Published paper

Cornell, S.J. and Ovaskainen, O. (2008) *Exact asymptotic analysis for metapopulation dynamics on correlated dynamic landscapes,* Theoretical Population Biology**,** Volume 74 (3), 209 -225.

Exact asymptotic analysis for metapopulation dynamics on correlated dynamic landscapes

Stephen J. Cornell

Institute of Integrative and Comparative Biology, University of Leeds, Leeds LS2 9JT, UK

Otso Ovaskainen

Department of Biological and Environmental Sciences, University of Helsinki, P.O. Box 65, FI-00014, Helsinki, Finland

---

Abstract

We compute the mean patch occupancy for a stochastic, spatially explicit patch-occupancy metapopulation model on a dynamic, correlated landscape, using a mathematically exact perturbation expansion about a mean-field limit that applies when dispersal range is large. Stochasticity in the metapopulation and landscape dynamics gives negative contributions to patch occupancy, the former being more important at high occupancy and the latter at low occupancy. Positive landscape correlations always benefit the metapopulation, but are only significant when the correlation length is comparable to, or smaller than, the dispersal range. Our analytical results allow us to consider the importance of spatial kernels in all generality. We find that the shape of the landscape correlation function is typically unimportant, and that the variance is overwhelmingly the most important property of the colonisation kernel. However, short-range singularities in either the colonisation kernel or landscape correlations can give rise to qualitatively different behaviour.

Key words: metapopulation, dynamic landscape, correlated landscape, perturbation theory, asymptotic analysis

---

Email addresses: `s.j.cornell@leeds.ac.uk` (Stephen J. Cornell), `otso.ovaskainen@helsinki.fi` (Otso Ovaskainen).

Preprint submitted to Elsevier Science *July 23, 2008*

1 Introduction

A fundamental prediction of metapopulation theory is that the persistence of a species inhabiting a fragmented landscape is facilitated by habitat which is both abundant and highly connected (Hanski and Ovaskainen, 2000). The key variable is the mean amount of habitat available for an individual through migration, which Hanski (1999) decomposed into four components: amount of habitat available now in the present patch, amount of connected habitat available now at the other patches, amount of future habitat available in the present patch, and amount of linked habitat available in the future at the other patches.

The availability of nearby patches is particularly important if habitat is sporadically destroyed by biotic and abiotic disturbances, since the species needs to find and colonise new patches in order to persist. Habitat turnover can be caused e.g. by natural disturbance dynamics, such as gap dynamics or fires in case of boreal forests (Wimberly, 2002; Pennanen et al., 2004). In many cases, habitat is intrinsically ephemeral, and needs to be constantly renewed for the species to persist. Examples include: butterflies that require vegetation at a particular successional stage (Thomas and Harrison, 1992; Wahlberg et al., 2002); beetles or fungi that devour dead wood (Schroeder et al., 2007); and pathogens, whose habitat is living hosts. The population dynamics of such species are strongly linked to the processes that generate and destroy their habitat. In this study, we analyse how the spatiotemporal properties of these disturbances influence metapopulation occupancy by considering a model for metapopulation dynamics in a dynamic and spatially correlated landscape.

Dynamic landscape metapopulation simulation models are increasingly used as practical tools to inform conservation efforts (Akcakaya et al., 2004; Verheyen et al., 2004; Wintle et al., 2005). In addition, there have been several theoretical studies (Fahrig, 1992; Keymer et al., 2000; Gu et al., 2002; Johst et al., 2002; Johst and Drechsler, 2003; Biedermann, 2004; Wimberly, 2006; North and Ovaskainen, 2007; Vuilleumier et al., 2007) investigating how landscape dynamics affects the properties of the metapopulation. As a first approximation, all other things being equal, a species is expected to do worse in a dynamic than in a static landscape. This is mainly due to the fact that, in a dynamic landscape, the species continuously needs to disperse to where the new habitat is, which leaves a larger fraction of habitats unoccupied at any given time (North and Ovaskainen, 2007). The turnover rate of habitat patches should be assessed relative to the rate of the species' colonization-extinction dynamics. In case of relatively quick patch turnover, local extinctions may be solely due to disappearance of suitable habitat, and the system can be described as a patch-tracking metapopulation (Wahlberg et al., 2002; Snäll et al., 2003; Lobel et al., 2006). Under some circumstances a species may actually survive better in a dynamic than in a static landscape. For example, if the rate of colonization is very rapid, but spatially very restricted, a species may not be able to colonize isolated habitat patches in a static landscape, whereas a dynamic landscape provides temporary bridges that facilitate colonization (Wimberly, 2006). Spatially correlated disturbances facilitate persistence by creating aggregated clusters of habitats, but they may also increase the risk of extinction by wiping the current population out simultaneously from large areas (Moloney and Levin, 1996). While longer-ranged dispersal should improve a species' persistence in the landscape, short-range dispersal can be advantageous if the habitat appears in a very clustered manner (Johst et al., 2002).

Most studies of metapopulations on dynamic landscapes rely on computer simulation, and therefore it is difficult to generalise their conclusions beyond the parameter range, dispersal kernels, etc. chosen for the study. A few studies have been based on analytical approximations, but these have usually been limited by omitting explicit spatial structure (Hastings, 2003; Ross, 2006a,b) or stochasticity (DeWoody et al., 2005), or because the approximation itself was based on an ad-hoc moment closure assumption (Ellner and Fussmann, 2003) and hence has an unknown domain of validity. Recently, however, we developed a perturbation technique for spatially explicit, stochastic metapopulation dynamics, where we exploited the fact that the 'mean-field' limit of infinite-range colonisation is exactly solvable, and treated the inverse of the colonisation length L as a small parameter (Ovaskainen and Cornell, 2006a). The perturbation coefficients are evaluated as explicit expressions in terms of the model parameters, so their properties can be analysed in a general fashion for a wide range of scenarios. If we truncate the expansion at finite order then we get an approximation to metapopulation dynamics which is comparable to a moment closure approach (Ovaskainen and Cornell, 2006b), but the perturbation expansion has two principal advantages. First, by evaluating higher order terms it is possible (in principle) to estimate the error incurred by truncating the expansion at lower order. Second, the truncated series calculates the corrections to mean-field in a way that is exact in the limit $L \to \infty$, and therefore any qualitative conclusions regarding the relative contribution of different factors in the dynamics are exact for a well-defined parameter range. It turns out that the perturbation series truncated to first order is a surprisingly good approximation to the true stochastic dynamics even when the colonisation length is comparable to the distance between neighbouring patches (Ovaskainen and Cornell, 2006a).

In the present manuscript, we extend the methods of Ovaskainen and Cornell (2006a)—which considered a static, correlated landscape—to the dynamic landscape case. We allow the patches to be created and destroyed by a stochastic process which is spatially correlated, so that the landscape is itself correlated in both time and space. We derive expressions for the mean patch occupancy in terms of the model parameters and shapes of the spatial kernels that determine the colonisation-extinction dynamics of the metapopulation and the creation-destruction dynamics of the habitat patches. This enables us to identify and compare the contributions due to two forms of stochasticity—that in the metapopulation dynamics, and that in the patch dynamics—and from the spatial and temporal correlations in the landscape. Our results are closed-form expressions that allow us to explore the patch occupancy for general choices of parameters and spatial kernels.

Our approach in the present manuscript differs slightly from Ovaskainen and Cornell (2006a), where we used the formalism of distributions to describe the landscapes and stochastic differential equations (SDEs) for the colonisation-extinction dynamics. The landscape dynamics now permit us to use the same SDE formalism for both the landscape and the metapopulation dynamics. Also, we develop the perturbation expansion from moment equations rather than directly from the SDEs themselves. This simplifies the analysis, as well as facilitating a discussion of higher orders in the expansion (Ovaskainen and Cornell, 2006b) (although we do not go beyond the first contribution to mean-field in the present study). We retain the SDE formalism as a convenient way to derive the moment equations, though other methods are possible (Bolker and Pacala, 1999). Our underlying model for the dynamic landscape reduces to the static landscape in the limit where patch turnover is slow, and the results of the present manuscript coincide with those of Ovaskainen and Cornell (2006a) in this limit.

We begin by defining our models for the landscape and metapopulation dynamics. We then develop the perturbation expansion and calculate general expressions for the first-order perturbation coefficients. We investigate the corrections to mean-field for a range of explicit choices for the spatial kernels governing colonisation and creation of patches, before discussing how these corrections depend on the general properties of these kernels. Finally we compare the predictions of the perturbation expansion truncated to first order with numerical simulations of the underlying stochastic system.

2 Dynamics of the landscape

Since we are interested in a metapopulation whose habitat is itself dynamic, we need to consider the processes by which the patch landscape is generated. The model presented here is appropriate for a species which lives on intrinsically ephemeral habitat, such as a saproxylic beetle or a butterfly that depends on vegetation that is not too high. The habitat patches for such a species (e.g. fallen dead trees, areas of short vegetation) are created by chance events (e.g. storms, fires) which we call *disturbances*. Each disturbance is localised at a particular time and space, and can generate one or more population patches in its vicinity. The habitat patches then degrade over time by decay or succession, and the local species population dies out. Note that, in this scenario, disturbances are beneficial (i.e. generate habitat) for the species, whereas in many studies the word ‘disturbance’ is used to denote processes that destroy habitat (Vuilleumier et al., 2007). Our model generates patch landscapes with positive spatial autocorrelations, since habitat is created in clusters. Dynamic landscapes without spatial correlations can be generated in our model by assuming that each disturbance never generates more than one habitat patch.

We assume that habitat is created in infinite d-dimensional space by localised disturbances that take place at a rate per unit area α. The disturbances themselves are assumed to be very short lived (relative to the lifetime of patches, and to the metapopulation dynamics which will later be superimposed on the landscape), and their effect is to bring into existence a Poisson-distributed number of pointlike habitat patches with mean ν. The l'th patch is created at a displacement $\boldsymbol{y}_l$ from the disturbance, where the $\boldsymbol{y}_l$ are iid random variables with probability density $K(\boldsymbol{y}_l)$. We assume that K is spherically symmetric and characterised by a single length scale, λ, and write

$$K(\boldsymbol{y}) = \frac{1}{\lambda^d} k(\frac{|\boldsymbol{y}|}{\lambda}), \tag{1}$$

where $k(y)$ is a dimensionless function of the dimensionless scalar argument y, with no explicit dependence on λ. Patches disappear independently by habitat degradation at a rate β.

2.1 SDE for the patch density

Suppose that at time t there is a set of patches at positions $\Pi(t) = \{\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3, \dots\}$. Let $Q(\boldsymbol{x}, t) = \sum_{\boldsymbol{X} \in \Pi(t)} \delta(\boldsymbol{x} - \boldsymbol{X})$ be the *patch density*, having the property that, for any region Ω, $\int_\Omega Q(\boldsymbol{y}) d\boldsymbol{y}$ is the number of patches contained within Ω. Q is an example of the type of point process considered in Appendix A, so we can derive a corresponding SDE and the properties of the noise terms:

- Since each disturbance creates on average ν patches, the rate of patch creation per unit area is
$$r_{\mathrm{C}}(\boldsymbol{x}, t) = \alpha\nu.$$
- Each patch dies independently a rate β, so the rate of patch destruction per unit area is
$$\begin{aligned} r_{\mathrm{D}}(\boldsymbol{x}, t) &= \sum_{X \in \Pi(t)} \beta\delta(\boldsymbol{x} - \boldsymbol{X}) \\ &= \beta Q. \end{aligned}$$

- To calculate the probability per unit time of creating two patches simultaneously, consider a disturbance that creates n patches and takes place in a volume $d\boldsymbol{y}$ in the vicinity of position $\boldsymbol{y}$. The joint probability that one particular patch is created in volume $d\boldsymbol{x}$ vicinity of $\boldsymbol{x}$, and another particular patch is created in volume $d\boldsymbol{x}'$ in the vicinity of $\boldsymbol{x}'$, is $K(\boldsymbol{x}-\boldsymbol{y})K(\boldsymbol{x}'-\boldsymbol{y})d\boldsymbol{x}\,d\boldsymbol{x}'$. Since any of the n patches is equally likely to be at $\boldsymbol{x}$, and given this any of the $(n-1)$ other patches is equally likely to be at $\boldsymbol{x}'$, the total probability that this disturbance produces one patch in the vicinity of $\boldsymbol{x}$ and another patch in the vicinity of $\boldsymbol{x}'$ is $n(n-1)K(\boldsymbol{x}-\boldsymbol{y})K(\boldsymbol{x}'-\boldsymbol{y})d\boldsymbol{x}\,d\boldsymbol{x}'$. We now average over disturbance positions $\boldsymbol{y}$ (probability per unit time $\alpha\, d\boldsymbol{y}$) and number n of patches per disturbance ($E(n(n-1))=\nu^2$,since n ∼Poi(ν)), so the rate of simultaneous production of patch pairs is

$$r_{\mathrm{CC}}(\boldsymbol{x},\boldsymbol{x}',t)=\alpha\nu^2\int K(\boldsymbol{x}-\boldsymbol{y})K(\boldsymbol{x}'-\boldsymbol{y})d\boldsymbol{y}.$$

- There are no processes that simultaneously destroy two patches, or that simultaneously destroy and create patches:

$$\begin{aligned} r_{\mathrm{DD}} &= 0 \\ r_{\mathrm{CD}} &= 0. \end{aligned}$$

Combining these results with Eqns. (A.2) and (A.4–A.6), we find that the patch density satisfies the following SDE

$$dQ(\boldsymbol{x},t)=r_Q dt+d\eta_Q, \tag{2}$$

where

$$\begin{aligned} r_Q &= \alpha\nu-\beta Q(\boldsymbol{x},t) \\ d\eta_Q &= d\eta_\alpha(\boldsymbol{x},t)-d\eta_\beta(\boldsymbol{x},t), \end{aligned}$$

and the covariances of the noise terms are

$$\begin{aligned} \overline{d\eta_\alpha(\boldsymbol{x},t)d\eta_\alpha(\boldsymbol{x}',t')} &= \alpha\nu\delta(t-t')dt\,dt'\Big(\delta(\boldsymbol{x}-\boldsymbol{x}') \\ &\qquad +\nu\int K(\boldsymbol{x}-\boldsymbol{y})K(\boldsymbol{x}'-\boldsymbol{y})d\boldsymbol{y}\Big) \\ \overline{d\eta_\beta(\boldsymbol{x},t)d\eta_\beta(\boldsymbol{x}',t')} &= \beta\overline{Q(\boldsymbol{x},t)}\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t')dt\,dt' \\ \overline{d\eta_\alpha(\boldsymbol{x},t)d\eta_\beta(x',t')} &= 0, \end{aligned}$$

where $\overline{A}$ represents the ensemble average of a quantity A over some, as yet unspecified, set of initial states. As explained in Appendix A.1, equal-time covariances for the noise are obtained from unequal-time covariances by replacing $\delta(t-t')dt\,dt'$ with dt. The equal-time covariance of the noise term $d\eta_Q$ is therefore

$$\overline{d\eta_Q(\boldsymbol{x},t)d\eta_Q(\boldsymbol{x}',t)} = V_{QQ}(\boldsymbol{x},\boldsymbol{x}',t)dt, \tag{3}$$

where

$$V_{QQ} = [\alpha\nu + \beta\overline{Q}]\delta(\boldsymbol{x}-\boldsymbol{x}') + \alpha\nu^2 \int K(\boldsymbol{x}-\boldsymbol{y})K(\boldsymbol{x}'-\boldsymbol{y})d\boldsymbol{y}.$$

2.2 Properties of the landscape

We now consider the properties of the landscape in the stationary state, which is obtained by evolving the dynamics for a long time. For the remainder of this manuscript we shall use $\overline{A}$ to mean the average of the quantity A in this stationary state, which can be obtained by starting from an arbitrary initial state at time t_0 in the limit $t_0 \to -\infty$. By performing such an ensemble average of (2) we see directly that, in the stationary state where $\overline{dQ} = 0$, the mean patch density is

$$Q_0 = \lim_{t\to\infty} \overline{Q(\boldsymbol{x},t)} = \alpha\nu/\beta.$$

If we write $Q = \overline{Q} + s_Q$, so that $\overline{s_Q(\boldsymbol{x}_1,t)r_Q(\boldsymbol{x}_2,t)} = \overline{s_Q(\boldsymbol{x}_1,t)[\alpha\nu - \beta(\overline{Q} + s(\boldsymbol{x}_2,t))]} = -\beta\overline{s_Q(\boldsymbol{x}_1,t)s_Q(\boldsymbol{x}_2,t)}$, then the equation for the second central moment of Q can be obtained directly from Eqn. (B.2):

$$\begin{aligned}\frac{d\overline{s_Q(\boldsymbol{x}_1,t)s_Q(\boldsymbol{x}_2,t)}}{dt} &= \overline{s_Q(\boldsymbol{x}_1,t)r_Q(\boldsymbol{x}_2,t)} + \overline{r_Q(\boldsymbol{x}_1,t)s_Q(\boldsymbol{x}_2,t)} + \overline{V_{QQ}(\boldsymbol{x}_1,\boldsymbol{x}_2,t)} \\ &= -2\beta\overline{s_Q(\boldsymbol{x}_1,t)s_Q(\boldsymbol{x}_2,t)} + \overline{V_{QQ}(\boldsymbol{x}_1,\boldsymbol{x}_2,t)}.\end{aligned}$$

The equilibrium spatial autocovariance of the patch density is obtained by setting $\frac{d\overline{s_Q(\boldsymbol{x}_1,t)s_Q(\boldsymbol{x}_2,t)}}{dt} = 0$:

$$\overline{s_Q(\boldsymbol{x}_1,t)s_Q(\boldsymbol{x}_2,t)} = Q_0[\delta(\boldsymbol{x}_1-\boldsymbol{x}_2) + Q_0 M(\boldsymbol{x}_1-\boldsymbol{x}_2)], \tag{4}$$

where

$$M(\boldsymbol{x}) = \frac{\nu}{2Q_0}\int K(\boldsymbol{x}+\boldsymbol{y})K(\boldsymbol{y})d\boldsymbol{y}$$

represents the spatial correlations in the landscape. We could equally well have derived Eqn. (4) by integrating Eqn. (2) (which is linear and can be solved by standard methods) starting at time $-\infty$ to give the patch density in equilibrium, and taking an ensemble average of $Q(\boldsymbol{x}_1,t)Q(\boldsymbol{x}_2,t)$. Notice that an uncorrelated landscape can be created by taking the limit $\nu \to 0$, in which case clusters of more than two patches are created with vanishing probability.

If we multiply Eqn. (4) by arbitrary functions $a(\boldsymbol{x}_1)$ and $b(\boldsymbol{x}_2)$ and integrate over $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, we recover Lemma 3.4 from Ovaskainen and Cornell (2006a).

This shows that M may be identified as the surplus probability density (relative to an uncorrelated landscape with the same mean patch density) that there is a patch at $\boldsymbol{x}_1$ given that there is one at $\boldsymbol{x}_2$. We can write the correlation function in terms of dimensionless functions by substituting for K using Eqn. (1), leading to

$$M(\boldsymbol{x}) = \frac{1}{Q_0\lambda^d} m(\frac{|\boldsymbol{x}|}{\lambda}),$$

where

$$m(|\boldsymbol{x}|) = \frac{\nu}{2} \int k(|\boldsymbol{x}+\boldsymbol{y}|)k(|\boldsymbol{y}|)d\boldsymbol{y} \tag{5}$$

is a dimensionless function with no explicit dependence on λ.

In Ovaskainen and Cornell (2006a), we simulated static, correlated landscapes by simulating the same procedure for long enough to generate an equilibrium. Eqn. (5) proves the result quoted in Appendix A of Ovaskainen and Cornell (2006a), which discussed how to create a static landscape with a given two-point correlation function.

3 Dynamics of the metapopulation

The species itself undergoes the standard colonisation-extinction dynamics of a stochastic patch-occupancy metapopulation, as in Ovaskainen and Cornell (2006a). We denote by $p_j(t) \in \{0,1\}$ the *occupancy* of the j'th patch, being unity if the patch is occupied at time t and zero otherwise, and impose the following dynamics:

- Extinctions, where an occupied patch becomes empty, take place independently at rate μ.
- An occupied patch at position $\boldsymbol{x}_j$ colonises an empty patch at $\boldsymbol{x}_i$ at a rate $C(\boldsymbol{x}_i - \boldsymbol{x}_j)$, where C is the *colonization kernel*, which is typically a decreasing function of the magnitude of its argument. We assume that colonization events are independent, so that the total colonization rate for the i'th patch is $(1-p_i)\sum_j p_j C(\boldsymbol{x}_i - \boldsymbol{x}_j)$, the factor $(1-p_i)$ ensuring that only empty patches can become colonised. Although patches cannot colonise themselves, there is no need to specify $j \neq i$ in the sum as the $j = i$ term contains a factor $p_i(1-p_i)$ which is always zero.

3.1 SDE for colonisation-extinction dynamics

We split the patch density $Q(\boldsymbol{x}, t)$ into unoccupied patches $u(\boldsymbol{x}, t) = \sum_i (1 - p_i(t))\delta(\boldsymbol{x} - \boldsymbol{x}_i)$ and occupied patches $q(\boldsymbol{x}, t) = \sum_i p_i(t)\delta(\boldsymbol{x} - \boldsymbol{x}_i)$, so that $Q = u + q$. We now have a marked point process (Stoyan and Penttinen, 2000), with events that lead to destruction or creation of one type of point (patch creation and destruction) and events that move a point from one class to the other (patch colonization and extinction). The SDE for the metapopulation dynamics may therefore be derived using a straightforward application of the methods of Appendix A. Since we have already derived an equation for Q (eqn. (2)), once we have derived an equation for q we can simply calculate u via $u = Q - q$.

The colonization rate for an empty patch at position x can be written in the form $\sum_j p_j C(\boldsymbol{x} - \boldsymbol{x}_j) = (C \star q)(\boldsymbol{x})$, where $\star$ denotes convolution (i.e. $(f \star g)(\boldsymbol{x}) \equiv \int f(\boldsymbol{y}) g(\boldsymbol{x} - \boldsymbol{y}) d\boldsymbol{y}$). The rate of creation of colonised patches per unit area is therefore $\kappa = \sum_i (1 - p_i)\delta(\boldsymbol{x} - \boldsymbol{x}_i) \sum_{j \neq i} C(\boldsymbol{x}_i - \boldsymbol{x}_j) p_j$, which may be written more concisely in the following form:

$$\kappa(u, q) = u(\boldsymbol{x})(C \star q)(\boldsymbol{x}). \tag{6}$$

The rate of extinction per unit area is $\sum_i \mu p_i \delta(\boldsymbol{x} - \boldsymbol{x}_i) = \mu q$. We also need to split the destruction of patches into empty and occupied patch components, which take place independently at rates βu and βq respectively.

Using the methods described in Appendix A, we can show that the density of occupied patches satisfies the stochastic differential equation

$$dq = r_q dt + d\eta_q, \tag{7}$$

where

$$\begin{aligned} r_q &= \kappa(u, q) - \mu q - \beta q \\ d\eta_q &= d\eta_\kappa - d\eta_{\mu q} - d\eta_{\beta q}. \end{aligned}$$

The rates κ, μq, and βq represent respectively the rates per unit area of colonisation of empty patches, extinction of occupied patches, and destruction of occupied patches, and $d\eta_\kappa$, $d\eta_{\mu q}$, $d\eta_{\beta q}$ respectively represent the noise associated with these processes.

The only process which takes place in a spatially correlated way is patch creation. However, this term does not enter directly into the dynamics of q, so the variances of the noise term are given by:

$$\begin{aligned} \overline{d\eta_{\beta q}(\boldsymbol{x}, t) d\eta_{\beta q}(\boldsymbol{x}', t')} &= \beta \overline{q} \delta(\boldsymbol{x} - \boldsymbol{x}')\delta(t - t') dt\, dt' \\ \overline{d\eta_\kappa(\boldsymbol{x}, t) d\eta_\kappa(\boldsymbol{x}', t')} &= \overline{\kappa} \delta(\boldsymbol{x} - \boldsymbol{x}')\delta(t - t') dt\, dt' \\ \overline{d\eta_{\mu q}(\boldsymbol{x}, t) d\eta_{\mu q}(\boldsymbol{x}', t')} &= \mu \overline{q} \delta(\boldsymbol{x} - \boldsymbol{x}')\delta(t - t') dt\, dt'. \end{aligned}$$

All cross correlations between the noises $d\eta_\alpha$, $d\eta_\kappa$, $d\eta_{\mu q}$, $d\eta_{\beta q}$ are zero, as they represent statistically independent processes. The cross-correlation between

$d\eta_\beta$, $d\eta_\kappa$ and $d\eta_{\mu q}$ are zero for the same reasons. However, there is a correlation between $d\eta_\beta$ and $d\eta_{\beta q}$, because the former represents destruction of all patches and the latter the destruction of occupied patches. This correlation is readily evaluated by noting that $d\eta_\beta = d\eta_{\beta u} + d\eta_{\beta q}$, where $d\eta_{\beta u}$ is the noise associated with destruction of unoccupied patches, which is statistically independent of the destruction of occupied patches, so $\overline{d\eta_{\beta u} d\eta_{\beta q}} = 0$. Therefore,

$$\overline{d\eta_\beta(\boldsymbol{x},t)d\eta_{\beta q}(\boldsymbol{x}',t')} = \overline{d\eta_{\beta q}(\boldsymbol{x},t)d\eta_{\beta q}(\boldsymbol{x}',t')}.$$

We can therefore write the equal-time covariances of the total noises for Q and q as Eqn. (3) and

$$\begin{aligned}\overline{d\eta_q(\boldsymbol{x},t)d\eta_q(\boldsymbol{x}',t)} &= V_{qq}dt\\ \overline{d\eta_Q(\boldsymbol{x},t)d\eta_q(\boldsymbol{x}',t)} &= V_{qQ}dt,\end{aligned}$$

where

$$\begin{aligned}V_{qq} &= [\overline{\kappa} + (\beta+\mu)\overline{q}]\delta(\boldsymbol{x}-\boldsymbol{x}')\\ V_{qQ} &= -\beta\overline{q}\delta(\boldsymbol{x}-\boldsymbol{x}').\end{aligned}$$

3.2 Patch occupancy at equilibrium

The principal quantity that interests us is the density of occupied patches at equilibrium, $\overline{q}$. By definition, $\overline{q}$ is independent of time, and it is also independent of space since the dynamical processes are statistically the same everywhere. We can obtain an expression for $\overline{q}$ by taking the average of Eqn. (7):

$$\begin{aligned}0 &= \overline{\kappa(u,q)} - (\mu+\beta)\overline{q}\\ \Rightarrow (\mu+\beta)\overline{q} &= \overline{(Q(\boldsymbol{x}) - q(\boldsymbol{x}))(C \star q)(\boldsymbol{x})},\end{aligned} \tag{8}$$

where we have used (6) and substituted $u = Q - q$. A mean-field approximation is obtained by replacing the average of a product by the product of the averages, so the mean-field solution to (8) is $\overline{q} \approx q_0$, with

$$q_0 = Q_0 - \frac{\mu+\beta}{\tilde{C}(0)}, \tag{9}$$

where $\tilde{C}(0) \equiv \int C(\boldsymbol{x})d\boldsymbol{x}$ is the integral over all space of C, i.e. the zero-frequency component of the Fourier transform of C. A negative value for the expression on the right hand side of Eqn. (9) corresponds to extinction. Here, and throughout this manuscript, we shall adopt the following convention for the definition of Fourier transforms:

$$\begin{aligned}\tilde{F}(\boldsymbol{\omega}) &= \int e^{2\pi i \boldsymbol{x}\cdot\boldsymbol{\omega}} F(\boldsymbol{x}) d\boldsymbol{x} \\ \Leftrightarrow F(\boldsymbol{x}) &= \int e^{-2\pi i \boldsymbol{x}\cdot\boldsymbol{\omega}} \tilde{F}(\boldsymbol{\omega}) d\boldsymbol{\omega},\end{aligned}$$

where $\boldsymbol{x}$ and $\boldsymbol{\omega}$ are d-dimensional vectors, and $d\boldsymbol{x}$, $d\boldsymbol{\omega}$ represent volume measures in d-dimensional space.

Let us define $q_d = \overline{q} - q_0$, which is the difference between the average over realisations of the density of occupied patches and the mean-field approximation, and $s_q = q - \overline{q}$, which is the 'stochastic' component of q (i.e. the difference between q and its mean) so that $q = q_0 + q_d + s_q$. Then (8) may be rearranged to give

$$\begin{aligned}q_d \tilde{C}(0)(q_0 + q_d) &= \overline{(s_Q - s_q)(C \star s_q)}. &(10)\\ &= C \star g_{qQ} - C \star g_{qq}, &(11)\end{aligned}$$

where $g_{qq}(\boldsymbol{x}) = \overline{s_q(\boldsymbol{y}) s_q(\boldsymbol{x}+\boldsymbol{y})}$ and $g_{qQ}(\boldsymbol{x}) = \overline{s_q(\boldsymbol{y}) s_Q(\boldsymbol{y}+\boldsymbol{x})}$ are second central moments, which are independent of $\boldsymbol{y}$ because of the translational invariance of the ensemble average. So far no approximations have been made, and Eqn. (11) is exact.

3.3 Equations for second moments

To make progress with Eqn (11), we apply the methods in Appendix B.1 to the SDEs (2) and (7) to obtain differential equations for g_{qq} and g_{qQ}

$$
\begin{aligned}
\frac{dg_{qq}(\boldsymbol{x}-\boldsymbol{x}',t)}{dt} &= \overline{R_q s'_q} + \overline{s_q R'_q} + V_{qq}(\boldsymbol{x},\boldsymbol{x}',t) \\
&= (\overline{Q}-\overline{q})\overline{(C\star s_q)s'_q} + \tilde{C}(0)\overline{q}\,\overline{(s_Q - s_q)s'_q} + \\
&\quad \overline{(s_Q - s_q)(C\star s_q)s'_q} - (\mu+\beta)\overline{s_q s'_q} + \\
&\quad (\overline{Q}-\overline{q})\overline{(C\star s'_q)s_q} + \tilde{C}(0)\overline{q}\,\overline{(s'_Q - s'_q)s_q} + \\
&\quad \overline{(s'_Q - s'_q)(C\star s'_q)s_q} - (\mu+\beta)\overline{s'_q s_q} + \\
&\quad [\overline{\kappa} + (\mu+\beta)\overline{q}]\delta(\boldsymbol{x}-\boldsymbol{x}') \\
&= 2(\overline{Q}-\overline{q})C\star g_{qq} + 2\tilde{C}(0)\overline{q}[g_{qQ}-g_{qq}] - 2(\mu+\beta)g_{qq} + \\
&\quad 2\overline{(s_Q - s_q)(C\star s_q)s'_q} + [\overline{\kappa} + (\mu+\beta)\overline{q}]\delta(\boldsymbol{x}-\boldsymbol{x}') \qquad (12) \\
\frac{dg_{qQ}(\boldsymbol{x}-\boldsymbol{x}',t)}{dt} &= \overline{R_q s'_Q} + \overline{s_q R'_Q} + V_{qQ} \\
&= (\overline{Q}-\overline{q})\overline{s_Q(C\star s'_q)} + \tilde{C}(0)\overline{q}\,\overline{s_Q(s'_Q - s'_q)} \\
&\quad -\beta\overline{s_Q s'_q} + \overline{s_Q(s'_Q - s'_q)(C\star s'_q)} \\
&\quad -(\mu+\beta)\overline{s_Q s'_q} + \beta\overline{q}\delta(x-x') \\
&= (\overline{Q}-\overline{q})C\star g_{qQ} + \tilde{C}(0)\overline{q}[g_{QQ}-g_{qQ}] - (\mu+2\beta)g_{qQ} + \\
&\quad \overline{s_Q(s'_Q - s'_q)(C\star s'_q)} + \beta\overline{q}\delta(\boldsymbol{x}-\boldsymbol{x}'), \qquad (13)
\end{aligned}
$$

where we have used the shorthand $s'_q \equiv s(\boldsymbol{x}',t)$, $r'_Q \equiv r_Q(\boldsymbol{x}',t)$, etc. and substituted for r_q, V_{qq} etc. in terms of $s_q = q - \overline{q}$ and $s_Q = Q - \overline{Q}$. We have also exploited the reflection symmetry of the ensemble average to write e.g. $g_{qq}(\boldsymbol{x}) = g_{qq}(-\boldsymbol{x}')$ and $\overline{s_q(C\star s_q)s'_q} = \overline{s_q(C\star s'_q)s'_q}$. The equations for the second central moments contain higher order moments (in this case the third), as is usually the case for nonlinear dynamics, and the infinite hierarchy of equations cannot be solved exactly.

4 Perturbation expansion for equilibrium patch occupancy

The mean field approximation in (9) would be exact for an infinite landscape where all patches interact equally, i.e. in the limit of N patches with $C \propto 1/N$ and $N \to \infty$. We instead assume that the colonization kernel is spherically symmetric and described by a single length scale L, so that

$$C(\boldsymbol{x}) = L^{-d}c(|\boldsymbol{x}|/L), \qquad (14)$$

where $c(x)$ is a monotonically decreasing integrable function of x with no explicit dependence on L, and the prefactor L^{-d} ensures that the total mean colonization effort $\int C(\boldsymbol{x})dx = \int c(|\boldsymbol{y}|)d\boldsymbol{y}$ is independent of L. Then the number of patches contributing significantly to colonization at a given point $\boldsymbol{x}$ is of the same order of magnitude as the number of patches within a distance L of $\boldsymbol{x}$, i.e. $Q_0 L^d$, which is large if L is large relative to the typical patch separation

${Q_0}^{-1/d}$. Therefore, we expect the mean-field approximation to become exact in the limit $L \to \infty$, where the mean density of occupied patches is (using Eqns. (9) and (14)) $q_0 = Q_0 - (\mu + \beta)/\tilde{c}(0)$, so that the mean fraction of patches that are occupied is $p_0 = q_0/Q_0$, i.e.

$$p_0 = 1 - \frac{\mu + \beta}{\tilde{c}(0)Q_0}. \tag{15}$$

Eqn. (15) is equivalent to Eqn. (12) in Keymer et al. (2000), who pointed out that the correct mean field theory for metapopulation dynamics in dynamic landscapes needs to take account of the extinctions due to occupied patches being removed.

In order for the correction q_d to mean field limit to be obtained as $L \to \infty$, the RHS of Eqn. (10) must vanish in this limit. We can understand this by interpreting $C \star s_q = \int C(\boldsymbol{x} - \boldsymbol{y}) s_q(\boldsymbol{y}) d\boldsymbol{y}$ as a smoothing of the fluctuation s_q by the kernel C. Since C in Eqn. (14) is normalised to unity and has width L, $C \star s_q$ is effectively an average of the fluctuations over a region of size L, which will therefore become small for large L. An alternative way of seeing this is given by expressing the convolutions in Eqn. (11) in Fourier space:

$$\begin{aligned} q_d\tilde{C}(0)(q_0 + q_d) &= \int \tilde{C}(\boldsymbol{\omega})\tilde{g}_{qQ}(\boldsymbol{\omega})d\boldsymbol{\omega} - \int \tilde{C}(\boldsymbol{\omega})\tilde{g}_{qq}(\boldsymbol{\omega})d\boldsymbol{\omega} \\ &= L^{-d}\int \tilde{c}(|\boldsymbol{y}|)\tilde{g}_{qQ}(\frac{\boldsymbol{y}}{L})d\boldsymbol{y} - L^{-d}\int \tilde{c}(|\boldsymbol{y}|)\tilde{g}_{qq}(\frac{\boldsymbol{y}}{L})d\boldsymbol{y}, \end{aligned} \tag{16}$$

where in the second line we have changed the integration variable to $\boldsymbol{y} = \boldsymbol{\omega} L$. q_d will be of order L^{-d} if the integrals on the RHS of Eqn. (16) do not diverge for large L, i.e. if the central second moments are of order unity when expressed as functions of ωL. This is the central idea that enables us to perform a perturbation expansion of $\overline{q}$ in L^{-d}: the first order term in q_d is given by the zeroth order term for the central second moments.

We use similar arguments to show that the third moment terms do not contribute to the zeroth order contribution to the second moments in Eqns. (12) and (13). These terms all contain a convolution of the fluctuation, $C \star s_q$, which vanished when $L \to \infty$. More formally, let $f(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = \overline{s_q(\boldsymbol{x}_1)s_q(\boldsymbol{x}_2)s_q(\boldsymbol{x}_3)}$, so that $\overline{s_q(\boldsymbol{x})(C \star s_q)(\boldsymbol{x})s_q(\boldsymbol{x}')} = \int f(\boldsymbol{x}, \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x}')C(\boldsymbol{y})d\boldsymbol{y}$. If f is of order unity, then we can show by expressing the convolution in Fourier space as we did in Eqn (16) that $\overline{s_q(\boldsymbol{x})(C \star s_q)(\boldsymbol{x})s_q(\boldsymbol{x}')}$ will be of order L^d. The same argument can be applied to all third moment terms in Eqns. (12) and (13).

To derive equations for the zeroth-order contributions $\tilde{g}^0_{qq}$ and $\tilde{g}^0_{qQ}$ to the second moments in equilibrium, we perform the following steps on Eqns. (12) and (13):

- Set $dg_{qq}/dt = dg_{qQ}/dt = 0$

- Drop all terms involving convolutions of third moments, because they only contribute to order L^{-2d} and higher
- Take Fourier transforms of the second moments
- Replace $\overline{Q} = Q_0$, and use the mean-field values $\overline{q} \to q_0$, $\overline{\kappa} \to (Q_0 - q_0)\tilde{C}(0)q_0 = (\mu + \beta)q_0$.

This leads to

$$\begin{aligned}
0 &= \{(Q_0 - q_0)\tilde{C}(\boldsymbol{\omega}) - q_0\tilde{C}(0) - \mu - \beta\}\tilde{g}^0_{qq}(\boldsymbol{\omega}) + q_0\tilde{C}(0)\tilde{g}^0_{qQ}(\boldsymbol{\omega}) + (\mu + \beta)q_0 \\
0 &= \{(Q_0 - q_0)\tilde{C}(\boldsymbol{\omega}) - q_0\tilde{C}(0) - \mu - 2\beta\}\tilde{g}^0_{qQ}(\boldsymbol{\omega}) + q_0\tilde{C}(0)\tilde{g}_{QQ}(\boldsymbol{\omega}) + \beta q_0,
\end{aligned}$$

from which $\tilde{g}^0_{qq}$ and $\tilde{g}^0_{qQ}$ can be solved by linear algebra in terms of $\tilde{g}_{QQ}$. The Fourier transform of Eqn (4) gives

$$\tilde{g}_{QQ}(\boldsymbol{\omega}) = Q_0(1 + Q_0\tilde{M}(\boldsymbol{\omega})),$$

and solving Eqns (12) and (13) leads, after some algebra, to

$$\tilde{g}^0_{qQ}(\boldsymbol{y}L) - \tilde{g}^0_{qq}(\boldsymbol{y}L) = \frac{\phi(|\boldsymbol{y}|)p_0Q_0(1 - p_0)\{(\phi(|\boldsymbol{y}|) - 1)Q_0\tilde{M}(\boldsymbol{y}) - p_0 - \gamma\}}{[\phi(|\boldsymbol{y}|) - (1 - p_0)][\phi(|\boldsymbol{y}|)(1 + \gamma) - (1 - p_0)]}, \tag{17}$$

where

$$\begin{aligned}
\phi(|\boldsymbol{y}|) &= \frac{\tilde{C}(0)}{\tilde{C}(\boldsymbol{y}L)} \\
\gamma &= \frac{\beta(1 - p_0)}{\mu + \beta}.
\end{aligned}$$

We next write the patch occupancy as $\overline{p} = \frac{\overline{q}}{Q_0} = p_0 + p_d$, where p_d is the correction relative to mean-field, and express p_d as a series in $\frac{1}{Q_0L^d}$, $p_d = \frac{1}{Q_0L^d}p^1_d + O(Q_0^{-2}L^{-2d})$. We take the terms of order $\frac{1}{Q_0L^d}$ on both sides of Eqn. (16) to give

$$\begin{aligned}
p^1_d &= \frac{1}{\tilde{C}(0)q_0}\int \tilde{c}(\boldsymbol{y})\{\tilde{g}^0_{qQ}(\frac{\boldsymbol{y}}{L}) - \tilde{g}^0_{qQ}(\frac{\boldsymbol{y}}{L})\}dy \\
&= p_{\mathrm{S}} + p_{\mathrm{D}} + p_{\mathrm{C}} + p_{\mathrm{DC}},
\end{aligned} \tag{18}$$

where

$$p_{\mathrm{S}} = -p_0(1-p_0)\int \frac{1}{[\phi(|\boldsymbol{y}|)-(1-p_0)]^2}d\boldsymbol{y} \tag{19}$$

$$p_{\mathrm{D}} = -\gamma(1-p_0)^2\int \frac{[\phi(|\boldsymbol{y}|)-1]}{[\phi(|\boldsymbol{y}|)(1+\gamma)-(1-p_0)][\phi(|\boldsymbol{y}|)-(1-p_0)]^2}d\boldsymbol{y} \tag{20}$$

$$p_{\mathrm{C}} = (1-p_0)\int \frac{\tilde{m}(\lambda y|\boldsymbol{y}|/L)[\phi(|\boldsymbol{y}|)-1]}{[\phi(|\boldsymbol{y}|)-(1-p_0)]^2}d\boldsymbol{y} \tag{21}$$

$$p_{\mathrm{DC}} = -\gamma(1-p_0)\int \frac{\tilde{m}(\lambda|\boldsymbol{y}|/L)\phi(y)[\phi(|\boldsymbol{y}|)-1]}{[\phi(|\boldsymbol{y}|)(1+\gamma)-(1-p_0)][\phi(|\boldsymbol{y}|)-(1-p_0)]^2}d\boldsymbol{y}. \tag{22}$$

Eqns. (19–22) are the central results of this paper. They may be interpreted as follows:

- p_{S} is the value of p_d^1 when $\beta = 0 = M$, i.e. when the landscape is static and uncorrelated. This coefficient gives the contribution due to the noise in the colonisation-extinction dynamics and the randomness of the landscape alone. It is identical to the result derived in Ovaskainen and Cornell (2006a), corresponding to the combination $\omega + \theta$ in the notation of that paper. p_{S} is always negative, since the integrand in Eqn (19) is positive. Therefore, the stochastic metapopulation always fares worse than the mean-field prediction.
- $p_{\mathrm{S}} + p_{\mathrm{D}}$ is the value of p_d^1 when $M = 0$ and $\beta \neq 0$, i.e. when the landscape is dynamic and uncorrelated. Therefore, p_{D} may be identified as the coefficient of the contribution due to the stochasticity in the patch creation-removal dynamics. We can write the integrand of Eqn. (20) as

  $$\frac{\phi-1}{[\phi(1+\gamma)-(1-p_0)][\phi-(1-p_0)]^2} = \frac{1-\frac{1}{\phi}}{[2+\gamma-p_0+(1-\frac{1}{\phi})(1-p_0)][\phi-(1-p_0)]^2}.$$

  Although $\tilde{C}(\boldsymbol{\omega})$, and therefore $\phi(y)$, can be negative, $C(\boldsymbol{x})$ is nowhere negative so $\sup_\omega \tilde{C}(\boldsymbol{\omega}) = \tilde{C}(0)$ and therefore $\frac{1}{\phi(y)} \leq 1 \forall y$. Since $(1-\frac{1}{\phi})$, γ, and $(1-p_0)$ are positive, the integrand in Eqn (20) is positive. Therefore, p_{D} is always negative, so the the metapopulation always does worse than would be predicted by mean-field or by considering the contribution due to the stochastic colonization-extinction dynamics alone. To understand why this is the case, consider the two situations where (i) a population goes extinct and (ii) a population dies because its patch vanishes, which must (on average) be balanced by the appearance of another patch somewhere in the landscape. Both cases result in an empty patch, but in case (i) the fact that the patch was previously occupied means that it is likely to be close to some other occupied patches, whereas in case (ii) the new patch is placed at random on the landscape with no relation to the position of occupied patches. The empty patch is therefore more likely to be colonised in case (i)—the static landscape—than in case (ii)—the dynamic landscape.
- $p_{\mathrm{S}} + p_{\mathrm{C}}$ is the value of p_d^1 when $\beta = 0$ and $M \neq 0$, i.e. when the landscape

is static but correlated. Therefore, p_{C} may be identified as the coefficient of the contribution due to the static correlations in the landscape structure. This quantity is identical to the quantity θ_c derived in Ovaskainen and Cornell (2006a). Taking the Fourier transform of Eqn. (5), we find $\tilde{m}(\boldsymbol{y}) = \frac{\nu}{2}\{\tilde{k}(\boldsymbol{y})\}^2 > 0$, i.e. landscapes generated by the process described in this manuscript have correlations with no negative Fourier components. Therefore, p_C is positive and proportional to the strength ν of the correlations, showing that a clustered landscape is beneficial to the species.

- $p_{\mathrm{S}} + p_{\mathrm{D}} + p_{\mathrm{C}} + p_{\mathrm{DC}}$ is the value of p_d^1 when β and M are both nonzero. Since we have already identified the significance of p_{S}, p_{D}, and p_{C}, we see that p_{DC} represents the interaction between spatial and temporal correlations in the dynamics, i.e. the residual term that is obtained when the contributions due to static spatial correlations and uncorrelated dynamics are removed. We might therefore expect the form of p_{DC} to depend on the mechanistic processes that create and destroy patches, and not just the spatial correlations in the patch landscape. It is interesting that this term is of the same order as p_{C} and p_{D}, whereas there is no interaction term of this order between colonization-extinction noise and spatial correlations (i.e. p_{C} is the same for deterministic and stochastic dynamics). If $\tilde{C}(\boldsymbol{\omega})$, and therefore $\phi(y)$, is everywhere positive, then the integrand in Eqn. (22) is positive, and p_{DC} will be negative. However, if the colonisation kernel is very abrupt in space, it is possible for $\tilde{C}$ to be negative for some frequencies. The only cases where we have found $p_{\mathrm{DC}} > 0$ are when either the colonisation kernel or the patch creation density is not monotonically decreasing with separation, i.e. there are some characteristic distances which are preferred.

5 2-dimensional examples

We shall now consider some specific examples of kernels in two dimensional (2D) space. We shall also use simulations to investigate how well the perturbation theory performs as an approximation when the colonisation length is finite. In 2D, the Fourier transform of a circularly symmetric function $G(|\boldsymbol{x}|)$ takes the form

$$\begin{aligned}\tilde{G}(\omega) &= \int_0^\infty \int_0^{2\pi} e^{2\pi i x\omega\cos\theta} G(x)\, d\theta\, x\, dx \\ &= 2\pi \int_0^\infty J_0(2\pi\omega x) G(x) x\, dx,\end{aligned}$$

where J_0 is a Bessel function of the first kind of order zero (Abramowitz and Stegun, 1964).

5.1 Colonisation kernels

We begin by considering dynamic landscapes without correlations. We shall consider the following four kernels:

$$c_0(x) = \frac{Q_0(1-p_0)}{\mu+\beta}\frac{K_0(x)}{2\pi} \tag{23}$$

$$c_1(x) = \frac{Q_0(1-p_0)}{\mu+\beta}\frac{xK_1(2^{\frac{1}{2}}x)}{2^{\frac{1}{2}}\pi} \tag{24}$$

$$c_G(x) = \frac{Q_0(1-p_0)}{\mu+\beta}\frac{e^{-\frac{x^2}{4}}}{4\pi} \tag{25}$$

$$c_F(x) = \frac{Q_0(1-p_0)}{\mu+\beta}\frac{1}{2\pi[1+x^2]^{\frac{3}{2}}}, \tag{26}$$

where K_0 and K_1 are modified Bessel functions of the second kind, respectively of zeroth and first order (Abramowitz and Stegun, 1964). All of the kernels have the same integral over 2D space,$\int c(x) 2\pi x \, dx = \frac{Q_0(1-p_0)}{\mu+\beta}$, to ensure that the mean-field patch occupancy is p_0 (see Eqn. (15)). The first three kernels are exponentially bounded at large argument, and have the same variance $\int x^2 c(x) 2\pi x \, dx = \frac{4Q_0(1-p_0)}{\mu+\beta}$. The fourth kernel c_F is 'fat-tailed', i.e. it is not exponentially bounded at large argument, and its variance is infinite.

The first kernel c_0 is appropriate if individuals perform random walks with an exponentially distributed number of steps (Bolker and Pacala, 1999; Ovaskainen and Cornell, 2003). It diverges logarithmically at small x, so to see how this singularity affects the results we also consider a kernel c_1 which is finite at $x = 0$. The Gaussian kernel c_G, implying normally distributed dispersal distances, is another popular choice. The kernel c_F (which corresponds to the Cauchy density in 2D (Chave and Leigh, 2002)) is biologically important because fat-tailed kernels are often found for wind-dispersed plants (Nathan, 2006), and in particular some studies report dispersal kernels that decay like $\sim x^{-3}$ at large distance x (Paradis et al., 2002).

The 2D Fourier transform of all four kernels can be calculated explicitly, and the functions $\phi(y) = \frac{\tilde{c}(0)}{\tilde{c}(y)}$ are

$$\begin{aligned}
\phi_0 &= 1 + 4\pi^2 y^2 \\
\phi_1 &= (1 + 2\pi^2 y^2)^2 \\
\phi_G &= e^{4\pi^2 y^2} \\
\phi_F &= e^{2\pi y}.
\end{aligned}$$

The first-order perturbation $p_S + p_D$ to the nonspatial limit for these colonisa-

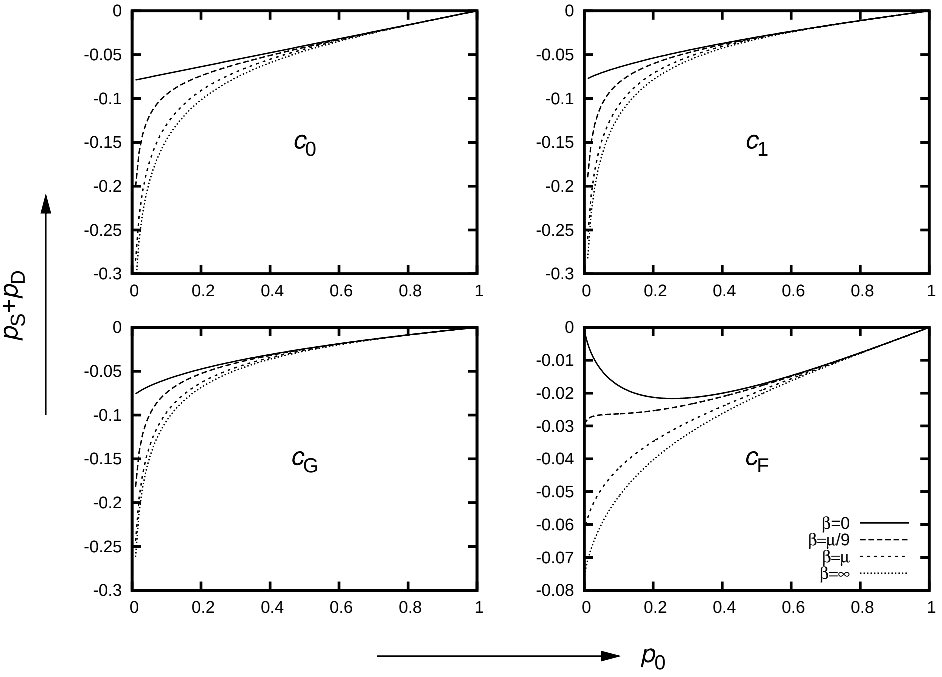

Figure 1. Coefficient of corrections to mean field for a dynamic, uncorrelated landscape, as a function of the mean-field occupancy p_0. Different curves correspond to different ratios of patch turnover to patch extinction rate. Solid line: static landscape $\beta = 0$; long dashes: $\beta = \mu/9$; short dashes: $\beta = \mu$; dotted line: maximally dynamic landscape, $\frac{\mu}{\beta} = 0$. The four panels correspond to the different colonisation kernels defined in Eqn. (23–26). Top left: colonisation kernel c_0; top right: colonisation kernel c_1; bottom left: Gaussian colonisation kernel c_G; bottom right: fat-tailed colonisation kernel c_{F}.

tion kernels is obtained by substituting these forms for ϕ in to Eqns. (19) and (20), the results of which are to be found in Appendix C. The results depend on only two variables: the metapopulation occupancy p_0 in the mean-field limit (see Eqn. (15)), which summarises the metapopulation colonisation-extinction dynamics, and the ratio $\frac{\beta}{\mu}$ of patch turnover to patch extinction rate.

The coefficient $p_{\mathrm{S}} + p_{\mathrm{D}}$ for the four different kernels is plotted as a function of p_0 in Figure 1. As noted in the previous section, the coefficient is always negative, and becomes more negative as the landscape is made more dynamic. Consider first the figures for the three exponentially-bounded kernels c_0, c_1, and c_{G} , which are remarkably similar. The curve for $\beta = 0$ (solid line) has the same value ≈ -0.0795 at $p_0 = 0$, then becomes less negative as p_0 is increased and becomes zero at $p_0 = 1$. Meanwhile, the curves for $\beta \neq 0$ (dotted and dashed curves) diverge logarithmically as $p_0 \to 0$, with apparently the same amplitude for all kernels. The reason for this behaviour can be understood by analysing Eqns. (19) and (20) in the limits $p_0 \to 0$ and $p_0 \to 1$ (Appendix D.1). We find that the limiting behaviour when $p_0 \to 0$ is

$$p_{\mathrm{S}} + p_{\mathrm{D}} \rightarrow \begin{cases} -\frac{1}{\pi V} & \text{for } \beta = 0 \\ -\frac{\log \frac{1}{p_0}}{\pi V} + \text{constant} & \text{for } \beta \neq 0, \end{cases}$$

where $V = \frac{\int x^2 c(x) d2\pi x\, dx}{\int c(x) 2\pi x\, dx}$ is the variance of the kernel c, and that $p_{\mathrm{S}} + p_{\mathrm{D}} \rightarrow 0$ when $p_0 = 1$. These expressions depend on the kernel's variance alone. Since p_{S} and p_{D} have the same values near $p_0 = 0$ and $p_0 = 1$ for kernels with the same variance, they are likely also to take similar values at intermediate values of p_0.

The curves for the fat-tailed kernel, however, are rather different. While we still have $p_{\mathrm{S}} + p_{\mathrm{D}} \rightarrow 0$ as $p_0 \rightarrow 1$, $p_{\mathrm{S}} + p_{\mathrm{D}}$ approaches a finite value at $p_0 = 0$, which is zero when $\beta = 0$. In Appendix D.2 we show in general that, if the colonisation kernel has infinite variance, then p_{S} is zero at $p_0 = 0$, and that for a dynamic landscape $p_{\mathrm{S}} + p_{\mathrm{D}}$ is finite and nonzero. Stochastic fluctuations therefore have a weaker effect than for the other kernels, which is reasonable because the fatter tail allows local extinctions to be re-colonised from very large distances.

5.2 Landscape correlations

In order to study the effect of landscape correlations on the metapopulation dynamics, we study three different forms for the function $k(r)$ governing the probability density of distances of new patches from disturbances:

$$k_1(x) = \frac{e^{-x}}{2\pi x}$$
$$k_2(x) = \frac{K_0(x)}{2\pi}$$
$$k_3(x) = \frac{e^{-x}}{2\pi},$$

where K_0 is a modified Bessel function of the second kind. These functions are all exponentially bounded at large distance, but have different behaviour as $x \rightarrow 0$: k_1 diverges as $\frac{1}{x}$, k_2 diverges logarithmically, and k_3 remains finite.

Using Eqn. (5), the Fourier transform of the correlation function for patch creation density k_n is $\tilde{m}_n = \frac{\nu}{2}(1 + 4\pi^2\lambda^2\omega^2)^{-n}$, which corresponds respectively to the following landscape correlation functions

$$m_1(x) = \frac{\nu K_0(x)}{4\pi} \tag{27}$$

$$m_2(x) = \frac{\nu x K_1(x)}{8\pi} \tag{28}$$

$$m_3(x) = \frac{\nu x^2 K_2(x)}{32\pi}, \tag{29}$$

where K_n is a modified Bessel function of the second kind of order n. This collection of landscapes allows us to investigate the cases where both, one, or neither out of the patch creation function and landscape correlation function have short-distance singularities.

The coefficient $p_{\mathrm{C}} + p_{\mathrm{DC}}$, which represents the additional corrections to the patch occupancy beyond that for an uncorrelated landscape with the same patch turnover, are listed in appendices C.1 and C.2 for the cases where the colonisation kernel is c_0 and c_1 (see eqns. (C.1–C.4)). Closed form expressions for $p_{\mathrm{C}} + p_{\mathrm{DC}}$ cannot be obtained for c_{G} and c_{F} with the above landscapes, but these cases can be obtained numerically. The coefficients are plotted for c_0, c_1, and c_{F} in Figures 2, 3, and 4 respectively (results for c_{G} are qualitatively similar to c_1).

A feature common to Figures 2, 3, and 4 is that the correction is always positive (as discussed in section 4), and decreases as the landscape correlation length λ is increased. This is because the landscape correlations represent an increased chance of finding other patches a distance approximately $\sim \lambda$ from any randomly chosen patch; since the colonisation kernel decreases monotonically with distance, there will be more successful colonisations if λ is smaller. The correction decreases sharply when $\lambda > L$, showing that landscape correlations are unimportant when colonisation events are unlikely to be between patches in the same cluster. The correction due to landscape correlations are reduced when the landscape is dynamic, but the shape of the curves is similar. This contrasts with Fig. 1, where the corrections due to a dynamic uncorrelated landscape differ markedly from those for a static, uncorrelated landscape. When p_0 is large, the corrections for dynamic and static landscapes do not differ much, just as was found for p_{S} and p_{D}.

The colonisation kernels c_1 and c_{F} that are finite at zero distance are insensitive to patch separation when this is very small, so the correction approaches a constant value when $\lambda \to 0$ for these kernels. However, for the kernel c_0 that has a logarithmic singularity at small distance, the correction diverges logarithmically when $\lambda \to 0$, since the colonisation rate increases without limit as patches are brought closer together. The shape of the curves is qualitatively very similar for the fat-tailed kernel c_{F} and exponentially-bounded kernel c_1, though the former decays more slowly for large λ. It is interesting to note that the results are very similar for all three landscape correlation functions, showing that a logarithmic singularity in the landscape correlation function

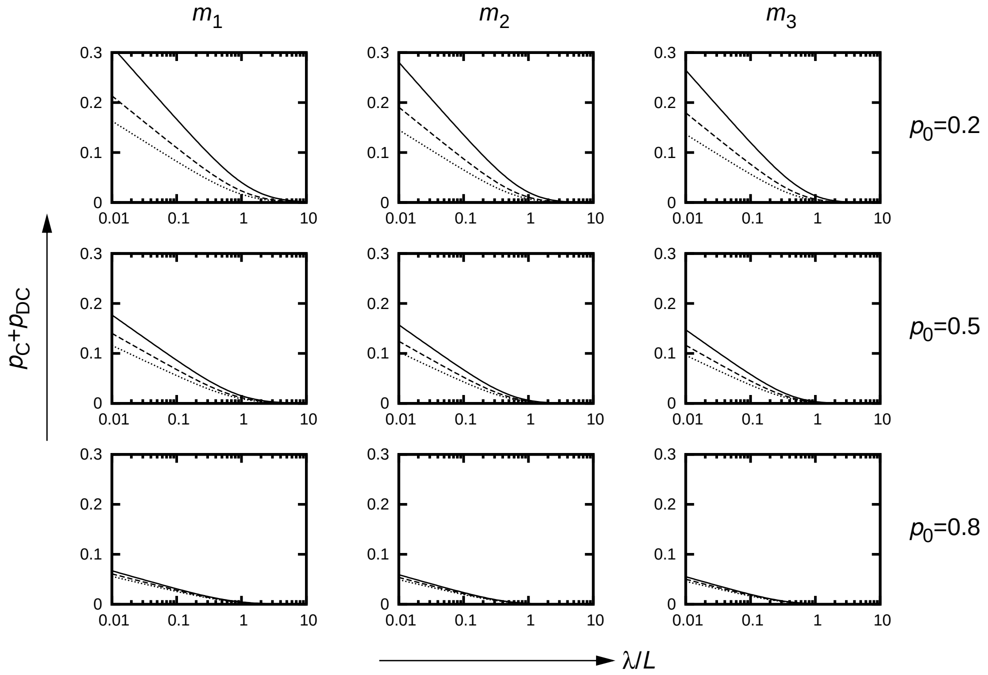

Figure 2. Coefficient $p_\mathrm{C} + p_\mathrm{DC}$ of corrections due to correlations in the landscape as a function of the relative landscape correlation length $\frac{\lambda}{L}$, when $\nu = 1$ and the colonisation kernel is c_0. The different columns correspond to different choices of landscape correlation functions as defined in Eqns. (27–29). Solid lines: static landscape; dashed lines: patch removal rate=patch extinction rate; dotted lines: maximally dynamic landscape.

does not have a strong effect on the metapopulation dynamics.

To this order in perturbation theory, the correction to the mean-field occupancy due to landscape correlations is proportional to ν, which represents the average patch cluster size at creation and hence the strength of the correlations. Since $p_\mathrm{S} + p_\mathrm{D}$ is always negative, and $p_\mathrm{C} + p_\mathrm{DC}$ is positive, the total correction to the mean-field is positive when $\nu > \nu^*$, and negative when $\nu < \nu^*$, where ν^* is the value of ν where the total coefficient is zero. From Eqn (18), the value of ν^* is

$$\nu^* = -\frac{\nu[p_\mathrm{S}(p_0) + p_\mathrm{D}(p_0, \gamma)]}{p_\mathrm{C}(p_0, \nu, \lambda/L) + p_\mathrm{DC}(p_0, \gamma, \nu, \lambda/L)}.$$

The quantity ν^* is a measure of the relative importance of landscape correlations and stochasticity for the metapopulation, a large value of ν^* implying that spatial correlations make a relatively weak contribution.

The behaviour of ν^* as a function of p_0 is plotted in Figure 5 for three colonisation kernels and three landscape correlation functions. We have arbitrarily chosen $\lambda = \frac{L}{10}$ in this figure; other values of λ would have the same shape but

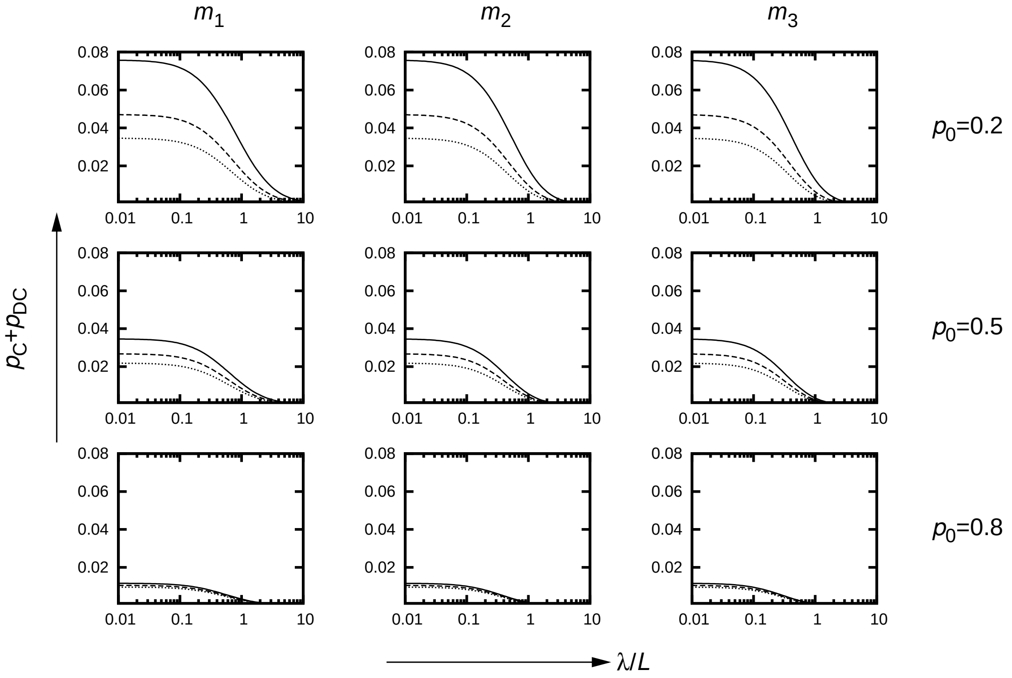

Figure 3. Coefficient $p_{\mathrm{C}}+p_{\mathrm{DC}}$ for the colonisation kernel c_1. Parameters and symbols are as for Fig. 2.

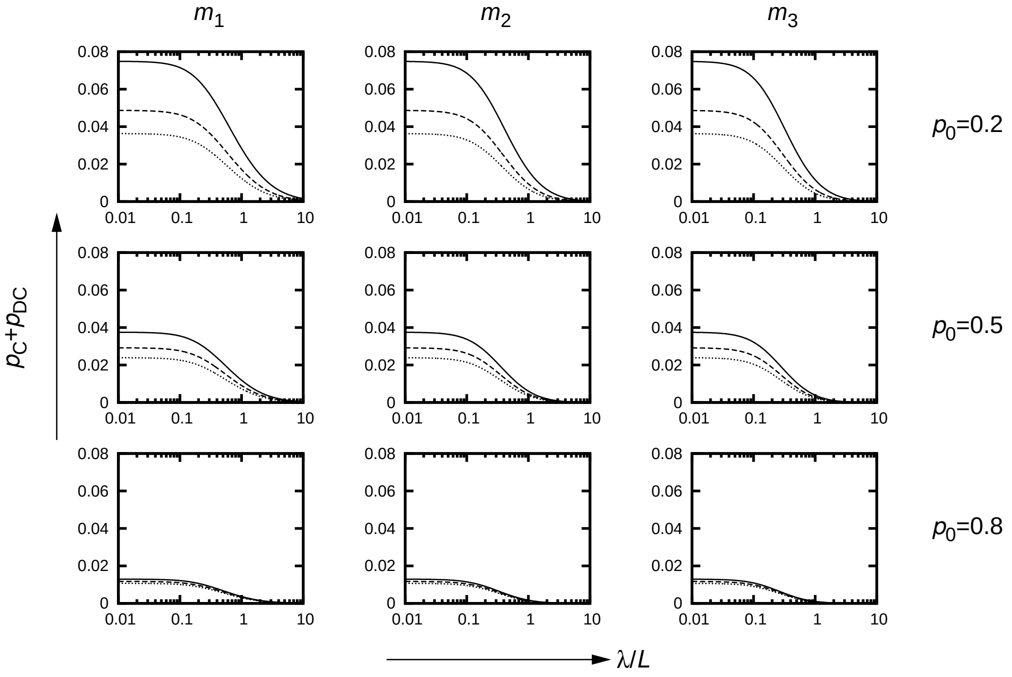

Figure 4. Coefficient $p_{\mathrm{C}}+p_{\mathrm{DC}}$ for the colonisation kernel c_{F}. Parameters and symbols are as for Fig. 2.

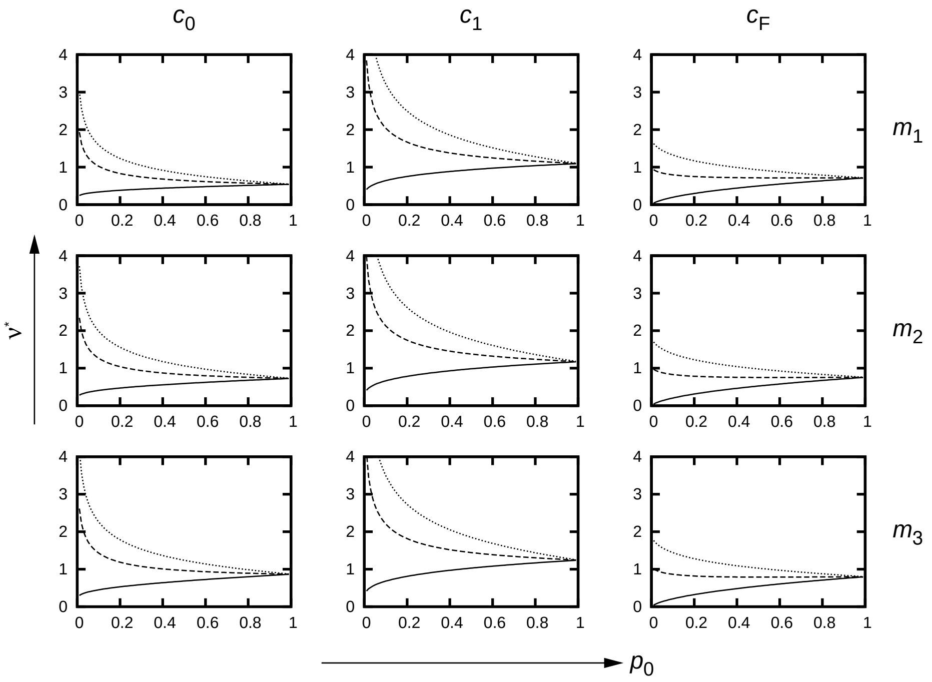

Figure 5. Plot of the threshold ν^*, denoting the value of ν where stochasticity and landscape correlations make an equal contribution to the metapopulation occupancy, as a function of p_0 for the three colonisation kernels c_0, c_1, c_F and three landscape correlation functions m_1, m_2, m_3. We have set $\lambda = \frac{L}{10}$. Solid line: static landscape $\beta = 0$, dashed line: $\beta = \mu$; dotted line: maximally dynamic landscape $\frac{\mu}{\beta} = 0$.

there would be an overall multiplier which can be obtained from Figs. 2, 3, and 4. This multiplier would depend on the colonisation kernel, with ν^* increasing if $\frac{\lambda}{L}$ is increased. Overall, it is interesting to note that correlations do not have to be particularly strong (ν about 3 or 4, or even less) to overcome the detrimental effect of stochasticity. As before, we find that the choice of landscape correlation function makes little difference to the results. Landscape correlations are relatively more important for static landscapes that for dynamic landscapes. For static landscapes, the relative importance of landscape correlations due to stochasticity decreases as p_0 is increased; for landscapes with high turnover, however, landscape correlations become relatively more important as p_0 is increased.

5.3 Anomalies due to short-range singularities

In the above discussions, we have tacitly assumed that the integrals in eqns. (19-22) are finite. It is straightforward to show that there are no singularities in the integrands at finite $\boldsymbol{y}$ provided $p_0 \neq 0$. However, the integrands are not necessarily integrable at large $\boldsymbol{y}$ for all choices of kernel.

The large-frequency behaviour of $\tilde{c}$ is determined by the short-distance behaviour in real space of the kernel $c(x)$. Suppose $c(x) \sim x^{\alpha}$ at small x, which implies $\tilde{c}(y) \sim y^{-2-\alpha}$ in 2 dimensions at large y. In order for $c(x)$ to be integrable (so that the total colonisation effort, $\tilde{c}(0)$, is finite), we must have $\alpha > -2$. Similarly, if the patch creation kernel behaves like $k(x) \sim x^{\theta}$ at small x, then we will have $\tilde{m}(y) = \tilde{k}^2(y) \sim y^{-4-2\theta}$ at large y. k must be integrable since it is a probability density, so $\theta > -2$.

Consider first the integrands in Eqns. (19) and (20), which determine p_{S} and p_{D}. In 2D, both of these behave like $y\phi^{-2} \sim y\tilde{c}^2 \sim y^{-3-2\alpha}$ at large y, which is not integrable when $-3-2\alpha \geq -1 \Rightarrow \alpha \leq -1$. Therefore, the perturbation expansion is ill-defined when $-2 < \alpha \leq -1$, even though the colonisation kernel is itself integrable.

The integrands that determine p_{C} and p_{DC} in Eqns. (21) and (22) behave as $y\tilde{m}(y)\tilde{c}(y) \sim y^{-5-\alpha-2\theta}$ at large y, so will not be integrable when $\theta < -2 - \frac{\alpha}{2}$. Thus, for the case $\alpha \geq -1$ where p_{S} and p_{D} are finite (so the perturbation expansion would be well-defined for an uncorrelated landscape), p_{C} and p_{DC} will be infinite for a correlated landscape when $-2 < \theta < -2 - \frac{\alpha}{2}$. This requires a combination of both sufficiently singular patch creation kernel and also a singular colonisation kernel with $\alpha < 0$. Note that the most singular choices in section 5.2, c_0 and k_1, have $\alpha = 0$ and $\theta = -1$, so are well inside the regime where the perturbation coefficients are finite.

5.4 Comparison with simulations

To verify the validity of Eqns. (C.1–C.4), we compare them with the results of numerical simulations on a finite landscape. We show results for the Bessel function kernel c_0 and the landscape correlation function m_1, as was the case in Ovaskainen and Cornell (2006a). The landscape dynamics described in Section 2 were implemented on a square of side 100 units, with parameter values that ensured that the mean patch density was unity, so that on average the landscape contained 10000 patches. We set mortality $\mu = 0$ so that all extinctions are caused by patch removal, i.e. the patch turnover dynamics is as fast as possible relative to colonization dynamics for the chosen value of p_0. For correlated landscapes, we chose correlation length $\lambda = 1$ and cluster sizes $\nu = 2$ and $\nu = 4$; these represent intermediate levels of correlation, the correction terms being larger when λ is small and ν is large. In order to mimic an infinite landscape and mitigate the effects of finite size effects, when calculating the colonization rate at a given focal point each other patch is mapped using periodic boundary conditions to its position within a square of side 100

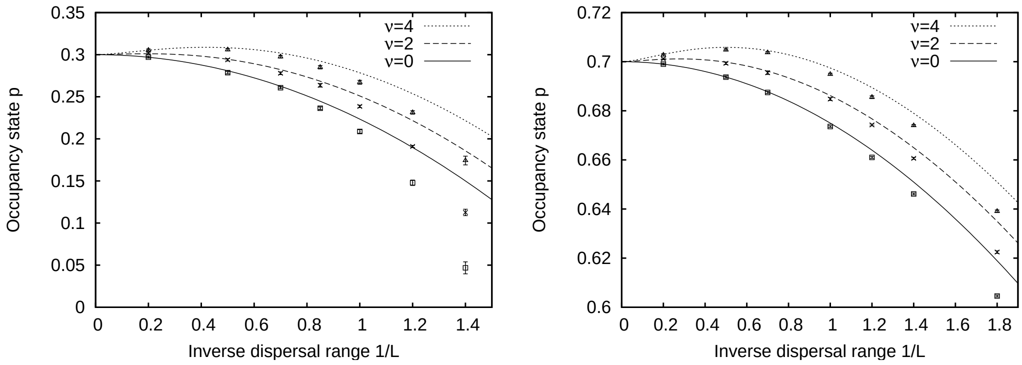

Figure 6. Equilibrium patch occupancy from simulations (points, average over 10 runs) and the second-order perturbation theory in Eqns. (C.1–C.4) (smooth lines), for a fully dynamic landscape ($\beta = 1$, $\mu = 0$). Colonization-extinction parameters corresponding to mean-field occupancy $p_0 = 0.3$ (left) and $p_0 = 0.7$ (right). Error bars represent 95% confidence intervals estimated from the variability between runs. The datasets are uncorrelated landscape (boxes, solid line); correlated landscape with $\lambda = 1$, $\nu = 2$ (crosses, dashed line); correlated landscape with $\lambda = 1$, $\nu = 4$ (triangles, dotted line).

centred on the focal point, with a correction term representing the infinite landscape outside the square calculated by integrating the colonization kernel as if the patch occupancy were equal to p_0. Patch extinction rate was set at $\beta = 1$, and the model was run for 300 time units with the first half discarded, which was long enough for equilibrium to be reached. Each point was averaged over 10 runs for $p_0 = 0.3$ and 20 runs for $p_0 = 0.7$.

The simulation results are shown in Figure 6. As expected, the perturbation theory captures the behaviour for large L, but becomes increasingly poor as L is decreased below about 1.4 for $p_0 = 0.3$ and about 1 for $p_0 = 0.7$. The discrepancy is also larger when the landscapes are more correlated, when the net correction to mean-field theory is actually smaller but contains competing terms due to dynamics and correlations. The departure from the perturbation results happens at a larger value of L than was the case for the static landscapes that were studied in Ovaskainen and Cornell (2006a). Nevertheless, the same qualitative result holds that L does not need to be particularly large for the perturbation theory to give reasonably good results. The reason for this is that the effective number of nearest neighbours is of order AL^2, with $A \approx 10$, as discussed in Ovaskainen and Cornell (2006a).

6 Discussion

We have calculated the correction to the mean patch occupancy for a metapopulation on a dynamic and correlated spatially-explicit landscape, relative to the mean-field limit where the colonisation length is infinite. Our method is based on a mathematical perturbation expansion, and is therefore exact in the limit where the colonisation length is large. The present manuscript generalises our earlier study of static landscapes, so that the results for the stochastic model in Ovaskainen and Cornell (2006a) are recovered as a special case. We have, however, adopted a complementary methodology, sacrificing some mathematical rigour for what we hope is a more intuitive explanation of our methods. Our results are explicit analytical expressions in terms of the parameters and kernels that define the model, so we have been able to discuss their properties in some generality.

We have assumed that habitat patches are generated in clusters by 'disturbances', and then degrade independently of each other. This is appropriate for a species whose habitat (such as short vegetation or dead trees) is ephemeral, and created by sporadic events (natural or unnatural). The spatial correlations in our patch landscapes are always positive, and we would need a quite different landscape model to generate landscapes with negative correlations (Johst and Drechsler, 2003). This means that we cannot obtain results for metapopulations on negatively correlated dynamic landscapes by simply setting the strength ν of the correlations to be negative, since the dynamics of the metapopulation depend on the spatiotemporal autocorrelations in the landscape, and not merely the spatial autocorrelation function M. By contrast, our earlier results for a static landscape are valid whether where the landscape spatial autocorrelation function is positive or negative (Ovaskainen and Cornell, 2006a).

Stochasticity in both metapopulation and landscape dynamics generate negative contributions to the metapopulation occupancy. These contributions depend on the variance of the colonisation kernel provided this is finite, but are otherwise somewhat insensitive to the shape of the kernel, which corroborates previous observations from simulations (Moilanen and Nieminen, 2003). If the variance is infinite the corrections remain finite but are smaller. When the patch occupancy is low, the corrections due to landscape stochasticity become much greater than those due to metapopulation stochasticity (compare the solid and dotted lines in Fig. 1). This is because small p_0 represents a poor colonisation ability on the part of the species, and hence an increased vulnerability to patches going extinct before colonisers have time to find new habitat.

The corrections due to landscape correlations depend on both the colonisation kernel and the landscape correlations. In practice the shape of the landscape correlation function is unimportant, however the corrections are sensitive to any (integrable) singularity at short distance in the colonisation kernel, reflecting an enhanced rate of re-colonisation of nearby patches in this case. The corrections are positive when the landscape is positively correlated (clusters of patches), and stronger if the landscape correlation length is shorter. The positive corrections due to landscape correlations tend to be weaker than the negative corrections due to stochasticity when the landscape is dynamic (see Fig. 5).

For uncorrelated landscapes, we find that the perturbation coefficient $p_{\mathrm{S}} + p_{\mathrm{D}}$ is negative, and as seen in Figure. 6 the metapopulation occupancy is greatest when $L = \infty$. However, landscape correlations make to a positive contribution to the perturbation coefficient, which depends on the ratio λ/L of the landscape correlation length to colonisation length. This leads to a peak in metapopulation occupancy at intermediate value of L for sufficiently correlated landscapes (see Figure. 6), which echoes the observation by Johst et al. (2002) that clustered habitat favours shorter dispersal distance. If the mean habitat quality in a landscape is too low to support a viable population of a species, then increasing landscape heterogeneity (keeping the mean quality fixed but increasing the relative difference between low- and high- quality areas) is beneficial for the species as it provides local expanses of habitat of sufficiently high quality. This benefit of heterogeneity, termed habitat association (Bolker, 2003), is counteracted by the cost of heterogeneity given by the loss of propagules to low quality regions. North and Ovaskainen (2007) used a first-order perturbation expansion to study population dynamics in landscapes with continuous spatio-temporal variation in habitat quality, and found that the trade-off between costs and benefits of habitat heterogeneity leads to a maximized population size when the scale of landscape correlation is comparable to the species dispersal distance.

We found that the perturbation expansion breaks down for certain combinations of colonisation kernels and patch creation kernels. When $C(\boldsymbol{x}) \sim |\boldsymbol{x}|^{\alpha}$ and $K(\boldsymbol{x}) \sim |\boldsymbol{x}|^{\theta}$ at short distance, the expression for the first-order perturbation coefficient is infinite when either (i) $-2 < \alpha \leq -1$, or (ii) $-1 < \alpha < 0$ and $-2 < \theta < -2 - \frac{\alpha}{2}$. Under these conditions, the kernels are still normalisable ($\int C(\boldsymbol{x}) d\boldsymbol{x}$ and $\int K(\boldsymbol{x}) d\boldsymbol{x}$ are finite) but there are strong short-range correlations in the patch occupancy, as can be seen from eqn. (17). It is not clear how the system behaves under these conditions; preliminary simulations on finite landscapes (not shown) suggest that the mean patch occupancy still approaches mean-field when $L \to \infty$, but the corrections are much larger than for better-behaved kernels, and they are no longer proportional to L^{-2}. Colonisation kernels arising from random walks have at worst a logarithmic singularity ($\alpha = 0$) in 2D (Bolker and Pacala, 1999; Ovaskainen and Cornell, 2003), so these anomalies are unlikely to be of biological importance.

Our perturbation technique is appropriate when the colonisation length L is less than the typical separation between patches, and this is indeed the case for some real metapopulations. Hanski and Thomas (1994) discuss metapopulation dynamics for three butterflies, concluding that the typical dispersal distance ($1/\tau$ in their notation) is of order 0.5–1 km, and provide maps of real metapopulations that show there to be several patches within a 1 km radius of most patches. However, further work would be needed to adapt our results (which were derived for a landscape of infinite extent) to a finite landscape, since the real metapopulations in Hanski and Thomas (1994) contain only 50–100 patches.

Although our expression for the patch occupancy is only exact in the limit where colonisation range is infinite, exact conclusions can nevertheless be drawn about the behaviour for a finite range of L values. From the perturbation expansion for the mean occupancy, $\overline{p} = p_0 + \frac{p_d^1}{Q_0 L^d} + O(L^{-2d} Q_0^{-2})$, we can identify the first-order perturbation coefficient as $p_d^1 = \frac{\partial \overline{p}}{\partial(\frac{1}{Q_0 L^d})}$. Thus, $\overline{p} - p_0$ will have the same sign as p_d^1 for a range of values of $\frac{1}{Q_0 L^d} > 0$. For instance, while the correction due to landscape stochasticity will not equal $\frac{p_{\mathrm{D}}}{Q_0 L^d}$ for any particular value of L, since $p_{\mathrm{D}} < 0$ we can be certain that the correction will be negative for all $L > L^*$, where L^* is some finite value. Since simulations agree reasonably well to the expansion truncating to first order when L equals the typical distance between neighboring patches, it seems that L^* is typically of order $Q_0^{-1/d}$, i.e. the typical dispersal distance is comparable in magnitude to the typical separation between patches. This means that our general discussion about the relative contributions from different sources to the corrections to mean-field will apply for a wide parameter range.

Nevertheless, there will undoubtedly be some phenomena that occur at small L which will not be revealed by our perturbation expansion. For instance, if the total amount of habitat is kept constant, then if clusters contain more patches then these clusters must be further apart and hence more prone to extinction. Fahrig (1992) observed that greater clustering of clumped habitat can be detrimental to a metapopulation, but our results predict that increasing ν (keeping other parameters constant) will always increases patch occupancy. Similarly, Wimberly (2006) found that habitat dynamics could benefit a metapopulation by providing ephemeral ‘stepping stones’ between isolated habitat, whereas to first order in perturbation theory we find that habitat dynamics is always detrimental. It seems likely that our perturbation theory may not be able to probe the limit of highly isolated clusters since it assumes that there are many potential patches within reach.

As in our previous paper (Ovaskainen and Cornell, 2006a), we made use of a SDE formalism to describe the dynamics of the metapopulation, though this time it was merely a convenient way to derive the moment equations. Xu et al. (2006) have derived an alternative SDE description of a dynamic-landscape metapopulation, but it is important to note that their model is fundamentally different to ours. Xu et al. assume that patches are sub-divided into many ‘micropatches’, and interpret the probability that a patch is occupied as the fraction of micropatches that are occupied. In the limit where there are many micropatches, the deviations in this probability from the deterministic, spatially-explicit Levins model are small. Consequently, the corresponding Ito SDE resembles the deterministic model with the addition of a small normally-distributed noise. By contrast, we have assumed that the patches are either vacant or occupied, and our SDE therefore has a shot noise which cannot be assumed small.

Hanski (1999) conjectured that in a dynamic landscape the threshold condition for persistence is obtained by replacing the amount of connected habitat by the amount of linked habitat, which he defined as the amount of connected habitat that an individual is expected to experience during its lifetime. Hanski (1999) was mainly concerned with the effect of habitat heterogeneity, as he assumed spatially correlated variation in patch quality, which is not included in the present model. However, we can use our model as a special case with respect to which Hanski's conjecture can be tested. In the case of a static landscape, Hanski assumed that all habitat is occupied, and defined the amount of connected habitat as the amount of habitat to which a randomly chosen individual has access. In the case of a dynamic landscape, he assumed again that all habitat is occupied, and defined the amount of linked habitat as the amount of habitat to which a randomly chosen individual has access during its lifetime. Using these definitions in the present modelling framework, we find that the amount of connected habitat in our static landscape is exactly the amount of linked habitat in our dynamic landscape, and thus according to Hanski's conjecture, the species should do equally well in the two cases. However, as shown by our results, this is not the case, but the species does worse in the dynamic landscape. The reason why Hanski's original conjecture does not hold here is that by not distinguishing between empty and occupied patches he relied on a mean-field model that ignores spatio-temporal correlations in patch occupancy. Indeed, the results of this study are based on deriving the amount of *empty* habitat that a randomly chosen individual has access to in its lifetime. Hence Hanski's conjecture is somewhat too optimistic, as it ignores that species track their changing habitats with a transient due to spatio-temporal correlations in patch occupancy.

Our results provide a firm mathematical foundation for understanding the contributions to metapopulation occupancy from a variety of mechanisms—stochasticity, landscape correlations, landscape dynamics. Our model omits several biological mechanisms which are thought to play an important role in metapopulation dynamics, such as variable patch quality (Thomas et al., 2001), spatially correlated extinctions (McCarthy and Lindmayer, 2000), and more complicated succession dynamics (Hastings, 2003; Ellner and Fussmann, 2003). However, our method is applicable for a general class of birth-death process (Ovaskainen and Cornell, 2006b), and could therefore be applied to such elaborations. While our perturbation method is only exact in the limit of long-range interactions, we have shown that it still affords a good approximation for intermediate interaction lengths, and it remains, to our knowledge, the only systematic method for exact computation of the properties of spatial, stochastic populations.

Acknowledgements

We thank Ilkka Hanski for stimulating discussions that gave the original impetus for this work. We thank Jenny Hodgson for constructive comments on the manuscript. OO thanks the Academy of Finland for funding (grants 213457 and 211173). SJC thanks the Metapopulation Research Group at the University of Helsinki for hospitality during part of this study.

A Stochastic differential equation for spatial birth-death processes

In this Appendix we show how to derive (in a non-rigorous way) a stochastic differential equation for a class of spatial point processes, starting from a mechanistic model. The spatial properties of a point process would be properly formulated using the theory of distributions as in Ovaskainen and Cornell (2006a), but we shall use a more cavalier notation treating functions and distributions on an even footing. Our aim is to show how SDEs can be obtained in an intuitive way from mechanistic models, rather than to provide the last word in mathematical rigour.

Suppose we have a stochastic spatial point process, i.e. one where pointlike objects appear and disappear in infinite, continuous d-dimensional space $\mathbb{R}^d$. We shall derive formulae that are valid for general d, though we are mostly interested in $d = 2$ and we shall, for example, use 'area' to mean d-dimensional volume. For convenience, let us label all points that ever existed or will ever exist by an index i, so that point i is created at position $\boldsymbol{X}_i$ at time $t_{\mathrm{C}i}$ and destroyed at time $t_{\mathrm{D}i} (> t_{\mathrm{C}i})$. Let $P(t) = \{i : t_{\mathrm{C}i} < t \leq t_{\mathrm{D}i}\}$ be the set of indices of the points in existence at time t. We define the *point density* $R(\boldsymbol{x}, t)$ as a quantity (strictly, a distribution) that has the property that the integral of $R(\boldsymbol{x}, t)$ over a region Ω gives the number of points within Ω at time t. We identify $R = \sum_{i \in P(t)} \delta(\boldsymbol{x} - \boldsymbol{X}_i)$, where δ is the Dirac delta distribution, so that R also has the property that $\int_\Omega F(\boldsymbol{x}) R(\boldsymbol{x}, t) dx = \sum_{i \in P(t): \boldsymbol{X}_i \in \Omega} F(\boldsymbol{X}_i)$ (the sum of the values of F at all the points existing at time t within region Ω) for any smooth function $F(\boldsymbol{x})$.

We assume that point creation and destruction events are Markov processes, i.e. the rates of these processes at a given time depend only on the set of points in existence at that time, but we allow for the possibility that these events have spatial correlations, for example allowing processes where several points are created or destroyed at once. Let $C(t_1, t_2) = \{i : t_1 < t_{Ci} \leq t_2\}$ and $D(t_1, t_2) = \{i : t_1 < t_{Di} \leq t_2\}$ denote the sets of indices of the points which are, respectively, created and destroyed in the time interval $(t_1, t_2]$. Then the point density at $t + dt$ satisfies the following equation:

$$R(\boldsymbol{x}, t + dt) = R(\boldsymbol{x}, t) + \sum_{i \in C(t,t+dt)} \delta(\boldsymbol{x} - \boldsymbol{X}_i) - \sum_{i \in D(t,t+dt)} \delta(\boldsymbol{x} - \boldsymbol{X}_i), \tag{A.1}$$

where $D(t, t + dt) \subset P(t)$ (i.e. only existing points can be destroyed).

We now introduce the *average over realizations.* We use the notation $\langle A \rangle_t$ as a shorthand for $E\left(A|\{R(\boldsymbol{x}, t), \boldsymbol{x} \in \mathbb{R}^d\}\right)$, i.e. the average of the quantity A over all possible stochastic evolutions starting with an initial state $R(\boldsymbol{x}, t)$ at time t, where the quantity A depends on the configuration at times later than t only. Then we can write Eqn. (A.1) as an SDE of the Ito type in the form

$$dR(\boldsymbol{x}, t) = (r_\mathrm{C} - r_\mathrm{D})dt + d\eta_\mathrm{C} - d\eta_\mathrm{D}, \tag{A.2}$$

where

$$\begin{aligned}
r_\mathrm{C}(\boldsymbol{x}, t)dt &= \left\langle \sum_{i \in C(t,t+dt)} \delta(\boldsymbol{x} - \boldsymbol{X}_i) \right\rangle_t \\
r_\mathrm{D}(\boldsymbol{x}, t)dt &= \left\langle \sum_{i \in D(t,t+dt)} \delta(\boldsymbol{x} - \boldsymbol{X}_i) \right\rangle_t \\
d\eta_\mathrm{C}(\boldsymbol{x}, t) &= \sum_{i \in C(t,t+dt)} \delta(\boldsymbol{x} - \boldsymbol{X}_i) - \left\langle \sum_{i \in C(t,t+dt)} \delta(\boldsymbol{x} - \boldsymbol{X}_i) \right\rangle_t \\
d\eta_\mathrm{D}(\boldsymbol{x}, t) &= \sum_{i \in D(t,t+dt)} \delta(\boldsymbol{x} - \boldsymbol{X}_i) - \left\langle \sum_{i \in D(t,t+dt)} \delta(\boldsymbol{x} - \boldsymbol{X}_i) \right\rangle_t ,
\end{aligned}$$

so that r_C and r_D are the instantaneous mean rates per unit area of point creation and destruction at time t, and $d\eta_\mathrm{C}$, $d\eta_\mathrm{D}$ the associated noise terms. Note that, by definition, $\langle d\eta_\mathrm{C} \rangle_t = \langle d\eta_\mathrm{D} \rangle_t = 0$. We shall now calculate some relevant properties of these rates and noises.

A.1 Creation

The quantity $r_\mathrm{C}(\boldsymbol{x}, t) = \frac{1}{dt} \left\langle \sum_{i \in C(t,t+dt)} \delta(\boldsymbol{x} - \boldsymbol{X}_i) \right\rangle_t$ represents the mean rate of point creation per unit area at time t, conditional on $R(\boldsymbol{x}, t)$. Note that r_C might be a smooth function of x, but could also be a distribution. For instance, if there were a discrete set of positions $S=\{\boldsymbol{Y}_j\}$ at which creation events take place respectively at rates $\{\zeta(\boldsymbol{Y}_j)\}$, then $\sum_{i\in C(t,t+dt)}\delta(\boldsymbol{x}-\boldsymbol{X}_i)=\sum_{\boldsymbol{Y}\in S}\xi(\boldsymbol{Y})\delta(\boldsymbol{x}-\boldsymbol{Y})$, where the $\xi(\boldsymbol{Y})$ are Bernouilli random variables with mean $\langle\xi(\boldsymbol{Y}_j)\rangle_t=\zeta(\boldsymbol{Y}_j)dt$, giving $r_{\mathrm{C}}(\boldsymbol{x},t)=\sum_{\boldsymbol{Y}\in S}\zeta(\boldsymbol{Y})\delta(\boldsymbol{x}-\boldsymbol{Y})$.

The covariance of the noise may be expressed as $\langle d\eta_{\mathrm{C}}(\boldsymbol{x},t)d\eta_{\mathrm{C}}(\boldsymbol{x}',t')\rangle_t$, where we have assumed without loss of generality that $t'\geq t$. Since the average of $d\eta_{\mathrm{C}}(\boldsymbol{x}',t')$ conditional on $R(\boldsymbol{x}',t')$ is zero, if $t\neq t'$ then $\langle d\eta_{\mathrm{C}}(\boldsymbol{x},t)d\eta_{\mathrm{C}}(\boldsymbol{x}',t')\rangle_t=0$ (more generally, the noise terms at two different times are statistically independent). For the case $t=t'$, we write (omitting in the notation the explicit dependence of C on t and dt)

$$\begin{aligned}\langle d\eta_{\mathrm{C}}(\boldsymbol{x},t)d\eta_{\mathrm{C}}(\boldsymbol{x}',t)\rangle_t=&\left\langle\sum_{i\in C}\sum_{i'\in C}\delta(\boldsymbol{x}-\boldsymbol{X}_i)\delta(\boldsymbol{x}'-\boldsymbol{X}_{i'})\right\rangle_t\\&-r_{\mathrm{C}}(\boldsymbol{x},t)r_{\mathrm{C}}(\boldsymbol{x}',t)(dt)^2\\=&\left\langle\sum_{i\in C}\delta(\boldsymbol{x}-\boldsymbol{X}_i)\delta(\boldsymbol{x}'-\boldsymbol{X}_i)\right\rangle_t\\&+\left\langle\sum_{i\in C}\sum_{i'\in C\setminus\{i\}}\delta(\boldsymbol{x}-\boldsymbol{X}_i)\delta(\boldsymbol{x}'-\boldsymbol{X}_{i'})\right\rangle_t\\&-r_{\mathrm{C}}(\boldsymbol{x},t)r_{\mathrm{C}}(\boldsymbol{x}',t)(dt)^2,\end{aligned}\tag{A.3}$$

where in the second term the index i' runs over all elements of C except i. The first term may be evaluated by noting that $\delta(\boldsymbol{x}-\boldsymbol{X}_i)\delta(\boldsymbol{x}'-\boldsymbol{X}_i)=\delta(\boldsymbol{x}-\boldsymbol{x}')\delta(\boldsymbol{x}-\boldsymbol{X}_i)$, so

$$\begin{aligned}\left\langle\sum_{i\in C}\delta(\boldsymbol{x}-\boldsymbol{X}_i)\delta(\boldsymbol{x}'-\boldsymbol{X}_{\boldsymbol{i}})\right\rangle_t&=\delta(\boldsymbol{x}-\boldsymbol{x}')\left\langle\sum_{i\in C}\delta(\boldsymbol{x}-\boldsymbol{X}_i)\right\rangle_t\\&=\delta(\boldsymbol{x}-\boldsymbol{x}')r_{\mathrm{C}}(\boldsymbol{x},t)dt.\end{aligned}$$

To evaluate the second term in (A.3), we define a quantity $r_{\mathrm{CC}}(\boldsymbol{x}_1,\boldsymbol{x}_2,t)$ as the probability per unit time per unit area-squared at which one point is created in the vicinity of position $\boldsymbol{x}_1$ simultaneously to the creation of another point in the vicinity of $\boldsymbol{x}_2$ (which will, like r_{C}, be a stochastic variable that depends on the current state $R(\boldsymbol{x},t)$). If point creation at different locations were statistically independent, the density of simultaneous point creations would be $r_{\mathrm{CC}}(\boldsymbol{x}_1,\boldsymbol{x}_2)dt=r_{\mathrm{C}}(\boldsymbol{x}_1)r_{\mathrm{C}}(\boldsymbol{x}_2)(dt)^2$, so r_{CC} would be infinitesimal. If, however, events can create more than one point simultaneously, then $r_{\mathrm{CC}}(\boldsymbol{x}_1,\boldsymbol{x}_2)$ will be finite. Since $r_{\mathrm{CC}}(\boldsymbol{x}_1,\boldsymbol{x}_2)dt$ is the probability density that points are created simultaneously at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in dt, we can use it to average any function of the position of two elements of $C(t,t+dt)$, e.g. $\left\langle\sum_{i\in C}\sum_{i'\in C\setminus\{i\}}G(\boldsymbol{X}_i,\boldsymbol{X}_{i'})\right\rangle_t\equiv\int\int r_{\mathrm{CC}}(\boldsymbol{X}_i,\boldsymbol{X}_{i'})dt\,G(\boldsymbol{X}_i,\boldsymbol{X}_{i'})d\boldsymbol{X}_id\boldsymbol{X}_{i'}$. The second term in Eqn. (A.3) is therefore

$$\left\langle \sum_{i\in C} \sum_{i'\in C\setminus\{i\}} \delta(\boldsymbol{x}-\boldsymbol{X}_i)\delta(\boldsymbol{x}'-\boldsymbol{X}_{i'}) \right\rangle_t =$$
$$\int\int r_{\mathrm{CC}}(\boldsymbol{X}_i,\boldsymbol{X}_{i'},t)dt\,\delta(\boldsymbol{x}-\boldsymbol{X}_i)\delta(\boldsymbol{x}'-\boldsymbol{X}_{i'})d\boldsymbol{X}_{i'}d\boldsymbol{X}_i$$
$$= r_{\mathrm{CC}}(\boldsymbol{x},\boldsymbol{x}',t)dt$$

(this is almost obvious because the term in the angled brackets represents the density of pair of points created simultaneously at x and x' in dt).

In the limit $dt \Longrightarrow 0$, terms of order $(dt)^2$ vanish, so have

$$\langle d\eta_{\mathrm{C}}(\boldsymbol{x},t)d\eta_{\mathrm{C}}(\boldsymbol{x}',t)\rangle_t = (r_{\mathrm{C}}(\boldsymbol{x},t)\delta(\boldsymbol{x}-\boldsymbol{x}') + r_{\mathrm{CC}}(\boldsymbol{x},\boldsymbol{x}',t))dt.$$

This result may be combined with the fact that the unequal-time covariance is zero in the following form:

$$\langle d\eta_{\mathrm{C}}(\boldsymbol{x},t)d\eta_{\mathrm{C}}(\boldsymbol{x}',t')\rangle_t = (r_{\mathrm{C}}(\boldsymbol{x},t)\delta(\boldsymbol{x}-\boldsymbol{x}') + r_{\mathrm{CC}}(\boldsymbol{x},\boldsymbol{x}',t))\,\delta(t-t')dt\,dt',$$

which may be obtained either by discretizing time then noting the limiting form for the Kronecker delta $\delta_{t,t'} \to \delta(t-t')dt'$, or more formally by integrating over t and t' using two general time intervals and noting the statistical independence of time creation events at different times.

A.2 Destruction

Point destruction events may only take place at positions where a point exists already, so the destruction rate r_{D} must be a distribution of the form $r_{\mathrm{D}} = \sum_{i\in P(t)} \theta_i \delta(\boldsymbol{x}-\boldsymbol{X}_i)$, where θ_i is the destruction rate of the i'th point. With this caveat, the derivation of the covariance of the noise follows along exactly the same lines as for creation, leading to

$$\langle d\eta_{\mathrm{D}}(\boldsymbol{x},t)d\eta_{\mathrm{D}}(\boldsymbol{x}',t')\rangle_t = (r_{\mathrm{D}}(\boldsymbol{x},t)\delta(\boldsymbol{x}-\boldsymbol{x}') + r_{\mathrm{DD}}(\boldsymbol{x},\boldsymbol{x}',t))\,\delta(t-t')dt\,dt',$$

where $r_{\mathrm{DD}}(\boldsymbol{x}_1,\boldsymbol{x}_2,t)$ is the probability per unit time of simultaneously destroying one patch in the vicinity of $\boldsymbol{x}_1$ and another patch in the vicinity of $\boldsymbol{x}_2$.

A.3 Ensemble average

We will often be interested in averaging over the initial condition as well as over stochastic evolutions from the initial state. We introduce a new notation

$\overline{A}$ to represent such an *ensemble average*, which is obtained by averaging the average over realizations $\langle A\rangle_{t_0}$, which depends upon $R(t_0)$, over a set of initial states whose properties may be specified statistically.

We may still use an SDE of the form (A.2) to describe the dynamics, and specify the relevant properties of the noise in terms of ensemble averages:

$$\overline{d\eta_{\mathrm{C}}(\boldsymbol{x},t)d\eta_{\mathrm{C}}(\boldsymbol{x}',t')} = \left(\delta(\boldsymbol{x}-\boldsymbol{x}')\overline{r_{\mathrm{C}}(\boldsymbol{x},t)} + \overline{r_{\mathrm{CC}}(\boldsymbol{x},\boldsymbol{x}',t)}\right)\delta(t-t')dt\,dt' \tag{A.4}$$

$$\overline{d\eta_{\mathrm{D}}(\boldsymbol{x},t)d\eta_{\mathrm{D}}(\boldsymbol{x}',t')} = \left(\delta(\boldsymbol{x}-\boldsymbol{x}')\overline{r_{\mathrm{D}}(\boldsymbol{x},t)} + \overline{r_{\mathrm{DD}}(\boldsymbol{x},\boldsymbol{x}',t)}\right)\delta(t-t')dt\,dt' \tag{A.5}$$

$$\overline{d\eta_{\mathrm{C}}(\boldsymbol{x},t)d\eta_{\mathrm{D}}(\boldsymbol{x}',t')} = 0, \tag{A.6}$$

where (A.6) holds if creation and destruction events are statistically independent (conditional on the current state). Note that there will also be biologically relevant models where this is not the case, for instance point motion may be modelled by simultaneous creation and destruction events at different locations, in which case there would be a term $[-r_{\mathrm{CD}}(\boldsymbol{x},\boldsymbol{x}',t)]$ on the right hand side of (A.6) representing the rate at which a point is destroyed at $\boldsymbol{x}'$ and created at $\boldsymbol{x}$. Note also that it is straightforward to generalise this approach to the case of two or more classes of points, for instance ‘occupied’ and ‘unoccupied’ patches as discussed in this paper, or where points represent individuals of different species.

If the dynamical processes only depend on relative rather than absolute point positions, then an ensemble average over a translationally independent set of initial states will give simple forms for the terms on the right hand side of Eqns. (A.4–A.6). The particular ensemble average that will interest us is a stationary state obtained from evolving the dynamics over a very long time. For a self-averaging system, the average will in this case not depend upon the initial condition, and we have $\overline{A} = \lim_{t_0\to-\infty}\langle A\rangle_{t_0}$. Note, however, that Eqns. (A.4–A.6) are quite general, and can for instance be used when away from equilibrium, or when starting from a single initial condition (when the ‘ensemble’ consists of a single state).

Eqn. (A.2) is not a conventional SDE because the noise terms are not Gaussian. Technically, they are ‘shot’ noise (Gardiner, 2004), so higher order correlation functions of the noise terms do not simply factorise into combinations of two-point functions (Ovaskainen and Cornell, 2006b). In this paper, we make use of second order correlations only, but higher order correlations could be calculated using methods similar to sections A.1 and A.2 if needed.

B Deriving moment equations from SDEs

Our motivation in studying SDEs is to obtain equations for ensemble averages. Suppose we have several species i with density q_i, satisfying a SDE

$$dq_i(\boldsymbol{x},t) = r_i dt + d\eta_i(\boldsymbol{x},t),$$

so that r_i is the rate of change in density (which will in general depend on the density of this and other species, at this and other points in space), and $d\eta_i$ is a shot noise with equal-time covariance

$$E(d\eta_i(\boldsymbol{x}_i,t)d\eta_j(\boldsymbol{x}_j,t)|\{q_k(\boldsymbol{x},t)\}) = V_{i,j}(\boldsymbol{x}_i,\boldsymbol{x}_j;\{q_k(\boldsymbol{x},t)\})dt$$

(note that the covariance is conditioned on the densities of all the species at time t). The equation of motion for the ensemble average density is

$$\frac{d\overline{q_i}}{dt} = \overline{r_i},$$

since $d\overline{\eta_i} = 0$ by definition. To obtain the equation for a second moment we need to use the Ito calculus:

$$\begin{aligned} d(q_iq_j) &= (q_i + dq_i)(q_j + dq_j) - q_iq_j \\ &= (q_i + r_idt + d\eta_i)(q_j + r_jdt + d\eta_j) - q_iq_j \\ &= q_i(r_jdt + d\eta_j) + (r_idt + d\eta_i)q_j + (r_idt + d\eta_i)(r_jdt + d\eta_j). \end{aligned}$$

We next note that, since the expectation of the noise at time t is zero, conditional on the densities at time t, we must have $E(q_jd\eta_i) = 0$, and similarly $E(r_jd\eta_i) = 0$ since r_j is just a function of the densities at time t. We consider here equal-time moments $(\overline{q_i(\boldsymbol{x}_i,t)q_i(\boldsymbol{x}_j,t)})$, but the procedure can be extended to unequal time moments $(\overline{q_i(\boldsymbol{x}_i,t)q_i(\boldsymbol{x}_j,t')})$. Taking Ensemble averages, we have

$$d\overline{q_iq_j} = \overline{q_ir_j}dt + \overline{r_iq_j}dt + \overline{d\eta_id\eta_j} + O(dt^2).$$

The differential equation for the second moment is then

$$\frac{d\overline{q_iq_j}}{dt} = \overline{q_ir_j} + \overline{r_iq_j} + \overline{V_{ij}(\boldsymbol{x}_i,\boldsymbol{x}_j;\{q_k(\boldsymbol{x},t)\})}. \tag{B.1}$$

Equations for higher order moments require expressions for higher-order moments of the noise such as $\overline{d\eta_id\eta_jd\eta_k}$ etc. (Ovaskainen and Cornell, 2006b).

B.1 Central second moment

Let us define $s_i = q_i - \overline{q_i}$, which represents the stochastic fluctuations in species i. By definition $\overline{s_i} = 0$, and ensemble averages of products of s_i are referred to as *central moments*. The differential equation for the second central moment (i.e. covariance) can be derived from Eqn. (B.1) as follows:

$$
\begin{aligned}
\frac{d\overline{s_i s_j}}{dt} &= \frac{d\overline{q_i q_j}}{dt} - \frac{d}{dt}(\overline{q_i}\,\overline{q_j}) \\
&= \overline{q_i r_j} + \overline{r_i q_j} + \overline{V_{ij}(x_i, x_j)} - (\frac{d}{dt}\overline{q_i})\overline{q_j} - \overline{q_i}\frac{d}{dt}(\overline{q_j}) \\
&= \overline{s_i r_j} + \overline{r_i s_j} + \overline{V_{ij}(x_i, x_j)}, \qquad \text{(B.2)}
\end{aligned}
$$

where in the last line we have used the fact that $\overline{q_i r_j} = \overline{(\overline{q_i} + s_i) r_j} = \overline{q_i r_j} + \overline{s_i r_j}$, and that $d\overline{q_i}/dt = \overline{r_i}$.

C Perturbation coefficients for specific 2D kernels

C.1 Zeroth-order Bessel function c_0

When c is given by c_0 from Eqn. (23), and when the landscape correlation function is given by $m = m_n$ for any of the alternatives $n =$ 1,2, 3 (see Eqns. (27–29)), Equations (19–22) can be integrated to give

$$
\begin{aligned}
p_{\mathrm{S}} &= -\frac{(1-p_0)}{4\pi} \qquad \text{(C.1)} \\
p_{\mathrm{S}} + p_{\mathrm{D}} &= -\frac{1}{4\pi}\left(1 + \frac{p_0}{\gamma}\right)\log\left(\frac{1+\frac{\gamma}{p_0}}{1+\gamma}\right) \qquad \text{(C.2)} \\
p_{\mathrm{C}} &= \begin{cases} \frac{\nu(1-p_0)(\kappa-1-\log\kappa)}{8\pi(\kappa-1)^2} & \text{for } n = 1 \\ \frac{\nu(1-p_0)[(1+\kappa)\log\kappa+2(1-\kappa)]}{8\pi(\kappa-1)^3} & \text{for } n = 2 \\ \frac{\nu(1-p_0)[(\kappa+5)(\kappa-1)-2(1+2\kappa)\log\kappa]}{16\pi(\kappa-1)^4} & \text{for } n = 3 \end{cases} \qquad \text{(C.3)} \\
p_{\mathrm{C}} + p_{\mathrm{DC}} &= \frac{\nu p_0}{8\pi\gamma}[\rho A_n(\rho\kappa) - A_n(\kappa)], \qquad \text{(C.4)}
\end{aligned}
$$

where

$$\begin{aligned}
\kappa &= \frac{p_0\lambda^2}{L^2} \\
\rho &= \frac{1+\frac{\gamma}{p_0}}{1+\gamma} \\
A_n(x) &= \begin{cases} \frac{\log x}{x-1} & \text{for } n=1 \\ \frac{x-1-\log x}{(x-1)^2} & \text{for } n=2 \\ \frac{(x-1)(x-3)+2\log x}{2(x-1)^3} & \text{for } n=3. \end{cases}
\end{aligned}$$

C.2 First-order Bessel function c_1

When c is given by c_1 from Eqn. (24), Equations (19–22) can be integrated to give

$$\begin{aligned}
p_{\mathrm{S}} &= -\frac{2-\frac{p_0}{(1-p_0)^{1/2}}\log\frac{1+(1-p_0)^{1/2}}{1-(1-p_0)^{1/2}}}{8\pi} \\
p_{\mathrm{S}}+p_{\mathrm{D}} &= -\frac{(p_0+\gamma)[\log\frac{1+(1-p_0)^{1/2}}{1-(1-p_0)^{1/2}}-(1+\gamma)^{1/2}\log\frac{(1+\gamma)^{1/2}+(1-p_0)^{1/2}}{(1+\gamma)^{1/2}-(1-p_0)^{1/2}}]}{4\pi(1-p_0)^{1/2}\gamma} \\
p_{\mathrm{C}}+p_{\mathrm{DC}} &= \frac{\nu}{8\pi\gamma}[\frac{\gamma+p_0}{1+\gamma}B_n(\lambda/L,\frac{1-p_0}{1+\gamma})-p_0B_n(\lambda/L,1-p_0)], \\
\text{where } B_n(l,a) &= \frac{1}{a}[A_n(2l^2(1-a^{1/2}))-A_n(2l^2(1+a^{1/2}))].
\end{aligned}$$

The coefficient p_{C} can most conveniently be obtained by taking the limit $\gamma \to 0$ of the expression for $p_{\mathrm{C}}+p_{\mathrm{DC}}$:

$$\begin{aligned}
p_{\mathrm{C}} &= \lim_{\gamma\to 0}\frac{\nu}{8\pi\gamma}[\frac{\gamma+p_0}{1+\gamma}B_n(\lambda/L,\frac{1-p_0}{1+\gamma})-p_0B_n(\lambda/L,1-p_0)] \\
&= \frac{\nu(1-p_0)}{8\pi}H_n(\lambda/L,1-p_0), \\
\text{where } H_n(l,a) &= B_n(l,a)-(1-a^2)\frac{\partial B_n(l,a)}{\partial a}.
\end{aligned}$$

C.3 Gaussian c_G

When c is given by c_{G} from Eqn. (25), Equations (19–20) can be integrated to give

$$
\begin{aligned}
p_{\mathrm{S}} &= -\frac{[1-p_0+p_0\log p_0]}{4\pi(1-p_0)} \\
p_{\mathrm{S}}+p_{\mathrm{D}} &= -\frac{(p_0+\gamma)}{4\pi}\frac{[(1+\gamma)\log\frac{p_0+\gamma}{1+\gamma}-\log p_0]}{\gamma(1-p_0)}.
\end{aligned}
$$

C.4 Fat-tailed c_F

When c is given by c_{F} from Eqn. (26), Equations (19–20) can be integrated to give

$$
\begin{aligned}
p_{\mathrm{S}} &= -\frac{p_0}{2\pi}[-\frac{\log(p_0)}{1-p_0}-D(1-p_0)] \\
p_{\mathrm{S}}+p_{\mathrm{D}} &= -\frac{(p_0+\gamma)[D(1-p_0)-D(\frac{1-p_0}{1+\gamma})]}{2\pi\gamma}, \\
\text{where } D(a) &= \frac{\frac{\pi^2}{6}-\mathrm{Li}_2(a)-\log(a)\log(1-a)}{a}
\end{aligned}
$$

and $\mathrm{Li}_2(x) \equiv \int_1^x \frac{\log(t)}{1-t}dt = \sum_{k=1}(-1)^k\frac{x^k}{k^2}$ is the dilogarithm function.

D Analysis of p_{S} and p_{D} when $p_0 \to 0$

D.1 Kernels with finite variance

The small frequency behaviour of $\phi(y)$ is related to the variance or other long-range behaviour of $C(x)$. For circularly symmetric kernels in 2D, we have

$$
\begin{aligned}
\tilde{C}(\boldsymbol{\omega}) &= \int e^{2\pi i\boldsymbol{x}\cdot\boldsymbol{\omega}}C(\boldsymbol{x})d\boldsymbol{x} \\
&= \int J_0(2\pi\omega x)C(x)2\pi x\,dx,
\end{aligned}
$$

where J_0 is a Bessel function of the first kind. For small argument, we have $J_0(y) = 1-\frac{1}{4}y^2+O(y^4)$, so

$$
\begin{aligned}
\tilde{C}(\omega) &= \int (1 - \frac{(2\pi\omega x)^2}{4}) C(x) 2\pi x \, dx + O(\omega^4) \\
&= \int C(x) 2\pi x \, dx - \frac{(2\pi\omega)^2}{4} \int x^2 C(x) 2\pi x \, dx + O(\omega^4) \\
&= \int C(x) 2\pi x \, dx [1 - \pi^2 \omega^2 \mathrm{Var}(C) + O(\omega^4)],
\end{aligned}
$$

where we have defined the variance of C as

$$
\mathrm{Var}(C) = \frac{\int x^2 C(x) 2\pi x \, dx}{\int C(x) 2\pi x \, dx}.
$$

Consider first the static landscape case, $\beta = 0$. The integral in Eqn. (19) in the limit $p_0 \to 0$ is dominated by the small-frequency behaviour of ϕ which, since $\phi = \tilde{c}(0)/\tilde{c}(y)$, is

$$
\phi(y) = 1 + \pi^2 y^2 V + \dots,
$$

where $V = \frac{\int x^2 c(x) 2\pi x \, dx}{\int c(x) 2\pi x dx}$ is the variance of the kernel c. When $p_0 = 0$, the integrand in Eqn. (19) has a non-integrable singularity at small y, but the coefficient can still be calculated in the limit $p_0 \to 0$:

$$
\begin{aligned}
p_{\mathrm{S}} &\to - \lim_{p_0 \to 0} p_0 \int \frac{1}{[\pi^2 V y^2 + p_0]^2} 2\pi y \, dy \\
&= -\frac{1}{\pi V}.
\end{aligned}
$$

For the three exponentially bounded kernels, we have $V = 4$ and hence $p_{\mathrm{S}} = -\frac{1}{4\pi} = -0.0795\dots$.

Considering now the dynamic case $\gamma \neq 0$, adding eqns. (19) and (20) gives

$$
p_{\mathrm{S}} + p_{\mathrm{D}} = -(p_0 + \gamma)(1 - p_0) \int \frac{1}{[\phi(|\boldsymbol{y}|)(1+\gamma) - (1-p_0)][\phi(y|\boldsymbol{y}|) - (1-p_0)]} d\boldsymbol{y}. \tag{D.1}
$$

The integral is again singular when $p_0 = 0$, but the asymptotic small p_0 limit requires some care because the factor $[\phi(|\boldsymbol{y}|)(1+\gamma) - (1-p_0)]^{-1}$ (which is not singular for $y \to 0$) needs to be included to ensure convergence of the integral at large y. We proceed by adding and subtracting terms with the same large-y behaviour.

$$\begin{aligned} p_{\mathrm{S}} + p_{\mathrm{D}} =& -(1-p_0)\int \frac{1}{[\phi(|\boldsymbol{y}|)-(1-p_0)]}[\frac{(p_0+\gamma)}{\phi(|\boldsymbol{y}|)(1+\gamma)-(1-p_0)} \\ & +\frac{1}{1+|\boldsymbol{y}|^2} - \frac{1}{1+|\boldsymbol{y}|^2}]d\boldsymbol{y} \\ =& -(1-p_0)\int \frac{1}{[\phi(|\boldsymbol{y}|)-(1-p_0)]}[\\ & \frac{(p_0+\gamma)(1+|\boldsymbol{y}|^2)-\phi(|\boldsymbol{y}|)(1+\gamma)+(1-p_0)}{[\phi(|\boldsymbol{y}|)(1+\gamma)-(1-p_0)][1+|\boldsymbol{y}|^2]} + \frac{1}{1+|\boldsymbol{y}|^2}]d\boldsymbol{y}. \end{aligned}$$

In the limit $p_0 \to 0$, the first term becomes

$$-\int \frac{1}{[\phi(|\boldsymbol{y}|)-1]}[\frac{1+\gamma(1+|\boldsymbol{y}|^2)-\phi(y|\boldsymbol{y}|)(1+\gamma)}{[\phi(|\boldsymbol{y}|)(1+\gamma)-1][1+|\boldsymbol{y}|^2]}]d\boldsymbol{y},$$

which is finite because the integrand is finite at small $|\boldsymbol{y}|$, and integrable at large y provided $\frac{1}{\phi^2}$ and $\frac{1}{\phi|\boldsymbol{y}|^2}$ are integrable [which are necessary conditions for p_{S} to be finite]. The second term, meanwhile, becomes

$$\begin{aligned} & -(1-p_0)\int \frac{1}{[\phi(|\boldsymbol{y}|)-(1-p_0)][1+|\boldsymbol{y}|^2]}d\boldsymbol{y} \\ \to_{(p_0\to 0)} & -\int \frac{1}{[\pi^2 V y^2 + p_0][1+y^2]} 2\pi y\, dy \\ = \quad & -\frac{\log\frac{1}{p_0}}{\pi V} + \text{const.} \end{aligned}$$

D.2 *Kernels with infinite variance*

For a kernel to have infinite variance in 2D, it must decay like a power law at large argument, $C(x) \to Ax^{-z}$ with $z < 4$. Meanwhile, we need $\tilde{C}(0)$ to be finite, which means $z < 2$. In the range $2 < z < 4$, we have

$$\begin{aligned} \tilde{C}(\omega) =& \int J_0(2\pi\omega x) C(x) 2\pi x\, dx \\ =& 1 - \omega^{z-2}\int([1-J_0(2\pi x)]x^{-z}2\pi x\, dx \\ & -\int([1-J_0(2\pi\omega x)][C(x)-x^{-z}]2\pi x\, dx. \end{aligned} \tag{D.2}$$

The coefficient $\int([1-J_0(2\pi x)]x^{-z}2\pi x\, dx$ is of order 1 when $2 < z < 4$, whereas the final term on the RHS of Eqn. (D.2) can be shown to be of higher order [the proof is somewhat tedious, necessitating splitting the domain of integration at a point $x^* = \omega^\alpha$ where $0 < \alpha < 1$, and the assumption $C(x) - x^{-z} \sim x^{-\epsilon}$ where $\epsilon > 0$]. This means that, as $y \to 0$,

$$\phi(y) = 1 + Ay^{z-2} + \ldots, \tag{D.3}$$

where A is a constant.

Considering first the static landscape case $\gamma = 0$. When $z < 3$, the integral in (D.3) is finite when $p_0 = 0$ and the small-y behaviour of ϕ is given by (D.3), which means that $\lim_{p_0 \to 0} \frac{p_{\mathrm{S}}}{p_0}$ is a constant. When $z > 3$, the integral in Eqn. (19) when $p_0 = 0$ is infinite, but the limit $p_0 \to 0$ Eqn. (19) can still be taken:

$$\begin{aligned} p_{\mathrm{S}} &\to - \lim_{p_0 \to 0} p_0(1 - p_0) \int \frac{2\pi x \, dx}{[x^{z-2} + p_0]^2} \\ &= -2\pi p_0^{\frac{4-z}{z-2}} \int \frac{2\pi \xi \, d\xi}{[\xi^{z-2} + 1]^2}, \end{aligned}$$

which approaches zero when $p_0 \to 0$, for $3 < z < 4$. In summary, we have $\lim_{p_0 \to 0} p_{\mathrm{S}} = 0$, and more specifically,

$$p_{\mathrm{S}} \sim p_0^{\min(1, \frac{4-z}{z-2})}.$$

For the dynamic landscape case $\gamma \neq 0$, the total coefficient $p_{\mathrm{S}} + p_{\mathrm{D}}$ at $p_0 = 0$, given by Eqn. (D.1), is

$$\lim_{p_0 \to 0} p_{\mathrm{S}} + p_{\mathrm{D}} = -2\pi\gamma \int \frac{y \, dy}{[\phi(y)(1 + \gamma) - 1][\phi(y) - 1]}.$$

When ϕ behaves like Eqn. (D.3) at small argument, and $z < 4$, this term is finite and nonzero.

References

Abramowitz, M., Stegun, I. A., 1964. Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables, ninth dover printing, tenth gpo printing Edition. Dover, New York.
URL `http://www.math.sfu.ca/ cbm/aands/`

Akcakaya, H. R., Radeloff, V., Mlandenoff, D. J., He, H. S., 2004. Integrating landscape and metapopulation modeling approaches: Viability of the sharp-tailed grouse in a dynamic landscape. Conservation Biology 18, 526–537.

Biedermann, R., 2004. Modelling the spatial dynamics and persistence of the leaf beetle *gonioctena olivacea* in dynamic habitats. Oikos 107, 645–653.

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

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

Chave, J., Leigh, E. G., 2002. A spatially explicit neutral model of β-diversity in tropical forests. Theoretical Population Biology 62, 153–168.

DeWoody, Y. D., Feng, Z., Swihart, R. K., 2005. Merging spatial and temporal structure within a metapopulation model. American Naturalist 166, 42–55.

Ellner, S. P., Fussmann, G., 2003. Effects of successional dynamics on metapopulation persistence. Ecology 84 (4), 882–889.

Fahrig, L., 1992. Relative importance of spatial and temporal scales in a patchy environment. Theoretical Population Biology 41, 300–314.

Gardiner, C. W., 2004. Handbook of Stochastic Methods, 3rd Edition. Springer.

Gu, W. D., Heikkilä, R., Hanski, I., 2002. Estimating the consequences of habitat fragmentation on extinction risk in dynamic landscapes. Landscape Ecology 17 (699-710).

Hanski, I., 1999. Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes. Oikos 87, 209–219.

Hanski, I., Ovaskainen, O., 2000. The metapopulation capacity of a fragmented landscape. Nature 404, 755–758.

Hanski, I., Thomas, C. D., 1994. Metapopulation dynamics and conservation: a spatially explicit model applied to butterflies. Biological Conservation 68, 167–180.

Hastings, A., 2003. Metapopulation persistence with age-dependent disturbance or succession. Science 301, 1525–1526.

Johst, K., Brandl, R., Eber, S., 2002. Metapopulation persistence in dynamic landscapes: the role of dispersal distance. Oikos 98, 263–270.

Johst, K., Drechsler, M., 2003. Are spatially correlated or uncorrelated disturbance regimes better for the survival of species? Oikos 103, 449–456.

Keymer, J. E., Marquet, P. A., Velasco-Hernández, J. X., Levin, S. A., 2000. Extinction thresholds and metapopulation persistence in dynamic landscapes. The American Naturalist 156, 478–494.

Lobel, S., Snall, T., Rydin, H., 2006. Species richness patterns and metapopulation processes - evidence from epiphyte communities in boreo-nemoral forests. Ecography 29, 169–182.

McCarthy, M. A., Lindmayer, D. B., 2000. Spatially-correlated extinction in a metapopulation model of leadbeater's possum. Biodiversity and Conservation 9, 57–63.

Moilanen, A., Nieminen, M., 2003. Simple connectivity measures in spatial ecology. Ecology 83, 1131–1145.

Moloney, K. A., Levin, S. A., 1996. The effects of disturbance architecture on landscape-level population dynamics. Ecology 77, 375–394.

Nathan, R., 2006. Long-distance dispersal of plants. Science 313, 786–788.

North, A., Ovaskainen, O., 2007. Interactions between dispersal, competition, and landscape heterogeneity. Oikos 116, 1106–1119.

Ovaskainen, O., Cornell, S. J., 2003. Biased movement at a boundary and conditional occupancy times for diffusion processes. Journal of Applied Probability 40, 557–580.

Ovaskainen, O., Cornell, S. J., 2006a. Asymptotically exact analysis of stochastic metapopulation dynamics with explicit spatial structure. Theoretical Population Biology 69, 13–33.

Ovaskainen, O., Cornell, S. J., 2006b. Space and stochasticity in population dynamics. Proceedings of the National Academy of Sciences USA 103, 12781–12786.

Paradis, E., Baillie, S. R., Sutherland, W. J., 2002. Modeling large-scale dispersal distances. Ecological Modelling 151, 279–292.

Pennanen, J., Greene, D. F., Fortin, M. J., Messier, C., 2004. Spatially explicit simulation of long-term boreal forest landscape dynamics: incorporating quantitative stand attributes. Ecological Modelling 180, 195–209.

Ross, J. V., 2006a. A stochastic metapopulation model accounting for habitat dynamics. Journal of Mathematical Biology 52, 788–806.

Ross, J. V., 2006b. Stochastic models for mainland-island metapopulations in static and dynamic landscapes. Bulletin of mathematical biology 68, 417–449.

Schroeder, L. M., Ranius, T., Ekbom, B., Larsson, S., 2007. Spatial occurrence of a habitat-tracking saproxylic beetle inhabiting a managed forest landscape. Ecological applications 17, 900–909.

Snäll, T., Riberiro, P. J., Rydin, H., 2003. Spatial occurrence and colonisations in patch-tracking metapopulations: local conditions versus dispersal. Oikos 103, 566–578.

Stoyan, D., Penttinen, A., 2000. Recent applications of point process methods in forestry statistics. Statistical Science 15, 61–78.

Thomas, C. S., Harrison, S., 1992. Spatial dynamics of a patchily distributed butterfly species. Journal of Animal Ecology 61, 437–446.

Thomas, J. A., Bourn, N. A. D., Clarke, R. T., Stewart, K. E., Simcox, D. J., Pearman, G. S., Curtis, R., Goodger, B., 2001. The quality and isolation of habitat patches both determine where butterflies persist in fragmented landscapes. Proceedings of the Royal Society B 268, 1791–1796.

Verheyen, K., Vellend, M., Van Calster, H., Peterken, G., Hermy, M., 2004. Metapopulation dynamics in changing landscapes: A new spatially realistic model for forest plants. Ecology 85, 3302–3312.

Vuilleumier, S., Wilcox, C., Cairns, B. J., Possingham, H. P., 2007. How patch configuration affects the impact of disturbances on metapopulation persistence. Theoretical Population Biology 72, 77–85.

Wahlberg, N., Klemetti, T., Hanski, I., 2002. Dynamic populations in a dynamic landscape: the metapopulation structure of the marsh fritillary butterfly. Ecography 25, 224–232.

Wimberly, M. C., 2002. Spatial simulation of historical landscape patterns in coastal forests of the pacific northwest. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 32, 1316–1328.

Wimberly, M. C., 2006. Species dynamics in disturbed landscapes: When does a shifting habitat mosaic enhance connectivity? Landscape Ecology 21, 35–46.

Wintle, B. A., Bekessy, S. A., Venier, L. A., Pearce, J. L., Chisholm, R. A., 2005. Utility of dynamic-landscape metapopulation models for sustainable forest management. Conservation Biology 19, 1930–1943.

Xu, D., Feng, Z., Allen, L. J. S., Swihart, R. K., 2006. A spatially structured metapopulation model with patch dynamics. Journal of Theoretical Biology 239, 469–481.