JOURNAL OF THE ROYAL SOCIETY INTERFACE

A unifying theory for 2d spatial redistribution kernels with applications in population spread modelling

Journal:	*Journal of the Royal Society Interface*
Manuscript ID	rsif-2020-0434.R2
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Koch, Dean; University of Alberta, Mathematical and Statistical Sciences Lewis, Mark; University of Alberta, Mathematical and Statistical Sciences; University of Alberta, Biological Sciences Lele, Subhash; University of Alberta, Mathematical and Statistical Sciences
Categories:	Life Sciences - Mathematics interface
Subject:	Biomathematics < CROSS-DISCIPLINARY SCIENCES, Computational biology < CROSS-DISCIPLINARY SCIENCES, Biogeography < CROSS-DISCIPLINARY SCIENCES
Keywords:	movement, redistribution, kernel, diffusion, fractal

Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:
This article does not present research with ethical considerations

Statement (if applicable):
CUST_IF_YES_ETHICS :No data available.

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:
Yes

Statement (if applicable):
All data and R code required to reproduce our analysis have been supplied as electronic supplementary material.

Conflict of interest

I/We declare we have no competing interests

Statement (if applicable):
CUST_STATE_CONFLICT :No data available.

Authors' contributions

This paper has multiple authors and our individual contributions were as below

Statement (if applicable):
Dean Koch wrote the first draft of the manuscript and participated in all stages of the project. Mark Lewis and Subhash Lele contributed to the model development and critically revised the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

A unifying theory for 2d spatial redistribution kernels with applications in population spread modelling

Dean C. Koch[1,*], Mark A. Lewis[2], Subhash R. Lele[3]

University of Alberta, Edmonton, Canada, T6G 2R3

Abstract

When building models to explain the dispersal patterns of organisms, ecologists often use an isotropic redistribution kernel to represent the distribution of movement distances based on phenomenological observations or biological considerations of the underlying physical movement mechanism. The Gaussian, 2d Laplace and Bessel kernels are common choices for 2-dimensional (2d) space. All three are special (or limiting) cases of a kernel family, the WMY, first derived by Yasuda from an assumption of 2d Fickian diffusion with gamma distributed settling times. We provide a novel derivation of this kernel family, using the simpler assumption of constant settling hazard, by means of a non-Fickian 2d diffusion equation representing movements through heterogeneous 2-dimensional media having a fractal structure. Our derivation reveals connections among a number of established redistribution kernels, unifying them under a single, flexible modeling framework. We demonstrate improvements in predictive performance in an established model for the spread of the mountain pine beetle upon replacing the Gaussian kernel by the WMY, and report similar results for a novel approximation, the pWMY, that substantially speeds computations in applications to large datasets.

Keywords: *movement*, *redistribution*, *kernel*, *diffusion*, *fractal*

*Corresponding author

[1]Dean C. Koch (dkoch@ualberta.ca) is postdoctoral fellow in the Department of Mathematical and Statistical Sciences at the University of Alberta (UofA)

[2]Mark A. Lewis (mark.lewis@ualberta.ca) is Professor of Mathematical and Statistical Sciences and Biological Sciences at UofA

[3]Subhash R. Lele (slele@ualberta.ca) is Professor of Mathematical and Statistical Sciences at UofA

1. Introduction

Ecologists are concerned with questions of an inherently spatial nature, since movement and environmental heterogeneity are ubiquitous features of the natural world. They are therefore often rewarded by new insights when the mechanism underlying a spatial effect can be worked into a process model [1]. *Redistribution kernels* are a popular means to this end, with applications as diverse as predator-prey interactions [2]; range expansion and invasion biology [3]; grouping/swarming behaviour [4]; chemical communication [5]; and cellular transport [6].

Statistical ecologists study many of the same questions, but with a focus on characterizing the randomness in measurements. This, too, is a spatial problem. Ecological datasets typically exhibit spatial autocorrelation (SAC) [7], which, if ignored, reduces the precision of estimators [8]. A common solution is to use phenomenological models, known as *covariograms* (or covariance kernels), that map separation distances to correlations [9].

Covariograms are qualitatively similar to redistribution kernels – both describe a tapering of relatedness with spatial scale, so it is perhaps unsurprising that the same function families often take on both roles. An example that we find particularly interesting is the function known variously as the Whittle-Matérn [11] or K-Bessel kernel [12]. Its versatility and mathematical elegance make it one of the most important covariograms in spatial statistics [13]. Yet, in spite of many advantages over more familiar alternatives (*eg.* the Gaussian), this function receives little attention in the context of movement modelling.

Up to a normalization constant (and a restriction on the shape parameter), the Whittle-Matérn is identical to the K-Bessel redistribution kernel first derived by Yasuda [12], and later by Yamamura [14] and Hapca *et al.* [15]. We focus in this paper on its potential as a movement model, offering a new mechanistic derivation. However, as the literature on covariograms contains decades worth of analysis into its mathematical properties, we refer to this function as the Whittle-Matérn-Yasuda (WMY) kernel.

1.1. Redistribution kernels

A redistribution kernel $D(\boldsymbol{x}, \boldsymbol{x}')$ maps pairs of coordinates (source $\boldsymbol{x}'$ and destination $\boldsymbol{x}$) to the probability density for the redistribution event from $\boldsymbol{x}'$ to $\boldsymbol{x}$. When modeling movement events having a random character, kernels provide a simple means of parametrizing the probability density function (PDF) for position following the movement event. For example, in process-based models, redistribution kernels are often derived from partial differential equations (PDE) representing the (time) evolution of the position PDF during a random walk [16].

The simplest such models are stationary, meaning their kernels are functions of the separation vector $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{x}'$ only, independent of location. When isotropy (radial symmetry) is also assumed, kernels can be defined more simply as a function of $r = |\boldsymbol{r}|$. Of these simpler kernels, the isotropic Gaussian (Table 1) is the most widely used. It solves a PDE for Fickian diffusion proceeding until a particular (fixed) time.

shape (κ)	kernel name	density* $D(r)$	mechanistic derivation
$(-1, 0)$	-	$(r/\rho)^\kappa K_\kappa(r/\rho)$	2d fractal diffusion with constant settling hazard (Section 2.1)
0	Bessel	$K_0(r/\rho)$	2d Fickian diffusion with constant settling hazard [17]
1/2	2d Laplace	$\exp(-r/\rho)$	2d turbulent diffusion with instantaneous settling [18]
∞	Gaussian	$\exp(-(r/\rho)^2)$	2d Fickian diffusion with instantaneous settling [19]

Table 1: Notable examples of stationary isotropic kernels from the WMY family, $\mathcal{D}(r; \kappa, \rho)$. All arise from 2d Fickian diffusion with gamma-distributed settling times [12]. The special cases listed here have been derived independently under various movement models. Alternatively, the full WMY family can be derived by repeated iterations of the kernel in the top row (Section 2.2). *for brevity the normalization $1/2\pi \int_{-\infty}^{\infty} rD \, \mathrm{d}r$ is omitted.

With 2-dimensional (2d) isotropic models, one must take care to distinguish between the full density $D(r) = D(r, \theta)$ of the random point (R, Θ) and the marginal density $D_r(r)$ of the random radius R (*ie.* distance). The two are related by

$$D_r(r) := \int_0^{2\pi} D(r, \theta) r \, \mathrm{d}\theta = 2\pi r D(r, \theta) = 2\pi r D(r), \tag{1}$$

where factor r in the integrand is the Jacobian of the polar transformation, so that $\int_{\mathbb{R}^2} D = \int_0^\infty D_r(r) \, \mathrm{d}r = 1$. This distinction between marginal and full densities is important, but sometimes unclear in the ecological literature. Our notation will always indicate the marginal by a subscript, as we do in (1).

The WMY is an important example of a stationary and isotropic kernel:

$$\mathcal{D}(r; \kappa, \rho) = A(\kappa, \rho) (r/\rho)^\kappa K_\kappa (r/\rho), \quad \text{with } 1/A(\kappa, \rho) = 2^{\kappa+1} \pi \rho^2 \Gamma(\kappa + 1), \tag{2}$$

where K_κ denotes the κ^{th} order modified Bessel function of the second kind (Appendix 1.1), $\rho > 0$ is a distance-scaling (range) parameter, and κ is a kurtosis (shape) parameter. The domain of κ depends on the application: For covariance, $\kappa > 0$ (to guarantee positive definiteness); and for redistribution, $\kappa > -1$ (to guarantee integrability). By studying κ we will see that the WMY is closely related to a number of other kernels in common use, generalizing some the most prominent ones (Table 1), while providing a more flexible range of tail behaviours (see Section 2.3).

This makes the WMY an unusually flexible model for the spread of populations. An important example from our research area is the mountain pine beetle (MPB). Dispersal flights of this forest pest allow populations to spill outward into neighbouring areas, and are therefore a key component of models for the progression of outbreaks moving across the landscape. We discuss the use of redistribution kernels in such applied situations in Section 3, showing how the WMY (and a computationally efficient approximation, the pWMY) can improve model fit when used in place of the less flexible Gaussian kernel in a popular model for MPB forest damage patterns.

The bulk of this paper, however, is devoted to developing a new process-based mechanism to explain how this kernel arises. To motivate the use of equation (2) more generally, we derive the WMY kernel as the solution to a partial differential equation (PDE) for diffusion through inhomogeneous habitat, with a constant hazard of settling

(Section 2). In the course of this derivation we highlight the many appealing mathematical properties of the WMY, and draw connections between the WMY and number of other, better-established, isotropic models for redistribution.

2. The WMY as a model for diffusion with settling

In Yasuda's 1975 paper [12], the WMY was derived as the settled density solution to a Fickian diffusion process with gamma-distributed settling times. Equation (2) can therefore be understood as an extension of the Bessel kernel (Table 1), with a more general characterization of stopping times. Both models, however, assume the random walk takes place in unrestricted 2d space. We will relax this assumption, by assuming instead that obstacles in the environment may inhibit movements to some degree.

In statistics, the WMY is derived from the stationary random field solution of a fractional stochastic PDE resembling a generalized Helmholtz equation [20]. This suggests that as a redistribution kernel, the WMY might also solve a similar deterministic PDE involving fractional derivatives. These exotic dynamical systems often appear in connection with the statistical mechanics of (non-Fickian) diffusion through complex media that hinder movement. They are studied in physics, for example, to explain the physical and chemical properties of porous substrates [21]. Here we adapt those results in the ecological context to find a versatile macro-scale description of random walks.

2.1. Diffusion over fractal media in 2d

To constrain the movements of our random walker, we view the domain of movement as a porous 2d medium with a fractal structure, rather than a full 2d space. This more faithfully reflects the type of movement behaviours that actively track patchy arrangements of habitat, or navigate around complex arrangements of physical obstacles.

Fractal environments are not an unfamiliar idea in ecology. Fractal aspects of tree crown cover, hydrological networks, and topography have been recognized for decades [22]. These structures often exhibit power laws under scaling that allow their pertinent features to be summarized by simple parameters, such as the Hausdorff dimension d_f [23]. The flight of a MPB, for example, is constrained within a complex network of

gaps in the forest vegetation. Like many aspects of forest structure, the porosity of this abstract medium can be conveniently described using d_f [24]. For example Jonckheere *et al.* improved predictions of light penetration in pine forests by estimating the d_f value in hemispherical images of scots pine canopy gaps, finding them to be highly fractal [25]. We will use d_f to describe the porosity of the patchy space navigated by random walker.

d_f summarizes space-filling properties. In full 2d space, the area enclosed in a disc of radius r scales as πr^2, whereas in an embedded fractal space it scales as πr^{d_f} [26]. Unlike the topological dimension (in our case, $d = 2$), Hausdorff dimension can assume non-integer values. One can define spaces with $0 < d_f < 2$ that nearly fill the plane, yet leave a complex arrangement of patches inaccessible, with the availability of habitat decreasing with d_f. [27], for example, used Hausdorff dimension to characterize the quality of marten habitat, estimating a d_f in the range of 1.7-1.9 for pine and spruce forests in their Utah study area.

In the context of random walks, a fractal medium offers less space for movement. This prompts some adjustments of the balance law behind Fickian diffusion: Suppose $u(r,t)$ is a PDF for occupancy within the available space at radius r, at time t. If all of $\mathbb{R}^2$ were available, the density within the annulus Ω would be measured by $2\pi \int_\Omega u(r,t) r \, \mathrm{d}r$. On a d_f-dimensional fractal it is:

$$2\pi \int_\Omega u(r,t)p(r)r \, \mathrm{d}r \quad \text{where } p(r) = r^{d_f-2}, \quad \text{and } 0 < d_f \leq 2. \tag{3}$$

The scaling function $p(r) = \pi r^{d_f}/\pi r^2$ is the proportion of the area inside radius r that is available for movement. For notational convenience we will suppress this dependence on r (and d_f) and simply write p. Function $u(r,t)$ is therefore an occupancy PDF with the (unusual) distance-scaled probability measure $pr \, \mathrm{d}r$. Under the more familiar Lebesgue measure, the PDF is $D(r,t) = u(r,t)p$ with measure $r \, \mathrm{d}r$.

In their 1985 paper [28], OShaughnessy and Procaccia described how the usual equation for Fickian diffusion may be modified in order to remain consistent with the scaling property (3). Their generalized 2d heat equation describes the time-evolution

of $D(r,t)/p = u(r,t)$:

$$\frac{\partial u}{\partial t} = \frac{1}{pr}\frac{\partial}{\partial r}\left(\alpha pr\frac{\partial u}{\partial r}\right) \quad \text{with } \alpha > 0,\ u(r,0) = \frac{\delta(r)}{2\pi pr}, \tag{4}$$

where $\delta(r)$ denotes the 1d Dirac delta function (and $\delta(r)/2\pi pr$ its 2d analogue), representing the initial departure of the disperser from the origin. Notice the diffusivity coefficient $\alpha p = \alpha r^{d_f-2}$ scales with distance, approximating correlations in movements due to geometrical constraints [21]. Under the condition $u \to 0$ as $r \to \infty$ (required for a valid PDF), the general solution to (4) is known. Switching to Lebesgue measure, it can be written:

$$D(r,t) = u(r,t)p = \left(\pi\Gamma(d_f/2)(4\alpha t)^{d_f/2}\right)^{-1} r^{d_f-2}\exp\left(-r^2/4\alpha t\right). \tag{5}$$

$D(r,t)$ is the PDF for the position (r,θ) at time t of a random walker that departs the origin at $t = 0$ and diffuses through a medium with Hausdorff dimension d_f. The effect of decreasing d_f is to make this density function fatter-tailed, with density shifted from the shoulders of the distribution towards its extremities.

Note that $u(r,t)$ is simply a Gaussian kernel renormalized for measure $pr\,\mathrm{d}r$. Indeed when $d_f = 2$ we have $D(r,t) = u(r,t)$ (since $p = 1$), and equation (4) is the classical 2d heat equation in radial coordinates, with the 2d Gaussian kernel its well-known solution [4]. When $d_f = 1$ equation (4) simplifies to become a 1d diffusion equation (since $pr = 1$), so in the resulting *marginal* density function $D_r = 2\pi u(r,t)pr = 2\pi u(r,t)$ we find the 1d Gaussian kernel.

2.1.1. Settling via constant hazard

To build a redistribution kernel directly from (4)-(5), one could simply assume movement proceeds until a particular fixed time t. However, in reality, the duration of dispersal is often stochastic. Unpredictable environmental factors such as temperature may compel the random walker to wait out unfavourable conditions [29]. Moreover, settling events can be prompted by chance encounters, such as the detection of a prey item [30] or mate [12].

Preferring a model that accounts for randomly cued settling events, we suggest a simple extension of (4) that introduces a constant settling hazard $\lambda > 0$. We then define

our kernel as the total settled density over all time. Thus we have $D = \int_0^\infty \lambda u p \, \mathrm{d}t$, with p defined as in (3), and:

$$\frac{\partial u}{\partial t} = \frac{\alpha}{pr}\frac{\partial}{\partial r}\left(pr\frac{\partial u}{\partial r}\right) - \lambda u \quad \text{with } u(r,0) = \frac{\delta(r)}{2\pi pr} \text{ and } \lim_{r\to\infty} u(r,t) = 0. \tag{6}$$

Here the disperser moves about the domain as in (4) but settles at a randomly determined time, drawn from an exponential distribution with mean $1/\lambda$. $D(r)$ now expresses the PDF for position at the time of settling. To find D, one can integrate the PDE (6) over all time and consider weak solutions u (Appendix 1.3.2). The resulting kernels are the singular members of the WMY family:

$$D(r) = \int_0^\infty \lambda u p \, \mathrm{d}t = \mathcal{D}(r;\kappa,\rho) \text{ where } \kappa = d_f/2 - 1 \text{ and } \rho^2 = \alpha/\lambda. \tag{7}$$

These WMY kernels ($-1 < \kappa \leq 0$) also emerge as the long-time limit of $u(r,t)p$ when the point source is stationary in time rather than instantaneous. To see this, we modify (6) by viewing λ as a mortality hazard and adding a source term $F = \lambda\delta(r)/2\pi pr$ to the right-hand-side of the PDE. Dispersers are therefore continuously introduced from the origin, and continously removed throughout $\mathbb{R}^2$ in a density-dependent manner. The steady state in this smokestack-like system is $\lim_{t\to\infty} u(r,t)p = \mathcal{D}(r;\kappa,\rho)$, with κ, ρ defined as in (7) (Appendix 1.3.1).

The (non-fractal) case of $d_f = 2$ produces a Bessel kernel, or $\mathcal{D}(r;0,\rho)$. This model has a long history in ecology, as a description of the movements of worms [17] and moths [31]; and later as a model for the biological activity of a chemical diffusing outward from the center of a petri dish [32]. More recently, approximations of the Bessel kernel have been used to model bark beetle dispersal flights [30, 49].

2.2. Multi-stage extensions

Redistribution events may naturally split into multiple stages. For example, diurnal periods of flight activity occur in many insect orders [33]. Moreover forest-dwelling insects like the MPB may initially fly in the unrestricted space above the canopy before switching to subcanopy dispersal [29]. In this section we look at two simple ways of extending (6) to model multi-stage processes.

2.2.1. Switching via constant hazard

The first idea is to connect each stage by a switching hazard with rate constant $\lambda > 0$. Writing $u_m(r,t)p$ for the density in the m^{th} stage, we assume an instantaneous point release of unit density initializes dispersers in the first stage, so that the dynamics of u_1 are the same as u in (6). As time progresses, this initial impulse trickles through subsequent stages, eventually exiting the n^{th} stage at rate λ (as settled density). We will be interested in the long-term settled density:

$$D(r) = U_n(r)p = \lambda \int_0^\infty u_n(r,t)p \, \mathrm{d}t, \text{ where} \tag{8a}$$

$$\frac{\partial u_1}{\partial t} = \frac{\alpha}{pr}\frac{\partial}{\partial r}\left(pr\frac{\partial u_1}{\partial r}\right) - \lambda u_1, \text{ with } u_1(r,0) = \frac{\delta(r)}{2\pi pr}, \tag{8b}$$

and, for $1 < m \leq n$,

$$\frac{\partial u_m}{\partial t} = \frac{\alpha}{pr}\frac{\partial}{\partial r}\left(pr\frac{\partial u_m}{\partial r}\right) + \lambda u_{m-1} - \lambda u_m, \text{ with } u_m(r,0) = 0. \tag{8c}$$

Assuming $\lim_{r\to\infty} u_m(r,t) = 0$, an analytic solution is available (Appendix 1.3.3):

$$D(r) = U(r)p = \left(2^{n-1+d_f/2}\pi\rho^2\Gamma(n)\Gamma(d_f/2)\right)^{-1}(r/\rho)^{n-2+d_f/2}\,K_{n-d_f/2}(r/\rho), \tag{9}$$

where $\rho^2 = \alpha/\lambda$. The PDF in (9) is simply the (renormalized) product of the WMY kernel $\mathcal{D}(r; n - d_f/2, \rho)$ with the scaling function p. Thus in the non-fractal case of $d_f = 2$ it produces the kernel family $D = U = \mathcal{D}(r; n-1, \rho)$. This extends the Bessel kernel ($n = 1$) to yield a sequence of PDFs that are concave and bounded in their approach to the origin. These are special cases of Yasuda's gamma-distributed settling time model [12], if we interpret settling time as a sum of exponentials (representing time spent in each of the stages).

For $d_f < 2$, however, the distribution (9) remains unbounded at the origin for all n. In general, by increasing n we shift density away from the tails, effectively stalling dispersers near the origin. Decreasing d_f has the opposite effect, producing fatter tails and a highly peaked shape.

This approach of linking PDEs for n-stage processes was suggested by Neubert, Kot and Lewis [2] to describe (non-fractal) diffusion with settling on the real line. In their 1d model, the solution is a product of a Laplace kernel and a polynomial of order $n - 1$. These same kernels were later proposed as extensions of the Laplace that are

robust to changes in sampling frequency [34]. We show in Appendix 1.3.3 that they emerge also from (8)-(9), as the *marginal* density functions (D_r) for $d_f = 1$.

2.2.2. Convolutions of WMY kernels

A simpler approach to the n-stage modeling problem is to suppose a long time delay separates stages. Thus within each stage, we assume the diffusion process (6) operates until (nearly) all density has settled. The $t \to \infty$ limit in (7) then becomes initial data for the next stage. Under this assumption, the n^{th} stage settled density can be written as a convolution of n WMY kernels.

This is because when a population independently undergoes the redistribution process represented by D twice in succession, the resulting composite kernel is the autoconvolution $D * D$. More generally if a population undergoes a sequence of n independent redistribution stages described by the kernels $D^{(1)}, D^{(2)}, \ldots D^{(n)}$ then their combined effect is $D = D^{(1)} * D^{(2)} * \cdots * D^{(n)}$.

Certain kernels have the property of closure under n-part convolutions. Chesson and Lee explained how this property aids interpretability, using it to develop redistribution kernels for lattice data [35]. The Gaussian kernel is an example. The WMY is another, provided the range parameter $\rho > 0$ is fixed in all stages (Appendix 1.2). In particular if $\mathcal{D}^{(m)} = \mathcal{D}(r; \kappa_m, \rho)$ is a 2d WMY,

$$\mathcal{D}^{(1)} * \mathcal{D}^{(2)} * \cdots * \mathcal{D}^{(n)} = \mathcal{D}\left(r; \kappa, \rho\right), \quad \text{where } \kappa = n - 1 + \sum_{m=1}^{n} \kappa_m. \tag{10}$$

Note that we may assume the shape parameters κ_m belong to $(-1, 0]$, since $0 < d_f \leq 2$. So by allowing arbitrary $n > 0$, the model (10) generates the complete set of thinner-tailed and bounded WMY kernels ($0 < \kappa < \infty$) not captured by the single stage model (6) in the previous section.

By adding stages we increase κ, shifting density away from the tails and origin, and towards the shoulders of the distribution. Thus as κ increases through zero, $\mathcal{D}$ becomes bounded at the origin; and as κ increases through $1/2$, its approach to the origin switches from convex to concave (Appendix 1.2). With further increases in κ, the WMY increasingly resembles a 2d Gaussian kernel; In fact if we parametrize $\rho^2 \propto 1/\kappa$, then $\mathcal{D}$ limits to a Gaussian kernel as $\kappa \to \infty$ [36].

When $\kappa = 1/2$ the WMY simplifies to $\mathcal{D}(r; 1/2, \rho) = (2\pi\rho^2)^{-1} \exp(-r/\rho)$, which has the same functional form as a (1d) Laplace kernel. The Laplace is the 1d analogue of the Bessel; it arises from diffusion with constant-hazard settling in $\mathbb{R}^1$ [4]. In the 2d setting, this exponential form arises from turbulent limnological diffusion [18]. However in ecological applications it is frequently invoked simply as a phenomenological model [*eg.* as in 37, 38]. Equations (6)-(7) and (10) provide a new mechanistic origin for this *2d Laplace* kernel. $\kappa = 1/2$ can arise, for example, in a two-stage process where the first stage of diffusion takes place in unrestricted 2d space (with $d_f = 2$) and the second in a fractal medium with $d_f = 1$ (so that $\kappa_1 = 0$ and $\kappa_2 = -1/2$).

Note that if all the κ_m are identical, (10) expresses that the WMY family is robust to changes in sampling frequency. In fact Schlägel and Lewis described the WMY family implicitly (in Fourier space) as their first example of a 2d kernel with this property [34]. Equation (2) provides its explicit (back-transformed) density function.

2.3. Flexibility in kurtosis

The foregoing derivations are meant to illustrate how WMY patterns of redistribution might arise in a very wide range of ecological systems. In ecology, inhomogeneous environments are the rule rather than the exception. It therefore makes sense to relax the assumption of unrestricted movement while retaining it as a special case (the Bessel), as we did in the model (6). The natural extension in (10) generates a wide spectrum of kernel shapes, ranging from highly singular and heavy-tailed examples like the Bessel, to the highly smooth and thin-tailed Gaussian kernel. Let us now consider how this flexibility in shape is also attractive from a phenomenological standpoint.

One of the more important factors to consider when modeling redistribution is the balance of peakedness and tailedness in a kernel, or its *kurtosis*. In $\mathbb{R}^1$, kurtosis is the standardized fourth moment. In $\mathbb{R}^2$, we use the bivariate kurtosis measure $k(D)$ suggested in [39],

$$k(D) = \frac{\int_0^\infty r^4 D_r \, \mathrm{d}r}{\left(\int_0^\infty r^2 D_r \, \mathrm{d}r\right)^2}. \tag{11}$$

This definition is similar to the univariate case, but differs in using moments of D_r (rather than D), and centering them on zero (rather than the nonzero marginal

mean of r). This concisely summarizes the relevant behaviour of isotropic kernels; Larger values of $k(D)$ indicate fatter tails (a higher density assigned to long-distance events) and a sharper peak at the origin. Note that while we use the term "fat-tailed" to describe kernels with higher kurtosis than the Gaussian (here, and throughout the paper), tail fatness is sometimes defined in terms more subtle than equation (11). A detailed discussion of this topic can be found in [10].

For the WMY, kurtosis is determined entirely by κ (Appendix 1.2):

$$k(\mathcal{D}) = 2\left(\frac{\kappa+2}{\kappa+1}\right), \text{ where } \kappa \in (-1, \infty). \tag{12}$$

The 2d Gaussian kernel sits at the low end of this range with a kurtosis of 2 (recall it is the $\kappa \to \infty$ limit of $\mathcal{D}$). The WMY family extends it to capture a wide range of leptokurtic (higher than Gaussian) alternatives. The 2d Laplace, for example, has kurtosis $10/3$; For the Bessel it is 4. The 1-stage diffusion model (7) spans $(4, \infty)$, with extremely fat-tailed examples emerging as d_f becomes small. This kind of flexibility in kurtosis is of particular importance in dynamical systems for population spread, where the tail behaviour can determine both the success and speed of invasions during a range expansion [3]. For example, a model that uses a (thin-tailed) Gaussian kernel out of mathematical convenience might fail to account for tail-like long-distance range expansion events.

3. The WMY in applications

Redistribution kernels are the central feature of the integrodifference equation (IDE), a model that combines growth (G) with spatially explicit dispersal (D) for a population density variable $N_t(\boldsymbol{x})$ indexed continuously in space and discretely in time [40]. D is stationary in the most well-studied case, and the IDE reads $N_t(\boldsymbol{x}) = (D * G(N_{t-1}))(\boldsymbol{x})$, where $*$ denotes convolution. A rich theory has emerged in connecting the mathematical properties of the function D with the spread dynamics of N_t as t becomes large [see *eg.* 3, 41]. Redistribution kernels are also often applied in analyses of data from individual and aggregated years [*eg.* 42, 39]. In this more general setting, $\tilde{N}(\boldsymbol{x}) = (D * N)(\boldsymbol{x})$ simply connects the pre (N) and post-dispersal ($\tilde{N}$) levels in a population.

In applications, populations are sampled at a finite set of locations, with centroids $\{\boldsymbol{x}_k\}$ $(k \in 1, \ldots m)$. The (continuous) kernel convolution $D * N$ must therefore be replaced with a discretized version. In place of D one specifies an $m \times m$ *redistribution matrix* $\boldsymbol{D}$, where the entry $[\boldsymbol{D}]_{ij}$ (a discretization of D, see Appendix 2.1) assigns density to the redistribution event from location $\boldsymbol{x}_j$ to $\boldsymbol{x}_i$. Thus $\tilde{\boldsymbol{n}} = \boldsymbol{D}\boldsymbol{n}$ connects the length-m pre ($\boldsymbol{n}$) and post-dispersal ($\tilde{\boldsymbol{n}}$) population vectors.

3.1. Separable approximations for gridded data

In many applied situations (*eg.* numerical likelihood maximization), the matrix-vector product $\boldsymbol{D}\boldsymbol{n}$ must be evaluated many times. The operation has $O(m^2)$ complexity, leading to issues of computability with large sample sizes (m). However, large spatial datasets are often gridded, and for this configuration there are remedies: Fast Fourier transforms (FFTs) reduce the complexity to $O(m \log m)$ [43]; and if D is spatially *separable*, Kronecker products can reduce complexity to $O(m_x^2 + m_y^2)$ (where the grid has dimensions $m_y \times m_x = m$). On square domains that is $O(m)$, a considerable improvement (Figure 1). The implementation of this trick is straightforward, but notationally awkward, so we relegate the details to Appendix 3.3.

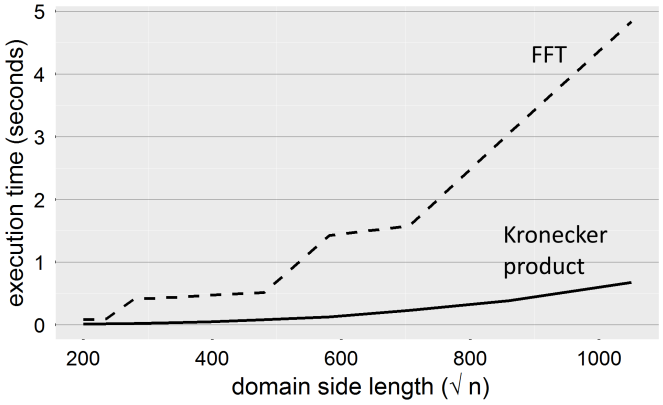

Figure 1: Time to compute the length-n post-dispersal vector $\boldsymbol{D}\boldsymbol{n}$, for redistribution matrices from the WMY (dashed line) and pWMY (solid line) kernels. The pWMY is computed using Kronecker products (Appendix 3.3), while the nonseparable WMY uses FFTs.

Separability is the property that $D(\boldsymbol{r})$ factors into a product of two 1d kernels,

$D_x(r_x)$ and $D_y(r_y)$, where $\boldsymbol{r} = (r_x, r_y)^T$, each depending on only one of the spatial dimension. The (2d) Gaussian kernel is an example: it has the form $D_x(r_x)D_y(r_y)$, where $D_x(r) = D_y(r) = (\sqrt{\pi}\rho)^{-1}\exp(-(r/\rho)^2)$ are 1d Gaussian kernels. The 2d WMY kernel, on the other hand, is not separable. In seeking a computationally simple alternative we propose the separable *product-WMY* (pWMY) kernel, $\mathcal{D}_\otimes$, in which the x and y components (D_x and D_y) are 1d WMY kernels. These have the same functional form as the 2d WMY (2), differing only in normalization constants. Thus,

$$\mathcal{D}_\otimes(\boldsymbol{r}; \kappa_x, \kappa_y, \rho_x, \rho_y) = \frac{\mathcal{D}(|r_x|; \kappa_x, \rho_x)\mathcal{D}(|r_y|; \kappa_y, \rho_y)}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathcal{D}(|r_x|; \kappa_x, \rho_x)\mathcal{D}(|r_y|; \kappa_y, \rho_y)\,\mathrm{d}r_x\,\mathrm{d}r_y}. \quad (13)$$

In Appendix 2.3 we report on a simulation study indicating that the pWMY closely approximates the WMY over much of its parameter space. Note that the 2d Gaussian is a limiting case of $\mathcal{D}_\otimes$, similar to $\mathcal{D}$. These (and other) computational properties of separable models are discussed in more detail in [44], where the pWMY is employed as a covariogram.

3.2. Extensions for geometric anisotropy

While mathematically pleasant, the assumption of isotropic redistribution is often unsatisfactory in ecological applications. Wind, for example, is directional, and wind-assisted migratory flights occur regularly in insects [29]. Indeed these were the cause of a recent range expansion of the MPB [45].

One of the simplest ways to incorporate anisotropy into an isotropic kernel is to compose it with an affine transformation of coordinates (Appendix 2.2). This effectively incorporates two stages of movement: drift (by a translation of coordinates) and directed diffusion (by a linear transformation of coordinates). In spatial statistics, such transformations are common extensions of the isotropic *covariance* kernel [36]. We see them far less often with redistribution kernels. However they are easy to implement, and their effect on D is easy to visualize and understand (Figure 2).

3.3. Case study: predicting MPB infestations with the red-top model

To demonstrate the efficacy of the WMY and pWMY in applied movement modelling, we use the example of the MPB. Populations of this tree-killing beetle exhibit

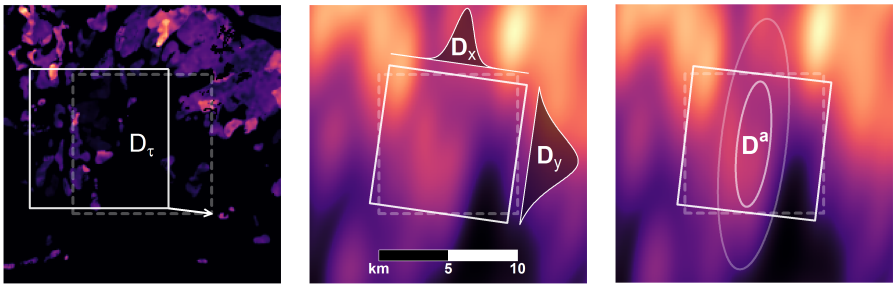

Figure 2: A geometric model for anisotropic redistribution. Source populations (left) drift via translation D_τ then diffuse according to a redistribution kernel D, which is rotated and stretched by a linear transformation of coordinates. Two examples of D are shown: the WMY (D^a, right), and an approximation, the pWMY (with components D_x and D_y, middle)

.

periods of elevated activity (outbreaks) in western North America that can persist for decades [46]. Pine mortality becomes visible on the landscape during outbreaks, with clusters of crown fade (red-tops) spreading through space from year to year, reflecting macro-scale movement patterns [47].

Annual aerial overview surveys (AOS) of MPB damage, carried out by the government of British Columbia (BC), provide a spatiotemporal record of outbreak spread. We analysed subsets of the AOS data using a slight modification of the red-top model of Heavilin and Powell [37]. Their IDE relates the pine mortality, $M_t(\boldsymbol{x})$, in a given year (t) to its value in the following year. Pine deaths are a proxy for the beetle population in this model, and dispersal across the landscape is represented by the kernel D. Our modified version adds a (spatially uniform) constant 1 to all populations (adjusting for the unobserved, but ubiquitous endemic beetle community), and accounts for model error by an additive Gaussian term on the *logit* (*ie.* log-odds) scale instead of the percent scale. The modified red-top model is thus written:

$$\text{logit}\,(M_{t+1}) = 2\log\,(1 + D * (M_t S_t)) - 2\log\,(\alpha) + Z_t\,, \tag{14}$$

where $S_t(\boldsymbol{x})$ is the observed density of susceptible (healthy) pine, and $Z_t(\boldsymbol{x})$ is a noise term representing process and measurement errors (Appendix 3.2). The growth pa-

rameter $\alpha > 0$ parametrizes a nonlinear growth function representing the dynamics of beetle-pine interactions. The reader will find a more detailed explanation of (14) and its discretization in Appendix 3.

3.3.1. Kernel selection: comparing the Gaussian, WMY and pWMY

In their model selection for D [37], Heavilin and Powell compared the Gaussian and 2d Laplace kernels (Table 1). We conducted a more exhaustive model selection (Figure 3), in order to emphasize three questions: Which WMY kernel (*ie.* which κ) best characterizes MPB spread? Is the product-WMY a useful surrogate for the WMY? And do the anisotropy extensions increase explanatory power?

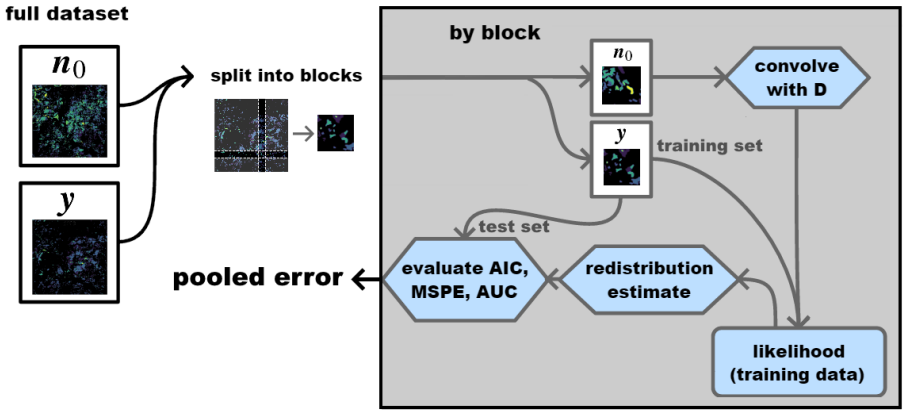

Figure 3: Flow chart for estimating model error. Data are split into blocks partitioned into test and training sets. Within-block likelihood is maximized on training data (for a given kernel), and predictions compared with test data. Blockwise errors were poooled and compared for six redistribution kernels.

Using numerical maximum likelihood, we fit the (discretized) modified red-top model independently to $M = 81$ nonoverlapping blocks of AOS data, each of size 10×10 km, with a grid resolution of 1 hectare (ha). Data were split into test and training sets, facilitating an evaluation of within-block model errors (Figure 3). To reduce the effect of spatial autocorrelation on these error estimates, we first fitted a Gaussian covariogram to the full dataset (containing all blocks), using it as a common (global) estimate of covariance in all of the within-block likelihood calculations (Appendix 3.4).

By pooling results over all M blocks, and repeating this process for different redistribution kernels D, we compared their overall explanatory power. We compared kernels from three families: the Gaussian; the WMY with $\kappa > 0$ (to exclude singularities at $r = 0$); and the pWMY (with $\kappa_x, \kappa_y > 0$). In each case we tested the kernel with and without the anisotropy extensions of Section 3.2, and measured model performance on test data (unseen by the model in training) using root-mean-square prediction error (RMSPE), log-likelihood (LL), and Akaike's Information Criterion (AIC) [53].

Figure 4 summarizes the results from all blocks and kernels, by centering each test statistic on that of the best performing kernel for that block. For all three kernel familes, the anisotropy extensions improved performance dramatically. Without these extensions, there was little difference among kernels, with the WMY favoured slightly overall. However, with the extensions, the WMY produced substantial improvements over the Gaussian, and the pWMY performed best overall.

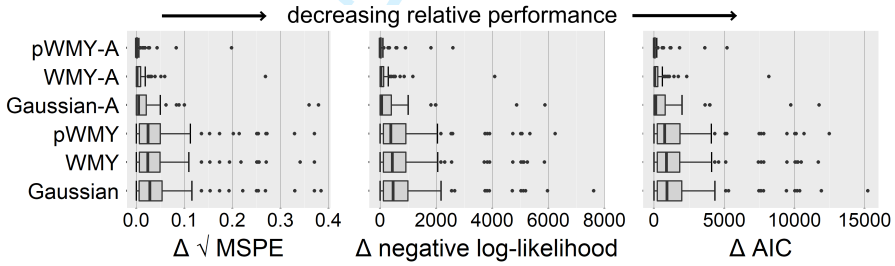

Figure 4: Model performance over 81 model-fitting trials for six redistribution kernels: The Gaussian, WMY, and pWMY; and their anisotropic extensions (Gaussian-A, WMY-A, and pWMY-A); which have 1, 2, 4, 5, 6, and 7 parameters, respectively. RMSPE, negative LL, and AIC were computed on test data in each block, and their differences (Δ) with the best model within that block are summarized as boxplots. Extreme outliers (points to the right of the boxplot whiskers) were associated with blocks having very low levels of MPB activity, where it appears there was insufficient information to reliably parametrize movement patterns.

For the isotropic WMY kernels, the fitted value of κ varied substantially among blocks. We observed a Laplace/Bessel-like spread pattern ($\hat{\kappa} \leq 1/2$) in around one quarter of the 81 tested blocks. For around a third of the blocks, the parameter estimate stalled at its upper bound ($\kappa = 25$), indicating a tendency towards Gaussian spread patterns ($\hat{\kappa} \to \infty$). On the remaining blocks it was intermediate. Thus we observed no

particular value (or range) of κ to be favoured overall in the point estimates.

4. Discussion

Radar evidence suggests that wind-assisted MPB flights, carrying the insects unusually long distances, are not uncommon [48]. One such long-distance event recently led the MPB to cross the Rocky Mountains, from BC into central Alberta – a range expansion with severe ecological and economic consequences [45]. Equation (12) shows how by using WMY kernels, the probability of such tail-like events in models can be more finely tuned using κ.

Previous kernel selections in MPB spread models have led to somewhat contradictory conclusions on this tail behaviour [37, 49]: One favoured the Gaussian kernel over a 2d Laplace, suggesting fixed-time settling events leading to thin-tailed movement patterns; the other favoured the (fat-tailed) Bessel over the Gaussian kernel, suggesting a movement mechanism with constant settling hazard (see Table 1).

This inconsistency can be resolved by simply dropping the assumptions of single-stage dispersal through homogeneous habitat. Our results on n-stage redistribution processes in fractal media (Section 2) give a mechanistic explanation of how disparities in dispersal patterns can arise simply as a result of geographical variation in habitat type (as represented by d_f), or in the number of times the dispersal event is interrupted and restarted, as might occur with local temperature swings [48].

This explanation is consistent with the wide range of fitted values for κ observed in our case study (Section 3.3.1), which indicated that no single (stationary) kernel adequately described the scope of MPB spread patterns exhibited in the AOS data. Note that since these observations were based on the maximum likelihood point estimates alone, their statistical significance cannot be verified here. However, it is very likely that movement patterns will vary locally due to habitat structure [5]. For this reason (among others) we also expect a certain directionality in movement patterns, so it is perhaps unsurprising that the anisotropy extensions of Section 3.2 improved model predictive performance across the board. This also indicated that adding directionality to the red-top model did not lead to issues of overfitting, despite the introduction of

3-4 additional parameters (see Appendix 3.5). In a companion paper, we discuss applications of these directed models, using them to estimate the size of cryptic endemic subpopulations of MPB [50].

Yasuda's model for ordinary Fickian diffusion with gamma-distributed settling times [12] provides an alternative mechanistic view on geographical variations in kernel shape. However it requires dropping the parsimonious hypothesis of a constant settling hazard, which, for the MPB, is supported by evidence from laboratory flight-mill experiments [51, 52].

More generally, the fractal diffusion kernels derived in Section 2 provide a plausible process-based description of random walks through complex media, and so have wide applicability beyond MPB dispersal. Recall that Section 2.1.1 extended the Bessel kernel (originally a model for nematodes) by relaxing the classical assumption of unrestricted movement space. Section 2.2.2 then showed how the full WMY family is recovered through n-stage iterations, where the Gaussian kernel (originally a model for muskrats) is approached in the limit $n \to \infty$. The WMY kernel unifies both extremes of qualitative behaviour under a common process model, defining a wide spectrum of possible redistribution patterns, suitable for a wide range of study organisms.

The WMY has for a long time been embraced in spatial statistics community as the covariogram of choice, for its mathematical elegance and flexibility. In light of the mathematical connections to movement processes outlined in this paper, we would encourage the ecological modelling community to embrace the WMY as the redistribution kernel of choice.

Flexibility in tail behaviour is also desireable in phenomenological applications for redistribution kernels. In Section 2.3, we explained how this notion of flexibility can be made more precise, using 2d kurtosis. The WMY admits an exceptionally large range of kurtosis values. To our knowledge, the only other comparable redistribution kernel is the 2Dt [39], a phenomenological model of seed shadow distributions. In fact the 2Dt and WMY are closely related: They are Fourier duals.

Unsurprisingly, the red-top model performed better on replacing the Gaussian with a more flexible WMY. More interesting was the superior performance of the pWMY. This phenomenological kernel takes the best of both worlds; it closely approximates the

(mechanistic) WMY kernel, matching its flexibility while retaining the computational advantages of separability.

Author Contributions

Dean Koch wrote the first draft of the manuscript and participated in all stages of the project. Mark Lewis and Subhash Lele contributed to the model development and critically revised the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

Acknowledgements

The authors thank the Lewis Research Group at the University of Alberta for providing expert advice and feedback.

Data Accessibility

All data and R code required to reproduce our analysis have been supplied as electronic supplementary material.

Funding

This work was supported by a grant to Mark A Lewis (MAL) from the Natural Science and Engineering Research Council of Canada [Grant No. NET GP 434810-12] to the TRIA Network, with contributions from Alberta Agriculture and Forestry, Foothills Research Institute, Manitoba Conservation and Water Stewardship, Natural Resources Canada-Canadian Forest Service, Northwest Territories Environment and Natural Resources, Ontario Ministry of Natural Resources and Forestry, Saskatchewan Ministry of Environment, West Fraser and Weyerhaeuser. MAL is also grateful for support through NSERC and the Canada Research Chair Program.

References

[1] Kareiva P. Space: The final frontier for ecological theory [A series of 5 papers]. Ecology. 1994;75(1):1.

[2] Neubert MG, Kot M, Lewis MA. Dispersal and pattern formation in a discrete-time predator-prey model. Theor Popul Biol. 1995;48(1):7–43. doi:10.1006/tpbi.1995.1020.

[3] Kot M, Lewis MA, Van Den Driessche P. Dispersal data and the spread of invading organisms. Ecology. 1996;77(7):2027–2042. doi:10.2307/2265698.

[4] Okubo A, Levin S. Diffusion and Ecological Problems: Modern Perspectives. vol. 14. 2nd ed. New York, NY: Springer-Verlag; 2001. doi:10.1007/978-1-4757-4978-6.

[5] Powell JA, Bentz BJ. Phenology and density-dependent dispersal predict patterns of mountain pine beetle (Dendroctonus ponderosae) impact. Ecol Modell. 2014;273:173–185. doi:10.1016/j.ecolmodel.2013.10.034.

[6] Painter KJ, Hillen T. From Random Walks to Fully Anisotropic Diffusion Models for Cell and Animal Movement. In: Stolarska M, Tarfulea N, editors. Model. Simul. Sci. Eng. Technol. Cham, Switzerland: Springer Nature; 2018. p. 103–141. doi:10.1007/978-3-319-96842-1.

[7] Kühn I, Dormann CF. Less than eight (and a half) misconceptions of spatial analysis. J Biogeogr. 2012;39(5):995–998. doi:10.1111/j.1365-2699.2012.02707.x.

[8] Legendre P. Spatial autocorrelation: trouble or new paradigm? Ecology. 1993;74(6):1659–1673. doi:10.2307/1939924.

[9] Chilès JP, Delfiner P. Geostatistics: Modeling Spatial Uncertainty: Second Edition. 2nd ed. Hoboken, NJ: John Wiley & Sons; 2012. doi:10.1002/9781118136188.

[10] Cooke, RM, Nieboer, D, Misiewicz, J. Fat-tailed Distributions: Data, Diagnostics and Dependence. Hoboken, NJ: John Wiley & Sons; 2014. 10.1002/9781119054207.

[11] Guttorp P, Gneiting T. Studies in the history of probability and statistics XLIX on the Matérn correlation family. Biometrika. 2005;93(4):989–995. doi:10.1093/biomet/93.4.989.

[12] Yasuda N. The random walk model of human migration. Theor Popul Biol. 1975;7(2):156–167. doi:10.1016/0040-5809(75)90011-8.

[13] Lindgren F, Rue H, Lindström J. An explicit link between gaussian fields and gaussian markov random fields: The stochastic partial differential equation approach. J R Stat Soc Ser B Stat Methodol. 2011;73(4):423–498. doi:10.1111/j.1467-9868.2011.00777.x.

[14] Yamamura K. Dispersal distance of heterogeneous populations. Popul Ecol. 2002;44(2):93–101. doi:10.1007/s101440200011.

[15] Hapca S, Crawford JW, Young IM. Anomalous diffusion of heterogeneous populations characterized by normal diffusion at the individual level. J R Soc Interface. 2009;6(30):111–122. doi:10.1098/rsif.2008.0261.

[16] Morgan BJT, Okubo A. Diffusion and Ecological Problems: Mathematical Models.. vol. 37. Heidelberg, Germany: Springer-Verlag Berlin; 1981. doi:10.2307/2530430.

[17] Broadbent SR, Kendall DG. The Random Walk of Trichostrongylus retortaeformis. Biometrics. 1953;9(4):460. doi:10.2307/3001437.

[18] Joseph J, Sendner H. Über die horizontale Diffusion im Meere. Dtsch Hydrogr Zeitschrift. 1958;11(2):49–77. doi:10.1007/BF02020293.

[19] Skellam JG. Random dispersal in theoretical populations. Biometrika. 1951;38(1-2):196–218. doi:10.1093/biomet/38.1-2.196.

[20] Whittle P. On Stationary Processes in the Plane. Biometrika. 1954;41(3/4):434. doi:10.2307/2332724.

[21] Metzler R, Glöckle WG, Nonnenmacher TF. Fractional model equation for anomalous diffusion. Phys A Stat Mech its Appl. 1994;211(1):13–24. doi:10.1016/0378-4371(94)90064-7.

[22] Goodchild MF, Mark DM. The Fractal Nature of Geographic Phenomena. Ann Assoc Am Geogr. 1987;77(2):265–278. doi:10.1111/j.1467-8306.1987.tb00158.x.

[23] Seuront L. Fractals and Multifractals in Ecology and Aquatic Science. Boca Raton, FL: CRC Press; 2009. doi:10.1201/9781420004243.

[24] Zeide B. Fractal geometry in forestry applications. For Ecol Manage. 1991;46(3-4):179–188. doi:10.1016/0378-1127(91)90230-S.

[25] Jonckheere I, Nackaerts K, Muys B, van Aardt J, Coppin P. A fractal dimension-based modelling approach for studying the effect of leaf distribution on LAI retrieval in forest canopies. Ecol Modell. 2006;197(1-2):179–195. doi:10.1016/j.ecolmodel.2006.02.036.

[26] Méndez V, Campos D, Bartumeus F. Stochastic Foundations in Movement Ecology. New York: Springer; 2014. doi:10.1007/978-3-642-39010-4.

[27] Hargis CD, Bissonette JA, Turner DL. The influence of forest fragmentation and landscape pattern on American martens. J Appl Ecol. 1999;36(1):157–172. doi:10.1046/j.1365-2664.1999.00377.x.

[28] O'Shaughnessy B, Procaccia I. Analytical Solutions for Diffusion on Fractal Objects. Phys Rev Lett. 1985;54(5):455–458. doi:10.1103/PhysRevLett.54.455.

[29] Jackson PL, Straussfogel D, Lindgren BS, Mitchell S, Murphy B. Radar observation and aerial capture of mountain pine beetle, Dendroctonus ponderosae Hopk. (Coleoptera: Scolytidae) in flight above the forest canopy. Can J For Res. 2008;38(8):2313–2327. doi:10.1139/X08-066.

[30] Turchin P, Thoeny WT. Quantifying dispersal of southern pine beetles with mark-recapture experiments and a diffusion model. Ecol Appl. 1993;3(1):187–198. doi:10.2307/1941801.

[31] Williams E. The Distribution of Larvae of Randomly Moving Insects. Aust J Biol Sci. 1961;14(4):598. doi:10.1071/bi9610598.

[32] Awerbuch TE, Samson R, Sinskey AJ. A quantitative model of diffusion bioassays. J Theor Biol. 1979;79(3). doi:10.1016/0022-5193(79)90350-3.

[33] Hu G, Lim KS, Reynolds DR, Reynolds AM, Chapman JW. Wind-related orientation patterns in diurnal, crepuscular and nocturnal high-altitude insect migrants. Front Behav Neurosci. 2016;10(5):1–8. doi:10.3389/fnbeh.2016.00032.

[34] Schlägel UE, Lewis MA. A framework for analyzing the robustness of movement models to variable step discretization. J Math Biol. 2016;73(4):815–845. doi:10.1007/s00285-016-0969-5.

[35] Chesson P, Lee CT. Families of discrete kernels for modeling dispersal. Theor Popul Biol. 2005;67(4):241–256. doi:10.1016/j.tpb.2004.12.002.

[36] Rathbun SL, Stein ML. Interpolation of Spatial Data: Some Theory for Kriging. New York, NY: Springer-Verlag; 2000. doi:10.2307/2669494.

[37] Heavilin J, Powell J. A novel method of fitting spatio-temporal models to data, with applications to the dynamics of mountain pine beetles. Nat Resour Model. 2008;21(4):489–524. doi:10.1111/j.1939-7445.2008.00021.x.

[38] Gilbert MA, White SM, Bullock JM, Gaffney EA. Speeding up the simulation of population spread models. Methods Ecol Evol. 2017;8(4):501–510. doi:10.1111/2041-210X.12684.

[39] Clark JS, Silman M, Kern R, Macklin E, Hillerislambers J. Seed dispersal near and far: Patterns across temperate and tropical forests. Ecology. 1999;80(5):1475–1494. doi:10.1890/0012-9658.

[40] Kot M, Schaffer WM. Discrete-time growth-dispersal models. Math Biosci. 1986;80(1):109–136.

[41] Lewis MA. Spread rate for a nonlinear stochastic invasion. J Math Biol. 2000;41(5):430–454. doi:10.1007/s002850000022.

[42] Lele S, Taper ML, Gage S. Statistical analysis of population dynamics in space and time using estimating functions. Ecology. 1998;79(5):1489–1502. doi:https://doi.org/10.1890/0012-9658(1998)079[1489:SAOPDI]2.0.CO;2.

[43] Andersen M. Properties of some density-dependent integrodifference equation population models. Math Biosci. 1991;104(1):135–157. doi:10.1016/0025-5564(91)90034-G.

[44] Koch D, Lele S, Lewis M. Computationally Simple Anisotropic Lattice Covariograms. Environ Ecol Stat (in review). 2020;.

[45] de la Giroday HMC, Carroll AL, Aukema BH. Breach of the northern Rocky Mountain geoclimatic barrier: Initiation of range expansion by the mountain pine beetle. J Biogeogr. 2012;39(6):1112–1123. doi:10.1111/j.1365-2699.2011.02673.x.

[46] Duan JJ, Taylor PB, Fuester RW. Biology and Life History of Balcha indica, an Ectoparasitoid Attacking the Emerald Ash Borer, Agrilus planipennis, in North America. In: Safranyik L, Wilson B, editors. J. Insect Sci.. vol. 11. Victoria, Canada: Canadian Forest Service; 2011. p. 1–9. doi:10.1673/031.011.12701.

[47] Chen H, Walton A. Mountain pine beetle dispersal: spatiotemporal patterns and role in the spread and expansion of the present outbreak. Ecosphere. 2011;2(6):art66. doi:10.1890/es10-00172.1.

[48] Ainslie B, Jackson PL. Investigation into mountain pine beetle above-canopy dispersion using weather radar and an atmospheric dispersion model. Aerobiologia (Bologna). 2011;27(1):51–65. doi:10.1007/s10453-010-9176-9.

[49] Goodsman DW, Koch D, Whitehouse C, Evenden ML, Cooke BJ, Lewis MA. Aggregation and a strong Allee effect in a cooperative outbreak insect. Ecol Appl. 2016;26(8):2621–2634. doi:10.1002/eap.1404.

[50] Koch D, Lewis M, Lele S. The signature of endemic populations in the spread of mountain pine beetle outbreaks. (in prep). 2020;.

[51] Evenden ML, Whitehouse CM, Sykes J. Factors Influencing Flight Capacity of the Mountain Pine Beetle (Coleoptera: Curculionidae: Scolytinae). Environ Entomol. 2014;43(1):187–196. doi:10.1603/en13244.

[52] Shegelski VA, Evenden ML, Sperling FAH. Morphological variation associated with dispersal capacity in a tree-killing bark beetle Dendroctonus ponderosae Hopkins. Agric For Entomol. 2019;21(1):79–87. doi:10.1111/afe.12305.

[53] Akaike, H Information theory and an extension of the maximum likelihood principle. vol. 3 New York, NY: Springer-Verlag; 1998. 10.1007/978-1-4612-1694-0.