At what scales does aggregated dispersal lead to coexistence?

Eric J. Pedersen, Department of Biology, McGill University, Montreal, QC, Canada, and Center for Limnology, University of Wisconsin-Madison, Madison, WI, USA eric.pedersen@wisc.edu

Frédéric Guichard, McGill University, Montreal, QC, Canada fred.guichard@mcgill.ca

Running title: Scale aggregated dispersal and coexistence

Keywords: aggregated dispersal, coexistence, dispersal, ecological scales, metacommunities, spatial ecology, stochasticity, stochastic dispersal

Statement of authorship: EP and FG conceived of the original idea for this study. EP developed the analytical results, moment closure and simulations. EP wrote the first draft of the manuscript, and both authors contributed substantially to revisions.

Article type: Letter

Number of words in the abstract: 137

Number of words in the text: 4905

Number of words in glossary text box: 120

Number of references: 49

5 figures, 0 tables, 1 text box

Corresponding Author:
Eric J. Pedersen
Center for Limnology
680 North Park Street
Madison, WI, USA, 53706
Phone: 608-262-3088
Fax: 608-265-2340
Email: eric.pedersen@wisc.edu

Abstract

Aggregation during dispersal from source to settlement sites can allow persistence of weak competitors, by creating conditions where stronger competitors are more likely to interact with conspecifics than with less competitive heterospecifics. However, different aggregation mechanisms across scales can lead to very different patterns of settlement. Little is known about what ecological conditions are required for this mechanism to work effectively. We derive a metacommunity approximation of aggregated dispersal that shows how three different scales interact to determine competitive outcomes: the spatial scale of aggregation, the spatial scale of interactions between individuals, and the time-scale of arrival rates of aggregations. We use stochastic simulations and a novel metacommunity approximation to show that an inferior competitor can invade only when the superior competitor is aggregated over short spatial scales, and aggregations of new settlers are small and rare.

Introduction

One of the most significant and longest lasting problems in ecology, dating back to its start as a quantitative discipline (Gause, 1932), is the paradox of coexistence: if two species have the same resource requirements and similar environmental tolerances, why does the species with higher fitness not drive the other to extinction (Hutchinson, 1961)? In general, unequal competitors can only coexist if there is some form of stabilizing mechanism: an ecological process which increases the weaker competitor's growth rate at low densities (Chesson, 2000*b*). One of the major factors stabilizing

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

species interactions are differential responses to spatial heterogeneity (Chesson, 2000*a*). The effectiveness of any spatial stabilizing mechanism at promoting coexistence is determined by the scales of dispersal and of interactions among competing species (Chesson et al., 2005). Short distance dispersal has been shown to affect intraspecific crowding and coexistence of species that interact over similarly local scales (Bolker and Pacala, 1999, Snyder and Chesson, 2003, 2004). However, clustering of conspecifics due to short range-dispersal by itself is not sufficient to allow a weaker competitor to invade a system(Chesson and Neuhauser, 2002). Instead, coexistence can occur if each species uses space in substantially different ways, either through endogenous spatial patterns of density (e.g. Bolker and Pacala, 1999, Snyder and Chesson, 2004), or through species-specific responses to environmental variation (e.g. Snyder and Chesson, 2003, Snyder, 2008). Also, dispersal can occur over distances that are orders of magnitude larger than the scales of species interactions (Kinlan and Gaines, 2003), which limits the application of dispersal as a mechanism of coexistence.

Many transport mechanisms associated with large sale dispersal, such as large marine current features (Siegel et al., 2008), can also lead to the aggregation of propagules in transit. These aggregated transport mechanisms create patterns of clustered settlement at scales much smaller than the scale of dispersal itself. Aggregated dispersal has been proposed as one factor driving coexistence of sedentary species in metacommunities (Potthoff et al., 2006, Berkley et al., 2010, Aiken and Navarrete, 2014). It has been demonstrated to enhance coexistence in field marine field plots (Edwards and Stachowicz, 2011), and given that many plant species face strongly density-dependent seed mortality (Harms et al., 2000), aggregated seed dispersal may play a significant

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

role in shaping plant communities (Muller-Landau and Hardesty, 2005, Potthoff et al., 2006). Aggregated long-distance dispersal can allow the coexistence of unequal competitors as long as the two species only rarely travel together in the same packets (Berkley et al., 2010). This works because the aggregated settlement of conspecifics results in higher intra-specific competition with no commensurate increase in inter-specific competition. This was previously shown to stabilize coexistence at small scales, such as insect herbivores competing for patchy plant resources (Ives and May, 1985). Aggregated dispersal can also allow two species to coexist even if they use space in the same way (Berkley et al., 2010) because it leads to each recruit settling near conspecifics, even if dispersal started from an area of low density. This is unlike non-aggregated dispersal where recruits will only experience conspecific clustering if they disperse from an area with high adult density, which are already difficult to invade (Chesson and Neuhauser, 2002).

Aggregated dispersal has been shown to stabilize coexistence over large scales in cases where competing species interact within patches that are connected by dispersal (that is, they form a metacommunity, Leibold et al., 2004) which are themselves the same size as propagule aggregations (Potthoff et al., 2006, Berkley et al., 2010). Given the wide range of possible aggregation mechanisms, such as marine currents (Siegel et al., 2008) or seed transport by wind and animal vectors (Muller-Landau and Hardesty, 2005), and the variety of spatial scales that individuals interact at, mismatches between the scale of aggregation and of interaction should be the rule. For instance, in the Southern California Bight, propagule aggregations can be nearly 100 km wide (Siegel et al., 2008), but benthic macro-algal species may only be interacting with neighbours up to 1

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

km along the coast (Cavanaugh et al., 2014). Further, many species, while competing for the same resources, may either interact at different spatial scales (Ritchie, 2009) or be dispersed by different processes with significantly different scales of aggregation (for instance, if the larvae of one species are active swimmers, while its competitors larvae are passive, the first species will likely be less aggregated when they settle (Harrison et al., 2013)).

Previous theoretical work on the role of spatial scale on coexistence with non-aggregated dispersal (e.g. Bolker and Pacala, 1999, Snyder and Chesson, 2004) provides a guide to how scales of aggregation and interactions may affect dynamics. However, it is built on the assumption that aggregated recruitment can only arise if source populations are already aggregated. Our goal is to understand the relative importance of aggregated dispersal and of species interactions for coexistence over a broad range of spatial scales. Towards this goal, we define key properties of propagule aggregation and of adult interactions to predict the effects of aggregated dispersal on coexistence. We first derive a single expression approximating settlement variability as a function of the scale of aggregation, the distribution of propagules among aggregations (packets), and of the spatial scale over which variability is measured. This approximation is useful both for incorporating aggregated dispersal into ecological models and for defining a set of metrics that can be used in the field to test model predictions. We use a combination of stochastic simulations and a novel moment-closure approximation to predict scales of aggregated dispersal that lead to coexistence. Our results show that aggregated dispersal can play a role in shaping community structure across a much wider range of spatial scales than has been previously shown.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Materials and Methods

Approximating aggregated dispersal

We approximate aggregation as a set of discrete aggregates, or packets (sensu Siegel et al., 2008, Berkley et al., 2010) of individuals. All aggregated dispersal mechanisms are then defined by three processes: how propagules are distributed between packets, where packets settle, and where propagules settle relative to the center of the packet. The outcome of all these processes will be a spatial distribution of settlers across a landscape. In mathematical terms, the pattern of settlers arriving on the landscape over a fixed period of time is a cluster point process (Illian et al., 2008). Cluster point processes are described by three functions: the intensity $\lambda_c(\chi, \mathbf{P})$ of cluster centers at a given location χ given a set of ecological state variables $\mathbf{P}$, the probability $p(n|\chi, \mathbf{P})$ of finding n points in a cluster at location χ, and the probability $\delta(\chi', \chi, \mathbf{P})$ of finding a point from a given cluster at location χ', given a cluster occurs at location χ.

Cluster point processes are general enough to describe any type of aggregated dispersal. However, for the sake of simplicity and tractability we focus on a subset of cluster point processes, called Neyman-Scott processes (Illian et al., 2008). Here, space is assumed to be homogeneous, so that packets settle at the same intensity ($\lambda_c(\chi, \mathbf{P}) = \lambda_c(\mathbf{P})$) and have the same properties ($p(n|\chi, \mathbf{P}) = p_c(n|\mathbf{P})$) at all points in the landscape. Finally, packets are assumed to be isotropic ($\delta(\chi', \chi, \mathbf{P}) = \delta(\epsilon)$ where ϵ is the distance between a location and the packet center). The first assumption is equivalent to assuming a propagule rain, where all sites are equally likely to get settlers. Therefore, all points will have a mean settlement intensity of $\lambda(\mathbf{P}) = \lambda_c(\mathbf{P})E(p_c(n|\mathbf{P}) = \lambda_c(\mathbf{P})\mu(\mathbf{P})$, where $\mu(\mathbf{P})$

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

is the mean number of propagules per packet.

We introduce an interaction scale by assuming that space is divided into circular patches of radius ν defined as the scale of interaction. Each patch will then have a volume Vol(ν), where $Vol(\nu)$ is a function that depends on the dimension of the space that individuals interact in: if space is one-dimensional, $Vol(\nu) = 2\nu$, and if space is two-dimensional, $Vol(\nu) = \pi\nu^2$. If we define s_i as the number of settlers in patch i, the mean number of settlers across all patches, $\bar{s}$ will be $\bar{s} = Vol(\nu)\lambda(\mathbf{P})$. The probability of finding s settlers in a given patch is approximately (Sheth and Saslaw, 1994, Illian et al., 2008):

$$p(s|\lambda,\kappa,\nu) = \frac{Vol(\nu)\lambda}{s!\kappa^{0.5}}[Vol(\nu)\lambda\kappa^{-0.5} + s(1-\kappa^{-0.5})]e^{-Vol(\nu)\lambda\kappa^{-0.5}-s(1-\kappa^{-0.5})} \quad (1)$$

where κ is a function summarizing all the effects of patch size, the distribution of propagules between packets, and the distribution of propagules within a packet. κ measures how aggregated settlement is, ranging from 1 where the number of settlers in each patch is randomly distributed following a Poisson distribution, to ∞. κ also defines the mean-variance relationship for this distribution, with $Var(s) = \lambda Vol(\nu)\kappa = \bar{s}\kappa$. While the expression for κ is complex, it can be closely approximated by a simple function (Appendix A):

$$\kappa \approx 1 + (\frac{\sigma^2 + \mu^2 - \mu}{\mu})\frac{a(\frac{\nu}{\omega})^b}{1 + a(\frac{\nu}{\omega})^b} \quad (2)$$

The parameters μ and σ^2 are the mean and variance of the distribution of propagules among packets, ω is the square root of the mean square distance of settlers from their

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

packet center (the standard deviation of the one dimension packet distribution), and a and b are unitless scaling coefficients. If species only interact in one-dimension, $a = 1$ and $b = 1.25$; in two dimensions, $a = 0.5$ and $b = 2$ (Appendix A). Equation (2) implies that κ increases with increasing mean packet density, among-packet variability in individual density, with the scale of interaction, and decreases with the scale of aggregation (Fig.). The first term in brackets captures the effects of the distribution of individuals among packets. This term can be simplified further: if all packets have the same number of settlers it equals $\mu - 1$, if settlers are Poisson distributed between packets it equals μ, and if they are negative binomial distributed it equals $\mu(1 + \frac{1}{k})$ (where k measures over-dispersion Bolker, 2008). We define μ as the time-scale of aggregation: as there are only a fixed number of propagules dispersing at a given time, if packets have higher mean densities, they must also arrive more infrequently. The second term in equation (2) captures the combined effects of the spatial scale of aggregation (ω) and of interaction (ν) on settlement variation.

Meta-community moment closure

To understand how scales, as defined by equation (2), will affect species persistence, we have to determine how spatial variability in propagule and adult densities affect the mean strength of local interactions (Chesson et al., 2005). To understand these interacting scales, we use a moment-closure approximation of a stochastic Lotka-Volterra meta-community model. Moment closure is a technique for approximating a complex stochastic system by reducing it to equations describing the dynamics of statistical summaries of the population (its moments), such as the mean

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

densities and spatial variances and covariances of all the species in the system (Bolker and Pacala, 1999, Keeling, 2000*b*). Here, we modify a moment closure derived for metapopulations (Keeling, 2000*b*) to incorporate both patch volume and aggregated settlement effects.

We start with a continuous time model with two species, x and y, interacting in a one-dimensional habitat. Fig. illustrates the basic processes assumed in our model, comparing dynamics in metacommunity without aggregation (Fig. A) and with aggregation (Fig. B). Starting with species x, we assume that each individual interacts with all the individuals of species x and y within a patch i of radius ν_x with local densities $\tilde{x_i} \equiv \frac{x_i}{2\nu_x}$ and $\tilde{y_i} \equiv \frac{y_i}{2\nu_x}$, where x_i and y_i indicate the number of individuals of species x and y in that patch. Species produce propagules at a constant per-capita rates r_x, which are released into a global propagule pool. Packets of propagules arrive at each site at a rate $\alpha \cdot \nu_x$ that increases linearly with patch size ν_x (as more packets are expected to arrive at a larger patch), and may vary with global density $\tilde{X}$, as higher global densities imply more propagules and propagules may divide into more packets at higher densities (Fig. C). Therefore, we give α as a function of $\tilde{X}$, $\alpha(\tilde{X})$. We also assume that each packet contains propagules of only one species. Given a packet of propagules settles with probability $p_s(s_x|\tilde{X}, \nu_x)$, s_x new individuals recruit into the population at the site; this shifts the population density at a site from $\tilde{x}$ to $\tilde{x} + \frac{s_x}{2\nu_x}$. Each propagule becomes a reproductive adult at settlement, and begins interacting with the other settlers and adults already in the patch. Individuals of species x in a patch die at a rate $2\nu_x(m + d_{x,x}\tilde{x} + d_{x,y}\tilde{y})$ where m is a density-independent mortality rate per unit area, $d_{x,x}$ is the intra-specific competition rate, and $d_{x,y}$ is the competitive effect of

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

y on x. The same rules described above apply for species y.

We approximate this system with multiplicative moments by using equations 1 and 2 to approximate variability in settlement in a given patch (see Appendix B for the derivation). This is equivalent to assuming that both x and y are log-normally distributed between patches (Keeling, 2000*b*). This yields a system of five equations for the mean densities $\tilde{X} = E(\tilde{x})$ and $\tilde{Y} = E(\tilde{y})$, the multiplicative variances $V_x \equiv \frac{E(\tilde{x}^2)}{\tilde{X}^2}$ and $V_y \equiv \frac{E(\tilde{y}^2)}{\tilde{Y}^2}$, and multiplicative covariance $C \equiv \frac{E(\tilde{x}\tilde{y})}{\tilde{X}\tilde{Y}}$. V_x and V_y range between 1, when all patches have the same density, and infinity. C ranges between zero, where the two species never co-occur in the same patch, and infinity. $C = 1$ when x and y are independently distributed over the landscape. The moment equations are:

$$\frac{d\tilde{X}}{dt} = r_x\tilde{X} - m_x\tilde{X} - d_{x,x}\hat{V}_x\tilde{X}^2 - d_{x,y}\hat{C}\tilde{X}\tilde{Y} \tag{3a}$$

$$\frac{d\tilde{Y}}{dt} = r_y\tilde{Y} - m_y\tilde{Y} - d_{y,y}\hat{V}_y\tilde{Y}^2 - d_{y,x}\hat{C}\tilde{X}\tilde{Y} \tag{3b}$$

$$\begin{aligned}\frac{d\hat{V}_x}{dt} =& 2r_x + r_x^2 + \frac{r_x\kappa(\alpha(\tilde{X}),\mu(\tilde{X}),\nu_x,\omega_x) + m_x}{2\nu_x\tilde{X}} + (\frac{d_{x,x}}{2\nu_x} - 2r_x)\hat{V}_x \\ & - 2d_{x,x}(\hat{V}_x - 1)\hat{V}_x^2\tilde{X} + \frac{d_{x,y}\tilde{Y}\hat{C}}{2\nu_x\tilde{X}} + 2d_{x,y}(1-\hat{C})\tilde{Y}\hat{V}_x\hat{C}\end{aligned} \tag{3c}$$

$$\begin{aligned}\frac{d\hat{V}_y}{dt} =& 2r_y + r_y^2 + \frac{r_y\kappa(\alpha(\tilde{Y}),\mu(\tilde{Y}),\nu_y,\omega_y) + m_y}{2\nu_y\tilde{Y}} + (\frac{d_{y,y}}{2\nu_y} - 2r_y)\hat{V}_y \\ & - 2d_{y,y}(\hat{V}_y - 1)\hat{V}_y^2\tilde{Y} + \frac{d_{y,x}\tilde{X}\hat{C}}{2\nu_y\tilde{Y}} + 2d_{y,x}(1-\hat{C})\tilde{X}\hat{V}_y\hat{C}\end{aligned} \tag{3d}$$

$$\frac{d\hat{C}}{dt} = (r_x + r_y)(1-\hat{C}) - (d_{x,x} + d_{y,x})(\hat{V}_x - 1)\tilde{X}\hat{C}^2 - (d_{y,y} + d_{x,y})(\hat{V}_y - 1)\tilde{Y}\hat{C}^2 \tag{3e}$$

Equations (3a) and (3b) are a modified form of the Lotka-Volterra equations where intra- and inter-specific competition rates are affected by the spatial distributions of x and y. Equations (3c) and (3d) show that either decreasing the size of the patches (the

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

spatial scale of interaction) or increasing κ (the amount of variability due to aggregated settlement), will increase V_x and V_y, as these parameters only contribute to positive terms in the equations. See table 1 for parameter definitions.

Individual-based simulations

Predictions from moment approximations can break down (Keeling, 2000*b*). We therefore compared our moment approximation from system (3) with results from an individual-based spatial simulation model. All simulations were run in R 3.0.3 (R Development Core Team, 2008), and written in c++ using the Rcpp library (Eddelbuettel et al., 2011). We ran simulations on a linear grid with 2048 patches with circular boundary conditions. Each patch i had an integer number of individuals of species x and y, and the simulation was run forward in discrete time with a time step length τ.

For all simulations, we assumed that both species in the system have identical density-independent mortality rates, and individuals increase one anotherś mortality equally via competition, regardless of species identity ($m = 0.01$, $d_{x,x} = d_{y,y} = d_{x,y} = d_{y,x} = d = 0.025$). To measure the effect of scale on coexistence, we varied the fitness inequality between the two species by altering fecundity rates, following the approach used by Berkley et al. (2010). We set $r_x = 0.11$ and $r_y = e \cdot 0.11$, where e measured the degree of fitness inequality. When $e = 1$, the two species would be ecologically neutral in a well mixed system. For $e > 1$, species y has higher fitness, and on average drives species x to extinction in a well mixed system. Therefore, e measures the strength of intra- to inter-specific competition in the well mixed system. Coexistence

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

with $e > 1$ can result either from increased intraspecific competition via V_x and V_y, or from reduced interspecific competition via C. The demographic parameters m, r_x, and d we set so that the weaker competitor would have and equilibrium population of four individuals per unit area under well-mixed conditions, to keep the total population size in each simulation small, allowing for faster simulations and more rapid extinction rates. For each time t, we simulated the following steps for each species (described here for species x for simplicity): (i) Calculate mean densities $\tilde{X}_t$ for each species at time t, and (ii) draw $n \sim Pois(2048 \cdot \tau\alpha(\tilde{X}))$ new packets from a Poisson distribution. (iii) For each packet j, draw $n_j \sim Pois(\frac{r\tilde{X}_t}{\alpha(2\nu\tilde{X}_t)})$ individuals, and set the spatial midpoint i_j of each packet from a uniform distribution. (iv) Distribute n_j settlers in packet j across the patches neighbouring i_j following a uniform distribution centered on i_j with standard deviation ω_x. This results in $s_{t+\tau,i,x}$ new settlers of species x in patch i at time $t + \tau$. (v) Calculate the number of individuals l dying in each patch i with $l_{t+\tau,i,x} \sim Pois(\tau(mx_{t,i} + dx_{t,i}\int_{i-\nu_x}^{i+\nu_x}\frac{x_{t,j}}{2\nu_x}dj + dx_{t,i}\int_{i-max(\nu_x,\nu_y)}^{i+max(\nu_x,\nu_y)}\frac{y_{t,j}}{2\nu_y}dj))$, where the integrals represent the interaction kernel: the death rate increases as the average density of x and y increase in an area of radius ν_x around i. (vi) Finally, combine births and deaths to obtain $x_{t+\tau,i} = x_{t,i} + s_{t+\tau,i,x} - min(x_{t,i}, l_{t+\tau,i.x})$. The minimum function prevents mortality from exceeding density in the patch at time t.

This a form of the τ-leap algorithm for approximating continuous-time stochastic systems (Gillespie, 2007), with a fixed τ step size. Each simulation was run for a length of 1000, with 32000 steps ($\tau \approx 0.03$). As this is a stochastic simulation with a finite carrying capacity, over long enough time periods both species will eventually go extinct. Therefore, we used the time when the inferior competitor (x) went globally extinct as

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

our metric of coexistence. Our results were quantitatively similar for simulations ran for lengths of 500 (not shown), indicating our results are robust to simulation time.

Results

Approximating spatial and temporal scales of settlement

Equation (2) implies that settlement variability depends heavily on the difference between the scale of aggregation (ω) and the scale of interaction (ν). Variability drops off substantially when $\frac{\nu}{\omega} < 1$. For example, in a one-dimensional system, equation (2) predicts that patch size corresponding to 10% of the scale of aggregation results in settlement variation at only 5% of its maximum value. However, when $\frac{\nu}{\omega} \gg 1$, increasing the scale of interaction or decreasing the scale of settlement only slightly increases variability; if patches are 100 times larger than aggregations, settlement variation will only be twice as high as when the two scales are equal.

Equation (2) also shows the importance of the temporal scale of aggregation for predicting settlement variability. Variability increases as each packet becomes denser (and therefore less frequent). Further, settlement variability depends on the relation between the number of individuals in a packet and the number of available propagules. For aggregation mechanisms such as eddies, where packets tend to arrive at a constant rate but the number of individuals in a packet increases with the number of available propagules (density-dependent packet size), the variance to mean ratio of settlement increases with population density. For aggregation mechanisms such as seed pods, the number of individuals per packet is independent from propagule density (fixed packet

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

size), and the variance to mean ratio remains constant across population densities. This means that rare species will tend to experience lower settlement variability than abundant species in the former case but not in the latter.

Coexistence in a metacommunity with aggregated dispersal

A species will generally only be able to persist if its average growth rate is positive at low density (in the absence of allee effects) (Chesson, 2000*b*). In our metacommunity model (system 3), setting x as the invading species, we can find its growth rate at low density by setting y to its single-species equilibrium density $\tilde{Y}^*$ and multiplicative variance V_y^*. We then assume there is only 1 individual of x in every n patches. This means that $\tilde{X} = \frac{1}{2n\nu_x}$, and $V_x = n$. Using equation (3a), the mean growth rate for x will be greater than zero if:

$$\begin{aligned} 0 &< \frac{r_x - m_x}{2n\nu_x} - d_{x,x}\frac{n}{4n^2\nu_x^2} - d_{x,y}\frac{C\tilde{Y}^*}{2n\nu_x} \\ C\tilde{Y}^* &< \frac{r_x - m_x - \frac{d_{x,x}}{2\nu_x}}{d_{x,y}} \end{aligned} \tag{4}$$

As expected, anything that reduces either $\tilde{Y}^*$ or the degree of spatial co-occurrence of the two species will promote coexistence. From equation (3), we can see that any factor that increases V_y would, all else equal, reduce both $\tilde{Y}^*$ and C. Note that the factor $d_{x,x}$ generally drops out in invasion analysis, as most models assume no self-competition for the invading population. This assumption is incompatible with the infinite population moment closure method we used because mean density would then becomes $\frac{1}{2\nu}$ and be allowed to increase even in very small patches. Simulations with and without

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

self-competition showed that our results are robust to this limitation of our approximation method (not shown).

Equation (4) reveals the influence of the spatial scale of interaction, ν, on V_y through two antagonistic mechanisms. Reducing ν directly increases V_y, by increasing the effect of demographic stochasticity on local population dynamics. However, when dispersal is aggregated, reducing ν decreases κ. This is because, when patch size is small, individuals effectively do not see the additional variability added by the arrival of a packet. Lower κ, in turn, acts to reduce V_y. These two opposing forces mean that changing interaction scales will not have a simple monotonic effect on coexistence. We now turn to numerical and stochastic simulations to resolve the net effect of interaction scale on coexistence.

Coexistence as a function of interaction and aggregation scale

In the absence of aggregated dispersal, both moment equations or stochastic simulations show very little effect of the spatial scale of interaction on coexistence. We did not observe any spatial scale where the two species could coexist when local inter-specific competition was higher than intra-specific competition (not shown).

In the presence of aggregated dispersal, coexistence depends heavily on the relative scales of interaction and aggregation of the two species. For both fixed density transport (Fig. 3A) and density-dependent transport (Fig. 3B), the inferior competitor is able to coexist if the superior competitor interacts at a smaller scale or is more densely aggregated within packets than the inferior competitor.

When both species interact and are aggregated at the same scales, the maximum

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

interaction strength allowing for coexistence occurs at intermediate scales of interaction and at high propagule density in packets, approximately where the scale of interaction equals the scale of aggregation (Fig. A&B dashed lines). The moment equations also predict that species will be able to coexist at higher levels of competition under density-dependent packets (Fig. B) compared with fixed-size packets (Fig. A).

The effect of aggregation scale on coexistence strongly depends on the scale of interaction (Fig. 5). When the aggregation scale (ω) is smaller than the scale of interaction (ν, Fig. 5 dotted line), reducing aggregation scales has no effect on coexistence. Only varying the mean number of individuals per packet will have an effect. However, when $\omega > \nu$, reducing the scale of aggregation or increasing the mean number of individuals per packet can promote coexistence.

The fixed density and density-dependent packet transport models showed very similar responses to parameter changes. However, under all conditions (Fig. 3,, and 5), extinction times were shorter and conditions for coexistence were more stringent for the fixed-density model relative to the density-dependent model. Further, with fixed packet sizes, the weaker competitor went extinct even when the moment approximation (equation (4)) predicted coexistence.

The coexistence criteria derived in equation (4) were able to accurately predict coexistence in the simulations except at high levels of fitness inequality and small interaction scales (fig. right). The mismatch between moment equations and simulations at high rates of competitive inequality may be due to the populations not following log-normal distributions at small scales or high fecundity, thus violating the assumptions used to construct the moment equations (Bolker and Pacala, 1997,

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Keeling, 2000*a*, Bolker, 2003). The moment approximation was also not able to predict the differing patterns of extinction between fixed and density-dependent packet models, as one of the assumptions made in the approximation was was that true extinction is not possible.

Discussion

Our work suggests it is possible to approximate and extend our understanding of coexistence under aggregated dispersal by considering three key scales: the spatial scale of interaction among settlers, and both the spatial and temporal scales of aggregation during dispersal. Our results broaden the predicted range of spatial scales allowing aggregated dispersal to work as a stabilizing mechanism of coexistence. Competitors interacting at scales one to two orders of magnitude larger than the scale of aggregation can still successfully coexist at low levels of competitive inequality. Coexistence is, however, strongly sensitive to the time-scale of aggregation. Increasing the frequency of arrival of aggregated individuals (packets) substantially reduces the region of fitness inequalities where both species persist. Our results also reveal the role of density-dependent aggregation on coexistence through its impact on the strength of intra-specific competition among settlers. Extending the nature and range of scales within theories of coexistence can improve their applicability to natural systems where multiple transport mechanisms mediate spatiotemporal patterns of dispersal and aggregation. By scaling up individual aggregation during dispersal to the spatial distribution of aggregated communities, we provide a theory of metacommunity

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

networks emerging from the movement and interaction among individuals, rather than as a imposed feature of the landscape.

Coexistence across scales of aggregation and interaction

Aggregation scale is related but distinct from dispersal scale in that it determines how closely propagules settle to one another rather than how far they settle from their parents. As decreasing the scale of aggregation will always increase local intraspecific interactions, coexistence should always be easier under smaller aggregation scales, whereas decreasing dispersal scales may strengthen or weaken stabilizing mechanisms (Bolker and Pacala, 1999, Snyder and Chesson, 2004).

Interaction scale has been identified previously as a key factor determining coexistence (Snyder and Chesson, 2003, 2004), but its effects tend to be ignored in metacommunity theory, where patches are typically treated as static and the same size for all species. We have shown that coexistence is easiest when the stronger competitor interacts at smaller spatial scales, and when species interact at scales smaller than they aggregate. The time-scale of aggregation in our model is a unique property of aggregated dispersal processes, and our work shows that coexistence is strongly sensitive to this scale. Increasing the frequency of arrival of aggregated individuals (packets) substantially reduces the region of fitness inequalities where both species persist. In marine systems, this predicts that ecological factors such as length of reproductive seasons or physical features such as eddy rotation time will have a larger impact on coexistence than short-range spatial mechanisms, such as small-scale ocean currents or post-settlement movement.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

The effects of the three scales on coexistence outcomes are not additive, a result predicted by equation (2). All three scales have thresholds which, if exceeded, prevent coexistence no matter the value of the other scales. Our work also highlights an important distinction between fixed density and density-dependent packet forming mechanisms. We have shown that global extinction rates were substantially higher and parameter regions allowing coexistence were smaller with fixed packet sizes. With density-dependent packets, aggregation will decline with global density. This in turn reduces intra-specific competition at low densities. However, when packet densities are fixed, new settlers will still settle in high density even when their global density is low, increasing their chance of extinction, as any factor that increases variability at low densities will also increase the rate of extinction due to stochastic fluctuations (Nisbet and Gurney, 1982). This illustrates the joint role variability plays in both coexistence and stochastic extinction, and the difficulty of separating their effects (Gravel et al., 2011). The effect of aggregated dispersal on extinction has been studied previously with regards to survival in systems with advective transport (Kolpas and Nisbet, 2010), diffusive transport (Williams and Hastings, 2013) and in the presence of allee effects (Rajakaruna et al., 2013), but all these approaches assumed density-dependent packet transport. Density-dependent packet formation will occur when aggregations are formed by correlated physical transport mechanisms such as eddy-driven dispersal. Fixed packet dispersal will occur when a given aggregation mechanism strongly controls the number of propagules able to move in a given packet, including many biological aggregation processes such as seed pods, egg clusters, or animals eating seeds and depositing them in faeces.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Where do we expect aggregated dispersal to play a role in shaping community structure?

Our results demonstrate that aggregated dispersal increases coexistence rates most strongly when individuals interact over small spatial scales, when each packet of settlers is small, and each packet carries a large number of propagules. As such, this mechanism will have substantially different effects on coexistence outcomes depending on the effective scales of interaction and aggregation in a given system. There are two types of systems where aggregated dispersal has been suggested to play a role in coexistence: larval aggregation in eddies (Potthoff et al., 2006, Berkley et al., 2010) and animal seed transport in terrestrial plant communities (Muller-Landau and Hardesty, 2005, Potthoff et al., 2006).

In marine systems the effects of aggregated dispersal on community composition will depend on two factors: eddy size and the scale of post-settlement species interactions. As eddies get larger, there will be less inter-eddy spaces, and thus the density of larvae in each packet will increase (Siegel et al., 2008), equivalent to increasing the time-scale of aggregation. As eddy size decreases strongly with increasing latitude (Chelton et al., 2011), we predict that the strength of this stabilizing mechanism will decrease in regions close to the poles. This may explain a striking empirical regularity: species richness declines sharply with latitude for marine organisms with a pelagic larval stage, but increases for species with no pelagic larval stage (Fernández et al., 2009). If aggregated dispersal is driving this pattern, we would also expect that the negative latitude-diversity gradient should be steepest for sessile or strongly territorial species relative to those that move over larger areas as adults, as sessile species will interact

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

over shorter spatial scales.

For terrestrial plant communities with animal dispersal, three factors will drive the strength of this stabilizing effect: how many seeds each disperser deposits at a time, post-deposition secondary dispersal, and the type of processes limiting plant establishment. The number of seeds a disperser deposits will determine the time-scale of aggregation, and should be related to its body size (Howe, 1989). Therefore, systems where larger animals are the primary seed dispersers should show higher diversity than those where dispersal by small animals or wind dominates. Also, any process that increases post-deposition spread, such as ants moving seeds (Passos and Oliveira, 2002) will reduce this stabilizing effect by increasing the scale of aggregation. Finally, the effect will be weakest for plants that need large areas to successfully establish, as the scale of interaction increases with plant size (Vogt et al., 2010).

Accounting for aggregation and scale in general meta-community models

Our aggregation approximation, equation (2), captured the dynamic effects of aggregation on population dynamics, and should be useful for modelling aggregated dispersal more generally. Aggregated dispersal has been shown to shape metacommunity dynamics beyond its effect on coexistence, by increasing extinction rates (Williams and Hastings, 2013), decreasing overall growth rates (Snyder et al., 2014), altering rates of spatial spread (Ellner and Schreiber, 2012), or reducing predation (Beckman et al., 2012).

Several mechanisms have recently been shown to promote coexistence in

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

metacommunities via species-specific patterns of connectivity. These include asymmetrical between-patch dispersal or variability of the strength of self-recruitment between competitors (Salomon et al., 2010, Figueiredo and Connolly, 2012, Aiken and Navarrete, 2014), irregular patch distribution coupled with interspecific variation in dispersal rates (Bode et al., 2011), or edge effects in the presence of advective dispersal (Aiken and Navarrete, 2014). These studies illustrate the usefulness of the metacommunity framework for understanding the effects of dispersal mechanisms on coexistence when dispersal takes place over large scales. By abstracting the system into patches and the pattern of connections between them, it is much easier to model complex patterns of connectivity or landscape structure relative to continuous models. However, there are relatively few natural systems where a single spatial scale of species interactions can be identified, and our work shows that the effectiveness of a given coexistence mechanism can be sensitive to assumptions about scales of interactions, and in particular about their variation among species. While this is known from prior theoretical work in local continuous spatial systems (Snyder and Chesson, 2003, 2004), it has been generally overlooked in the study of coexistence across metacommunities that are meant to capture a broad range of spatial scales. The approach we used, making patch size a species-specific parameter, is generally extensible to any metacommunity model and captures one aspect of interaction scale: shorter scale of non-linear interactions can enhance the effect of stochastic forces relative to deterministic processes.

Our approach can be seen as part of a broader mechanistic approach for integrating dispersal mechanisms to metacommunity theories. Rather than starting with the

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

assumption of a patch network, we predict this network by scaling up individual aggregated dispersal to spatio-temporal patterns of settlement, and by approximating metacommunity dynamics with a species specific scale of interactions. This approach, described by Black and McKane (2012) as deriving a population-level model from an individual-based model, allowed us to not only determine which scales were critical for coexistence, but also to identify the limits of the metacommunity as a useful model of spatial dynamics.

To include aggregated dispersal into metacommunity theory, we have to recognize that the choice of patch size (and thus interaction scale) will strongly affect dynamic outcomes. Our work shows how aggregated dispersal can be incorporated into metacommunity models from first principles, and what key processes need to be measured for a given aggregation process to understand its dynamic effects.

Acknowledgments

We would like to thank Janine Illian for many insightful discussions on point processes and spatial statistics. We also thanks Carly Ziter, Patrick Thompson, Michael Becker, Justin Marleau, Brian Leung, Gregor Fussmann, David Green, and Sean Connolly for their feedback on prior versions of this work. We thank Calcul Québec for the use of the CLUMEC cluster for individual-based simulations. EJP and FG are pleased to acknowledge support from the Natural Science and Engineering Research Council (NSERC) through its support to the Canadian Healthy Oceans Network.

References

Aiken, C. M., and S. A. Navarrete. 2014. Coexistence of competitors in marine metacommunities: Environmental variability, edge effects, and the dispersal niche. Ecology **95**:2289–2302.

Beckman, N. G., C. Neuhauser, and H. C. Muller-Landau. 2012. The interacting effects of clumped seed dispersal and distance- and density-dependent mortality on seedling recruitment patterns. Journal of Ecology **100**:862–873.

Berkley, H. A., B. E. Kendall, S. Mitarai, and D. A. Siegel. 2010. Turbulent dispersal promotes species coexistence. Ecology Letters **13**:360–371.

Black, A. J., and A. J. McKane. 2012. Stochastic formulation of ecological models and their applications. Trends in Ecology & Evolution **27**:337–345.

Bode, M., L. Bode, and P. R. Armsworth. 2011. Different dispersal abilities allow reef fish to coexist. Proceedings of the National Academy of Sciences **108**:16317–16321.

Bolker, B., and S. W. Pacala. 1997. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. Theoretical Population Biology **52**:179–197.

Bolker, B. M. 2003. Combining endogenous and exogenous spatial variability in analytical population models. Theoretical Population Biology **64**:255–270.

Bolker, B. M. 2008. Ecological models and data in R. Princeton University Press, Princeton, N.J., U.S.A.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Bolker, B. M., and S. W. Pacala. 1999. Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. The American Naturalist **153**:575–602.

Cavanaugh, K. C., D. A. Siegel, P. T. Raimondi, and F. Alberto. 2014. Patch definition in metapopulation analysis: A graph theory approach to solve the mega-patch problem. Ecology **95**:316–328.

Chelton, D. B., M. G. Schlax, and R. M. Samelson. 2011. Global observations of nonlinear mesoscale eddies. Progress In Oceanography **91**:167–216.

Chesson, P. 2000*a*. General theory of competitive coexistence in spatially-varying environments. Theoretical Population Biology **58**:211–237.

Chesson, P. 2000*b*. Mechanisms of maintenance of species diversity. Annual Review of Ecology and Systematics **31**:343–366.

Chesson, P., M. J. Donahue, B. Melbourne, and A. L. W. Sears, 2005. Scale transition theory for understanding mechanisms in metacommunities. *in* M. Holyoak, M. A. Leibold, and R. D. Holt, editors. Metacommunities: Spatial Dynamics and Ecological Communities. University of Chicago Press.

Chesson, P., and C. Neuhauser. 2002. Intraspecific aggregation and species coexistence. Trends in Ecology & Evolution **17**:210–211.

Eddelbuettel, D., R. François, J. Allaire, J. Chambers, D. Bates, and K. Ushey. 2011. Rcpp: Seamless R and C++ integration. Journal of Statistical Software **40**:1–18.

Edwards, K., and J. Stachowicz. 2011. Spatially stochastic settlement and the coexistence of benthic marine animals. Ecology **92**:1094–1103.

Ellner, S. P., and S. J. Schreiber. 2012. Temporally variable dispersal and demography can accelerate the spread of invading species. Theoretical Population Biology **82**:283–298.

Fernández, M., A. Astorga, S. A. Navarrete, C. Valdovinos, and P. A. Marquet. 2009. Deconstructing latitudinal species richness patterns in the ocean: Does larval development hold the clue? Ecology Letters **12**:601–611.

Figueiredo, J., and S. R. Connolly. 2012. Dispersal-mediated coexistence under recruitment limitation and displacement competition. Ecological Modelling **243**:133–142.

Gause, G. F. 1932. Experimental studies on the struggle for existence. Journal of Experimental Biology **9**:389–402.

Gillespie, D. 2007. Stochastic simulation of chemical kinetics. Annual Review of Physical Chemistry **58**:35–55.

Gravel, D., F. Guichard, and M. E. Hochberg. 2011. Species coexistence in a variable world. Ecology Letters **14**:828–839.

Harms, K. E., S. J. Wright, O. Calderón, A. Hernández, and E. A. Herre. 2000. Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. Nature **404**:493–495.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Harrison, C. S., D. A. Siegel, and S. Mitarai. 2013. Filamentation and eddy-eddy interactions in marine larval accumulation and transport. Marine Ecology Progress Series **472**:27–44.

Howe, H. F. 1989. Scatter-and clump-dispersal and seedling demography: Hypothesis and implications. Oecologia **79**:417–426.

Hutchinson, G. E. 1961. The paradox of the plankton. American Naturalist **95**:137–145.

Illian, J., A. Penttinen, H. Stoyan, and D. Stoyan. 2008. Statistical analysis and modelling of spatial point patterns. John Wiley, West Sussex, U.K.

Ives, A. R., and R. M. May. 1985. Competition within and between species in a patchy environment: Relations between microscopic and macroscopic models. Journal of Theoretical Biology **115**:65–92.

Keeling, M. J. 2000*a*. Metapopulation moments: Coupling, stochasticity and persistence. Journal of Animal Ecology **69**:725–736.

Keeling, M. J. 2000*b*. Multiplicative moments and measures of persistence in ecology. Journal of Theoretical Biology **205**:269–281.

Kinlan, B. P., and S. D. Gaines. 2003. Propagule dispersal in marine and terrestrial environments: A community perspective. Ecology **84**:2007–2020.

Kolpas, A., and R. Nisbet. 2010. Effects of demographic stochasticity on population persistence in advective media. Bulletin of Mathematical Biology **72**:1254–1270.

Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D. Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau, and A. Gonzalez. 2004. The metacommunity concept: A framework for multi-scale community ecology. Ecology Letters **7**:601–613.

Muller-Landau, H. C., and B. D. Hardesty, 2005. Seed dispersal of woody plants in tropical forests: Concepts, examples and future directions. Pages 267–309 *in* D. F. R. P. Burslem, M. A. Pinard, and S. E. Hartley, editors. Biotic Interactions In The Tropics: Their Role In The Maintenance Of Species Diversity. Cambridge University Press.

Nisbet, R. M., and W. S. C. Gurney. 1982. Modelling fluctuating populations. The Blackburn Press, Caldwell, N.J.

Passos, L., and P. S. Oliveira. 2002. Ants affect the distribution and performance of seedlings of *Clusia criuva*, a primarily bird-dispersed rain forest tree. Journal of Ecology **90**:517–528.

Potthoff, M., K. Johst, J. Gutt, and C. Wissel. 2006. Clumped dispersal and species coexistence. Ecological Modelling **198**:247–254.

R Development Core Team, 2008. R: A language and environment for statistical computing. URL `http://www.R-project.org`.

Rajakaruna, H., A. Potapov, and M. Lewis. 2013. Impact of stochasticity in immigration and reintroduction on colonizing and extirpating populations. Theoretical Population Biology **85**:38–48.

Ritchie, M. E. 2009. Scale, heterogeneity, and the structure and diversity of ecological communities. Princeton University Press, Princeton, N.J.

Salomon, Y., S. R. Connolly, and L. Bode. 2010. Effects of asymmetric dispersal on the coexistence of competing species. Ecology Letters **13**:432–441.

Sheth, R. K., and W. C. Saslaw. 1994. Synthesizing the observed distribution of galaxies. The Astrophysical Journal **437**:35–55.

Siegel, D. A., S. Mitarai, C. J. Costello, S. D. Gaines, B. E. Kendall, R. R. Warner, and K. B. Winters. 2008. The stochastic nature of larval connectivity among nearshore marine populations. Proceedings of the National Academy of Sciences **105**:8974.

Snyder, R. E. 2008. When does environmental variation most influence species coexistence? Theoretical Ecology **1**:129–139.

Snyder, R. E., and P. Chesson. 2003. Local dispersal can facilitate coexistence in the presence of permanent spatial heterogeneity. Ecology Letters **6**:301–309.

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

Snyder, R. E., C. B. Paris, and A. C. Vaz. 2014. How much do marine connectivity fluctuations matter? The American Naturalist **184**:523–530.

Vogt, D. R., D. J. Murrell, and P. Stoll. 2010. Testing spatial theories of plant coexistence: No consistent differences in intra- and interspecific interaction distances. The American Naturalist **175**:73–84.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

578 Williams, P. D., and A. Hastings. 2013. Stochastic dispersal and population persistence
579 in marine organisms. The American Naturalist **182**:271–282.

Table 1: Parameters of packet-transport approximation and moment closure model.

α	The mean arrival rate of packets at any given point in space.
μ	The mean number of individuals per packet.
σ	The standard deviation of the number of individuals between packets.
ω_x, ω_y	The scale of aggregation.
ν_x, ν_y	The scale that each species interacts with its neighbours at.
κ	Degree of settlement aggregation at a given scale over a given period of time. Ranges from 1 to ∞.
r_x, r_y	Instantaneous per-area settlement rates for species x and y.
e	Ratio of per-capita fecundity of species y to x (r_y/r_x). Measures fitness inequality between the two species.
m_x, m_y	Instantaneous per-area density-independent mortality rates for species x and y.
$d_{x,x}, d_{y,y}$, $d_{x,y}, d_{y,x}$	Interaction rates. Measures the degree to which mortality rate of species x within a patch increases with the density of species y.
V_x, V_y	Multiplicative variance of species x and y.
C	Multiplicative covariance between species x and y.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

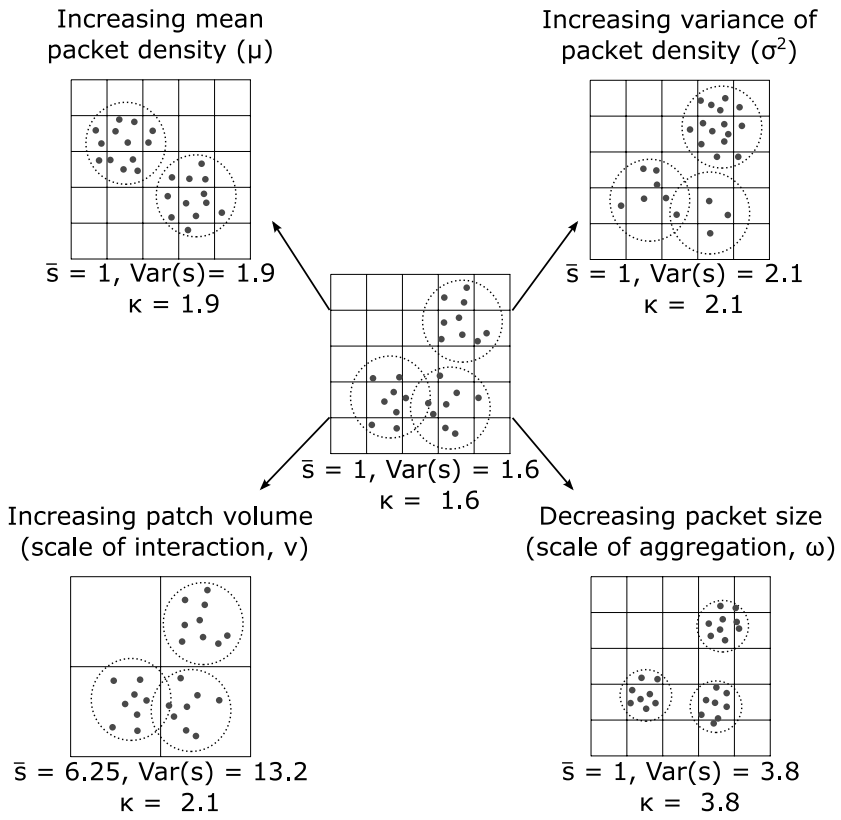

Figure 1: Factors that affect κ, the ratio of variance to mean of settlement under aggregated dispersal. Settlers (black dots) are aggregated into packets (dashed circles), and interact within patches (grid cells). κ is affected by the average ($\bar{s}$) and variance ($Var(s)$) in the density of settlers arriving in each patch, which are in turn affected by time scale of packet arrival (mean packet density μ, top left), variability of densities between packets (σ^2, top right), the scale of interactions (patch size v, bottom left), and the scale of aggregated dispersal (packet size ω, bottom right).

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

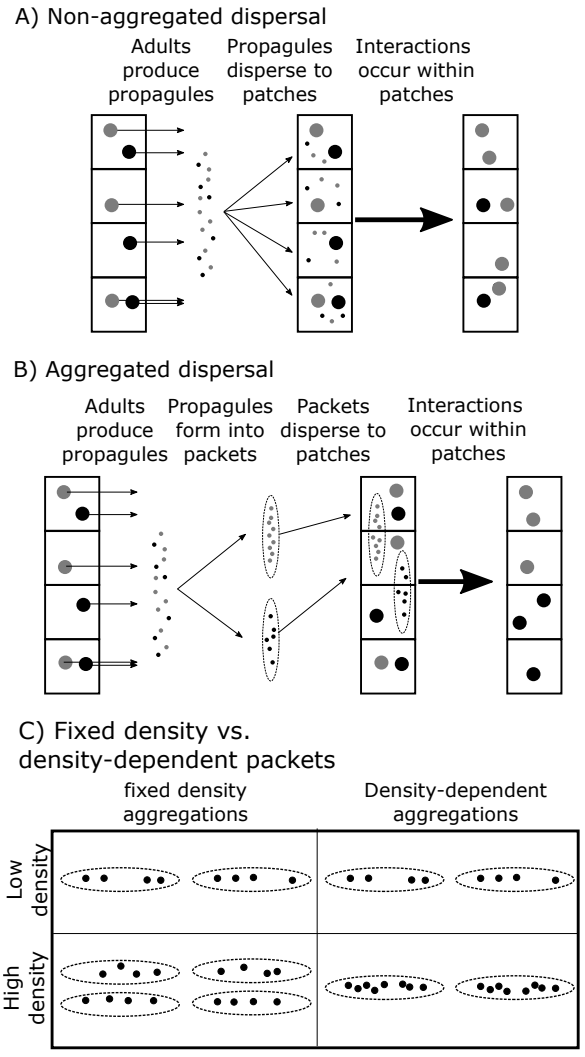

Figure 2: Population dynamics. A) Non-aggregated dispersal. Adults (large circles) of species x (black) and y (grey) produce propagules (small circles) and release them into a global propagule pool. These are randomly distributed among sites. New residents face mortality from interactions within their patch. B) Aggregated dispersal. propagules aggregate into packets (dashed ovals), and packets settle at random locations. C) Two models of propagule aggregation. Either propagules are aggregated into a fixed number of packets (top), or each packet holds a fixed number of propagules (bottom).

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

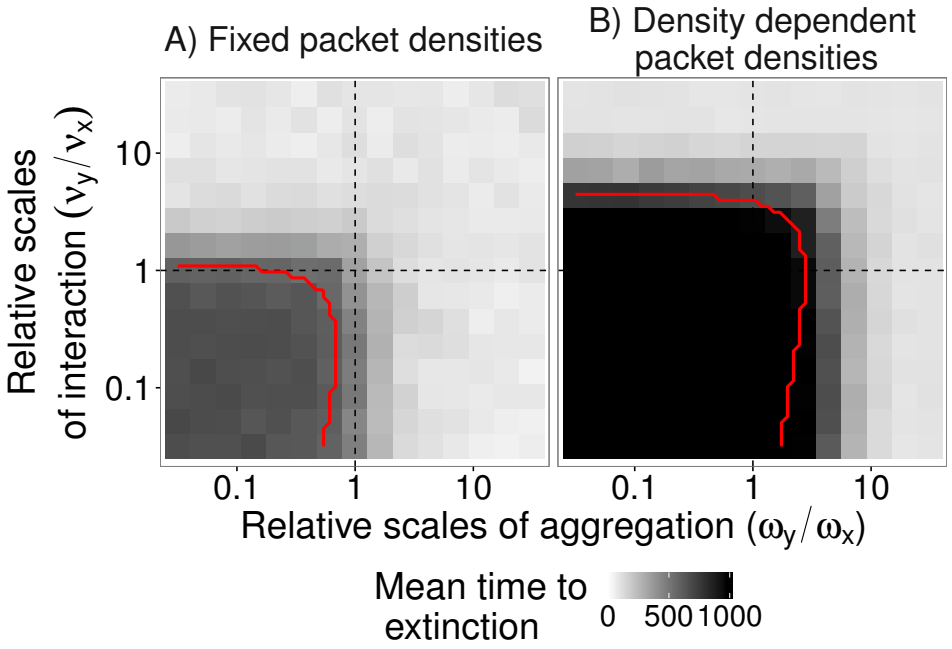

Figure 3: Regions of coexistence with interspecific variability in aggregation ($\frac{\omega_2}{\omega_1}$) and interaction ($\frac{\nu_2}{\nu_1}$) scales with aggregated dispersal with fixed packet sizes (A&C) and density-dependent packet sizes (B&D). The scale parameters for species x are held constant ($\nu_x = 10^{1.5}$ and $\omega_x = 10^{1.5}$, chosen to allow the widest range of relative scales of aggregation and interaction given the length of the simulation domain); ν_y and ω_y are allowed to vary. For all simulations, species y is 4 times more fecund than species x ($r_x = 0.11$, $r_y = 0.44$). Shading indicates mean time to extinction for the weaker competitor. The area below the red line corresponds to coexistence as predicted by equation (4). Dashed lines indicate equal scales of either aggregation or interaction between species.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

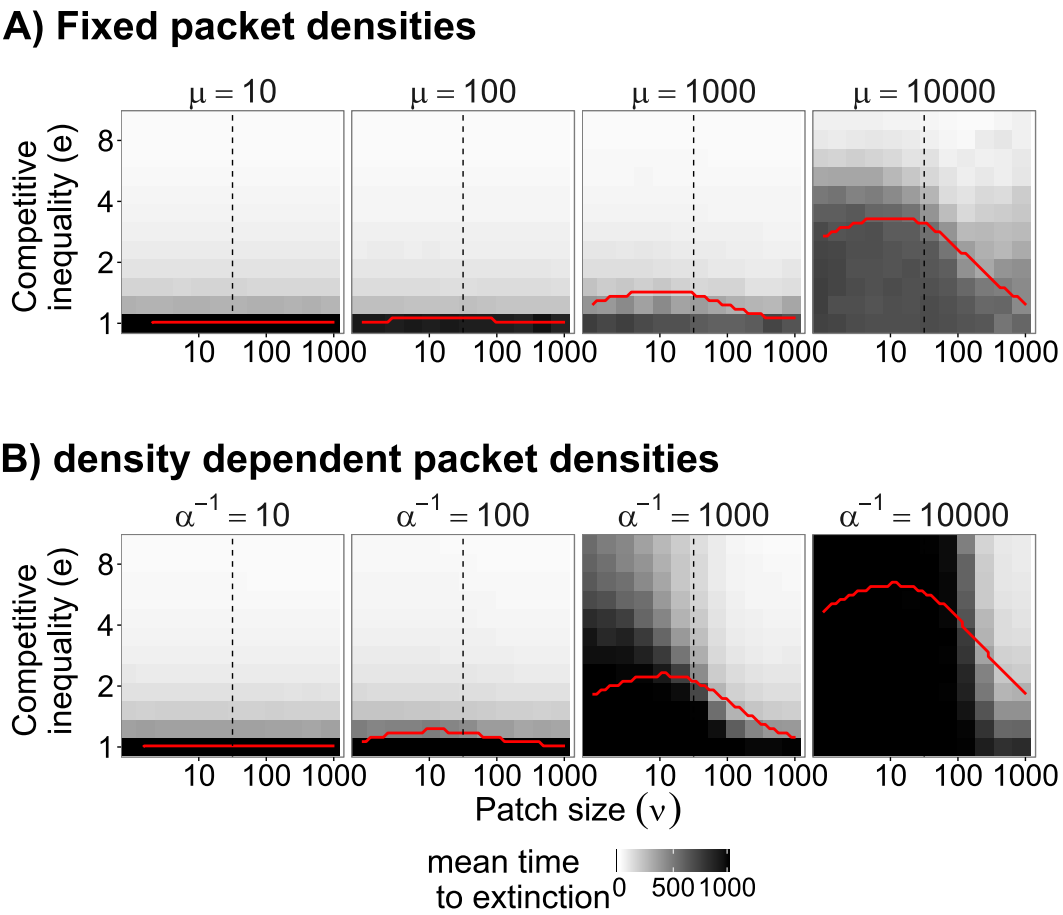

Figure 4: Regions of coexistence with aggregated dispersal and (A) with fixed packet sizes or (B) density-dependent packet sizes when both species interact and aggregate on the same scales. From left to right, plots represent increasing mean packet densities (μ) for the fixed density plots (A) and increasing time between packet arrivals (α^{-1}) for the density dependent plots (B). The dashed line indicates equal scales of interaction and of aggregation ($\omega = 10^{1.5}$). Shading indicates mean time to extinction for the weaker competitor.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

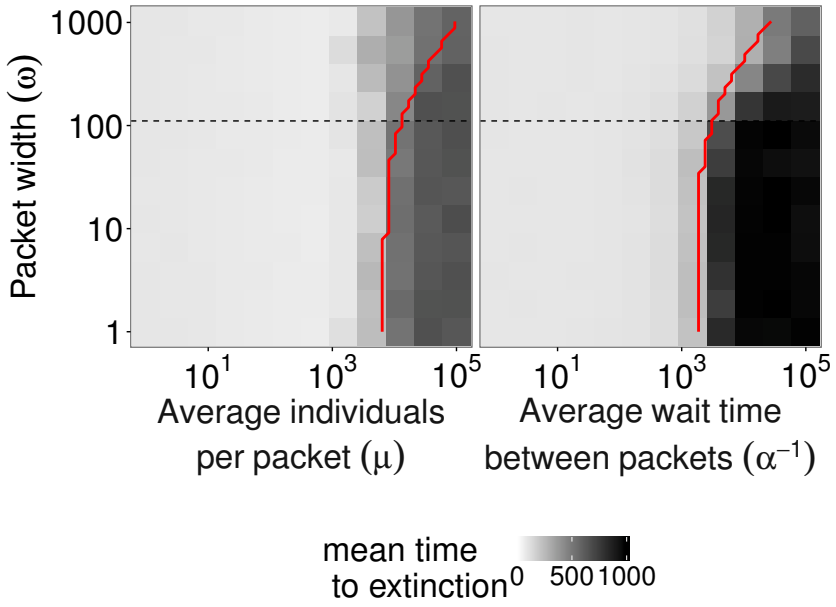

Figure 5: Regions of coexistence with aggregated dispersal and fixed (A) or density-dependent (B) packet sizes, as a function of the scale of aggregation ω and of mean packet density μ (A), or temporal scale of settlement α^{-1} (B). Shading indicates mean time to extinction for the weaker competitor, and the dashed line indicates equal scales of aggregation and of interaction between species.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Appendix A: Deriving the aggregation parameter κ

To understand how aggregation-driven clustering affects population dynamics, it is necessary to distill the aggregation process down to the essential elements which will affect population dynamics when altered.

We start with a abstract version of our dispersal process: settlers are treated as points arriving on a (one or two-dimensional) landscape, aggregated in packets (which may overlap with one another in space). All the mechanisms of the aggregation process are assumed to affect only two properties: the distribution of settlers between packets, and the spatial arrangement of settlers around the packet center. This describes what is called a spatio-temporal point process (Illian et al., 2008).

We first assume that the scale over which packets travel is much bigger than the size of the packets themselves, so that edge effects are not an issue. Further, we assume we are looking at a small enough region of space that packets are equally likely to arrive at any point in space, and that the distribution of settlers between packets is the same everywhere; that is, space is homogeneous. We also assume that all that affects the probability of finding a settler at a given location in a specific packet is how far that settler is from the packet center (the isotropic assumption).

We can then define $\alpha(t)$, the intensity[1] of packets arriving at any given point in time in any given location. This may or may not fluctuate over time or with population density. For instance, if the aggregation is caused by meso-scale eddies, the number of packets arriving at any given point in time should not vary with population density, but it will

[1]The intensity is a property of point processes, and is defined as the value that when integrated over a finite segment of space and time equals the mean number of individuals expected to be found in that segment.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

likely vary with season or latitude. On the other hand, if aggregation mechanism is seeds travelling in seed heads, the number of packets arriving will vary with the population density and fecundity of the source population.

If we measure settlement aggregation over a short enough period of time, then $\alpha(t)$ should not vary substantially in that period. Given that, we can then define $\lambda_c(t,\tau) = \int_t^{t+\tau} \alpha(t)dt$ as the intensity of clusters per unit area in the finite interval $t \to t+\tau$. This turns the model from a spatio-temporal point process to a spatial point process, for which significantly more is known. For the remainder of the derivation, for simplicity of notation, we will drop the time-dependent terms, and simply refer to λ_c. The resulting point pattern describes the pattern of new setters arriving in that time interval. If packets each have a mean of μ individuals, we can also define $\lambda = \mu\lambda_c$ as the intensity of settlers on the landscape in that time period.

Finally, we need to determine how we are going to measure clustering on the landscape. Here, we assume space is broken into a number of patches all the same shape, W. These patches have a length scale, ν, which we define as the radius of a d - dimensional ball with the same volume as W. The function $Vol_d(\nu)$ defines the volume of any shape with a length-scale ν in a system of dimension d. In one dimension, $Vol_d(\nu) = 2\nu$, in two dimensions $Vol_d(\nu) = \pi\nu^2$. From this point on, we drop the subscript d to simplify notation. Given our assumptions about homogeneity and isotropy, each patch will have an expected number of individuals $\lambda Vol(\nu) = \lambda_c \mu Vol(\nu)$, regardless of the shape of either the packets or patches. Given this, we want to know the variance of counts between patches.

Given our assumptions, we can determine what the ratio of variance to mean of the

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

count of points, N, in a sample area of volume ν will be (derived from Illian et al. (2008), page 226. (See Table A1 and A2 for the definition of terms in this derivation):

$$\kappa = \frac{Var(N)}{\lambda Vol(\nu)} = 1 + \frac{\lambda}{Vol(\nu)} db_d \int_0^\infty \bar{\gamma}_W(r)(g(r)-1)r^{d-1}dr \tag{5}$$

For any Neyman-Scott cluster-point process defined by the above parameters and functions, $g(r) = 1 + \frac{1}{\lambda\mu}\sum_{n=2}^{\infty} p_n n(n-1)\frac{f_d(r)}{db_d r^{d-1}}$ (Illian et al., 2008) . Therefore:

$$\kappa = 1 + db_d \frac{\lambda}{Vol(\nu)} \int_0^\infty \bar{\gamma}_W(r) r^{d-1}(1 + \frac{1}{\lambda\mu}\sum_{n=2}^{\infty} p_n n(n-1)\frac{f_d(r)}{db_d r^{d-1}} - 1)dr \tag{6a}$$

$$= 1 + \frac{db_d}{db_d}\frac{\lambda}{Vol(\nu)\lambda} \int_0^\infty \bar{\gamma}_W(r)\frac{r^{d-1}}{r^{d-1}}(\frac{1}{\mu}\sum_{n=2}^{\infty} p_n n(n-1) f_d(r))dr \tag{6b}$$

$$= 1 + \frac{1}{Vol(\nu)} \int_0^\infty \bar{\gamma}_W(r)(\frac{1}{\mu}\sum_{n=2}^{\infty} p_n n(n-1) f_d(r))dr \tag{6c}$$

$$= 1 + \frac{\sum_{n=2}^{\infty} p_n n(n-1)}{Vol(\nu)\mu} \int_0^\infty \bar{\gamma}_W(r) f_d(r) dr \tag{6d}$$

$$= 1 + \frac{\sigma^2 + \mu^2 - \mu}{\mu} \int_0^\infty \bar{\gamma}_W(r) f_d(r) Vol(\nu)^{-1} dr \tag{6e}$$

There are two important things to note about this equation. First, that the intensity of clusters on the landscape, $\lambda_c(t,\tau)$, does not enter into the equation. This means that κ will not vary with the length of the period in which we measure settlement clustering, as long as all the parameters of the aggregation process stay the same over that length of time. Second, this formulation cleanly separates all the interacting factors that determine how the aggregation process will affect settlement variability. The terms inside the integral, which we will refer to as the scale function $S(W, f_d)$, capture the interacting effects of dimension, packet shape, and patch shape and volume. The terms

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

outside the integral, which we will refer to as the mean function, $M(\mu, \sigma)$, capture the distribution of points between packets. The whole function can then be described as $\kappa = 1 + M(\mu, \sigma) \cdot S(W, f_d)$.

The form of the scale function, $S(W, f_d)$, is still quite complex. However, given the definitions of the functions, there are several inferences we can make about its properties: 1) As $f_d(r)$ is a probability distribution, and therefore has to be positive and integrate to one, and given the definitions $\bar{\gamma}_W(r)$ and $Vol(\nu)$, $\frac{\bar{\gamma}_W(r)}{Vol(\nu)}$ has to be between zero and one, then $S(W, f_d)$ must be between zero and one. 2) If a given shape is very small relative to the scale of the packet, $\bar{\gamma}_W(r)$ will drop to zero quickly and $S(W, f_d)$ will be close to zero. 3) If W has a large volume, $\bar{\gamma}_W(r)$ will only drop off slowly with r, and therefore $S(W, f_d)$ will be close to one. Therefore, we know that the function $S(W, f_d)$ has to be between zero and one, and for a given patch and packet shape has to go to zero as $\nu \to 0$ to one as $\nu \to \infty$. 4) For a given packet distribution, $f_d(r)$, and shape, W, if both the packet and patch were scaled by the same factor (either stretched or shrank in space), $S(W, f_d)$ would take the same value, meaning that $S(W, f_d)$ depends only on the ratio of these scales. Given these facts, it is possible to build a simpler approximation of equation 6e for one and two dimensional systems.

One-dimensional approximation

In the case of a one-dimensional pattern (settlement on a line), $\bar{\gamma}_W(r)$ is simply equal to $max(2\nu - r, 0)$. We look at three different packet shapes to calculate $f_d(r)$. 1) The uniform distribution, as a representative short-tailed distribution, 2) the Gaussian (or normal) distribution, and 3) the Laplacian distribution, as a representative

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

heavier-tailed distribution. For all three of these distributions, $f_d(r)$ is a one-parameter function, as any change in the location of the distribution (its mean value) will not affect the distances between points drawn from that distribution. These distributions, their $f_d(r)$ functions, and the value for the integral $S(W, f_d)$ are given in Table .

As table shows, the resulting functions for $S(W, f_d)$ for different distributions are very complex. However, all three functions share two properties: the interaction scale ν and packet scale ω only enter through their ratio, so the same scale effect will result if both interaction scale and packet scales are multiplied by the same value. Second, all three functions are increasing sigmoidal functions of the ratio $\frac{\nu}{\omega}$. That is, if either the scale of interaction increases or the scale of clustering decreases, S will increase. Further, by regressing log(S/(1-S)) on log($\frac{\nu}{\omega}$), we were able to show that all three scale functions were closely fit by the function $\frac{\frac{\nu}{\omega}^{1.25}}{1+\frac{\nu}{\omega}^{1.25}}$. This is illustrated in Fig. A1, showing the scale functions for each distribution, along side the approximate function.

Two-dimensional approximation

This approach becomes more complicated in two dimensions, but the overall result is the same. In two dimensions, patches are no longer defined by just their size, but also have a shape. Here we focus on circular patches, but simulations of point processes show that the results are quantitatively very similar for square patches (results not shown). For circular patches, the isotropised set covariance function $\bar{\gamma}_W(r)$ is (Illian et al., 2008):

$$\bar{\gamma}_W(r) = 2\nu^2 acos(\frac{r}{2\nu}) - \frac{r}{2}\sqrt{4\nu^2 - r^2}$$

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

We look at two packet distributions: the circular uniform and the symmetric Gaussian. For the circular uniform distribution, all points in the packet are distributed uniformly in a circle of radius $\sqrt{2}\omega$ around the center point. For the symmetric Gaussian, points are distributed around the packet center so that both the x and y coordinates are uncorrelated, and each is distributed following a Gaussian distribution with a variance of $\frac{\omega^2}{2}$.

For these distributions, the $f_d(r)$ functions are (modified from (Illian et al., 2008)):

$$\text{Circular uniform } f_d(r) = \begin{cases} \frac{2r}{\pi\omega^2}(acos(\frac{r}{2\sqrt{2}\omega}) - \frac{r}{2\sqrt{2}\omega}\sqrt{1-\frac{r^2}{8\omega^2}}) & \text{if } r \leq 2\sqrt{2}\omega \\ 0 & \text{if } r \leq 2\sqrt{2}\omega \end{cases}$$

$$\text{Symmetrical Gaussian } f_d(r) = \frac{r}{\omega^2}e^{-\frac{r^2}{2\omega^2}}$$

For both these cases, the joint scale function, $S(W, f_d)$ is too complicated to derive a closed form integral. However, they can be solved via numerical integration. As before, we fit $\log(\frac{S}{1-S})$ to $log(\frac{\nu}{\omega})$ using linear regression, using only values of $\frac{\nu}{\omega}$ below 1, as the overall variance κ will be more sensitive to changes in this range of $\frac{\nu}{\omega}$. This gave us the two-dimensional approximation to $S(W, f_d) = \frac{\frac{1}{2}(\frac{\nu}{\omega})^2}{1+\frac{1}{2}(\frac{\nu}{\omega})^2}$. As seen in Fig. A2, this approximation works well for both distributions, across several orders of magnitude variation in $\frac{\nu}{\omega}$.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Table A1: Definitions of variables used for the derivation of κ.

Parameter	Definition
λ_c	The intensity of cluster centers on the landscape.
p_n	The probability of finding n points in a given packet.
μ	The mean number of individuals in each packet. $\mu = \sum_{n=1}^{\infty} p_n n$
σ^2	The variance of points between packets. $\sigma^2 = \sum_{n=1}^{\infty} p_n (n-\mu)^2$
λ	$= \lambda_c \mu$: The average density of individuals on the landscape.
ω	The square root of the mean squared distance from the packet center for a given type of packet. For one-dimensional packet distributions, this is the standard deviation of the distribution.
W	The shape of the patch.
ν	length scale of the patch, equal to the radius of a d - dimensional ball with the same volume as W.
d	the dimension of the system. For the simulations and moment equations in this paper, d=1.
b_d	the volume of a ball with unit radius and dimension d. Equals 2 in one dimension, π in two dimensions.
r	The Euclidean distance of a point from a given focal point.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Table A2: Definitions of functions used for the derivation of κ. From Illian et al. (2008).

Function	Definition
$Vol(\nu)$	The volume of any shape with a length-scale ν.
$g(r)$	The pair correlation function for a spatial point process. Defined as the first derivative with respect to r of the mean number of points found in a ball of radius r around a focal individual, divided by the mean number points expected to find in a random ball of radius r. Varies between zero and infinity. If the settlers are randomly distributed across the landscape, g(r) =1 for all r.
$f_d(r)$	The distribution function of distances between points in a given packet. This is a probability distribution function describing the probability of any two points in a packet being exactly r distance away from one another after settling. This function defines the shape of the packet.
$\bar{\gamma}_W(r)$	The isotropised set covariance of a patch of shape W. For a given patch shape, this function describes the area of the overlapping between the shape and the same shape shifted r units away, averaging over all possible directions it could be shifted in. For a given patch shape W this will be uniquely defined, and captures the effect of patch shape on the variance.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Table A3: One-dimensional density functions used to model different packet shapes, and their derived inter-point distance and scale functions. Inter-point probability distances $f_d(r)$ were calculated by taking the product of the density function at a point x and the same density function at a point shifted away from x by a distance r, and integrating over the whole domain of x. Scale functions were calculated using the integral formula in the text. In all cases, distributions were parameterized so ω equaled the standard deviation of that distribution.

Distribution	Probability density ($p_X(x)$)	Inter-point distance density ($f_d(r)$)	Scale function ($S(W, f_d)$)
Uniform	$\frac{1}{2\sqrt{3}\omega}$ if $\|x\| < \sqrt{(}3)\omega$, 0 otherwise	$\frac{1}{\sqrt{3}\omega} - \frac{r}{6\omega^2}$ if $r < 2\sqrt{(}3)\omega$, 0 otherwise	$\frac{3\sqrt{3}\nu\omega^{-1} - \nu^2\omega^{-2}}{9}$ if $\nu\omega^{-1} \leq \sqrt{(}3)$, $1 - \frac{1}{\sqrt{\nu\omega^{-1}}}$ otherwise
Gaussian	$\frac{1}{\omega\sqrt{2\pi}} e^{\frac{-x^2}{2\omega^2}}$	$\frac{1}{\omega\sqrt{\pi}} e^{\frac{-x^2}{4\omega^2}}$	$erf(\nu\omega^{-1}) + \frac{e^{-\nu^2\omega^{-2}} - 1}{\nu\omega^{-1}\sqrt{\pi}}$
Laplacian	$\frac{1}{\sqrt{2}\omega} e^{\frac{-\|x\|\sqrt{2}}{\omega}}$	$\frac{\frac{\omega}{\sqrt{2}} + r}{\omega^2} e^{-\sqrt{2}r\omega^{-1}}$	$\frac{3\nu}{4\sqrt{2}\omega}(e^{-2\sqrt{2}\nu\omega^{-1}} - 1)$ $+ \frac{1}{2} e^{-2\sqrt{2}\nu\omega^{-1}} + 1$

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

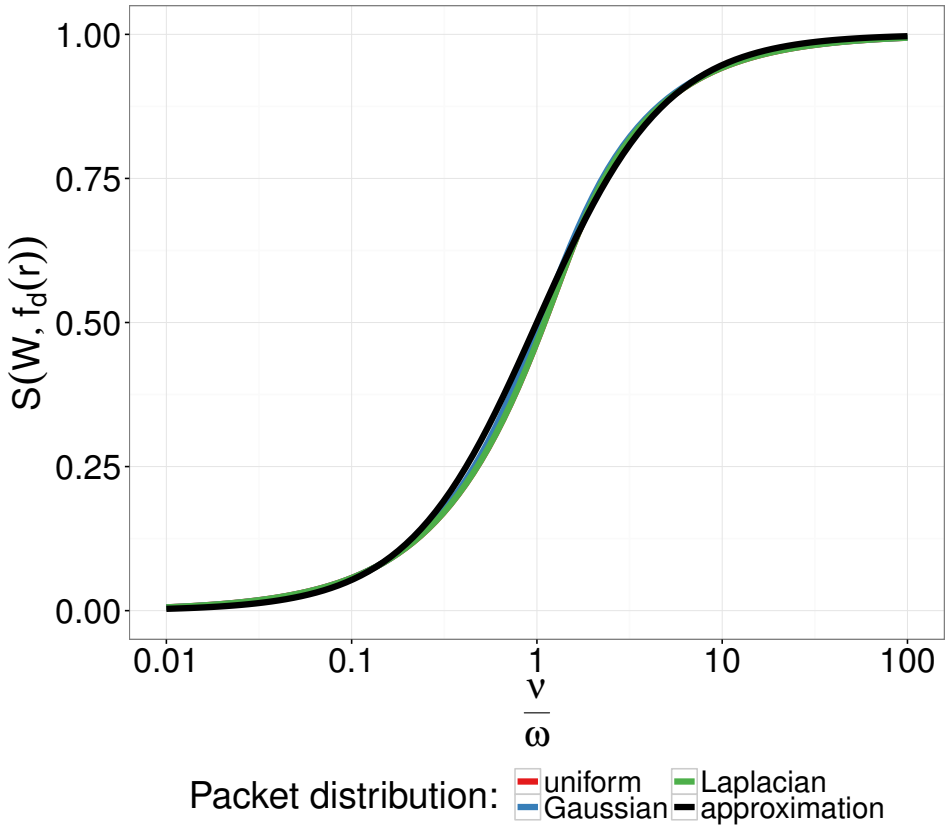

Figure A1: Scale functions $S(W, f_d)$ for three one-dimensional packet distributions, and the approximate scale function, plotted against $\frac{\nu}{\omega}$.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

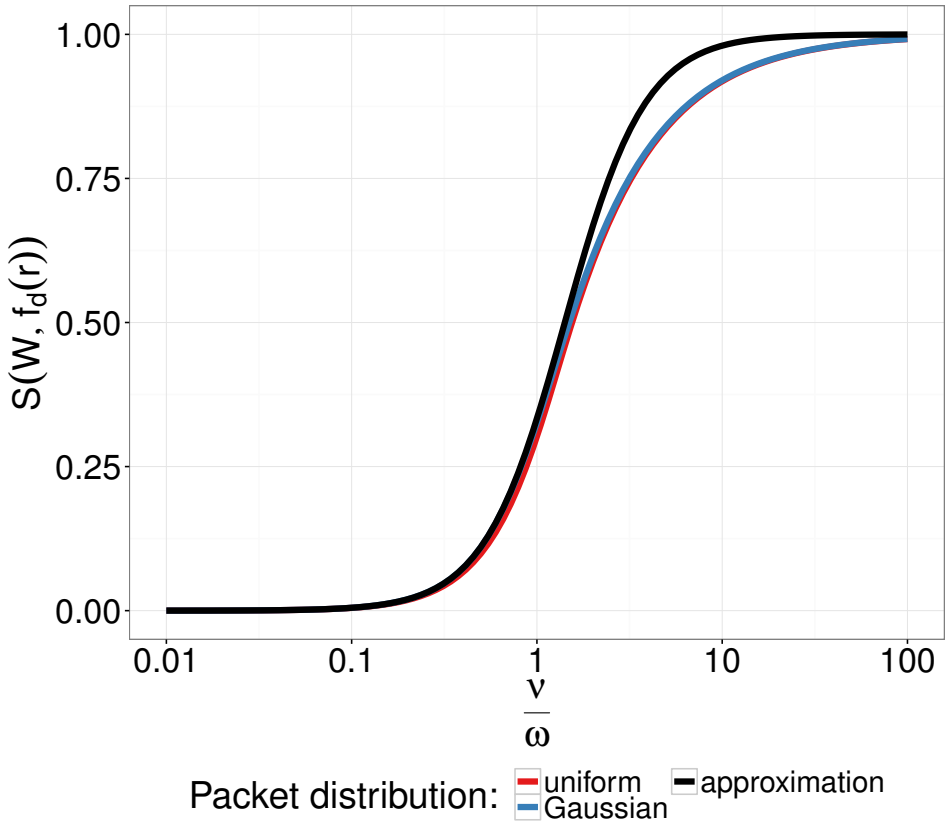

Figure A2: Scale functions $S(W, f_d)$ for two two-dimensional packet distributions, and the approximate scale function, plotted against $\frac{\nu}{\omega}$.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Appendix B: Deriving the moment closure for the aggregated dispersal meta-community model

Our moment closure model is an extended form of the meta-population by Keeling (2000*b*), incorporating patch size and competition. Whereas Keeling modeled the dynamics of numbers of individuals in patches, we model densities to enable us to include a measure of patch scale. Further, we incorporate a second competing species. We denote our two species x and y. We use x_i (y_i) to denote the number of individuals of species x (y) in patch i, $\tilde{x_i}$ ($\tilde{y_i}$) to denote the mean density of species x (y) in patch i (that is, x_i (y_i) scaled by patch volume), and $\tilde{X}$ ($\tilde{Y}$) to denote the mean density per unit area of x (y) across all patches.

We assume species interact in a D-dimensional space. We also assume that each species views this space as a set of patches, so than an individual of species x or y interacts only with the other individuals of its own and its competitor species that are in its patch. However, we do not assume that the patch structure is necessarily the same for both species. Instead, each species interacts in patches of radius ν_x or ν_y.

For the derivation of moment equations, we assume that the first species, y always interacts over a larger spatial scale, and that the scale overlaps η patches of the smaller-ranged species (x), so that $\nu_y = \nu_x \cdot \eta$. To take into account these different scales,any variable indexed ij refers to the jth smaller patch within the ith larger-scale patch. Any variable simply indexed i refers to the sum or mean of all the patches $\{ij\}$ in i. This means that patches ij will have volume $Vol_D(\nu_x)$ and patch i will have volume $Vol_D(\nu_y) = Vol_D(\nu_x \cdot \eta)$. This is shown in Fig. B1. These two assumptions are

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

only for convenience of the derivation; the resulting equations are identical if ν_x is larger or smaller than ν_y, and our simulation results show that the assumption of strict patch nesting does not affect population dynamics.

We use x_{ij} (y_{ij} to denote the number of individuals of species x (y) in patch ij. We use x_i (y_i) to denote the total number of individuals in all the smaller patches within i: $x_i \equiv \sum_j x_{ij}$. Equivalently, $\tilde{x}_{ij}$ ($\tilde{y}_{ij}$) and $\tilde{x}_i$ ($\tilde{y}_i$) respectively denote the density of x in the smaller patch j within patch i ($\tilde{x}_{ij} \equiv x_{ij}/Vol(\nu_x)$), and the average density in patch i ($\tilde{x}_i \equiv \sum_{j=1}^{\eta} s_{ij}/Vol(\nu_y)$).

As we assumed that species x interacts at a smaller scale than y, we assume individuals of species y are well-mixed and constantly moving between smaller patches within the larger patch, so that $\tilde{y}_{ij} = \tilde{y}_i$. We also define $\epsilon_{ij} \equiv \tilde{x}_i - \tilde{x}_{ij} \equiv \frac{\sum_{j=1}^{\eta} x_{ij}}{Vol(\nu_y)} - \frac{x_{ij}}{Vol(\nu_x)}$ as the deviation of the density of species x in patch ij from the overall density of the larger patch i.

We assume that both species live in patches, packets arrive that can overlap multiple patches, and packets are defined by the following equations (defined for species x, but the same equations hold for species y):

- Packets arrives at each site with a rate $a(\tilde{X}, \nu_x)$, and the probability that a number of individuals s arrive at a site given that a packet arrived there is $p_x(s|\tilde{X}, \nu_x)$.

- The average number of individuals arriving at a site per unit time has a mean rate of $r_x Vol(\nu_x)\tilde{X}$, and a variance $r_x Vol(\nu_x)\tilde{X}\kappa(\tilde{X}, \nu_x, \sigma_x)$, where σ_x is the root-mean square distance each of each recruit from its packet center for species x and $\kappa(\tilde{X}, \nu_x, \sigma_x)$ is a function varying from 1 to ∞ (see Appendix A for the definition

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

of κ).

In addition to the birth rates defined above, per-unit-density death rates are defined for x in patch ij as: $mVol(\nu_x)\tilde{x}_{ij} + d_{x,x}Vol(\nu_x)\tilde{x}_{ij}^2 + d_{x,y}Vol(\nu_x)\tilde{x}_{ij}\tilde{y}_i$, and for y as $mVol(\nu_y)\tilde{y}_i + d_{y,y}Vol(\nu_y)\tilde{y}_i^2 + d_{y,x}Vol(\nu_y)\tilde{x}_i\tilde{y}_i$.

Note that all demographic parameters are multiplied by the patch volume, as they are measured in units of individuals per unit volume (or individuals^2 per unit volume, in the case of the interaction parameters).

The moment closure approach relies on being able to describe a master equation for the dynamics. This is a system of equations describing the transition rates between different possible states of the system. Here, the state for patch ij is given by $\{\tilde{x}_{ij}(t), \tilde{y}_i(t)\}$. A death of x in patch ij would lead the system to transition to $\{\tilde{x}_{ij}(t) - \frac{1}{Vol(\nu_x)}, \tilde{y}_i(t)\}$, and a death of species y would lead to $\{\tilde{x}_{ij}(t), \tilde{y}_i(t) - \frac{1}{Vol(\nu_y)}\}$. If s new individuals of species x arrived in the patch ij, the system would transition to $\{\tilde{x}_{ij}(t) + \frac{s}{Vol(\nu_x)}, \tilde{y}_i(t)\}$, and if s individuals of y arriving would lead to $\{\tilde{x}_{ij}(t), \tilde{y}_i(t) + \frac{n}{Vol(\nu_y)}\}$. Given the birth and death rules described above, the master equation for this system will be:

$$
\begin{aligned}
\frac{dD_{\tilde{x},\tilde{y}}}{dt} = \; & a_x(\tilde{X},\nu_x)\sum p_x(s_x|\tilde{X},\nu_x)(\tilde{x} - \frac{s_x}{Vol(\nu_x)})D_{\tilde{x}-\frac{s_x}{Vol(\nu_x)},\tilde{y}} \\
& +Vol(\nu_x)\cdot(m + d_{x,x}\tilde{x} + d_{x,y}\tilde{y})(\tilde{x} + \frac{1}{Vol(\nu_x)})D_{\tilde{x}+\frac{1}{Vol(\nu_x)},\tilde{y}} \\
& -Vol(\nu_x)\cdot(r(\tilde{X}) + m + d_{x,x}\tilde{x}_i + d_{x,y}\tilde{y}_i)\tilde{x}D_{\tilde{x},\tilde{y}} \\
& +a_y(\tilde{Y},\nu_y)\sum p_y(s_y|\tilde{Y},\nu_y)(\tilde{y} - \frac{s_y}{Vol(\nu_y)})D_{\tilde{x},\tilde{y}-\frac{s_y}{Vol(\nu_y)}} \\
& +Vol(\nu_y)\cdot(m + d_{y,y}\tilde{y} + d_{y,x}\tilde{x})(\tilde{y} + \frac{1}{Vol(\nu_y)})D_{\tilde{x},\tilde{y}+\frac{1}{Vol(\nu_y)}} \\
& -Vol(\nu_y)\cdot(r(\tilde{Y}) + m + d_{y,y}\tilde{y} + d_{y,x}\tilde{x})\tilde{y}D_{\tilde{x},\tilde{y}}
\end{aligned}
$$

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Here, $D_{\tilde{x},\tilde{y}}$ represents the fraction of patches in the state $\{\tilde{x}_{ij}, \tilde{y}_i\}$ at a given time. Positives term in the equation represent the stochastic rate at which different possible states transition to $\{\tilde{x}_{t,ij}, \tilde{y}_{t,i}\}$ due to settlement (from states at lower densities than $\{\tilde{x}_{t,ij}, \tilde{y}_{t,i}\}$) or deaths (from states with higher densities). Negative terms represent the stochastic rate at which sites in this state transition to other states, either through settlement of new recruits or death.

Since this is an infinitely large set of coupled equations (one for every possible density the system can be in), we need to simplify it to gain any sort of theoretical understanding. To do this, we transform it into a set of equations describing the how the multiplicative moments change over time (Keeling, 2000*b*). This approach models the population change in terms of the statistical moments of the population. First, the average densities of the sites in the system: $\langle \tilde{x_{ij}} \rangle = \tilde{X}$, $\langle \tilde{y}_i \rangle = \tilde{Y}$. Second the second order multiplicative moments: $\hat{V}_x \tilde{X}^2 = \langle \tilde{x_{ij}}^2 \rangle$ (variation in density in x between small-scale sites), and $\hat{V}_y \tilde{Y}^2 = \langle \tilde{y_i}^2 \rangle$ (variation in density in y between broad-scale sites), and $C\tilde{X}\tilde{Y} = \langle \tilde{x_{ij}} \tilde{y}_i \rangle$ (multiplicative covariance in density in x and y at the broader spatial scale).

We finally assume that any higher order moments do not have separate dynamics, but are instead determined by products of all pair-wise combinations of lower-order moments: $\langle \tilde{x}\tilde{x}\tilde{y} \rangle = \frac{\langle \tilde{x}\tilde{y} \rangle \langle \tilde{x}\tilde{y} \rangle \langle \tilde{x}\tilde{x} \rangle}{\tilde{X}^2 \tilde{Y}} = C^2 V_x \tilde{X}^2 \tilde{Y}$, $\langle \tilde{x}\tilde{y}\tilde{y} \rangle = C^2 V_y \tilde{Y}^2 \tilde{X}$, $\langle \tilde{x}^3 \rangle = V_x^3 \tilde{X}^3$ and $\langle \tilde{y}^3 \rangle = V_y^3 \tilde{Y}^3$ (Keeling, 2000*b*). Capital letters denote population moments of their respective populations.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Dynamics for mean densities

The mean dynamics for species x and y can be described as:

$$\begin{aligned}\frac{d\tilde{X}}{dt} =&\langle a_x(\tilde{X}, Vol(\nu))\sum_{n=1}^{\infty} p_x(n|\tilde{X}, Vol(\nu_x))(\frac{n}{Vol(\nu_x)})+\\ &(mVol(\nu_x)\tilde{x}_{ij} + d_{x,x}Vol(\nu_x)\tilde{x}_{ij}^2 + d_{x,y}Vol(\nu_x)\tilde{x}_{ij}\tilde{y}_i)(\frac{-1}{Vol(\nu_x)})\rangle\\ \frac{d\tilde{Y}}{dt} =&\langle a_y(\tilde{Y}, Vol(\nu_y))\sum_{n=1}^{\infty} p_x(n|\tilde{X}, Vol(\nu_y))(\frac{n}{Vol(\nu_y)})+\\ &(mVol(\nu_y)\tilde{y}_i + d_{y,y}Vol(\nu_y)\tilde{y}_i^2 + d_{y,x}Vol(\nu_y)\tilde{x}_i\tilde{y}_i)(\frac{-1}{Vol(\nu_y)})\rangle\end{aligned}$$

Taking the averages over sites:

$$\begin{aligned}\frac{d\tilde{X}}{dt} =&\frac{r_xVol(\nu_x)\tilde{X}}{Vol(\nu_x)} - \frac{m_xVol(\nu_x)\tilde{X}}{Vol(\nu_x)} - \langle\frac{d_{x,x}Vol(\nu_x)}{Vol(\nu_x)}\tilde{x}_{ij}^2 + \frac{d_{x,y}Vol(\nu_x)}{Vol(\nu_x)}(\tilde{x}_i - \epsilon_{ij})\tilde{y}_i\rangle\\ =&r_x\tilde{X} - m_x\tilde{X} - d_{x,x}\hat{V}_x\tilde{X}^2 - d_{x,y}\hat{C}\tilde{X}\tilde{Y} + \langle d_{x,y}\epsilon_{ij}\tilde{y}_i\rangle\\ =&r_x\tilde{X} - m_x\tilde{X} - d_{x,x}\hat{V}_x\tilde{X}^2 - d_{x,y}\hat{C}\tilde{X}\tilde{Y}\end{aligned}$$

$$\begin{aligned}\frac{d\tilde{Y}}{dt} =&\frac{r_yVol(\nu_y)\tilde{Y}}{Vol(\nu_y)} - \frac{m_yVol(\nu_y)\tilde{Y}}{Vol(\nu_y)} - \frac{d_{y,y}Vol(\nu_y)}{Vol(\nu_y)}\hat{V}_y\tilde{Y}^2 - \frac{d_{y,x}Vol(\nu_y)}{Vol(\nu_y)}\hat{C}\tilde{X}\tilde{Y}\\ =&r_y\tilde{Y} - m_y\tilde{Y} - d_{y,y}\hat{V}_y\tilde{Y}^2 - d_{y,x}\hat{C}\tilde{X}\tilde{Y}\end{aligned}$$

The above derivation depends on the fact that $\sum_j^{\eta}\epsilon_{ij} \equiv 0$ for all i, so expectations involving $\epsilon_{ij} \cdot \tilde{y}_i$ drop out.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

Dynamics for variance functions

That gives three extra equations we have derive the dynamics of: $\frac{d\hat{V}_x}{dt}, \frac{d\hat{V}_y}{dt}$, and $\frac{d\hat{C}}{dt}$. We can use the following relations to get these (Keeling, 2000*b*):

$$\frac{d\langle\tilde{x}_{ij}^2\rangle}{dt} = \frac{d\hat{V}_x\tilde{X}^2}{dt} = \tilde{X}^2\frac{d\hat{V}_x}{dt} + 2\tilde{X}\hat{V}_x\frac{d\tilde{X}}{dt}$$

$$\frac{d\langle\tilde{y}_i^2\rangle}{dt} = \frac{d\hat{V}_y\tilde{Y}^2}{dt} = \tilde{Y}^2\frac{d\hat{V}_y}{dt} + 2\tilde{Y}\hat{V}_y\frac{d\tilde{Y}}{dt}$$

$$\frac{d\langle\tilde{x}_i\tilde{y}_i\rangle}{dt} = \frac{d\hat{C}\tilde{X}\tilde{Y}}{dt} = \tilde{X}\tilde{Y}\frac{d\hat{C}}{dt} + \tilde{Y}\hat{C}\frac{d\tilde{X}}{dt} + \tilde{X}\hat{C}\frac{d\tilde{Y}}{dt}$$

The equation for $\hat{V}_x$:

$$\tilde{X}^2\frac{d\hat{V}_x}{dt} = \frac{d\langle\tilde{x}_i^2\rangle}{dt} - 2\tilde{X}\hat{V}_x\frac{d\tilde{X}}{dt}$$

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

$$\begin{aligned}
\tilde{X}^2\frac{d\hat{V}_x}{dt} =&\langle a_x(\tilde{X},Vol(\nu_x))\sum_{n=1}^{\infty}p_x(n|\tilde{X},Vol(\nu_x))(\frac{2\tilde{x}_{ij}n}{Vol(\nu_x)}+\frac{n^2}{Vol(\nu_x)^2})\\
&+(m_xVol(\nu_x)\tilde{x}_{ij}+d_{x,x}Vol(\nu_x)\tilde{x}_{ij}^2+d_{x,y}Vol(\nu_x)\tilde{x}_{ij}\tilde{y}_{ij})(-\frac{2\tilde{x}_{ij}}{Vol(\nu_x)}+\frac{1}{Vol(\nu_x)^2})\rangle\\
&-2\tilde{X}\hat{V}_x\frac{d\tilde{X}}{dt}\\
=&\langle\frac{2r_xVol(\nu_x)\tilde{X}\tilde{x}_{ij}}{Vol(\nu_x)}+\frac{r_xVol(\nu_x)\tilde{X}\kappa(\tilde{X},\nu_x,\sigma_x)}{Vol(\nu_x)^2}+\frac{{r_x}^2Vol(\nu_x)^2\tilde{X}^2}{Vol(\nu_x)^2}-\\
&\frac{2m_xVol(\nu_x)\tilde{x}_{ij}^2}{Vol(\nu_x)}+\frac{m_xVol(\nu_x)\tilde{x}_{ij}}{Vol(\nu_x)^2}-\frac{2d_{x,x}Vol(\nu_x)\tilde{x}_{ij}^3}{Vol(\nu_x)}+\\
&\frac{d_{x,x}Vol(\nu_x)\tilde{x}_{ij}^2}{Vol(\nu_x)^2}-\frac{2d_{x,y}Vol(\nu_x)\tilde{x}_{ij}^2\tilde{y}_{ij}}{Vol(\nu_x)}+\frac{d_{x,y}Vol(\nu_x)\tilde{x}_{ij}\tilde{y}_{ij}}{Vol(\nu_x)^2}\rangle-2\tilde{X}\hat{V}_x\frac{d\tilde{X}}{dt}\\
=&2r_x\tilde{X}^2+\frac{r_x\tilde{X}\kappa(\tilde{X},\nu_x,\sigma_x)}{Vol(\nu_x)}+r_x^2\tilde{X}^2-2m_x\hat{V}_x\tilde{X}^2+\frac{m_x\tilde{X}}{Vol(\nu_x)}\\
&-2d_{x,x}\tilde{X}^3\hat{V}_x^3+\frac{d_{x,x}\tilde{X}^2\hat{V}_x}{Vol(\nu_x)}-2d_{x,y}\tilde{X}^2\tilde{Y}\hat{V}_x\hat{C}^2+\frac{d_{x,y}\tilde{X}\tilde{Y}\hat{C}}{Vol(\nu_x)}\\
&-2\tilde{X}\hat{V}_x(r_x\tilde{X}-m_x\tilde{X}-d_{x,x}(\hat{V}_x\tilde{X}^2)-d_{x,y}\hat{C}\tilde{X}\tilde{Y})
\end{aligned}$$

$$\begin{aligned}
=&(\frac{r_x\kappa(\tilde{X},\nu_x,\sigma_x)+m_x}{Vol(\nu_x)})\tilde{X}+(2r_x+r_x^2)\tilde{X}^2+(\frac{d_{x,x}}{Vol(\nu_x)}-2r_x)\tilde{X}^2\hat{V}_x\\
&-2d_{x,x}(\hat{V}_x-1)\hat{V}_x^2\tilde{X}^3+\frac{d_{x,y}\tilde{X}\tilde{Y}\hat{C}}{Vol(\nu_x)}-2d_{x,y}(\hat{C}-1)\tilde{X}^2\tilde{Y}\hat{V}_x\hat{C}
\end{aligned}$$

778 Dividing out $\tilde{X}^2$, we get:

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

$$\frac{d\hat{V}_x}{dt} = 2r_x + r_x^2 + \frac{r_x\kappa(\tilde{X},\nu_x,\sigma_x) + m_x}{Vol(\nu_x)\tilde{X}} + (\frac{d_{x,x}}{Vol(\nu_x)} - 2r_x)\hat{V}_x$$
$$- 2d_{x,x}(\hat{V}_x - 1)\hat{V}_x^2\tilde{X} + \frac{d_{x,y}\tilde{Y}\hat{C}}{Vol(\nu_x)\tilde{X}} - 2d_{x,y}(\hat{C} - 1)\tilde{Y}\hat{V}_x\hat{C}$$

Equivalently,

$$\frac{d\hat{V}_y}{dt} = 2r_y + r_y^2 + \frac{r_y\kappa(\tilde{Y},\nu_y,\sigma_y) + m_y}{Vol(\nu_y)\tilde{Y}} + (\frac{d_{y,y}}{Vol(\nu_y)} - 2r_y)\hat{V}_y$$
$$- 2d_{y,y}(\hat{V}_y - 1)\hat{V}_y^2\tilde{Y} + \frac{d_{y,x}\tilde{X}\hat{C}}{Vol(\nu_y)\tilde{Y}} - 2d_{y,x}(\hat{C} - 1)\tilde{X}\hat{V}_y\hat{C}$$

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

The equation for $\hat{C}$:

$$\begin{aligned}
\tilde{X}\tilde{Y}\frac{d\hat{C}}{dt} =& \frac{d\langle\tilde{x}_i\tilde{y}_i\rangle}{dt} - \tilde{Y}\hat{C}\frac{d\tilde{X}}{dt} - \tilde{X}\hat{C}\frac{d\tilde{Y}}{dt} \\
=& \langle a_x(\tilde{X}, Vol(\nu_y))\sum_{n=1}^{\infty} p_x(n|\tilde{X}, Vol(\nu_y))\frac{n\tilde{y}_i}{Vol(\nu_y)} \\
&+ a_y(\tilde{Y}, Vol(\nu_y))\sum_{n=1}^{\infty} p_y(n|\tilde{Y}, Vol(\nu_y))\frac{n\tilde{x}_i}{Vol(\nu_y)} \\
&+ (m_x Vol(\nu_y)\tilde{x}_i + d_{x,x}Vol(\nu_y)\tilde{x}_i^2 + d_{x,y}Vol(\nu_y)\tilde{x}_i\tilde{y}_i)(\frac{-\tilde{y}_i}{Vol(\nu_y)}) \\
&+ (m_y Vol(\nu_y)\tilde{y}_i + d_{y,y}Vol(\nu_y)\tilde{y}_i^2 + d_{y,x}Vol(\nu_y)\tilde{x}_i\tilde{y}_i)(\frac{-\tilde{x}_i}{Vol(\nu_y)})\rangle \\
&- \tilde{Y}\hat{C}\frac{d\tilde{X}}{dt} - \tilde{X}\hat{C}\frac{d\tilde{Y}}{dt} \\
=& r_x\tilde{X}\tilde{Y} + r_y\tilde{X}\tilde{Y} - m_x\tilde{X}\tilde{Y}\hat{C} - d_{x,x}\tilde{X}^2\tilde{Y}\hat{V}_x\hat{C}^2 - d_{x,y}\tilde{X}\tilde{Y}^2\hat{V}_y\hat{C}^2 \\
&- m_y\tilde{X}\tilde{Y}\hat{C} - d_{y,y}\tilde{X}\tilde{Y}^2\hat{V}_y\hat{C}^2 - d_{y,x}\tilde{X}^2\tilde{Y}\hat{V}_x\hat{C}^2 \\
&- r_x\tilde{X}\tilde{Y}\hat{C} + m_x\tilde{X}\tilde{Y}\hat{C} + d_{x,x}\tilde{X}^2\tilde{Y}\hat{C} + d_{x,y}\tilde{X}\tilde{Y}^2\hat{C}^2 \\
&- r_y\tilde{X}\tilde{Y}\hat{C} + m_y\tilde{X}\tilde{Y}\hat{C} + d_{y,y}\tilde{X}\tilde{Y}^2\hat{C} + d_{y,x}\tilde{X}^2\tilde{Y}\hat{C}^2 \\
=& (r_x + r_y)(1 - \hat{C})\tilde{X}\tilde{Y} - (d_{x,x} + d_{y,x})(\hat{V}_x - 1)\tilde{X}^2\tilde{Y}\hat{C}^2 - (d_{y,y} + d_{x,y})(\hat{V}_y - 1)\tilde{X}\tilde{Y}^2\hat{C}^2
\end{aligned}$$

Dividing $\tilde{X}\tilde{Y}$ out, we get:

$$\frac{d\hat{C}}{dt} = (r_x + r_y)(1 - \hat{C}) - (d_{x,x} + d_{y,x})(\hat{V}_x - 1)\tilde{X}\hat{C}^2 - (d_{y,y} + d_{x,y})(\hat{V}_y - 1)\tilde{Y}\hat{C}^2$$

The whole system:

Combined together, this gives us the dynamics for all five state variables:

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

$$
\begin{aligned}
\frac{d\tilde{X}}{dt} &= r_x\tilde{X} - m_x\tilde{X} - d_{x,x}\hat{V}_x\tilde{X}^2 - d_{x,y}\hat{C}\tilde{X}\tilde{Y} \\
\frac{d\tilde{Y}}{dt} &= r_y\tilde{Y} - m_y\tilde{Y} - d_{y,y}\hat{V}_y\tilde{Y}^2 - d_{y,x}\hat{C}\tilde{X}\tilde{Y} \\
\frac{d\hat{V}_x}{dt} &= 2r_x + r_x^2 + \frac{r_x\kappa(\tilde{X},\nu_x,\sigma_x) + m_x}{Vol(\nu_x)\tilde{X}} + (\frac{d_{x,x}}{Vol(\nu_x)} - 2r_x)\hat{V}_x \\
&\quad -2d_{x,x}(\hat{V}_x - 1)\hat{V}_x^2\tilde{X} + \frac{d_{x,y}\tilde{Y}\hat{C}}{Vol(\nu_x)\tilde{X}} + 2d_{x,y}(1-\hat{C})\tilde{Y}\hat{V}_x\hat{C} \\
\frac{d\hat{V}_y}{dt} &= 2r_y + r_y^2 + \frac{r_y\kappa(\tilde{Y},\nu_y,\sigma_y) + m_y}{Vol(\nu_y)\tilde{Y}} + (\frac{d_{y,y}}{Vol(\nu_y)} - 2r_y)\hat{V}_y \\
&\quad -2d_{y,y}(\hat{V}_y - 1)\hat{V}_y^2\tilde{Y} + \frac{d_{y,x}\tilde{X}\hat{C}}{Vol(\nu_y)\tilde{Y}} + 2d_{y,x}(1-\hat{C})\tilde{X}\hat{V}_y\hat{C} \\
\frac{d\hat{C}}{dt} &= (r_x + r_y)(1-\hat{C}) - (d_{x,x} + d_{y,x})(\hat{V}_x - 1)\tilde{X}\hat{C}^2 - (d_{y,y} + d_{x,y})(\hat{V}_y - 1)\tilde{Y}\hat{C}^2
\end{aligned}
$$

If we substitute in $Vol(\nu_x) = 2\nu_x$ ($Vol(\nu_y) = 2\nu_y$), the one-dimensional case, we get the results shown in equation (3).

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016

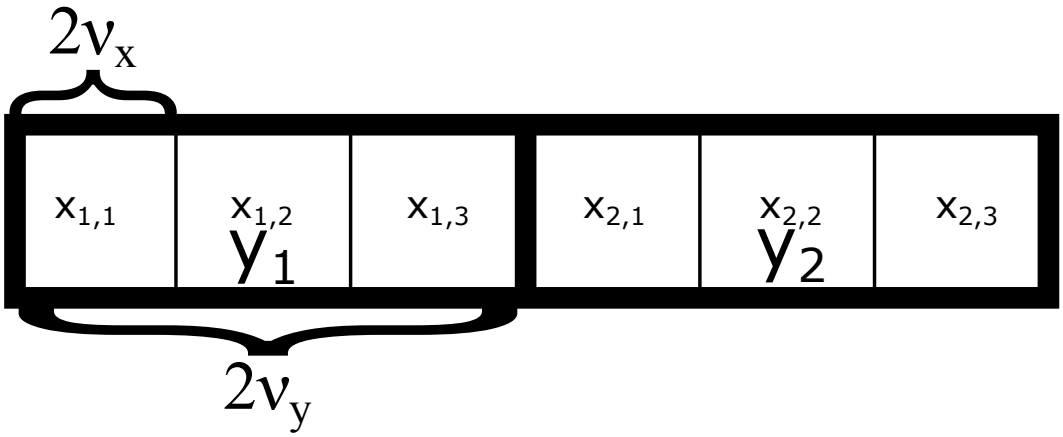

Figure B1: Multi-scale patch configuration in one dimension.

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1734v1 | CC-BY 4.0 Open Access | rec: 9 Feb 2016, publ: 9 Feb 2016