MODELLING THE DEMOGRAPHY OF PLANTS, PATCHES, AND METACOMMUNITIES

DANIEL S. FALSTER†, RICHARD G. FITZJOHN, ÅKE BRÄNNSTRÖM, ULF DIECKMANN, AND MARK WESTOBY

†Department of Biological Sciences, Macquarie University, Sydney, Australia

daniel.falster@mq.edu.au, rich.fitzjohn@gmail.com

CONTENTS

1   INTRODUCTION

This document outlines methods used to model demography in the PLANT package. We first outline the system of equations being solved. These equations were primarily developed elsewhere, in particular in Falster *et al.* (2011) and Falster *et al.* (2015), following general principles laid out in De Roos, Taljapurkar & Caswell (1997), Kohyama (1993), and Moorcroft, Hurtt & Pacala (2001). They are presented here so that users can understand the full system of equations being solved within PLANT. We then describe the numerical techniques used to solve the equations. Table 1 provides a list of names and definitions used throughout this document.

2   METACOMMUNITY DYNAMICS

The following material is written assuming there exists a physiological sub-model that takes as inputs a plant's traits x, height H, and light environment E_a, and returns rates of growth, mortality, and fecundity. The physiological model must describe how much leaf area the plant contributes to shading other plants in the patch. Specifically, the physiological model is responsible for calculating the following variables from Table 1: $A_l(H)$, $Q(z, H)$, $g(x, H, E_a)$, $f(x, H, E_a)$, $d(x, H, E_a)$, and $S_G(x, H_0, E_{a0})$. All are other variables in Table 1 are calculated in the demographic model using the equations provided below.

2.1   *Individual plants*

We first consider the dynamics of an individual plant. Throughout, we refer to a plant as having traits x and size H, with the latter given by its height. The plant grows in a light environment E, a function describing the distribution of light with respect to height. Ultimately, E depends on the composition of plants in a patch, and thus on the patch age a. To indicate this dependence, we write E_a. The functions $g(x, H, E_a)$, $f(x, H, E_a)$, and $d(x, H, E_a)$ denote the growth, death, and fecundity rates of the plant. Then,

$$H(x, a_0, a) = H_0 + \int_{a_0}^{a} g(x, H(x, a_0, a'), E_{a'}) \, da' \tag{1}$$

is the trajectory of plant height,

$$S_I(x, a_0, a) = S_G(x, H_0, E_{a0}) \exp\left(-\int_{a_0}^{a} d(x, H(x, a_0, a'), E_{a'}) \, da'\right) \tag{2}$$

is the probability of plant survival within the patch, where $S_G(x, H_0, E_{a0})$ is the probability a seed germinates successfully, and

$$\tilde{R}(x, a_0, a) = \int_{a_0}^{a} f(x, H(x, a_0, a'), E_{a'}) \, S_I(x, a_0, a') \, da' \tag{3}$$

is the plant's cumulative seed output, from its birth at patch age $a = a_0$ until patch age a.

   The notational complexity required in Eqs. 1-3 potentially obscures an important point: these equations are simply the general, nonlinear solutions to integrating growth, mortality, and fecundity over time.

2.2   *Patches of competing plants*

We now consider a patch of competing plants. At any age a, the patch is described by the size-density distribution $N(H|x, a)$ of plants with traits x and height H. In a finite-sized patch,

N is described by a collection of points, each indicating the height of one individual, whereas in a very (infinitely) large patch, N is a continuous function. In either case, the demographic behaviour of the plants within the patch is given by Eqs. 1-3. Integrating the dynamics over time is complicated by two other factors: (i) plants interact, thereby altering E_a with age, and (ii) new individuals may establish, expanding the system of equations.

In the current version of PLANT, plants interact by shading one another. Following standards biophysical principles, we let the canopy openness $E_a(z)$ at height z in a patch of age a decline exponentially with the total amount of leaf area above z,

$$E_a(z) = \exp\left(-k_I \sum_{i=1}^{K} \int_0^\infty A_1(H)\, Q(z,H)\, N(H|x_i,a)\, \mathrm{d}H\right),$$ (4)

where $A_1(H)$ is the total leaf area of a plant with height H, $Q(z,H)$ is the fraction of this leaf area situated above height z for plants of height H, k_I is the light extinction coefficient, and K is the number of species in the metacommunity.

Assuming patches are sufficiently large, the dynamics of N can be modelled deterministically via the following partial differential equation (PDE) (Kohyama, 1993; De Roos, Taljapurkar & Caswell, 1997; Moorcroft, Hurtt & Pacala, 2001),

$$\frac{\partial}{\partial a} N(H|x,a) = -d(x,H,E_a)\, N(H|x,a) - \frac{\partial}{\partial H}\left[g(x,H,E_a)\, N(H|x,a)\right].$$ (5)

See section 4.2 for the derivation of this PDE.

Eq. 5 has two boundary conditions. The first links the flux of individuals across the lower bound (H_0) of the size-density distribution to the rate Y_x at which seeds arrive in the patch,

$$N(H_0|x,a_0) = \begin{cases} \frac{Y_x\, S_G(x,H_0,E_{a0})}{g(x,H_0,E_{a0})} & \text{if } g(x,H_0,E_{a0}) > 0 \\ 0 & \text{otherwise.} \end{cases}$$ (6)

The function $S_G(x,H_0,E_{a0})$ denotes survival through germination and must be chosen such that $S_G(x,H_0,E_{a0})/g(x,H_0,E_{a0}) \to 0$ as $g(x,H_0,E_{a0}) \to 0$, to ensure a smooth decline in initial density as conditions deteriorate (Falster $et\ al.$, 2011).

The second boundary condition of Eq. 5 specifies the initial size-density distribution in patches right after a disturbance, i.e., for $a = 0$. Throughout, we consider only situations starting with an empty patch,

$$N(H|x,0) = 0,$$ (7)

although non–zero distributions could alternatively be specified (e.g., Kohyama, 1993; Moorcroft, Hurtt & Pacala, 2001).

2.3 Age structure of patches

We now consider the distribution of patch age a in the metacommunity. With a denoting the time since last disturbance, $P(a)$ denotes the frequency-density of patch age a and $d_P(a)$ is the age-dependent probability that a patch of age a is transformed into a patch of age o by a disturbance event. Here we focus on situations where the age structure has reached an equilibrium. See section 4.1 for the derivation and non-equilibrium situations. The dynamics of P are given by (McKendrick, 1926; Foerster, 1959)

$$\frac{\mathrm{d}}{\mathrm{d}a} P(a) = -d_P(a)\, P(a),$$ (8)

with the boundary condition

$$P(0) = \int_0^\infty d_\mathrm{P}(a)\, P(a)\, \mathrm{d}a.$$ (9)

The probability that a patch remains undisturbed from patch age a_0 to patch age a is Then given by

$$S_\mathrm{P}(a_0, a) = \exp\left(-\int_{a_0}^a d_\mathrm{P}(a')\, \mathrm{d}a'\right).$$ (10)

These equations lead to an equilibrium distribution of patch ages,

$$P(a) = P(0) S_\mathrm{P}(0, a),$$ (11)

where

$$P(0) = \frac{1}{\int_0^\infty S_\mathrm{P}(0, a)\mathrm{d}a}$$ (12)

is the average disturbance frequency of a patch and, at the same time, the frequency-density of patches of age 0.

The default approach in PLANT is to assume that $d_\mathrm{P}(a)$ increases linearly with patch age, which leads to a Weibull distribution for P, specified by a single parameter $\bar{a}$ measuring the mean interval between disturbances (see section 4.1 for details).

2.4   *Trait-, size-, and patch-structured metacommunities*

We consider a large area of habitat where: (i) disturbances (such as fires, storms, landslides, floods, or disease outbreaks) strike patches of the habitat on a stochastic basis, killing all individuals within the affected patches; (ii) individuals compete for resources within patches, with a spatial scale of competitive interactions that renders negligible such interactions between individuals in adjacent patches; and (iii) there is high connectivity via dispersal between all patches in the habitat, allowing empty patches to be quickly re-colonised. Such a system can be modelled as a metapopulation (for a single species) or metacommunity (for multiple species). The dynamics of such a metacommunity are described by the PDEs in Eqs. 5 and 8.

The seed rain of each species in the metacommunity is given by the rate at which seeds are produced across all patches,

$$Y_x = \int_0^\infty \int_0^\infty P(a)\, S_\mathrm{D}\, f(x, H, E_a)\, N(H|x, a)\, \mathrm{d}H\, \mathrm{d}a,$$ (13)

where S_D is the survival probability of seeds during dispersal.

A convenient feature of Eqs. 5 - 7 is that the dynamics of a single patch scale up to give the dynamics of the entire metacommunity. Note that the rate Y_x at which offspring of individuals with traits x arrive from the disperser pool is constant when the metacommunity is at equilibrium. Combined with the assumption that all patches have the same initial (empty) size-density distribution after a disturbance, the assumption of constant seed rains ensures that all patches show the same temporal behaviour, the only difference between them being their ages.

To model the temporal dynamics of an archetypal patch, we need only a value for Y_x. The numerical challenge is therefore to find the right value for Y_x, by solving Eqs. 6 and 13 simultaneously, for all species with traits x in the metacommunity.

2.5 *Emergent properties of metacommunities*

Summary statistics of a metacommunity are obtained by integrating over the size-density distribution, weighting by the frequency-density $P(a)$. In particular, the average density of individuals per unit ground area across the metacommunity is

$$\hat{N}(x) = \int_0^\infty \int_0^\infty P(a) \, N(H|x,a) \, \mathrm{d}a \, \mathrm{d}H, \tag{14}$$

and the average size-density of plants with height H is

$$\bar{N}(x,H) = \int_0^\infty P(a) \, N(H|x,a) \, \mathrm{d}a. \tag{15}$$

Averages for other individual-level quantities can also be calculated. Denoting by $w(x,H,E_a)$ a quantity of interest, either a demographic rate (growth, mortality) or a state (plant height, leaf area, light environment), the average of w for plants with height H and traits x is

$$\bar{w}(x,H) = \frac{1}{\bar{N}(x,H)} \int_0^\infty P(a) \, N(H|x,a) \, w(x,H,E_a) \, \mathrm{d}a. \tag{16}$$

The average of w across all individuals of the species is

$$\hat{w}(x) = \frac{1}{\hat{N}(x)} \int_0^\infty \int_0^\infty P(a) \, N(H|x,a) \, w(x,H,E_a) \, \mathrm{d}a \, \mathrm{d}H. \tag{17}$$

When calculating the average mortality rate, one must decide whether mortality due to patch disturbance is included. Non-disturbance mortality is obtained by setting $w(x,H,E_a) = d(x,H,E_a)$, while the total mortality due to growth processes and disturbance is obtained by setting $w(x,H,E_a) = d(x,H,E_a) + d_P(a)S_P(0,a)$.

Similarly, we can integrate over the size-density distribution to extract aggregate features of the vegetation within a patch,

$$W(a) = \sum_{i=1}^N \int_0^\infty N(H|x_i,a) \, w(x_i,H,E_a) \, \mathrm{d}H, \tag{18}$$

and across the entire metacommunity,

$$\hat{W} = \int_0^\infty P(a) \, W(a) \, \mathrm{d}a. \tag{19}$$

2.6 *Invasion fitness*

We now consider how to estimate the fitness of a rare mutant individual with traits x' growing in the light environment of a resident community with traits x. We focus on phenotype-dependent components of fitness – describing the aggregate consequences of a given set of traits for growth, fecundity, and mortality – taking into account the non-linear effects of competition, but ignoring the underlying genetic basis for trait inheritance and expression. We also adhere to standard conventions in such analyses by assuming that the mutant phenotype is sufficiently rare to have a negligible effect on the light environment where it is growing (Geritz *et al.*, 1998). The approach for calculating fitness implemented here follows the approach described by Falster *et al.* (2015).

In general, invasion fitness is defined as the long-term per capita growth rate of a rare mutant phenotype in the environment determined by a resident phenotype (Metz, Nisbet &

Geritz, 1992). Calculating per capita growth rates, however, is particularly challenging in a structured metacommunity model (Gyllenberg & Metz, 2001; Metz & Gyllenberg, 2001). As an alternative measure of invasion fitness in metacommunities, we can use the basic reproduction ratio measuring the expected number of new dispersers arising from a single dispersal event. For metacommunities at demographic equilibrium, evolutionary inferences made using basic reproduction ratios are equivalent to those made using per capita growth rates (Gyllenberg & Metz, 2001; Metz & Gyllenberg, 2001).

We denote by $R\left(x', x\right)$ the basic reproduction ratio of individuals with traits x' growing in the competitive light environment of the resident traits x. Recalling that patches of age a have frequency-density $P(a)$ in the landscape, it follows that any seed with traits x' has a probability of $P(a)$ of landing in a patch of age a. The basic reproduction ratio for individuals with traits x' is then

$$R\left(x', x\right) = \int_0^\infty P(a)\ \tilde{R}\left(x', a, \infty\right)\ \mathrm{d}a, \tag{20}$$

where $\tilde{R}\left(x', a_0, a\right)$ is the expected number of dispersing offspring produced by a single dispersing seed with traits x' arriving in a patch of age a_0 up until age a (Gyllenberg & Metz, 2001; Metz & Gyllenberg, 2001). In turn, $\tilde{R}\left(x', a, \infty\right)$ is calculated by integrating an individual's fecundity over the expected lifetime of a patch, taking into account competitive shading from residents with traits x, the individual's probability of surviving, and its traits,

$$\tilde{R}(x', a_0, \infty) = \int_{a_0}^\infty S_\mathrm{D}\, f(x', H(x', a_0, a), E_a)\, S_\mathrm{I}(x', a_0, a)\, S_\mathrm{P}(a_0, a)\, \mathrm{d}a. \tag{21}$$

3   NUMERICAL METHODS FOR SOLVING METACOMMUNITY DYNAMICS

The equations below describe how we solve the trait-, size-, and patch-structured population dynamics specified in the previous section. Our approach to solving for the size-density distribution is based on the characteristic method (Angulo & López-Marcos, 2004). The characteristic method is related to the Escalator Boxcar Train technique (De Roos, 1988; De Roos, Taljapurkar & Caswell, 1997; De Roos, Diekmann & Metz, 1992), but not identical to it.

When simulating an individual plant, or the development of a patch, we need to solve for the size, survival, and seed output of individual plants. When solving for the size-density distribution in a large patch, we also need to estimate the average abundance of individuals. Each of these problems is formulated as an initial-value ODE problem (IVP), which can be solved using an ODE stepper.

3.1   *Approach*

SIZE — The size of an individual is obtained via Eq. 1, which is solved via the IVP

$$\frac{dy}{dt} = g(x, y, t),$$

$$y(0) = H_0.$$

SURVIVAL — The probability of an individual surviving from patch age a_0 to patch age a is obtained via Eq. 2, which is solved via the IVP

$$\frac{dy}{dt} = d(x, H_i(t), E_t),$$

$$y(0) = -\ln\left(S_{\mathrm{G}}(x, H_0, E_{a0})\right).$$

Survival is then

$$S_{\mathrm{I}}(x, a_0, a) = \exp\left(-y(a)\right).$$

SEED PRODUCTION — The lifetime seed production of individuals is obtained via Eq. 21, which is solved via the IVP

$$\frac{dy}{dt} = S_{\mathrm{D}}\, f(x, H_i(t), E_t)\, S_{\mathrm{I}}(x, a_0, t)\, S_{\mathrm{P}}(a_0, t),$$

$$y(0) = 0,$$

where S_{I} is calculated as described above and S_{P} is calculated as in Eq. 10.

SIZE-DENSITY OF INDIVIDUALS — By integrating along the characteristics of Eq. 5, the size-density of individuals born with height H_0 and traits x at patch age a_0 is given by (De Roos, Taljapurkar & Caswell, 1997)

$$N(H|x, a) = N(H_0|x, a_0) \exp\left(-\int_{a_0}^{a}\left[\frac{\partial g(x, H(x, a_0, a'), E_{a'})}{\partial H} + d(x, H(x, a_0, a'), E_{a'})\right] da'\right).$$
(22)

Eq. 22 states that the size-density N at a specific patch age a is the product of the size-density at patch age a_0 adjusted for changes through growth and mortality. Size-density decreases through time because of mortality, as in a typical survival equation, but also changes because of growth. If growth is slowing with size, (i.e., $\partial g/\partial H < 0$), size-density will increase since the characteristics compress. Conversely, size-density will increases if $\partial g/\partial H > 0$.

Denoting by $[H_0, H_+)$ the range of heights attainable by any individual, our algorithm for solving metacommunity dynamics proceeds by sub-dividing this interval into a series of cohorts with heights $H_0 < H_1 < \ldots < H_k$ at the initial points of the characteristic curves. These cohorts are then transported along the characteristics of Eq. 5. The placement of cohorts is controlled indirectly, via the schedule of patch ages at which new cohorts are introduced into the metacommunity. We then track the demography of each such cohort.

The integral in Eq. 22 is solved via the IVP

$$\frac{dy}{dt} = \frac{\partial g(x, H_i(t), E_t)}{\partial H} + d(x, H_i(t), E_t),$$

$$y(0) = -\ln\left(N(H_0|x, a_0)\, S_{\mathrm{G}}(x, H_0, E_{a0})\right),$$

from which we obtain the size-density

$$N(H_0|x, a_0) = \exp(-y(a)).$$

3.2 *Controls on approximation error*

We now outline how to control the error of the approximate solutions to the system of equations described above.

In our algorithm, numerical solutions are required to address a variety of problems:

- To estimate the amount of light at a given height in a patch requires numerically integrating over the size-density distribution within that patch.

- To calculate the assimilation of a plant requires numerically integrating photosynthesis over this light profile.

- To simulate patch dynamics requires numerically identifying a vector of patch ages at which new cohorts are introduced, and then numerically stepping the equations for each cohort forward in time to estimate their size, survival, and fecundity at different subsequent patch ages.

- To solve for the initial height of a plant given its seed mass, and for the equilibrium seed rains across the metacommunity, requires numerical root finding.

As with all numerical techniques, solutions to each of these problems are accurate only up to a specified level. These levels are controlled via parameters in the PLANT code. Below, we provide a brief overview of the different numerical techniques being applied and outline how error tolerance can be increased or decreased.

We refer to various control parameters that can be found within the `control` object. For a worked example illustrating how to modify these control parameters, see the section `parameters` of Appendix S3.

INITIAL PLANT HEIGHTS — When a seed germinates, it produces a seedling of given height. The height of these seedlings is assumed to vary with the seed mass. Because there is no analytical solution relating seedling height to seed mass – at least when using our default physiological model – we must solve for this height numerically. The calculation is performed by the function `height_seed` within the physiological model, using the Boost library's one-dimensional `bisect` routine (Schäling, 2014; Eddelbuettel, Emerson & Kane, 2015). The accuracy of the solution is controlled by the parameter `plant_seed_tol`.

STEPPING OF ODES — All of the IVPs outlined above must be stepped through time. For this, PLANT uses the embedded Runge-Kutta Cash-Karp 4-5 algorithm (Cash & Karp, 1990), with code ported directly from the GNU Scientific Library (Galassi, 2009). The accuracy of the solver is controlled by two control parameters for relative and absolute accuracy, `ode_tol_rel` and `ode_tol_abs`.

APPROXIMATION OF SIZE-DENSITY DISTRIBUTION VIA THE CHARACTERISTIC METHOD — Errors in the approximation of the size-density distribution arise from two sources: (i) coarse stepping of cohorts through time and (ii) poor spacing of cohorts across the attainable size range.

As described above, the stepping of the ODE solver is controlled by two control parameters for relative and absolute accuracy, `ode_tol_rel` and `ode_tol_abs`.

A second factor controlling the accuracy with which cohorts are stepped through time is the accuracy of the derivative calculation according to Eq. 22, calculated via standard finite differencing (Abramowitz & Stegun, 2012). When the parameter `cohort_gradient_richardson` is TRUE a Richardson extrapolation (Stoer & Bulirsch, 2002) is used to refine the estimate, up to depth `cohort_gradient_richardson`. The overall accuracy of the derivative is controlled by `cohort_gradient_eps`.

The primary factor controlling the spacing of cohorts is the schedule of cohort introduction times. Because the system of equations to be integrated is deterministic, the schedule of cohort introduction times determines the spacing of cohorts throughout the entire development of a patch. Poor cohort spacing introduces error because various emergent properties – such as total leaf area, biomass, or seed output – are estimated by integrating over the size-density distribution. The accuracy of these integrations declines directly with the inappropriate spacing of cohorts. Thus, our algorithm aims to build an appropriately refined schedule, which allows the required integrations to be performed with the desired accuracy at every

time point. At the same time, for reasons of computational efficiency, we want as few cohorts as possible.

A suitable schedule is found using the function `build_schedule`. This function takes an initial vector of introduction times and considers for each cohort whether removing that cohort causes the error introduced when integrating two specified functions over the size-density distribution to jump over the allowable error threshold `schedule_eps`. This calculation is repeated for every time step in the development of the patch. A new cohort is introduced immediately prior to any cohort failing these tests. The dynamics of the patch are then simulated again and the process is repeated, until all integrations at all time points have an error below the tolerable limit `schedule_eps`. Decreasing `schedule_eps` demands higher accuracy from the solver, and thus increases the number of cohorts being introduced. Note that we are assessing whether removing an existing cohort causes the error to jump above the threshold limit, and using this to decide whether an extra cohort – in addition to the one used in the test – should be introduced. Thus, the actual error is likely to be lower than, but at most equal to, `schedule_eps`. The general idea of adaptively refining the vector of introduction times was first applied by Falster *et al.* (2011), and was described further by Falster *et al.* (2015).

To determine the error associated with a given cohort, we integrate two different functions over the size-density distribution, within the function `run_scm_error`. We then assess how much removing the focal cohort increases the error in these two integrations. The first integration, performed by the function `area_leaf_error`, determines how much the removal of the focal cohort increases the error in the estimate of the total leaf area in the patch. The second integration, performed by the function `seed_rain_error`, determines how much the removal of the focal cohort increases the error in the estimate of the total seed production from the patch. The relative error in each integration is then calculated using the `local_error_integration` function.

For a worked example illustrating the `build_schedule` function, see the section `cohort_spacing` of Appendix S3.

CALCULATION OF LIGHT ENVIRONMENT AND INFLUENCE ON ASSIMILATION — To progress with solving the system of ODEs requires that we calculate the amount of shading on each of the cohorts, from all other plants in the patch.

Calculating the canopy openness $E_a(z)$ at a given height z in a patch of age a requires that we integrate over the size-density distribution (Eq. 4). This integration is performed using the trapezium rule, within the function `area_leaf_above` in `species.h`. The main factor controlling the accuracy of the integration is the spacing of cohorts. The cohort introduction times determining the spacing of cohorts are adaptively refined as described above. This implies that also the trapezium integration within the `area_leaf_above` function is adaptively refined via the `build_schedule` function.

The cost of calculating $E_a(z)$ linearly increases with the number of cohorts in the meta-community. Since the same calculation must be repeated for every cohort, the overall computational cost of a step increases as $O(k^2)$, where k is the total number of cohorts across all species. This disproportionate increase in computational cost with the number of cohorts is highly undesirable.

We reduce the computational cost from $O(k^2)$ to $O(k)$ by approximating $E_a(z)$ with a spline. Eq. 4 describes a function monotonically increasing with size. This function is easily approximated using a piecewise continuous spline fitted to a limited number of points. Once fitted, the spline can be used to estimate any additional evaluations of competitive effect. Since spline evaluations are computationally cheaper than integrating over the size-density

distribution, this approach reduces the overall cost of stepping the resident population. A new spline is constructed for each time step.

The accuracy of the spline interpolation depends on the number of points used in its construction and on their placement along the size axis. We select the number and locations of points via an adaptive algorithm. Starting with an initial set of 33 points, we assess how much each point contributes to the accuracy of the spline fit at the location of each cohort, first via exact calculation, and second by linearly interpolating from adjacent cohorts. The absolute difference in these values is compared to the control parameter `environment_light_tol`. If the error is greater than this tolerance, the interval is bisected and the is process repeated (see `adaptive_interpolator.h` for details).

INTEGRATION OVER LIGHT ENVIRONMENT — Plants have leaf area distributed over a range of heights. Estimating a plant's assimilation at each time step thus requires integrating leaf-level rates over the plant. The integration is performed using Gaussian quadrature, using the QUADPACK routines (Piessens *et al.*, 1983) adapted from the GNU Scientific Library (Galassi, 2009); see `qag.h` for details. If the control parameter `plant_assimilation_adaptive` is `TRUE`, the integration is performed using adaptive refinement with an accuracy controlled by the parameter `plant_assimilation_tol`.

SOLVING FOR SEED RAINS — For a single species, solving for Y_x is a straightforward one-dimensional root-finding problem, which can be solved with accuracy `equilibrium_eps` via a simple bisection algorithm (see `equlibrium.R` for details).

Solving for seed rains in metacommunities with multiple species is significantly harder, because there is no generally applicable method for multi-dimensional root finding. In PLANT, we have therefore implemented several different approaches (see `equlibrium.R` for details).

4   APPENDICES

4.1   *Derivation of* PDE *describing age-structured dynamics*

We consider patches of habitat that are subject to an intermittent disturbance, with the age of a patch measuring the time since the last disturbance. Denoting by $P(a, t)$ the frequency-density of patch age a at time t and by $d_P(a)$ the age-dependent probability that a patch of age a is transformed into a patch of age 0 through disturbance, the dynamics of $P(a, t)$ are given by

$$\frac{\partial}{\partial t} P(a, t) = -\frac{\partial}{\partial a} P(a, t) - d_P(a, t) P(a, t),$$

with boundary condition

$$P(0, t) = \int_0^\infty d_P(a, t) P(a, t) \, da.$$

The frequency-density of patches of age $a < a'$ is given by $\int_0^{a'} P(a, t) \, da$, with $\int_0^\infty P(a, t) \, da = 1$. If $\frac{\partial}{\partial t} d_P(a, t) = 0$, then $P(a)$ will approach an equilibrium solution given by

$$P(a) = P(0) \, S_P(0, a),$$

where

$$S_P(0, a) = \exp\left(-\int_0^a d_P(a') \, da'\right)$$

is the probability that a patch remains undisturbed for duration a, and

$$P(0) = \frac{1}{\int_0^\infty S_P(0,a)\,\mathrm{d}a}$$

is the frequency-density of patches of age 0. The rate of disturbance for patches of age a is given by $\frac{\partial(1 - S_P(0,a))}{\partial a} = -\frac{\partial S_P(0,a)}{\partial a}$, while the expected lifetime of patches is $-\int_0^\infty a\frac{\partial}{\partial a}S_P(0,a)\,\mathrm{d}a = \int_0^\infty S_P(0,a)\,\mathrm{d}a = \frac{1}{P(0)}$ (first step made using integration by parts).

An equilibrium distribution of patch ages may be achieved under a variety of conditions, for example, if $d_P(a,t)$ depends on patch age a but not on time t. The rate of disturbance may also depend on features of the vegetation in the patch, rather than on patch age directly, in which case an equilibrium distribution of patch ages can still arise, provided the vegetation is also assumed to be at equilibrium.

EXPONENTIAL DISTRIBUTION — If the rate d_P of patch disturbance is constant with respect to patch age, then the rates at which patches of age a are disturbed follow an exponential distribution, $-\partial S_P(0,a)/\partial a = d_P \exp(-d_P a)$. The distribution of patch ages is then given by

$$S_P(0,a) = \exp\left(-d_P a\right), \; P(0) = d_P.$$

WEIBULL DISTRIBUTION — If the rate of patch disturbance changes with patch age according to $d_P(a) = \lambda\psi a^{\psi-1}$, then the rates at which patches of age a are disturbed follow a Weibull distribution, $-\partial S_P(0,a)/\partial a = \lambda\psi a^{\psi-1}e^{-\lambda a^\psi}$. $\psi > 1$ implies that the probability of disturbance increases with patch age, while $\psi < 1$ implies that it decreases with patch age. For $\psi = 1$, we obtain the exponential distribution, a special case of the Weibull distribution. The Weibull distribution results in the following distribution of patch ages,

$$S_P(0,a) = \exp(-\lambda a^\psi), \; P(0) = \frac{\psi\lambda^{\frac{1}{\psi}}}{\Gamma\left(\frac{1}{\psi}\right)},$$

where $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}\,dt$ is the gamma function. We can also specify this distribution by the mean disturbance interval $\bar{a} = \frac{1}{P(0)}$. From this, we can calculate the relevant value for

$$\lambda = \left(\frac{\Gamma\left(\frac{1}{\psi}\right)}{\psi\bar{a}}\right)^\psi.$$

The default in PLANT is to assume $\psi = 2$, such that d_P increases as a linear function of patch age. The distribution of patch ages is then specified by a single parameter, $\bar{a}$.

4.2 *Derivation of PDE describing size-structured dynamics*

To model the metacommunity, we use a PDE describing the dynamics for a thin slice ΔH. This PDE can be derived as follows, following De Roos, Taljapurkar & Caswell (1997). Note that in this sub-section the dependencies on the traits x are deliberately not made explicit. Assuming that all rates are constant within the interval ΔH, the total number of individuals within the interval spanned by $[H - 0.5\Delta H, H + 0.5\Delta H)$ is $N(H,a)\Delta H$. The flux of individuals in and out of this size interval can be expressed as

$$g(H - 0.5\Delta H, a)\,N(H - 0.5\Delta H, a) - g(H + 0.5\Delta H, a)\,N(H + 0.5\Delta H, a) \atop -d(H,a)\,N(H,a)\Delta H \qquad . \tag{23}$$

The first two terms describe the flux in and out of the considered size interval through growth, while the last term describes losses through mortality. The change in the density of individuals within this size interval during a time step Δa is thus

$$
\begin{aligned}
N(H, a + \Delta a)\Delta H - N(H, a)\Delta H = \ & g(H - 0.5\Delta H, a)\, N(H - 0.5\Delta H, a)\Delta a \\
& -g(H + 0.5\Delta H, a)\, N(H + 0.5\Delta H, a)\Delta a \\
& -d(H, a)\, N(H, a)\Delta H \Delta a.
\end{aligned}
\tag{24}
$$

By rearranging, we obtain

$$
\frac{N(H, a + \Delta a) - N(H, a)}{\Delta a} = \ -d(H, a)\, N(H, a) \\
- \frac{g(H + 0.5\Delta H, a)\, N(H + 0.5\Delta H, a) - g(H - 0.5\Delta H, a)\, N(H - 0.5\Delta H, a)}{\Delta H}.
\tag{25}
$$

The left-hand side above corresponds to the derivative of N as $\Delta a \rightarrow 0$. For thin slices, $\Delta H \rightarrow 0$, this yields

$$
\frac{\partial}{\partial a} N(H, a) = -d(H, a)\, N(H, a) - \frac{\partial}{\partial H}\left(g(H, a)\, N(H, a)\right).
\tag{26}
$$

To complete the model, this PDE must be supplemented with boundary conditions that specify the density at the lower end of heights for all a, as well as the initial distribution $N(H, 0)$. The former is derived by integrating the PDE with respect to H over the interval (H_0, H_∞), yielding

$$
\frac{\partial}{\partial a} \int_{H_0}^{H_\infty} N(H, a)\, \mathrm{d}H = g(H_0, a)\, N(H_0, a) - g(H_\infty, a)\, N(H_\infty, a) - \int_{H_0}^{H_\infty} d(H, a)\, N(H, a)\, \mathrm{d}H.
\tag{27}
$$

The left-hand size above is evidently the rate of change of the total density of individuals in the population, while the right-hand side is the population's total death rate. Further, we can assume that $N(H_\infty, a) = 0$. Thus, to balance total births and deaths, $g(H_0, a)\, N(H_0, a)$ must equal the population's total birth rate B, yielding the boundary condition

$$
g(H_0, a)\, N(H_0, a) = B.
\tag{28}
$$

4.3 *Converting density from one size unit to another*

The quantification of size-density depends on the chosen size unit (Eq. 5). Thus, what if we want to express size-density with respect to another size unit? A relation between the two corresponding values of size-density can be derived by noting that the total density of individuals within a given size range must be equal. We consider a situation in which size-density is expressed in units of size M, but we want it in units of size H. First, we require a one-to-one function that translates the first size unit into the second, $H = \hat{H}(M)$. Then the following must hold

$$
\int_{M_1}^{M_2} N(M|x, a)\, \mathrm{d}M = \int_{\hat{H}(M_1)}^{\hat{H}(M_2)} N'(H|x, a)\, \mathrm{d}H.
\tag{29}
$$

For very small size intervals, this equation is equivalent to

$$
(M_2 - M_1)\, N(M_1|x, a) = \left(\hat{H}(M_2) - \hat{H}(M_1)\right) N'(\hat{H}(M_1)|x, a).
\tag{30}
$$

Rearranging gives

$$
N'(\hat{H}(M_1)|x, a) = N(M_1|x, a)\, \frac{M_2 - M_1}{\hat{H}(M_2) - \hat{H}(M_1)}.
\tag{31}
$$

Noting that the second term on the right-hand side is simply the definition of $\frac{\mathrm{d}M}{\mathrm{d}H}$ evaluated at M_1, we have

$$N'(H|x,a) = N(M|x,a)\,\frac{\mathrm{d}M}{\mathrm{d}H}.\tag{32}$$

TABLE 1: Variable names and definitions in demographic model.

Symbol	Unit	Description
Patch state variables		
K		Number of species in the metacommunity
a, a'	yr	Patch age (time since disturbance)
a_0	yr	Patch age when plant germinates
E_a		Profile of canopy openness within a patch of age a
$E_a(z)$		Canopy openness at height z within a patch of age a
Plant state variables		
x		Vector of traits for a species
H	m	Height of a plant
H_0	m	Height of a seedling after germination
z	m	Height in canopy
$A_1(H)$	m^{-2}	Leaf area of a plant with height H
$Q(z, H)$		Fraction of leaf area above height z for a plant with height H
Abundance measures		
$P(a)$	yr^{-1}	Frequency-density of patches of age a
Y_x	$m^{-2} \, yr^{-1}$	Global seed rain for a species with traits x
$N(H|x, a)$	$m^{-1} \, m^{-2}$	Density of plants at height H per unit ground area for given traits x and patch age a
Demographic rates		
$g(x, H, E_a)$	$m \, yr^{-1}$	Height growth rate of a plant with traits x and height H in the light environment E_a in a patch of age a
$f(x, H, E_a)$	yr^{-1}	Seed production rate of a plant with traits x and height H in the light environment E_a in a patch of age a
$d(x, H, E_a)$	yr^{-1}	Instantaneous mortality rate of a plant with traits x and height H in the light environment E_a in a patch of age a
$d_P(a)$	yr^{-1}	Instantaneous disturbance rate of a patch of age a
Demographic outcomes		
$H(x, a_0, a)$	m	Height of a plant with traits x in a patch of age a having germinated at patch age a_0
S_D		Probability a seed survives dispersal
$S_G(x, H_0, E_{a0})$		Probability the seed of a plant with traits x germinates successfully at height H_0 in the light environment E_{a0} in a patch of age a_0
$S_I(x, a_0, a)$		Probability a plant survives from patch age a_0 to patch age a
$S_P(a_0, a)$		Probability a patch remains undisturbed from patch age a_0 to patch age a
$\tilde{R}(x, a_0, a)$		Cumulative seed output for plants with traits x from patch age a_0 to patch age a
$R(x', x)$		Basic reproduction ratio for a mutant plant with traits x' growing in a metapopulation of resident plants with traits x
Miscellaneous constants		
k_I		Light extinction coefficient
$\bar{a}$	yr	Mean interval between patch disturbances

REFERENCES

Abramowitz, M. & Stegun, I.A. (2012) *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Courier Corporation.

Angulo, O. & López-Marcos, J.C. (2004) Numerical integration of fully nonlinear size-structured population models. *Applied Numerical Mathematics*, **50**, 291–327.

Cash, J.R. & Karp, A.H. (1990) A variable order Runge-Kutta method for initial value problems with rapidly varying right-hand sides. *ACM Transactions on Mathematical Software*, **16**, 201–222.

De Roos, A.M. (1988) Numerical methods for structured population models: the Escalator Boxcar Train. *Numerical Methods for Partial Differential Equations*, **4**, 173–195.

De Roos, A.M., Diekmann, O. & Metz, J.A.J. (1992) Studying the dynamics of structured population models: a versatile technique and its application to daphnia. *American Naturalist*, **139**, 123.

De Roos, A.M., Taljapurkar, S. & Caswell, H. (1997) A gentle introduction to physiologically structured population models. *Structured population models in marine, terrestrial and freshwater systems*, pp. 119–204. Chapman & Hall, New York.

Eddelbuettel, D., Emerson, J.W. & Kane, M.J. (2015) *BH: Boost C++ Header Files*. R package version 1.55.0-3.

Falster, D.S., Brännström, Å., Dieckmann, U. & Westoby, M. (2011) Influence of four major plant traits on average height, leaf-area cover, net primary productivity, and biomass density in single-species forests: a theoretical investigation. *Journal of Ecology*, **99**, 148–164.

Falster, D.S., Brännström, Å., Westoby, M. & Dieckmann, U. (2015) Multi-trait eco-evolutionary dynamics explain niche diversity and evolved neutrality in forests. *bioRxiv*, p. 014605.

Foerster, H.V. (1959) Some remarks on changing populations. F. Stohlman, ed., *The kinetics of cellular proliferation*, pp. 382–407. Grune et Stratton, New York.

Galassi, M., ed. (2009) *GNU scientific library: reference manual*. A GNU manual. Network Theory, s.l., 3. ed., for gsl version 1.12 edition.

Geritz, S.A.H., Kisdi, É., Meszéna, G. & Metz, J.A.J. (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, **12**, 35–57.

Gyllenberg, M. & Metz, J.A.J. (2001) On fitness in structured metapopulations. *Journal of Mathematical Biology*, **43**, 545–560.

Kohyama, T. (1993) Size-structured tree populations in gap-dynamic forest: the forest architecture hypothesis for the stable coexistence of species. *Journal of Ecology*, **81**, 131–143.

McKendrick, A.G. (1926) Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society*, **44**, 1–34.

Metz, J.A.J. & Gyllenberg, M. (2001) How should we define fitness in structured metapopulation models? Including an application to the calculation of evolutionarily stable dispersal strategies. *Proceedings of the Royal Society of London - Series B: Biological Sciences*, **268**, 499–508.

Metz, J.A.J., Nisbet, R.M. & Geritz, S.A.H. (1992) How should we define 'fitness' for general ecological scenarios? *Trends in Ecology and Evolution*, **7**, 198–202.

Moorcroft, P.R., Hurtt, G.C. & Pacala, S.W. (2001) A method for scaling vegetation dynamics: the Ecosystem Demography model (ED). *Ecological Monographs*, **71**, 557–586.

Piessens, R., Doncker-Kapenga, E.D., Überhuber, C. & Kahaner, D. (1983) *QUADPACK, a subroutine package for automatic integration*. Springer, Berlin Heidelberg.

Schäling, B. (2014) *The Boost C++ libraries*. XML Press, 2nd edition.

Stoer, J. & Bulirsch, R. (2002) *Introduction to numerical analysis*. Springer Science & Business Media.