{y}} K(\mathbf{x}-\mathbf{y})K(\mathbf{y}), \tag{24}$$

which defines the probability distribution for two steps of the kernel K, i.e. the distribution of $\mathbf{X}_1+\mathbf{X}_2$, where $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent and have the same distribution as $\mathbf{X}$. Closure of a family under convolution means that the kernel defining the outcome of multiple steps is a member of the same family as the kernel for one step, and therefore can be conveniently interpreted in terms of parameter changes within that family. In addition, the unit of time in which one step takes place is often somewhat arbitrary in models, and so it would be desirable if the family in use did not depend on the

choice of time unit. Thus, a single step ought to be expressible as a convolution of steps within the same family so that it can be interpreted as the outcome of a sequence of steps over shorter time units.

Our kernel construction, where one step is composed of N steps of the stepping-stone kernel, means that these convolution requirements for the kernel are convolution requirements for the family to which N belongs: two steps of the kernel are $N_1 + N_2$ steps of the stepping-stone kernel, where N_1 and N_2 are independent and have the distribution of N. For a single step of the kernel to be represented in terms of steps over shorter periods of time, we require that for any k, N can always be represented as a sum

$$N = N_1 + N_2 + \cdots + N_k, \tag{25}$$

where the N_j are independent and identically distributed, i.e. the distribution of N is a k-fold convolution of some distribution with itself. This requirement on N means that the distribution of N is by definition infinitely divisible (Billingsley, 1986; Feller, 1971). If a discrete random variable $N_{(1)}$ has probability generating function $\pi_{(1)}$, the probability generating function of the k-fold convolution of the distribution of $N_{(1)}$ is

$$\pi_{(k)}(\theta) = [\pi_{(1)}(\theta)]^k. \tag{26}$$

If the distribution of $N_{(1)}$ is infinitely divisible, k can be fractional, and then Eq. (26) defines an infinitely divisible family of distributions for N where k is the single parameter of the family. An infinitely divisible family of kernels results when $\pi_{(k)}$ is substituted for π_N in the characteristic function (17) of the kernel.

Infinitely divisible distributions for N are easily specified. They have probability generating functions of the form

$$\pi_{(k)}(\theta) = e^{k(\pi^*(\theta)-1)}, \tag{27}$$

where π^* is the probability generating function of any nonnegative discrete random variable (Johnson et al., 1992). These distributions are the "generalized Poisson distributions," and for each probability generating function π^*, we obtain a one-parameter family of kernels, with parameter k, where k simultaneously affects both shape and scale, as we shall see below. To obtain two-parameter families with one parameter affecting the shape or kurtosis, and another the scale, we choose a subset of generalized Poisson distributions, the Poisson mixture distributions (Johnson et al., 1992).

With Poisson mixtures, infinite divisibility of N is inherited from infinite divisibility of U. Many well-known families of nonnegative continuous random variables have the property of infinite divisibility. Given an infinitely divisible random variable $U_{(1)}$ with Laplace transform $\psi_{(1)}$, there is by definition a family of random variables $\{U_{(k)}, k>0\}$, with Laplace transforms $\{\psi_{(k)}\}$ satisfying the relationship

$$\psi_{(k)}(\lambda) = [\psi_{(1)}(\lambda)]^k. \tag{28}$$

By Eq. (21), this Laplace transform induces a one-parameter infinitely divisible family $\{N_{(k)}, k>0\}$ for N, with probability generating function $\pi_{(k)}(\theta) = \psi_{(k)}(1-\theta)$. However, as we shall see shortly, we can introduce a scale parameter, and remove the scaling property of k by choosing for the U family the random variables $U_{(k,\mu)} = \mu U_{(k)}/k$. When the mean of $U_{(k,\mu)}$ exists, it is independent of k, and the family can be defined so that $\mu = E[U_{(k,\mu)}]$. Now $U_{(k,\mu)}$ has the Laplace transform

$$\psi_{(k,\mu)}(\lambda) = \left[\psi_{(1)}\left(\frac{\lambda\mu}{k}\right)\right]^k. \tag{29}$$

The corresponding N family has probability generating function

$$\pi_{(k,\mu)}(\theta) = \left[\psi_{(1)}\left(\frac{\mu}{k}(1-\theta)\right)\right]^k. \tag{30}$$

The Fourier transform of the kernel becomes

$$\tilde{K}_{(k,\mu)}(\mathbf{t}) = \left[\psi_{(1)}\left(\frac{\mu}{k}[1-\cos(\mathbf{t})]\right)\right]^k. \tag{31}$$

With these families, both N and U have the same mean, when it is finite. This common mean is $\mu E[U_{(1)}]$, which is just μ if $U_{(1)}$ is chosen to have mean 1, for example by replacing any candidate $U_{(1)}$ with $U_{(1)}/E[U_{(1)}]$. Most important, however, the common mean of N and U is independent of k. When the mean exists, the law of large numbers ensures that the spread of both U and N about their mean decreases as k increases. The next section on moments shows that the mean number of stepping-stone steps, $E[N]$, determines the scale of $\mathbf{X}$. The spread in the number of steps about this mean, which is determined by k, governs the shape of kernel. Thus, defining the kernel family in terms of the random variables $U_{(k,\mu)}$ allows μ and k to be interpreted as shape and scale parameters of the kernel, at least in the case where the mean of N is finite. Where this mean is infinite, for example in models of extreme leptokurtosis, shape and scale become difficult to separate, as we shall see below in Section 4.3.

We name these kernels according the name of the distribution of N, or U when N does not have a specific name. We choose three pairs of families of infinitely divisible U and N distributions for illustration and which we believe can be valuable for applications.

4.1. Negative binomial kernels

In this family, U has a gamma distribution (Johnson et al., 1994), which means that N is negative binomial (Johnson et al., 1992). The base level $U_{(1)}$ random variable can be chosen to be exponential (a special case

of the gamma) with mean 1, i.e.

$$P(U_{(1)} > u) = e^{-u}, \tag{32}$$

and the parameter μ, defined above, then corresponds to the mean of U and N. The negative binomial probabilities, and corresponding probability generating function, are given in Table 2. The shape parameter k is the usual negative binomial clumping parameter, of common use in ecology to characterize the degree to which negatively binomially distributed counts from quadrats indicate spatial clumping of organisms (Southwood and Henderson, 2000). The negative binomial provides an archetypical two parameter family where the parameters are the shape and scaling parameters, μ and k, discussed here. The $U_{(k,\mu)}$ distribution is the gamma distribution with mean μ and shape parameter k.

Although the widely familiar negative binomial is noted for its strong skew, which also means that it has high variance for a given value of the mean, it has some limitations from the point of view of modeling dispersal: the tails of the distribution are bounded by a negative exponential, implying the same property for the tails of the dispersal kernel (Appendix B) and also that all moments of the kernel are finite. Thus, the negative binomial kernel may not be sufficiently leptokurtic for some dispersal in nature, which appears not to have exponentially bounded tails (Clark et al., 1999; Portnoy and Willson, 1993), with most moments effectively infinite.

4.2. Sichel kernels

For these kernels, the distribution of U is the generalized inverse Gaussian distribution (Johnson et al., 1994), which gives rise to distributions of N known as Sichel distributions (Johnson et al., 1992). They are more complicated than negative binomials, but allow a very wide range of leptokurtosis. In particular, limiting forms within this family do not have exponentially bounded tails.

The density function of the generalized inverse Gaussian can be written as

$$f_U(u) = \frac{u^{\gamma-1} e^{-\frac{1}{2}\omega(\xi^{-1}u + \xi u^{-1})}}{2\xi^{\gamma} K_\gamma(\omega)}, \tag{33}$$

(Johnson et al., 1994) where the K_γ is the modified Bessel function of the third kind (Abramowitz and Stegun, 1964; Johnson et al., 1992) defined implicitly here by the fact that (33) must integrate to 1 over the range 0 to ∞. The parameters ξ and ω are positive, and γ is an unrestricted real number. The parameter ξ is a scale parameter, but is not equal to the mean. The mean μ, is instead given by the formula

$$\mu = \xi \frac{K_{\gamma+1}(\omega)}{K_\gamma(\omega)}. \tag{34}$$

For fixed positive γ, the limit of the generalized inverse Gaussian as $\xi \to 0$, with $\omega = 2\gamma\xi$, is the gamma distribution with mean 1 and shape parameter $k = \gamma$. On the other hand, if γ is negative and fixed, the limit as $\xi \to \infty$, with $\omega = 2|\gamma|/\xi$, is the distribution of the reciprocal of a gamma with mean 1 and $k = -\gamma$. Thus, for positive γ in the limit as $\xi \to 0$, the Sichel kernels are the negative binomial kernels. For negative γ in the limit as $\xi \to \infty$ they are the Poisson reciprocal gamma kernels. The reciprocal gamma has the property that its tails are not exponentially bounded. They are instead asymptotically proportional to the power $u^{\gamma-1}$. The tails of N are then asymptotically proportional to $n^{\gamma-1}$ and the tails of the kernel are asymptotically proportional to $|x|^{2\gamma-1}$ (Appendix B). This power-law tail is also found in the continuous-space analogue of this kernel, the t-distribution kernel that Clark et al. (1999) used to fit dispersal of tree seeds. These t-distribution kernels are derived as reciprocal gamma mixtures of the normal kernel, which our Poisson kernel approximates on lattices.

Fig. 1 illustrates these Sichel kernels and their convergence on the Poisson reciprocal gamma and negative binomial limits. The variance and kurtosis are the same for each kernel in the figure, and so the differences between these kernels represent shape differences available within the Sichel family not captured by the kurtosis measure κ alone. From perspective of the tails of the kernel (Fig. 1(b)) these shape differences are manifest as different degrees of deviation from a negative exponential. The kernels show a clear gradation in tail behavior from nearly negative exponential to a power function as one moves through these Sichel kernels from the negative binomial to the Poisson reciprocal gamma.

One complication with the Sichel kernels is that although U is infinitely divisible (Barndorff-Nielsen et al., 1989), the kth convolutions of the distributions are not known distributions except in the limiting gamma case and the case where $\gamma = -\frac{1}{2}$ (the inverse Gaussian distributions, Johnson et al., 1994). Thus, the prescription above (formulae (29)–(31)) for generating two-parameter families of kernels, preserving convolution, will lead to kernels that are not Sichel kernels for $k \neq 1$ except in the case $\gamma = -\frac{1}{2}$, and in the negative binomial limit. The new kernels might be called generalized Sichel kernels. To use the formulae (29)–(31) to generate these generalized Sichel kernels, $\psi_{(1)}$ is given as

$$\psi_{(1)}(\lambda) = \frac{K_\gamma(\omega\sqrt{1 + 2\xi\lambda/\omega})}{K_\gamma(\omega)(\sqrt{1 + 2\xi\lambda/\omega})^{\gamma}} \tag{35}$$

with $\xi = K_\gamma(\omega)/K_{\gamma+1}(\omega)$ so that $\mu = 1$ in Eq. (34).

4.3. Stable kernels

The final kernels that we consider are generated by distributions of U called stable distributions (Feller,

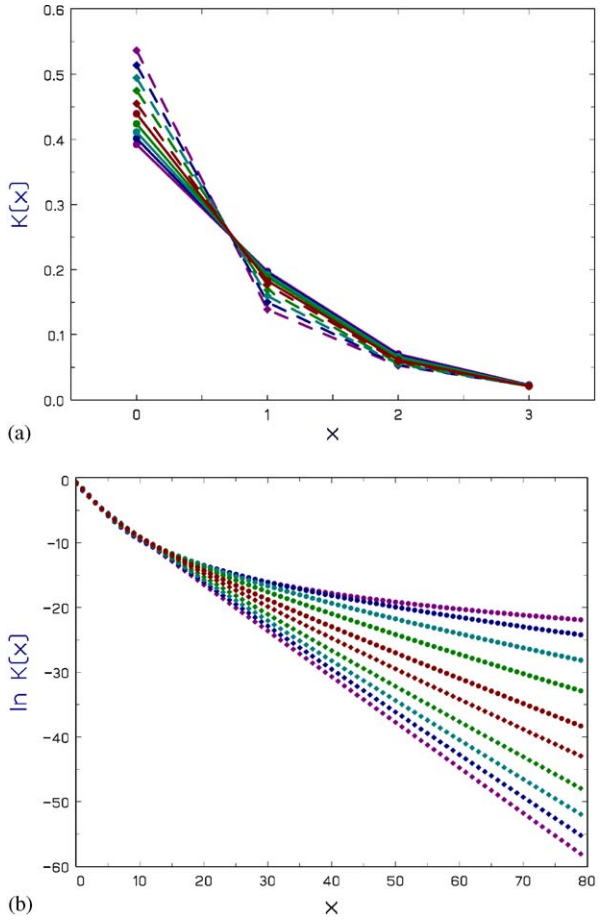

Fig. 1. Sichel kernels converging on Poisson reciprocal gamma and negative binomial limits. Kernel variance, μ, and kurtosis, κ, were fixed at 2 and 6.5, respectively. The parameter ω was varied through the values 0.54, 0.47, 0.35, 0.2 and 0, with 0 corresponding to the Poisson-reciprocal gamma (for $\gamma<0$) and negative binomial ($\gamma>0$). For each ω value there are two solutions for γ (with a corresponding ξ value), for the given μ and κ. The smaller (more negative) γ solution is indicated by the circles and solid lines. Diamonds and dashed lines correspond to the larger solution. (a) $K(x)$ for small x. (b) $\ln K(x)$ demonstrating tail behavior. In both cases smaller values of ω correspond to more peripheral curves, with the outer-most curves being the limiting cases for $\omega=0$, and the central-most curves having $\omega=5.4$.

1971). These distributions are for modeling extreme leptokurtosis because they all give infinite dispersal variance. The Poisson reciprocal gamma kernels discussed above intersect with this family for the special case $\gamma=-\frac{1}{2}$. Indeed, this family provides an alternative to the Poisson reciprocal gamma family for representing extreme leptokurtosis. Its chief advantages are its very simple Fourier transforms, and the property that convolutions of the U distribution do remain in the original family.

The Laplace transform of the base-level random variable $U_{(1)}$ can be taken as

$$\psi_{(1)}(\lambda)=e^{-\lambda^{\alpha}}, \tag{36}$$

where the parameter α lies between 0 and 1. Feller (1971, p. 583) shows that the density function is asymptotically proportional to $u^{-\alpha-1}$. Coincidence with the Poisson reciprocal gamma occurs for $\alpha=-\gamma=\frac{1}{2}$. The Laplace transform for the random variable $U_{(k,\xi)}=\xi U_{(k)}/k$ for generating a kernel with scale parameter ξ and shape (convolution) parameter k is

$$\psi_{(k,\xi)}=e^{-k^{1-\alpha}\xi^{\alpha}\lambda^{\alpha}}, \tag{37}$$

where the scale parameter is chosen as ξ instead of μ to emphasize that the mean of this distribution does not exist, and so the scale parameter cannot be set equal to the mean. Note that here k and ξ do not have separate effects on this Laplace transform. They both multiply λ^{α}, but their effects are of different magnitude unless $\alpha=\frac{1}{2}$. In contrast to cases where U has a finite mean, the distribution of U does not become more highly concentrated about a central value as k is increased. Instead, it has an effect equivalent to increasing the scale of U. This feature comes from the fact that, for cases of infinite variance (Feller, 1971), stable distributions replace the normal distribution in the central limit theorem. Although k does not function as a shape parameter for stable distribution kernels, the parameter α is an independent shape parameter, governing the tails of U. Presumably the tails of the kernel are asymptotically proportional to $|x|^{-2\alpha-1}$, but we have proved this proposition only for the case $\alpha=\frac{1}{2}$ where these stable kernels coincide with reciprocal gamma kernels (Appendix B).

5. Moments

The first few moments of the these new kernels are easily derived in the cases where they are finite. The $2n$th order moments of $\mathbf{X}$ can be expressed in terms of the nth and lower order moments of $\mathbf{N}$. The $2n$th order moments of $\mathbf{X}$ are finite if and only if the nth order moments of N are finite (Appendix C). In all cases, because the stepping-stone kernel is symmetric about 0, the components X_i are symmetric about 0. Hence, $E[\mathbf{X}]=0$ whenever $EN<\infty$. Note also that $E[\mathbf{X}|N]=0$ because the stepping-stone steps $\mathbf{S}_i$ are independent of N, and have zero mean.

The second order moments can be derived from the conditional covariance formula, $\mathrm{Cov}(X_i,X_j)=E\{\mathrm{Cov}(X_i,X_j|N)\}+\mathrm{Cov}(E[X_i|N],E[X_j|N])$, which here reduces to $\mathrm{Cov}(X_i,X_j)=E\{\mathrm{Cov}(X_i,X_j|N)\}$. As S_{ij} has variance 1, and $\mathrm{Cov}(S_{ik},S_{jl})$ equals zero whenever $i\neq j$ or $k\neq l$, it follows that $\mathrm{Cov}(X_i,X_j|N)=0$ for $i\neq j$, but $V(X_i|N)=\mathrm{Cov}(X_i,X_i|N)=N$. Thus,

$$\mathrm{Cov}(X_i,Xj)=0,\quad i\neq j \tag{38}$$

and

$$V(X_i) = E[N]. \tag{39}$$

That the covariance between components of $\mathbf{X}$ should be zero is a minimal expectation from a distribution that has properties of radial symmetry. Symmetry through right angle rotations guarantees this outcome if the moments exist. Eq. (39) for the variance means that spread of the distribution is determined by the mean of N. If we define R as the Euclidean distance dispersed ($R = |\mathbf{X}|$), then because $E[Xi] = 0$, Eq. (39) implies that, for dimension d,

$$E[R^2] = dE[N]. \tag{40}$$

In particular, the root mean square distance dispersed ($\sqrt{E[R^2]}$) is $\sqrt{dE[N]}$. Thus, we can think of $E[N]$ as controlling dispersal distance in this sense of root mean square. We can think of this as the scale of dispersal. Thus, the parameter μ ($= E[N] = E[U]$, when finite), in the kernel families defined here, is a scale parameter.

The symmetry of these distributions means that third moments, when they exist, are all zero. The fourth moments of the components of $\mathbf{X}$ can be evaluated using the conditional expectation formula $E[X_i^4] = E\{E[X_i^4|N]\}$, noting that

$$E[X_i^4|N] = E\left[\left(\sum_{j=1}^{N} S_{ij}\right)^4 \middle| N\right]. \tag{41}$$

This conditional expectation is just the fourth central moment of the binomial (Johnson et al., 1992), which is here equal to $3N^2 - 2N$, and is also given by the recurrence relation (C.4) of Appendix C. Taking the expectation of this conditional expectation, and using the relationship $E[N^2] = V(N) + (E[N])^2$, we see that

$$E[X_i^4] = 3V(N) + 3(E[N])^2 - 2E[N]. \tag{42}$$

It follows that the kurtosis coefficient κ (expression (5)) is

$$\kappa = 3\varphi_N^2 - \frac{2}{E[N]}, \tag{43}$$

where φ_N is the coefficient of variation of N, i.e. the ratio of the standard deviation of N to its mean.

The second part of this expression, $-2/E[N]$, is the kurtosis of the binomial kernel. Thus, the kurtosis of our general kernel exceeds that of the binomial kernel according to the spread of the distribution of N, as measured by its coefficient of variation. This outcome makes intuitive sense because high variation in the number of stepping-stone steps means that both small and large values of N are likely, with small values contributing to the peak in the distribution of $\mathbf{X}$ near zero, and large values of N fattening the tails by allowing $\mathbf{X}$ to take large values.

In the case of the Poisson, $\varphi_N^2 = 1/E[N]$, and so κ is $1/E[N] = 1/\mu$, as stated earlier. For infinitely divisible Poisson mixture families, $\varphi_N^2 = \varphi_{(1)}^2/k + 1/\mu$, where $\varphi_{(1)}^2$ is the coefficient of variation of $U_{(1)}$. Thus, in these families,

$$\kappa = \frac{3\varphi_{(1)}^2}{k} + \frac{1}{\mu}. \tag{44}$$

Except when the scale parameter μ is small, this expression will be dominated by the first term involving k. Thus, we can see that k functions as a shape parameter. Is inevitable, however, that scale will also have a role in shape because the distribution is discrete. If μ is small, this discreteness is pronounced, but for μ large, the kernel approximates a continuous function. We can think of k as controlling shape beyond that accounted for the discreteness of the kernel and a Poisson level of variation in N.

For multidimensional distributions, Mardia (1970) has suggested a measure of kurtosis that combines the univariate kurtosis measure κ above with the covariance between X_i^2 and X_j^2 divided by the product of the variances, for $i \neq j$. However, $\mathrm{Cov}(X_i^2, X_j^2)$ contains separate information of value too. If the components were independent, as they are for the binomial, this covariance would be zero. A nonzero value, however, reflects the statistical dependence induced between the coordinates by the randomness in N, the number of stepping-stone steps in an $\mathbf{X}$ step. The magnitude of N influences the magnitudes of X_i^2 and X_j^2 and induces covariance between them, even though the unsquared coordinates, X_i and X_j, which have direction as well as magnitude, are uncorrelated. Thus, $\mathrm{Cov}(X_i^2, X_j^2)$ is an important measure of deviation from a Gaussian kernel.

To obtain the covariance, we first evaluate the mixed moment $E[X_i^2 X_j^2]$ using conditional expectations, noting that, given N, the components are independent. Thus,

$$\begin{aligned} E[X_i^2 X_j^2] &= E\{E[X_i^2 X_j^2|N]\} \\ &= E\{E[X_i^2|N]E[X_j^2|N]\} \\ &= E[N^2], \end{aligned} \tag{45}$$

where, in the last step, N is the conditional variance of a component, as noted above. Now $\mathrm{Cov}(X_i^2, X_j^2) = E[X_i^2 X_j^2] - E[X_i^2]E[X_j^2]$, and $E[N^2] = V(N) + (E[N])^2$. Thus, we see that

$$\mathrm{Cov}(X_i^2, X_j^2) = V(N). \tag{46}$$

Making this measure independent of the scale, so that its magnitude is more interpretable, we obtain the measure

$$\frac{\mathrm{Cov}(X_i^2, X_j^2)}{V(X_i)V(X_j)} = \varphi_N^2. \tag{47}$$

Thus, we see that the coefficient of variation of N arises here as well. This measure shows that whenever there is any variance in N, the components of $\mathbf{X}$ are not independent, even though they are uncorrelated. More-

over, the degree of this dependence as measured by (47) is closely related to the unvariate kurtosis κ of any component. In case of the Poisson kernel, $\varphi_N^2 = 1/\mu = \kappa$ and so we see that for large μ, this dependence, along with univariate kurtosis, κ, is weak. For most of the other kernels presented here, however, there is strong dependence between components.

In two-dimensions, the results here for $\mathbf{X}$ are readily translated into results for $\mathbf{Z}$ by means of formula (9). As formula (9) is a rotation followed by a contraction by $\sqrt{2}$, $E|\mathbf{Z}|^2 = \frac{1}{2}E|\mathbf{X}|^2 = E[N]$, and therefore

$$V(Z_i) = \tfrac{1}{2}E[N]. \tag{48}$$

Combining the kurtosis calculation (43) with the relative covariance (47) leads to the kurtosis for $\mathbf{Z}$ of

$$\kappa_{\mathbf{Z}} = 3\varphi_N^2 - \frac{1}{E[N]}, \tag{49}$$

which is slightly larger than the kurtosis of $\mathbf{X}$. Mardia's (1970) multivariate kurtosis is invariant under rotation, and for $d = 2$ is a linear function of the sum of the relative covariance (47) and κ. Thus, the relative covariance for $\mathbf{Z}$ must counterbalance the increase in κ to give

$$\frac{\mathrm{Cov}(Z_i^2, Z_j^2)}{V(Z_i)V(Z_j)} = \varphi_N^2 - \frac{1}{E[N]}. \tag{50}$$

These results demonstrate that these kernels do have some rotational asymmetry through 45°, but that asymmetry will be small for modest mean numbers of steps, $E[N]$. For strongly leptokurtic kernels, this small asymmetry will be swamped by the symmetric contribution of φ_N^2. As remarked earlier, exact rotational symmetry is impossible on the integer lattice, and so the small degrees of asymmetry here indicate success rather than failure of our kernel construction. For Poisson kernels, the covariance (50) is zero, as it should be because, as we noted from the characteristic function, the coordinates of $\mathbf{Z}$ are independent in the Poisson case. This compares well with the slight dependence observed for the Poisson $\mathbf{X}$.

6. Formulae for calculating the probabilities

The actual probabilities for a kernel are needed for various numerical calculations, for example finding expected values of nonlinear functions that are not polynomials, exponential or trigonometric functions. However, to use these kernels for simulation, the probabilities are not needed. Instead one generates values of $\mathbf{X}$, by generating values of N, and adding up N independently generated values of the individual stepping-stone steps, $\mathbf{S}_i$. Some potential distributions for N (e.g. negative binomial) are included in standard simulation packages. Otherwise, the number of steps, N, would be generated by first generating U, and substituting U as the mean in a Poisson random number generator. If a given simulation package does not include a random number generator for the particular U distribution needed, then acceptance-rejection methods might be used (Johnson et al., 1994).

Probabilities can be obtained from the Fourier transform, $\tilde{K}(\mathbf{t}) = \pi_N(\cos(\mathbf{t}))$, by using the inverse Fourier transform:

$$\begin{aligned} K(\mathbf{x}) &= \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} \pi_N(\cos(\mathbf{t}))e^{-i\mathbf{t}\cdot\mathbf{x}}\,d\mathbf{t} \\ &= \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} \pi_N(\cos(\mathbf{t}))\prod_{j=1}^{d}\cos(t_j x_j)\,d\mathbf{t}, \end{aligned} \tag{51}$$

where the last step uses the fact that $\cos(\mathbf{t})$, and hence $\pi_N(\cos(\mathbf{t}))$, is an even function of each component of $\mathbf{t}$. The probabilities for these kernels can be obtained by standard numerical procedures. Formula (51) might also yield closed-form solutions, but at the present time we are unaware of any for these families of kernels.

An alternative to Fourier inversion is direct calculation of the probabilities by averaging the binomial kernel (4) over the distribution of N:

$$K(\mathbf{x}) = \sum_{n=0}^{\infty} B_{n,d}(\mathbf{x})P(N = n), \tag{52}$$

which in the one-dimensional case becomes

$$K(x) = \sum \binom{n}{\frac{1}{2}[n-x]} \left(\frac{1}{2}\right)^n P(N = n)$$

where the sum is over $n \geqslant |x|$, for even values of $n - x$. On making the substitution $n = 2m + |x|$, this formula becomes

$$K(x) = \sum_{m=0}^{\infty} \binom{2m+|x|}{m} \left(\frac{1}{2}\right)^{2m+|x|} P(N = 2m + |x|). \tag{53}$$

We have found the most difficult problem in the application of this formula to be the calculation of the summand to sufficient accuracy for very large m, which is needed in highly leptokurtic cases. However, such numerical problems are surmountable, and specific procedures for calculations are outlined in Appendix D.

7. Discussion

Because measured dispersal in nature often deviates markedly from Gaussian, families of dispersal kernels that allow various departures from the Gaussian case are needed for spatial population studies. Such families of kernels are available for continuous space (Clark et al., 1999; Yamamura, 2002). However, models where populations are distributed on integer lattices have an

important theoretical role, and so families of dispersal kernels for integer lattices, which we provide here, are needed too.

The particular technique that we use, where a random number of steps of one kernel converts it into another, has also been exploited implicitly by Clark et al. (1999) and explicitly by Yamamura (2002) in their constructions of kernels for continuous space. Our negative binomial kernels are the discrete counterpart of the kernels of Yamamura (2002), which involve gamma distributed step duration (the continuous-time counterpart of the number of steps) in a Gaussian diffusion process. On the other hand, Clark et al. (1999) in effect use a reciprocal of a gamma for step duration in the same Gaussian diffusion process to obtain t-distribution kernels that have asymptotic power function tails, a feature that describes well the observed dispersal of many plant seeds (Clark et al., 1999; Portnoy and Willson, 1993).

The t-distribution kernels of Clark et al. (1999) have been especially useful in modeling dispersal of seeds of forest trees. They include kernels with extreme leptokurtosis, but Gaussian kernels occur as a limiting case. We have proved here that our Poisson reciprocal gamma kernels have the same asymptotic tail properties as the continuous kernels of Clark et al. (1999). The Poisson reciprocal gamma kernels are a limiting case of the Sichel kernels, which include also the negative binomial kernels as another limiting case. These Poisson reciprocal gamma kernels are thus part of a large flexible family. However, the kernels constructed here appear all to have either asymptotically negative exponential tails, or asymptotic power function tails. These characteristics ultimately derive from the tail properties of the distribution of U, which is used to generate the N distribution. Some dispersal studies have used kernels with tails that are negative exponentials of powers (e.g. Brown and Bolker, 2004), which are not covered by the families explicitly considered here. However, a Weibull distribution (Johnson et al., 1994), which does have an asymptotically exponential power tail, might be chosen for U. The results here lead us to speculate that the tails of the resulting dispersal kernel will also have asymptotically exponential power tails.

By randomizing the number of steps in a simpler kernel, these kernel constructions potentially have biological or physical interpretations. For example, a population will be heterogeneous in the extent to which different individuals are prone to dispersal. Each individual might have essentially a Gaussian dispersal kernel, with the scale of dispersal varying from individual to individual, and corresponding to different numbers of steps of a kernel with unit scale. In our construction, a natural interpretation is that each individual has its own Poisson kernel, with its own value of U defining its variance. Averaging the Poisson kernels over the distribution of U then gives the kernel at the population level. However, with this particular interpretation, convolutions of the kernel cannot represent multiple steps of the same individual because they would not be statistically independent. Either a single step should represent dispersal over the life of the individual, or dependence over time should be taken into account in modeling the effects of multiple steps (Chesson, 1978).

An alternative interpretation is that the value of U represents a chance interaction between an individual and physical dispersal processes or biological vectors (Clark et al., 1999; Higgins and Richardson, 1999). For example, a seed that is released from a tree on a windy day will tend to be dispersed further than a seed released on a calm day, or a fruit discovered by a monkey near sunset is likely to have its seeds defecated by the monkey nearer the point at which it was found than the seeds of a fruit consumed earlier in the day (Chapman and Russo, in press). These explanations are all nonadaptive explanations. However, it has been suggested also that leptokurtic dispersal kernels might be favored by natural selection in landscapes with positive spatial autocorrelation in habitat favorability (Hovestadt et al., 2001).

The Sichel family of kernels developed here provides a highly flexible family with potentially broad application to modeling dispersal. Its main disadvantage is its relatively complex Fourier transform. In cases where tails are exponentially bounded, the simpler negative binomial limits of the Sichels might be preferred over general Sichel kernels, and for extreme leptokurtosis, the stable kernels provide an alternative to Sichels with their very simple Fourier transforms.

The Fourier transforms and moments of these kernels are in general the simplest features to determine. Fortunately, these features are prominent in many applications. Direct estimation of the moment generating function, a variant of the characteristic function, has been proposed (Clark et al., 2001) and put into practice (Caswell et al., 2003). Moments, and the relationships between the moments of N and the moments of the kernel, are particularly valuable in characterizing the properties of these kernels as discussed here, while the Fourier transform and the moment generating functions have proved most useful in modeling. The actual values, $K(\mathbf{x})$, of the kernel itself are in general more difficult to obtain than these other features, and require numerical methods. However, the formulae for them are straightforward, and the experience discussed in Appendix D has shown that relatively simple procedures can overcome numerical difficulties. Simulating dispersal with these kernels is a straightforward matter of generating random numbers of stepping stone steps.

Acknowledgments

Thanks are due to two anonymous reviewers for comments that have improved the manuscript.

This work was supported by NSF grant DEB-9981926 (PC), NSF DEB 9602226 (Research Training Grant in Nonlinear Dynamics in Biology, CTL) and UCD Graduate Studies Dissertation Year Fellowship (CTL).

Appendix A. Long range symmetry

The Poisson kernel satisfies two distinct forms of long-distance radial symmetry. A Poisson distribution with mean μ can be written as an independent sum of μ Poisson distributions of mean 1, if μ is an integer, and so the Poisson kernel of variance μ can be represented as the μ-fold convolution of Poisson kernels of mean 1. From the central limit theorem, the distribution of $\mathbf{X}/\sqrt{\mu}$ converges on the spherical Gaussian distribution with variance 1 for large μ. Thus, if A is any compact set in d-dimensional space, and H is an orthogonal matrix, then $P(\mathbf{X}/\sqrt{\mu} \in HA)$ is asymptotically independent of H for large μ. A different form of asymptotic radial symmetry applies to the tails, and it holds for all values of $E[N]$. Thus, we keep μ fixed. Focusing on formula (52), we note that for large $\mathbf{x}, K(\mathbf{x})$ is increasingly dominated by binomial contributions with large numbers of steps. The central limit theorem implies that these binomial contributions converge on the spherical multivariate normal, which is radially symmetric. Thus, we see that the tails of the Poisson kernel, and indeed any other kernel in which the distribution of N is unbounded, must approach radial symmetry in its tails.

Appendix B. Tail relationships

To study the behavior of the tails of dispersal kernels, we shall consider only the univariate case ($d = 1$). However, the marginal distributions of the components of $\mathbf{X}$ are the same for all d, and the radial symmetry properties of the kernel (Appendix A) mean that the tail behavior of the marginal distributions is indicative of tail behavior in general. The Pascal's triangle representation of binomial coefficients implies that the binomial kernel, $B_{n,1}(x)$, is monotone decreasing in $|x|$. Because averages of monotone decreasing functions are monotone decreasing, the kernel families presented here all have this same property.

It is not difficult to see that if a probability distribution is discrete, or is continuous and has asymptotically monotonic tails, then its moment generating function is finite in an open neighborhood of zero if and only if its tails are asymptotically bounded by a negative exponential. The relationships between various distributional transforms of X, N, and U given in the section on Fourier transforms imply that the moment generating functions of X, N, and U satisfy the following identities

$$M_X(t) = M_N(\ln(\cosh(t))) = M_U(\cosh(t) - 1). \tag{B.1}$$

Thus, finiteness of any of these in an open neighborhood of zero implies finiteness of the others. This means that asymptotic exponential boundedness of the tails of N or U is necessary and sufficient for asymptotic exponential boundedness of the tails of X. This result means that the negative binomial kernel inherits exponentially bounded tails from the gamma distribution for U, and the Sichel kernels also inherit this property from the generalized inverse Gaussian, except in the limit when the generalized inverse Gaussian becomes the reciprocal gamma.

In the case of a reciprocal gamma distribution, the tails of U are asymptotically proportional to the power $u^{\gamma-1}$, where γ is negative. We now show that the tail of $K(x)$ is asymptotically proportional to $|x|^{2\gamma-1}$. We begin by noting from (33) that the modified Bessel function $K_v(\omega)$ can be defined by the formula

$$K_v(\omega) = \frac{1}{2}\int_0^\infty u^{v-1} e^{-\frac{1}{2}\omega(u^{-1}+u)}\, du \tag{B.2}$$

for $\omega > 0$ and all real v. Change of variable also provides the integral formula

$$\int_0^\infty u^{v-1} e^{-\frac{1}{2}(au^{-1}+bu)}\, du = 2\left(\frac{a}{b}\right)^{v/2} K_v(\sqrt{ab}). \tag{B.3}$$

From (B.2) we deduce the asymptotic formula

$$K_v(\omega) \sim \tfrac{1}{2}(\tfrac{1}{2}\omega)^{-v}\Gamma(v), \tag{B.4}$$

where $\sim$ means that the ratio of the LHS to RHS approaches 1 as v approaches infinity. Comparison of (B.2) to the integral formula defining the gamma function immediately establishes the RHS of (B.4) as a strict upper bound of the LHS. Truncating the integral at an arbitrary positive value above zero can be used to establish adequate lower error bounds on (B.4) sufficient for rigorous demonstration of the result whose proof we outline here.

The density function of the reciprocal gamma is

$$\frac{\xi^{-\gamma}}{\Gamma(-\gamma)}\, u^{\gamma-1} e^{-\xi u^{-1}}, \tag{B.5}$$

where the shape parameter γ is negative and ξ is a positive scale parameter. Using the fact that all Poisson mixture kernels can be obtained by averaging the Poisson kernel $P_{U,1}(x) = e^{-U} I_{|x|}(U)$ (Eq. (8)), over the distribution of U, we see that

$$K(x) = \frac{\xi^{-\gamma}}{\Gamma(-\gamma)} \int_0^\infty e^{-u} I_{|x|}(u) u^{\gamma-1} e^{-\xi u^{-1}}\, du. \tag{B.6}$$

Now the modified Bessel function $I_{|x|}(u)$ is given by the integral formula

$$I_{|x|}(u) = \frac{(\frac{1}{2}u)^{|x|}}{\sqrt{\pi}\Gamma(|x|+\frac{1}{2})}\int_{-1}^{1}(1-t^2)^{|x|-\frac{1}{2}}e^{ut}\,dt$$

(Abramowitz and Stegun, 1964, formula 9.6.16). Substituting in (B.6) and changing the order of the integration gives the expression

$$K(x) = \frac{\xi^{-\gamma}(\frac{1}{2})^{|x|}}{\sqrt{\pi}\Gamma(-\gamma)\Gamma(|x|+\frac{1}{2})}\int_{-1}^{1}(1-t^2)^{|x|-\frac{1}{2}} \times \int_0^\infty u^{|x|+\gamma-1}e^{-(1-t)u-\xi/u}\,du\,dt. \quad \text{(B.7)}$$

Using the integral formula (B.3) for the inner integral of (B.7), and substituting the asymptotic formula (B.4), we see that $K(x)$ is asymptotically

$$\frac{\xi^{-\gamma}(\frac{1}{2})^{|x|}\Gamma(|x|+\gamma)}{\sqrt{\pi}\Gamma(-\gamma)\Gamma(|x|+\frac{1}{2})}\int_{-1}^{1}(1-t)^{-\gamma-\frac{1}{2}}(1+t)^{|x|-\frac{1}{2}}\,dt. \quad \text{(B.8)}$$

The substitution $v=(1+t)/2$ in (B.8) shows that the integral is equal to $2^{|x|-\gamma}B(|x|+\frac{1}{2},-\gamma+\frac{1}{2})$ where B is the beta function (Johnson et al., 1992, p. 7). Because the beta function obeys the formula $B(a,b)=\Gamma(a)\Gamma(b)/\Gamma(a,b)$, we see that

$$K(x) \sim \frac{(2\xi)^{-\gamma}\Gamma(|x|+\gamma)\Gamma(-\gamma+\frac{1}{2})}{\sqrt{\pi}\Gamma(-\gamma)\Gamma(|x|-\gamma+1)}. \quad \text{(B.9)}$$

Now as $|x|\to\infty$, the ratio of gamma functions $\Gamma(|x|+\gamma)/\Gamma(|x|-\gamma+1)$ is asymptotically $|x|^{2\gamma-1}$ (Johnson et al., 1992, p. 6). Thus,

$$K(x) \sim \frac{(2\xi)^{-\gamma}\Gamma(-\gamma+\frac{1}{2})}{\sqrt{\pi}\Gamma(-\gamma)}|x|^{2\gamma-1}. \quad \text{(B.10)}$$

In a similar way, one can show that the $P(N=n)$ is asymptotically proportional to $n^{\gamma-1}$.

Appendix C. Moment relationships

The moments of **X** can be expressed generally in terms of the moments of N. The text gives details for first and second moments. Here we show how higher moments can be related. We restrict attention to one dimension with only the necessary detail to demonstrate that the moments of **X** of order $2n$ are finite if and only if the nth moment of N is finite for positive integer n. The Holder inequality extends this result to multiple dimensions because it shows that mixed moments of order $2n$ are finite whenever the univariate moments of order $2n$ are finite. The symmetry properties of **X** mean that all moments of odd order are zero whenever they exist regardless of the dimension of **X**.

Because X^{2n} is nonnegative, the $2n$th moment of X can be calculated according the conditional expectation formula

$$E[X^{2n}] = E\{E[X^{2n}|N]\}, \quad \text{(C.1)}$$

regardless of whether this moment is finite. Representing X as a sum of stepping-stone steps (formula (2)), we see that

$$E[X^{2n}|N] = \sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\cdots\sum_{i_{2n}=1}^{N}E[S_{i_1}S_{i_2}\cdots S_{i_{2n}}].$$

Because $E[S_i^m]$ is 0 for odd m and 1 for even m, and the S_i are mutually independent, $E[S_{i_1}S_{i_2}\cdots S_{i_{2n}}]$ equals 1 if every integer in the subscript sequence $\{i_1,i_2,\ldots,i_n\}$ appears an even number of times. Else $E[S_{i_1}S_{i_2}\cdots S_{i_{2n}}]=0$. Thus, if $f(m,2n)$ is the number of distinct sequences of $2n$ elements from the set $\{1,2,\ldots,m\}$ with even appearances of elements ("even sequences"), then

$$E[X^{2n}|N] = f(N,2n), \quad \text{(C.2)}$$

and

$$E[X^{2n}] = E[f(N,2n)]. \quad \text{(C.3)}$$

Now $f(m,2n)$ satisfies the recurrence relationship

$$f(m,2n) = 1 + \sum_{j=1}^{m-1}\sum_{k=1}^{n}\binom{2n}{2k}f(m-j,2[n-k]), \quad \text{(C.4)}$$

where $f(m,0)=1$. This formula is derived as the sum over j and k of the number of even sequences for which j occurs $2k$ times, but $1,\ldots,j-1$ do not occur, exhausting the set of even sequences of length $2n$, from m elements. The combinatorial notation

$$\binom{2n}{2k}$$

is the number of ways the $2k$ appearances of j can be arranged in the $2n$ possible places, and $f(m-j,2[n-k])$ is the number of even sequences of length $2n-2k$ for indices other than $1,\ldots,j$. The 1 in front of the sum counts the single sequence containing $j=m$ only.

Using induction, this iteration implies that $f(m,2n)$ is a polynomial in m of order n because it is a sum from $j=1$ to $m-1$ of a polynomial of order $n-1$ in $(m-j)$, according to finite difference calculus (e.g. Johnson et al., 1992). Eq. (C.3) now yields the conclusion that the $2n$th moment of X is finite whenever the nth moment of N is finite. The reverse conclusion can be seen from the fact that finiteness of the $2n$th moment of X implies finiteness of all lower order moments of X, which allows finiteness of each moment of N up to n to be established by induction from (C.3).

Appendix D. Tips for calculating the kernel

In the text formulae, (52) and (53) define kernel calculations in terms of the distribution of N. Although straightforward in principle, these formulae require high-precision calculations for accurate results. Of most difficulty is calculating $P(N=n)$ in cases where N has a high coefficient of variation. For the case of negative binomial N in the one-dimensional case, we have obtained good results from formula (53) by combining the binomial kernel calculation and $P(N=n)$ from Table 2 into one long product, as follows:

$$\begin{aligned}\binom{2m+|x|}{m}\left(\frac{1}{2}\right)^{2m+|x|} & P(N=2m+|x|)\\ &= \left(1+\frac{\mu}{k}\right)^{-k}\prod_{j=1}^{m+|x|}\left(\frac{\frac{1}{2}\mu}{k+\mu}\right)\left(\frac{k+m-1}{j}+1\right)\\ &\quad\times\prod_{j=1}^{m}\left(\frac{\frac{1}{2}\mu}{k+\mu}\right)\left(\frac{k-1}{j}+1\right). \qquad \text{(D.1)}\end{aligned}$$

Such products can lead to overflow and underflow problems in numerical calculations even though the actual value of the product is neither exceptionally large nor exceptional small. To prevent large intermediate values from arising in these calculations, we changed the order of these products by folding the sequence over as follows: We created a sequence $\{a_i\}$ of $2m+|x|$ numbers from the terms after each product symbol in (D.1). We reversed the second half of this sequence and interleaved it with the first half to obtain a new sequence $\{b_i\}=\{a_1, a_{2m+|x|}, a_2, a_{2m+|x|-1}, \ldots\}$. We repeated this process several times until we achieved the desired accuracy of the final product. We judged the accuracy of the calculations by the precision of the total sum of the probabilities, the variance and kurtosis, all of which are known analytically. Three foldings of the sequence were generally sufficient for values of k as small as 0.02 for 10-figure accuracy for the sum of the probabilities, and 8-figure accuracy for the variance and kurtosis using Gauss 6.0 (Aptech Systems, Inc).

For Sichel distributions, which include the negative binomial as a limiting case, the following recurrence formula allows the efficient calculation of $P(N=n)$, providing an alternative to the above methods:

$$\begin{aligned}P(N=n) = \frac{2\xi\omega}{\alpha^2}\left(\frac{\gamma+n-1}{n}\right)P(N=n-1)\\ +\left(\frac{\xi\omega}{\alpha}\right)^2\frac{P(N=n-2)}{n(n-1)}, \qquad \text{(D.2)}\end{aligned}$$

(Johnson et al., 1992). The first two probabilities are calculated from the formula in Table 2, and hence only three evaluations of the Bessel functions K_ν are needed. We have used this formula in conjunction with a table of binomial kernels for efficient and accurate calculation of Sichel kernels in one and two dimensions. The table of binomial kernels was calculated by the recurrence relationship $P(X=x+2)=(n-x)P(X=x)/(n+x+2)$. For n in the thousands, this relationship was initialized at $x=0$ or 1, with this initial calculation aided by the sequence folding technique described above.

References

Abramowitz, M., Stegun, I.A., 1964. Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, Applied Mathematics. National Bureau of Standards, Washington, DC, 1046pp.

Barndorff-Nielsen, O.E., Blaesild, P., Halgreen, C., 1989. First hitting time models for the generalized inverse Gaussian distribution. Stochastic Process. Appl. 7, 49–54.

Barton, N., Depaulis, F., Etheridge, A., 2002. Neutral evolution in spatially continuous populations. Theoret. Population Biol. 61, 31–48.

Billingsley, P., 1986. Probability and Measure. Wiley, New York.

Bolker, B., Pacala, S., 1999. Spatial moment equations for plant competition: understanding spatial strategies and the advantages of short dispersal. Amer. Naturalist 153, 575–602.

Bolker, B.M., Pacala, S.W., Neuhauser, C., 2003. Spatial dynamics in model plant communities: what do we really know? American Naturalist 162, 135–148.

Briggs, C., Hoopes, M.F., 2004. Stabilizing effects in a spatial parasitoid–host and predator–prey models: a review. Theoret. Population Biol. 65, 299–315.

Brown, D.H., Bolker, B.M., 2004. The effects of disease dispersal and host clustering on the epidemic threshold in plants. Bull. Math. Biol. 66, 341–371.

Cain, M.L., Damman, H., Muir, A., 1998. Seed dispersal and the Holocene migration of woodland herbs. Ecol. Monogr. 68, 325–347.

Caswell, H., Lensink, R., Neubert, M.G., 2003. Demography and dispersal: life table response experiments for invasion speed. Ecology 84, 1968–1978.

Chapman, C.A., Russo, S.E., in press. Primate seed dispersal: linking behavioral ecology with forest community structure. In: Campbell, C.J., Fuentes, A.F., MacKinnon, K.C., Panger, M., Bearder, S. (Eds.), Primates in Perspective. Oxford University Press, Oxford.

Chesson, P.L., 1978. Models for animal movements. Ph.D. Thesis, The University of Adelaide, Adelaide.

Clark, J., Silman, M., Kern, R., Macklin, E., HilleRisLambers, J., 1999. Seed dispersal near and far: patterns across temperate and tropical forests. Ecology 80, 1475–1494.

Clark, J., Horvath, L., Lewis, M., 2001. On the estimation of spread rate for a biological population. Statist. Probab. Lett. 51, 225–234.

Comins, H.N., Hassell, M.P., May, R.M., 1992. The spatial dynamics of host-parasitoid systems. J. Anim. Ecol. 61, 735–748.

De Roos, A.M., McCauley, E., Wilson, W.G., 1998. Pattern formation and the spatial scale of interaction between predators and their prey. Theoret. Population Biol. 53, 108–130.

Dieckmann, U., Law, R., Metz, J.A.J., 2000. The Geometry of Ecological Interactions: Simplifying Spatial Complexity, Cambridge Studies in Adaptive Dynamics. Cambridge University Press, Cambridge, 564pp.

Fang, K.T., Kotz, S., Ng, K.W., 1990. Symmetric Multivariate and Related Distributions. Chapman & Hall, London.

Feller, W., 1971, An Introduction to Probability Theory and its Applications, second ed., vol. II. Wiley, New York.

Hassell, M.P., Comins, H.N., May, R.M., 1994. Species coexistence and self-organizing spatial dynamics. Nature 370, 290–292.

Higgins, S.I., Richardson, D.M., 1999. Predicting plant migration rates in a changing world: the role of long-distance dispersal. Amer. Naturalist 153, 464–475.

Hovestadt, T., Messner, S., Poethke, H.J., 2001. Evolution of reduced dispersal mortality and "fat-tailed" dispersal kernels in autocorrelated landscapes. Proc. Roy. Soc. London, Seri. B 268, 385–391.

Ibrahim, K.M., Nichols, R.A., Hewitt, G.M., 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. Heredity 77, 282–291.

Johnson, N.J., Kotz, S., Kemp, A.W., 1992. Univariate Discrete Distributions. Wiley, New York.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. Continuous Univariate Distributions, vol. 1. Wiley, New York.

Kareiva, P., Wennergren, U., 1995. Connecting landscape patterns to ecosystem and population processes. Nature 373, 299–302.

Kimura, M., Weiss, G.H., 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49, 561–576.

Kinezaki, N., Kawasaki, K., Takasu, F., Shigesada, N., 2003. Modeling biological invasions into periodically fragmented environments. Theoret. Population Biol. 64, 291–302.

Klausmeier, C.A., 1999. Regular and irregular patterns in semiarid vegetation. Science 284, 1826–1828.

Kot, M., Lewis, M.A., van den Driessche, P., 1996. Dispersal data and the spread of invading organisms. Ecology 77, 2027–2042.

Lande, R., 1991. Isolation by distance in a quantitative trait. Genetics 128, 443–452.

Lande, R., Engen, S., Saether, B.E., 1999. Spatial scale of population synchrony: environmental correlation versus dispersal and density regulation. Amer. Naturalist 154, 271–281.

Levin, D.A., Kerster, H.W., 1975. The effects of gene dispersal on the dynamics and statics of gene substitution in plants. Heredity 35, 317–336.

Levin, S.A., 1992. The problem of pattern and scale in ecology. Ecology 73, 1943–1967.

Lewis, M.A., Pacala, S., 2000. Modeling and analysis of stochastic invasion processes. J. Math. Biol. 41, 387–429.

Malécot, G., 1969. The Mathematics of Heredity. Freeman, San Francisco.

Mardia, K.V., 1970. Measures of multivariate skewness and kurtosis with application. Biometrika 57, 519–530.

Murrell, D.J., Law, R., 2003. Heteromyopia and the spatial coexistence of similar competitors. Ecol. Lett. 6, 48–59.

Nichols, R.A., Hewitt, G.M., 1994. The genetic consequences of long-distance dispersal during colonization. Heredity 72, 312–317.

Ouborg, N.J., Piquot, Y., Van Groenendael, J.M., 1999. Population genetics, molecular markers and the study of dispersal in plants. J. Ecol. 87, 551–568.

Portnoy, S., Willson, M.F., 1993. Seed dispersal curves: behavior of the tail of the distribution. Evol. Ecol. 7, 25–44.

Roughgarden, J., 1977. Patchiness in the spatial distribution of a population caused by stochastic fluctuations in resources. Oikos 29, 52–59.

Roughgarden, J., 1978. Influence of competition on patchiness in a random environment. Theoret. Population Biol. 185–203.

Rousset, F., 2000. Genetic differentiation between individuals. J. Evol. Biol. 13, 58–62.

Snyder, R.E., Chesson, P., 2003. Local dispersal can facilitate coexistence in the presence of permanent spatial heterogeneity. Ecol. Lett. 6, 301–309.

Snyder, R.E., Chesson, P., 2004. How the spatial scales of dispersal, competition, and environmental heterogeneity interact to affect coexistence. Amer. Naturalist 164, 633–650.

Southwood, R., Henderson, P.A., 2000. Ecological Methods. Blackwell Science, Oxford.

Thomson, N.A., Ellner, S.P., 2003. Pair-edge approximation for heterogeneous lattice population models. Theor. Population Biol. 64, 271–280.

Wilson, W.G., Harrison, S.P., Hastings, A., McCann, K., 1999. Exploring stable pattern formation in models of tussock moth populations. J. Animal Ecol. 68, 94–107.

Yamamura, K., 2002. Dispersal distance of heterogeneous populations. Population Ecol. 44, 93–101.