Supplement A: Modeling and computation
Pulsed-resource mast systems and the movement, demographic storage, and diet breadth of consumers

JAMES S. CLARK[1,2], CHASE NUÑEZ[1], and BRADLEY TOMASEK[1]

May 3, 2018

[1] *Nicholas School of the Environment, Duke University, Durham NC 27708, USA*
[2] *Department of Statistical Science, Duke University, Durham NC 27708, USA*

Contents

S1   General theory

The spatially-explicit analysis for field data can be generalized to an arbitrary landscape, where explicit tree locations are unspecified. Consider a stand where trees that differ in fecundity are specified only in terms of their spatial contagion and fecundity class, perhaps summarized by tree size–a distance-size distribution. Fecundity differences may be linked to the distribution of tree-to-tree distances, as when fecund trees tend to be large and also sparse. Let r represent distance, and $r_1, r_2, \ldots$ represent the ranked distances from a randomly selected tree to others in the forest. Spatial contagion is described by density functions $p_k(r) \cdot dr$ for the probability that the kth closest tree occurs at distance interval $(r, r + dr)$ from a random tree. We consider a landscape that is homogeneous, in the sense that spatial aggregation exists

at a fine scale relative to the area of the entire landscape. In this sense, 'homogeneous' means that density $p_k(r)$ does not depend on location, despite the possibility of clumped host distributions.

The probability that the kth closest neighbor will be at least distance r is $P_k(r) = [r_k \geq r]$. The corresponding density for nearest neighbors is $p_1(r) = -dP_1/dr$. The density for nearest trees from a random spatial location is

$$q_1(r) = \frac{1}{\mu} P_1(r)$$

where $\mu = \int_0^\infty P_1(r) dr$ is the mean distance between nearest neighbors. This is the translation of inter-tree distances to space-to-tree distances.

The density of distances to the kth neighbor is translated to a renewal intensity for the expected number of trees at distance r. From a randomly selected location, the renewal intensity is

$$m(r) = \sum_{k=0}^\infty q_k(r)$$

where $m(r)dr$ is the expected number of trees in $(r, r + dr)$. On a homogeneous landscape (density $q_k(r)$ does not depend on location), $m(r)$ is a non-decreasing function that converges to an asymptote as $q_k(r)$ tends to zero with higher values of k.

Thus far, the model takes no account of the fact that fecundity may increase with tree size, and large trees tend to be sparse. To account for this effect, the renewal intensity can be determined for each fecundity class g, or m_g. Let $E(f_g)$ represent expected seed production for trees in fecundity class g. The expected intensity for seed arrivals emanating from trees in class g at all distances is

$$E(\lambda) = \sum_g E(f_g) \int_0^\infty m_g(r) S(r) dr$$

The landscape (conditional) variance is

$$Var(E(y|f)) = Var(\lambda|f) = \sum_g E^2(f_g) \int_0^\infty m_g(r) S(r) dr - (E(\lambda))^2$$

The variance in counts is

$$Var(y) = Var(E(y|f)) + E(Var(y|f))$$
$$= \sum_g E^2(f_g) \int_0^\infty m_g(r)S(r)dr - (E(\lambda))\left((E(\lambda)) - 1\right)$$

the last term incorporating the Poisson sampling variance. The resource score in the section *Variance for a generalist consumer* is available for this generic landscape based on these expressions for mean and variance. The basic elements of this analysis are detailed in the next section.

S2   GENERATIVE MODEL AND PREDICTION

S2.1   *Model elements*

Elements summarized in MODEL DEVELOPMENT of the main text are described with additional detail here, with variable definitions from table S2.1:

- An *observation* consists of covariates for trees and seed traps, including $\{\mathbf{X}_{j,t}, \mathbf{W}_{j,t}, \mathbf{V}_{j,t}, \mathbf{Y}_{j,t}, \mathbf{z}_{j,t}, \{\mathbf{c}_{ij,t}\}_{i=1}^{n_j}, \{A_{sj,t}\}_{s=1}^{S_{j,t}}\}$. There are $i = 1, \ldots, n_{j,t}$ trees on plot $j, \ldots, J$ in years $t, \ldots, T_j$. There are $s = 1, \ldots, S_{j,t}$ seed traps in plot j in year t. The observation matrices are:

  - *Predictors* that explain maturation occupy the $n_{j,t} \times q^v$ design matrix $\mathbf{V}_{j,t}$. Predictors that explain fecundity occupy the $n_{j,t} \times q^x$ design matrix $\mathbf{X}_{j,t}$. If there are random individual effects, they occupy the random effects design $\mathbf{W}_{j,t}$. The different species h are treated as factor levels in $\mathbf{V}_{j,t}, \mathbf{X}_{j,t}$, and $\mathbf{W}_{j,t}$ including all interactions with predictors. Species labels are absorbed into design matrices.

  - The length-$n_{j,t}$ *maturation* vector $\mathbf{z}_{j,t}$ holds observed maturation states (often unknown).

  - The length-2 observed fecundity vector $\mathbf{c}_{ij,t} = (c_{ij,t,min}, c_{ij,t,max})$ holds minimum and maximum estimates of seeds observed on trees. If fruiting structures containing multiple seeds are counted (e.g., cones), then $\mathbf{c}_{ij,t}$ is expressed on a seed basis (cones $\times$ seeds per cone) (Ladeau and Clark, 2006).

4

- The $S_{j,t} \times M$ *response* matrix $\mathbf{Y}_{j,t}$ holds seed counts. It has one row for each seed-trap-year and one column for each seed type m.

- *Sample effort* $A_{sj,t}$ is the trap area (m^{-2}) times deployment time (proportion of the year, traps are occasionally damaged).

- Trap-to-tree distances enter the $S_{j,t} \times n_{j,t}$ *redistribution kernel* matrix $\mathbf{S}_{j,t}$. The distance from seed trap (sj) to tree (ij) is $d_{sij} = |\mathbf{s}_{sj,t} - \mathbf{s}_{ij,t}|$. The t subscript allows for ingrowth of new individuals, for mortality loss, and for the addition or loss of seed traps over time. $\mathbf{S}$ incorporates dispersal parameters $u_g, g = 1, \ldots G$, where g can correspond to species h or to random groups. The redistribution kernel has elements

$$\mathbf{S}_{si} = \frac{u_{h[i]}}{\pi \left(u_{h[i]} + d_{si}^2 \right)^2} \tag{S2.1}$$

for distance d_{si} and fitted dispersal parameter $u_{h[i]}$, where subscript $h[i]$ references the parameter value for the species h to which tree i belongs. This is a two-dimensional Student's t distribution (Clark et al., 1999).

S2.2   *Posterior distribution*

Parameters can include random effects, year effects, and lag effects (table S2.1). There are additional latent states, maturation and conditional fecundity. In its most basic form (main effects only) the model has parameters

$$\boldsymbol{\theta} = \left\{ \boldsymbol{\beta}^x, \boldsymbol{\beta}^v, \sigma^2, \mathbf{u}, \mathbf{M} \right\} \tag{S2.2}$$

(table S2.1). The posterior distribution is

Table S2.1: Terms for conditional fecundity, $\log \psi_{i,t} \sim N(\mu_{i,t}, \sigma^2)$ and variable list in MASTIF.

Term(s) in $\mu_{i,t}$	Definitions
$\mathbf{x}'_{i,t}\boldsymbol{\beta}^x$	A. Fixed effects
$\gamma_t + \gamma_{g[i],t}$	B. Fixed year, random by group
$\sum_{l=1}^{p}(\alpha_l + \alpha_{g[i],l})\psi_{i,t-l}$	C. AR(p) fixed lag, random by group
$\mathbf{w}'_{i,t}\boldsymbol{\beta}^w_i$	D. Random individual effect
Variables/parameters (dimension)	(units)
$A_{s,t}$	sample effort (m^2 yr)
$\mathbf{A}_\alpha$ ($G \times G$)	random group covariance for AR(p)
α_l	AR(p) lag effect $l = 1, \ldots, p$
$\mathbf{B}_w$ ($q^w \times q^w$)	random individual covariance (x^{-2})
$\boldsymbol{\beta}^v$ ($q^v \times 1$)	maturation coefficients (v^{-1})
$\boldsymbol{\beta}^x$ ($q^x \times 1$), $\boldsymbol{\beta}^w_i$ ($q^w \times 1$)	fecundity fixed, random coefficients (x^{-1})
$\mathbf{c}_{i,t}$ (2×1)	(min, max) observed fecundity
$f_{i,t}$	fecundity (seeds)
$g_{sh,t}$	seed mass on the ground (g m^{-2})
$\lambda_{sm,t}$	expected seed density (m^{-2})
$\psi_{i,t}$	conditional fecundity (seeds)
γ_t	year effect $t = 1, \ldots, T$
$\mathbf{m}_h$ ($M \times 1$), $\mathbf{M}$ ($H \times M$)	seed-type composition for species h
$q_{i,t}$	seed mass on the tree (kg tree^{-1})
$\rho_{i,t}$	maturation status
$\mathbf{S}$ ($S \times n$)	redistribution kernel matrix (m^{-2})
σ^2	log fecundity residual variance
τ^2	random year variance
u_g or u_h, $\mathbf{u}$ ($G \times 1$)	random dispersal parameter (m^2)
$\tilde{u}, \tilde{U}$	fixed dispersal mean (m^2) and variance (m^4)
$\mathbf{v}_{i,t}$ ($q^v \times 1$)	maturation design (x)
$\mathbf{x}_{i,t}$ ($q^x \times 1$), $\mathbf{w}_{i,t}$ ($q^w \times 1$)	fecundity fixed, random design (x)
$y_{sm,t}$	seed count
$z_{i,t}$	observed maturation status

Subscripts reference trap s and seed type m or individual i and species h in year t. Plot subscript j is omitted for clarity. The notation $g[i]$ refers to the membership of individual i in group g. There are n trees, S traps, H species, M seed types, and G random groups. The dimension of non-scalar quantities is indicated. Variables lacking units at right are dimensionless.

$$[\boldsymbol{\theta}|\mathbf{X}, \mathbf{W}, \mathbf{V}, \mathbf{Y}, \mathbf{z}, \{A_{sj,t}\}] \propto \prod_{s,m,j} \prod_{t=1}^{T_{sj}} Poi\left(y_{smj,t}|A_{sj,t}\lambda_{smj,t}\right) \tag{S2.3}$$

$$\times \prod_{i,j} \prod_{t=1}^{T_{i,j}} [\mathbf{z}_{ij}|\rho_{ij,t}][\rho_{ij,t}|\mathbf{v}'_{ij,t}\boldsymbol{\beta}^v, \rho_{ij,t-1}, \rho_{ij,t+1}] \tag{S2.4}$$

$$\times \prod_{i,j} \prod_{t=1}^{T_j} [\mathbf{c}_{ij,t}|\psi_{ij,t}][\psi_{ij,t}|\mathbf{x}'_{ij,t}\boldsymbol{\beta}^x, \sigma^2, \rho_{ij,t}] \tag{S2.5}$$

$$\times [\boldsymbol{\theta}]$$

where $\lambda_{smj,t}$ is an element of $S_{j,t} \times M$ matrix $\boldsymbol{\Lambda}_{j,t} = \mathbf{S}_{j,t}(\mathbf{u})\mathbf{F}_{j,t}(\psi_{ij,t}, \rho_{ij,t}, \mathbf{M})$. The $n_{j,t} \times M$ matrix $\mathbf{F}_{j,t}$ has elements given by eq. (S2.9). $\mathbf{S}(\mathbf{u})$ is the $S_{j,t} \times n_{j,t}$ kernel matrix, and $A_{sj,t}$ is the effort, trap area $\times$ deployment time (fraction of fruiting season). The distributions for eq. (S2.4) and eq. (S2.5) are given in section S2.3. The prior distribution is discussed in section S2.5.

S2.3   *Dynamic interpretation*

The process is a multivariate dynamic (state-space) model for conditional fecundity and maturation, with a joint distribution of these latent states, $[\psi_{ij,t}, \rho_{ij,t}]$. Both state variables respond to environmental conditions.

Dynamics start with maturation. The maturation observation model recognizes uncertainty in the assignment of maturation (fruits are often unobservable in crowded canopies) and the fact that trees are not observed in many years. Let $z_{ij,t}$ be the observed status, which can be mature (fruits observed, $z_{ij,t} = 1$), uncertain (fruits not observed, canopy obscure), and immature (entire canopy visible in the fruiting season and fruits not observed, $z_{ij,t} = 0$). Observations can be missing in many years (Clark et al., 2004). $t_{ij,l}$ is the last year in which individual ij was observed to be immature. $t_{ij,m}$ is the first year ij was observed in the mature state. True maturation status is the indicator $\rho_{ij,t} \in \{0, 1\}$, with $\rho_{ij,t} = 1$ being the event that individual ij is mature in year t. Maturation is a one-way process, $[\rho_{ij,t+1} = 1|\rho_{ij,t} = 1] = 1$, and $[\rho_{ij,t} = 1|\rho_{ij,t+1} = 0] = 0$. Status is known to be mature any time after first observed to be mature and to be immature any time before the last time it is established to have been immature. Between these times, the status is

unknown and modeled with a probit:

$$
\begin{aligned}
z_{ij,t_m} = 0 &\to & \rho_{ij,t} = 0, \forall t \le t_m \\
z_{ij,t_l} = 1 &\to & \rho_{ij,t} = 1, \forall t \ge t_l \\
t_l < t < t_m &\to & [\rho_{ij,t} = 1 | \rho_{ij,t-1} = 0] = \Phi\left(\mathbf{v}'_{ijt}\beta^v\right)
\end{aligned}
\tag{S2.6}
$$

where t_l is an observation year earlier than t, and t_m is an observation year after year t. This relationship between $z_{ij,t}$ and $\rho_{ij,t}$ must be conditioned on past and future imputed states,

$$
\rho_{ij,t} \sim Bernoulli(p_{ij,t})
\tag{S2.7}
$$

where

$$
p_{ij,t} = \begin{cases} \rho_{ij,t-1} + (1 - \rho_{ij,t-1})\rho_{ij,t+1}\Phi\left(\mathbf{v}'_{ij,t}\boldsymbol{\beta}^v\right) & \text{missing } z_{ij,t} \\ z_{ij,t} & \text{otherwise} \end{cases}
$$

where $\boldsymbol{\Phi}(\cdot)$ is the standard normal distribution function. For Gibbs sampling, maturation states used to evaluate $p_{ij,t}$ are currently imputed states.

Conditional fecundity is continuous, $\psi_{ij,t} \in (0, \infty)$, and depends on maturation status,

$$
\psi_{ij,t} \in \begin{cases} (1, \infty) & \rho_{ij,t} = 1 \\ (0, 1] & \rho_{ij,t} = 0 \end{cases}
$$

Note that fecundity here means the capacity to produce at least one seed. For eq. (S2.5), mature individuals produce seeds at log fecundity

$$
\log(\psi_{ij,t}) \sim N\left(\mathbf{x}'_{ij,t}\beta^x, \sigma^2\right) I(\psi_{ij,t} \le 1)^{1-\rho_{ij,t}} I(\psi_{ij,t} \in \mathbf{c}_{ij,t})^{\rho_{ij,t}}
\tag{S2.8}
$$

If there is no fecundity observation for the tree year, with bounds $\mathbf{c}_{ij,t} = (c_{ij,t,min}, c_{ij,t,max})$, then $\mathbf{c}_{ij,t}|\rho_{ij,t} = 0$ is the interval $(0, 1)$, and $\mathbf{c}_{ij,t}|\rho_{ij,t} = 1$ is the interval $[1, c_{max})$, where the upper bound is set to a large value determined by prior knowledge for the species. Additional terms in the fecundity submodel are summarized in table S2.1.

S2.4   *Multiple seed types*

The observation model includes the uncertain assignment of seeds to species. Seed collections are classified as those that can be confidently assigned to a

species and those that can only be assigned to genus or even family. There are $h = 1, \ldots, H$ species that might potentially contribute to $m = 1, \ldots, M$ seed types in $H \times M$ matrix $\mathbf{M}$. For example, seed types in a data set might include three *Pinus* species plus a larger category, *Pinus* spp., which includes all seeds that could not be confidently identified to species. A prior seed-type composition matrix for $H = 3$ species might be organized like this:

$$\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3]'$$

$$= \begin{array}{cccc} P.\ echinata & P.\ taeda & P.\ virginiana & P.\ \mathrm{spp.} \end{array}$$

$$= \left[\begin{array}{cccc} 0.1 & 0 & 0 & 0.9 \\ 0 & 0.2 & 0 & 0.8 \\ 0 & 0 & 0.2 & 0.8 \end{array} \right] \begin{array}{l} h = P.\ echinata \\ h = P.\ taeda \\ h = P.\ virginiana \end{array}$$

A row $\mathbf{m}_h$ of matrix $\mathbf{M}$ represents the fraction of seeds produced by species h that are counted in each seed-type class m. The rows accommodate observation errors, the fraction of species-h seed that is misclassified as seed type m. The vector corresponding to individual i is designated with the notation $\mathbf{m}_{h[i]}$. The length-M fecundity vector on the seed-type basis is

$$\mathbf{F}_{ij,t} | \psi_{ij,t}, \rho_{ij,t} = \mathbf{m}_{h[i]} \rho_{ij,t} \psi_{ij,t} \tag{S2.9}$$

where $h[i]$ indicates the species for individual i. Of course, $\mathbf{M}$ must be estimated.

By Bayes' theorem we obtain from the fitted model the inverse $M \times H$ probability matrix $\mathbf{H}$ that an unknown seed of type m was produced by species h. The posterior estimate of total seed produced on the plot is $f_{hj,t} = \sum_{i \in h} \rho_{ij,t} \psi_{ij,t}$. Row m of the inverse matrix is

$$\mathbf{h}_{mj,t} = [h = 1, \ldots, H | \mathbf{m}_h]_{j,t} = \int \int \frac{\mathbf{m}_h f_{hj,t}}{\sum_h \mathbf{m}_h f_{hj,t}} [f_{hj,t}, \mathbf{m}_h] df_{hj,t} d\mathbf{m}_h \tag{S2.10}$$

where the integral is taken over the posterior distribution of $f_{hj,t}$ and $\mathbf{m}_h$.

S2.5 *Prior parameter values*

The prior parameter distribution depends on which components are included in the model. *Maturation* and *conditional fecundity* parameters have the

prior distributions $MVN(\boldsymbol{\beta}^v|\mathbf{0}, 10 \times \mathbf{I}_{q^v})$, and $MVN(\boldsymbol{\beta}^x|\mathbf{0}, 10 \times \mathbf{I}_q)$, both non-informative–design matrices are centered and standardized, so the prior is 10 standard deviations. Random effects have the prior distribution $\boldsymbol{\beta}_i^w \sim MVN(\mathbf{0}, \mathbf{B}_w)$ and $\mathbf{B}_w \sim IW(\mathbf{I}_{q^w}, df)$ with degrees of freedom $df = q^w + \sqrt{n_+}$, the second term being rounded to the next integer.

Year effects have the prior distribution $\prod_t N(\gamma_t|0, 10)$. For random groups, $\gamma_{g,t} \sim N(\gamma_t, \tau_t^2)$, with group variance $\tau_t^2 \sim IG(2,1)$. *Autoregressive lag terms* have the prior distribution $(\alpha_1, \ldots, \alpha_L) \sim MVN(\mathbf{0}, \mathbf{I}_p)$. If there are G random groups, they have the prior $(\alpha_{g,1}, \ldots, \alpha_{g,L}) \sim MVN(\mathbf{0}_G, \mathbf{A}_\alpha)$, with $IW(\mathbf{A}_\alpha|\mathbf{I}_p, p+1)$.

Seed composition vectors $\mathbf{m}_h$ in the observation error matrix $\mathbf{M}$ have a Dirichlet prior distribution.

Dispersal parameters have the prior distribution $N(u_h|\tilde{u}, \tilde{U})$. The next stage is $\tilde{u} \sim N(u^*, U^*)$ and $\tilde{U} \sim IG(v_1, v_2)$, where parameter values (u^*, U^*, v_1, v_2) are selected based on understanding of the species dispersal properties.

S2.6   *In- and out-of-sample prediction*

Seed data are generated by dispersal from mature and fecund trees. Data can be predicted either from the estimates of these latent states and dispersal or from the full posterior distribution of parameters. Let $\boldsymbol{\theta}_1 = \{\boldsymbol{\psi}, \boldsymbol{\rho}, \mathbf{u}, \mathbf{M}\}$ be the set of latent states, the dispersal kernel, and error matrix. The predictive distribution of seed data $\mathbf{Y}^*$ is

$$[\mathbf{Y}^*] = \int [\mathbf{Y}^*|\mathbf{F}(\boldsymbol{\psi}, \boldsymbol{\rho}), \mathbf{S}(\mathbf{u}), \mathbf{M}] \, [\boldsymbol{\theta}_1|\mathbf{Y}, \mathbf{X}, \mathbf{V}, \mathbf{z}, \mathbf{c}] d\boldsymbol{\theta}_1 \qquad (S2.11)$$

From right to left, the interpretation of the integrand is i) the posterior estimates of states, dispersal, and observation error, and ii) the likelihood (variables are defined in table S2.1). This prediction is available for trees and years that are part of the model fitting. It can be used to predict seed at locations where seed rain was not observed on observed plots and years. It cannot be used to predict outside sample plots and years, because it requires estimates for the states of the trees themselves.

Predicting out-of-sample plots and years requires the full posterior distribution of parameters $\boldsymbol{\theta}_2 = \{\boldsymbol{\beta}^x, \boldsymbol{\beta}^v, \sigma^2, \mathbf{u}, \mathbf{M}\}$, marginalizing the latent states and parameters related to seed observations, $\{\boldsymbol{\psi}, \boldsymbol{\rho}\}$,

$$[\mathbf{Y}^*|\mathbf{X}^*, \mathbf{V}^*] = \int [\mathbf{Y}^*|\mathbf{F}(\boldsymbol{\psi}, \boldsymbol{\rho}), \mathbf{S}^*(\mathbf{u}), \mathbf{M}]$$
$$\times [\psi|\boldsymbol{\beta}^x, \sigma^2, \mathbf{X}^*] [\rho|\boldsymbol{\beta}^v, \mathbf{V}^*]$$
$$\times [\boldsymbol{\theta}_2|\mathbf{Y}, \mathbf{X}, \mathbf{V}, \mathbf{z}, \mathbf{c}] \, d\boldsymbol{\theta}_2 d\boldsymbol{\psi} d\boldsymbol{\rho} \qquad (S2.12)$$

From right to left, we have the posterior distribution of parameters, maturation status, conditional fecundity, and likelihood. Predictions from eq. (S2.11) will be more accurate than eq. (S2.12), but, again, eq. (S2.11) is only available for in-sample prediction, which relies on the capacity to predict maturation and fecundity from environmental variables as opposed to simply estimating them and on scenarios for $\mathbf{X}^*$ and $\mathbf{V}^*$. The fully generative eq. (S2.12) can be used for fore- and backcasting.

S3   DATA SIMULATION

The simulator takes numbers of plots, trees, traps, and years as means for stochastic generation of sample size. The simulator follows these steps:

- Generate random species identities, random diameters, and random locations for trees. Generate random locations of traps and evaluate the $S \times n$ distance matrix for traps to trees.

- Draw coefficients for maturation, conditional fecundity, and dispersal parameters from a range that can generate patterns like observed data. Both maturation and conditional fecundity have an intercept and a slope for log diameter.

- Draw maturation status of each tree year from the probit submodel, subject to the constraint that this is a one-way transition (eq. (S2.6)).

- For mature individuals, draw conditional fecundity from eq. (S2.8).

- Assuming that only a fraction of seeds can be assigned to species, construct the $\mathbf{m}_h$ vector for each species h and distribute seed production for each individual according to its species identity.

- Evaluate expected densities at seed trap locations (eq. (1)).

11

- Draw seed counts from the likelihood.

Code is provided in the Supplemental vignette as part of the package MAS-TIF. For fig. 5, the specific algorithm is:

- Simulate a $n \times T$ fecundity matrix $\mathbf{F} \sim MVN(\boldsymbol{\psi}, \mathbf{C}_n)$.

- Simulate random tree coordinates $\mathbf{s}^{\mathcal{N}} = (\mathbf{s}_{i1}, \mathbf{s}_{i2})_{i=1}^{n}$ and random trap coordinates, $\mathbf{s}^{\mathcal{S}} = (\mathbf{s}_{s1}, \mathbf{s}_{s2})_{s=1}^{S}$.

- Generate a distance matrix from the tree locations to the trap locations and evaluate a corresponding kernel matrix $\mathbf{S}(|\mathbf{s}^{\mathcal{S}} - \mathbf{s}^{\mathcal{N}}|; u)$ with dispersal parameter u.

- Evaluate intensities $\boldsymbol{\Lambda}_{\mathcal{S}}|\mathbf{f}_{\mathcal{N}}$ from eq. (1) and draw counts from $y_{s,t} \sim Poi(\lambda_{s,t})$. The variance in simulated $y_{s,t}$ is the horizontal axis in $fig.$ 5. The vertical axis shows the prediction from eq. (10) and its two terms.

S4   Means and covariances

This section summarizes means and covariances observed by a consumer in the canopy and on the forest floor.

S4.1   *Space-lag covariance*

The *Tree-time and space-time covariance* of the main text starts with a space-time covariance matrix and moves immediately to summaries thereof. Here we provide additional explanation that is relevant to the interpretation of fig. 9. In this section a matrix $\mathbf{L}$ can represent either host tree covariance $\mathbf{C}$ or spatial covariance $\mathbf{G}$.

We cannot estimate a $nT \times nT$ space-time covariance $\mathbf{L}$, because each (i, i', t, t') combination is observed once. However, we can estimate a lag matrix $\mathbf{L}$ consisting of vector blocks and elements $L_{ii',l}$ for lag $l = 1, l = -p, \ldots, 0, \ldots, p$. For the simplest example of two host trees or locations and one lag $(i = 1, i' = 2, p = 1)$, this structure is

$$\mathbf{L} = \left[\begin{array}{cc} \mathrm{AC}_1 & \mathrm{CC}_{12} \\ \mathrm{CC}_{21} & \mathrm{AC}_2 \end{array} \right] = \left[\begin{array}{ccc|ccc} L_{11,-1} & L_{11,0} & L_{11,+1} & L_{12,-1} & L_{12,0} & L_{12,+1} \\ L_{21,-1} & L_{21,0} & L_{21,+1} & L_{22,-1} & L_{22,0} & L_{22,+1} \end{array} \right]$$

$$\text{(S4.13)}$$

Blocks contain elements $L_{ii',l}$. Blocks along the diagonal are the symmetric auto-covariance (AC) vectors, $L_{ii,-l} = L_{ii,+l}$. Off-diagonal blocks are cross-covariance vectors (CC), which are asymmetric, because they depend on which individual leads the other. However, each element in $CC_{ii'}$ has a counterpart in its mirror block $CC_{i'i}$, where $L_{ii',-l} = L_{i'i,+l}$. In other words, $\mathbf{C}_L$ becomes block symmetric upon reversing the individual labels and the signs of lag indices in the upper (or lower) block.

Matrix $\mathbf{L}$ offers all combinations of tree-tree-lag covariances. Tree-to-tree covariances for a given lag are held in off-diagonal blocks (CC_{21} in eq. (S4.13)). Center elements of the off-diagonal blocks hold the covariances between trees in the same year ($L_{21,0}$). These elements determine the resource heterogeneity that consumers abide by moving between trees, depending on their foraging ambits. The masting phenomenon is identified with episodic seed production in lagged year l', ($L_{ii,l'} > 0, l' > 1$) and synchronicity ($L_{ii',0} > 0$). The plots in fig. 9 show columns of $\mathbf{L}$ for sequential lags $0, 1, \ldots$.

Tree-to-lag covariances determine the advantages of movement between years, as when a good year for this tree might be followed by a good year for a different tree next year ($L_{ii',\pm1} > 0$). A consumer might abide pulsed availability ($L_{ii,\pm1} < 0$) if it has the capacity to switch hosts, and hosts are not synchronized ($L_{ii',0} < 0$). The same principles apply to ground foragers with limited spatial ambit. We use the lag matrix to evaluate the advantages of movement between host trees between years.

S4.2 *Entropy*

To quantify the heterogeneity experienced by a consumer across trees, space, years, and hosts we evaluated entropy,

$$\frac{1}{2}\left[\log(2\pi) + \frac{\log|\mathbf{V}|}{d}\right] \tag{S4.14}$$

where $|\mathbf{V}|$ is the determinant of a $d \times d$ matrix $\mathbf{V}$ that represents one of the covariance matrices $\mathbf{C}_L$ (tree-to-tree over years), $\mathbf{C}_n$ (year-to-year over trees), $\mathbf{G}_T$ (spatial over years) or $\mathbf{G}_S$ (year-to-year over space). This entropy calculation is reported on a per-dimension basis (per d) due to the multiplicative effect of integrating volume over d dimensions.

S5 STABILITY OF THE AR(p) PROCESS

The sequence of fecundities for each tree can be evaluated for its stability properties. Eigenvalues for time series quantify the tendency for quasi-periodic behavior. For the AR(p) model, we construct the matrix

$$
\mathbf{E} = \begin{bmatrix}
\psi_1 & \psi_2 & \ldots & \psi_{p-1} & \psi_p \\
1 & 0 & \ldots & 0 & 0 \\
0 & 1 & \ldots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & 0 & 0 \\
0 & 0 & \ldots & 1 & 0
\end{bmatrix}
$$

(West and Harrison, 1997). If there are random groups, then there is a matrix $\mathbf{E}_g$ for each group, combining the (global) fixed effect and the random effect for that group. A quasi-periodic process has complex eigenvalues. A stationary AR(p) process has all eigenvalues of less than unit modulus (real plus imaginary parts). Eigenvalues are evaluated for the fecundity of each mature tree and summarized by plot and species. Eigenvalues for the AR(4) model are shown in fig. S5.1.

S6 DISPERSAL KERNEL

Seeds arriving in seed traps need not all derive from within the inventory plot (Clark et al., 1998; Muller-Landau et al., 2008; Clark et al., 2004). This possibility suggests an intercept proportional to basal area in eq. (1) as a rough accommodation of long-distance dispersal. We do not include an intercept here, because it can have a large impact on estimates, without actually being sensitive to seeds outside the plot. An intercept can have a large impact, because it provides an alternative to the dispersal kernel anytime data are noisy, which is always the case when seed recovery is low. It is insensitive to LDD otherwise, because the tail of the kernel has no impact on estimates when seed except in cases where seed is rare (Clark et al., 1999).

An approximate influence of LLD is available from eq. (S2.1), with mean dispersal distance obtained by integrating arc-wise and with distance,

$$
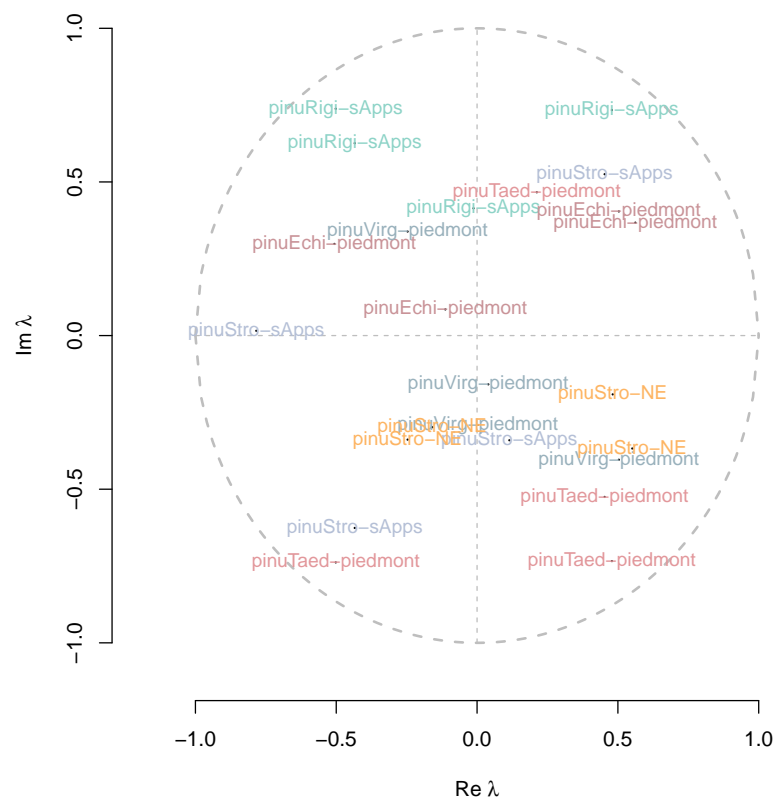\bar{d} = \frac{\pi \sqrt{u}}{2} \tag{S6.15}
$$

14

Figure S5.1: Eigenvalues, real and imaginary parts, lie within the unit circle.

Table S6.2: Data summary by plot for the southeastern Piedmont (DFBW, DFEE, DFEW, DFHW), the southern Appalachians (CW118, CW218, MHF, MHP), and New England (HFBW, HFS).

	CW118	CW218	DFBW	DFEE	DFEW	DFHW	HFBW	HFS	MHF	MHP	total
Trees											
pinuEchi	0	0	47	0	20	1	0	0	0	0	68
pinuRigi	32	0	0	0	0	0	0	0	0	0	32
pinuStro	3	29	0	0	0	0	172	143	82	24	453
pinuTaed	0	0	350	135	66	192	0	0	0	0	743
pinuVirg	0	0	10	7	11	1	0	0	0	0	29
Tree-years											
pinuEchi	0	0	1169	0	388	26	0	0	0	0	1583
pinuRigi	1062	0	0	0	0	0	0	0	0	0	1062
pinuStro	79	715	0	0	0	0	2332	1793	1552	464	6935
pinuTaed	0	0	8680	2413	1250	4870	0	0	0	0	17213
pinuVirg	0	0	249	126	212	26	0	0	0	0	613
Seeds											
pinuRigi	219	7	0	0	0	0	0	0	0	0	226
pinuStro	0	0	0	0	0	0	68	15	5	11	99
pinuTaed	0	0	0	0	0	0	0	0	0	2	2
pinuUNKN	190	9	20479	49	1419	8853	0	0	520	179	31698
Traps	20	20	148	9	69	66	36	36	35	36	475
Trap-years	520	520	2516	90	828	1188	216	180	420	432	6910

Seed arriving at a location from beyond distance R is

$$S(R) = \int_R^\infty \oint_{2\pi} \frac{u}{\pi(u+r^2)^2} dr = 2u \int_R^\infty \frac{r}{(u+r^2)^2} dr = \frac{u}{u+2R^2} \quad \text{(S6.16)}$$

(Clark et al., 1999), and a distance-specific intensity

$$l(r) = -\frac{d\log S}{dr} = \frac{4r}{u+2r^2} \quad \text{(S6.17)}$$

For perspective, if the average seed of a well-dispersed species moves 10 m from the parent in a forest of infinite area that is everywhere the same in composition, then $< 1\%$ of all seed derives from beyond 50 m. The insensitivity to distant sources is shown in the analysis of (Clark et al., 1998). It is consistent with the low scores associated of seed traps that collect few seeds.

In cases where fleshy-fruits are dispersed by songbirds, there is typically still substantial dispersal close to the source.

In addition to an undesirable effect on estimates, the intercept is costly. If we want to estimate how much seed might be arriving at a location from outside a plot, we would not integrate arcwise, but rather would require Reimann sums for areas that differ in every direction from every seed trap. They would necessarily be truncated at some distance beyond which inputs would be assumed negligible. For MCMC, this would require reevaluation at each iteration, an extremely costly algorithm.

S7 DATA SUMMARY

Data come from 11 inventory plots initiated between the years of 1991 and 2011 and sampled until the present. Habitats range from southeastern Piedmont mixed pine-hardwoods (DFEE, DFEW, DFBW, BFHW), Appalachian foothills (CW118, CW119, MHF, MHP), and New England mixed hardwoods (HFBW, HFS). Numbers of trees, tree-years, seed and cone types, seed traps, trap-years are listed in table S6.2. The plots from table S6.2 are shown with seed production in fig. S7.2. As a fraction of basal area, seed production is comparable across plots, evaluated as

$$\frac{\text{seeds} \, \frac{1}{\text{ha yr}}}{\text{BA} \, \frac{\text{m}^2}{\text{ha}}} = \frac{\text{seeds}}{\text{BA m}^2\text{yr}} \quad \text{(S7.18)}$$

In portions of plots with abundant trees, seed productivity extends to > 1000 seeds per m^2 of basal area.
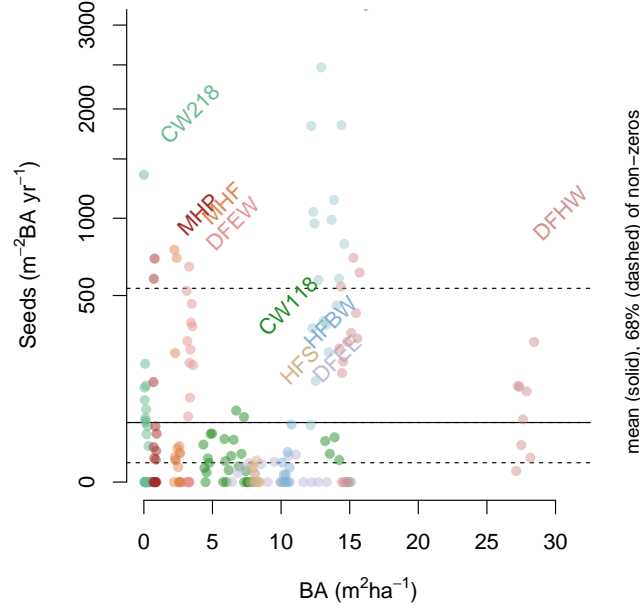

Figure S7.2: Each dot represents a single plot-year mean, evaluated from eq. (S7.18). Horizontal lines identify the median and ± 1 standard deviation of non-zero values.

S8 DIAGNOSTICS

S8.1 *Simulated data*

Plots for simulation examples include coefficients, latent states, and predictions. An example MCMC chain for the dispersal parameter in the example described in the main text is shown converging to $u = 279 \pm 1.33$, CI $= (276, 282)$ m^2 in fig. S8.3. The value used to simulate data is 253 m^2. Inaccuracies are expected due to the small number of trees of a given species per plot, ranging from 4 to 23 in this example. This concentration of seed production into a few individuals is typical of real data.
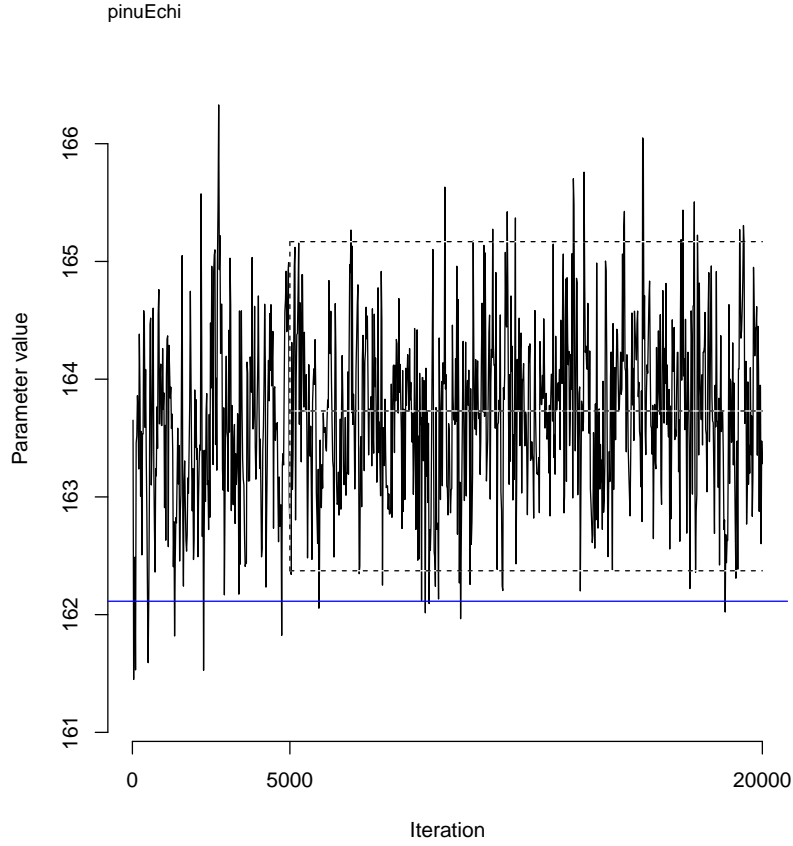
Figure S8.3: MCMC chain for the dispersal parameter fitted to data simulated with $u = 162$ in the simulation example. Outlined are the post-burnin values used for estimates.

Maturation, fecundity, and seed composition parameters are plotted against true values in fig. S8.4. Wide credible intervals in fig. S8.4a, b reflect the stochasticity and small numbers of trees on each plot. The error matrix $\mathbf{M}$ is recovered precisely (fig. S8.4c).

Relatively wide predictive intervals for seed counts from the full posterior (fig. S8.5a) marginalize the uncertain states. As expected, the estimated fecundities and maturation statuses predict seed data better (fig. S8.5b) than does the full posterior fig. S8.5a. The fecundity estimates recover the true values (fig. S8.5c).

The full MASTIF package is available in R for experiments with simulated data, showing posterior estimates, predictive diagnostics, DIC, root mean square predictive error (RMSPE), and mapped estimates of latent states.
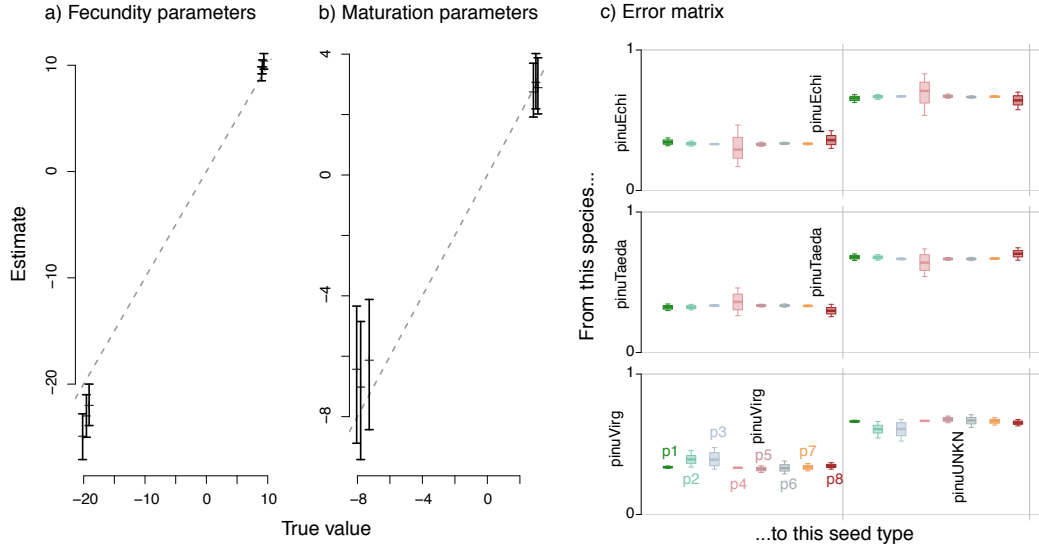
19

Figure S8.4: Parameter estimates plotted against values used to simulate data for conditional fecundity (a) and maturation (b) of three species. In both panels, low values are intercepts and high values are slopes. c) ADD TO FIG. Estimates of the composition vectors $\mathbf{m}_h$ in simulation, where h refers to three species that can be classified as $M = 4$ seed types. For each species, only 20% were recorded to the correct species, the remainder counted as "unknown". Model fitting recovers estimates precisely 2/3 of seed in the undifferentiated genus class.

The summary diagnostics provided here can be supplemented with direct experiments by the user.

S8.2 *Application*

The $H \times M$ error matrix $\mathbf{M}$ has only a few fitted entries (fig. S8.6a), because most are unambiguous, as when a plot comes from a location where only one species occurs or all seeds are identified only to genus level (DFBW, DFEE, DFEW, DFHW). Exceptions are CW118, where *P. rigida* occurs, but there are also *P. strobus* in the region, and MHF and MHP, where *P. strobus* occurs, but other pines occur in the region. The estimates in fig. S8.6a are elements $\mathbf{M}_{h,m}$, the probability that a seed produced by species h (rows) is counted as species m (columns).

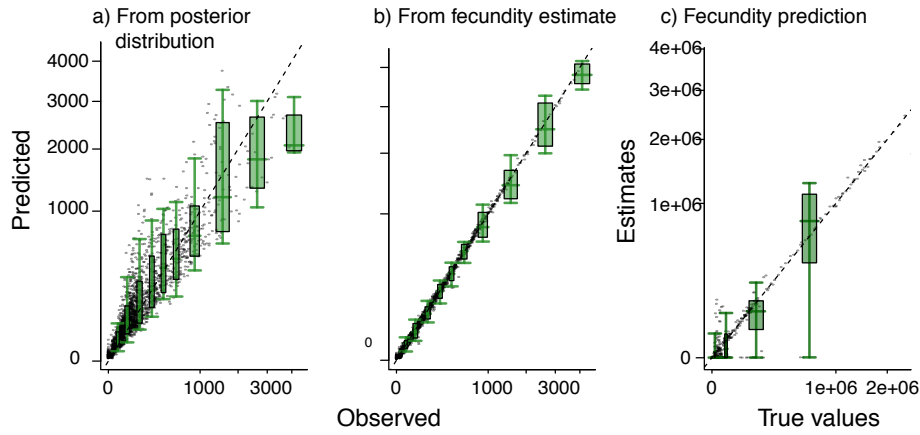The probability that a random seed of type m ('unknown') was produced

Figure S8.5: Seed data prediction from the full posterior distribution (a) and from the maturation/fecundity estimates (b) compared with simulated observations ('Observed') for the simulated example. In (c) is the comparison of fecundity estimates with values used to simulate data ('True values').
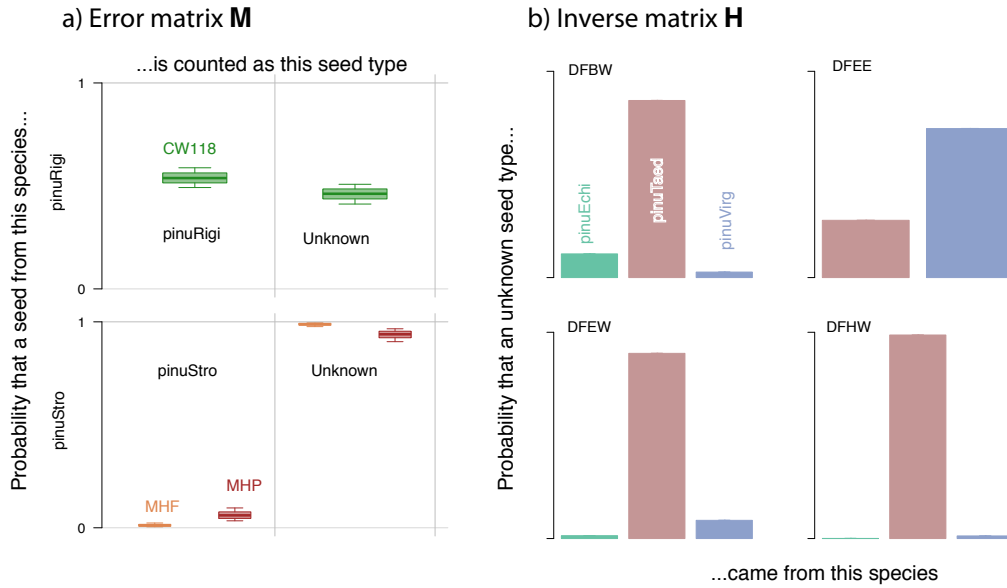

Figure S8.6: a) The error matrix **M** shows the probability that a seed produced by each species (rows) will be counted as each seed type (column). b) The inverse probability in **H** is shown as the probability that an unknown seed type m (rows) was produced by species h (columns).

by species h combine error estimates in **M** with species abundances on each plot. On the plots where all seeds are identified to genus level, the $M \times H$ matrix **H** is ambiguous. **H** is the inverse probability, obtained by Bayes' theorem (eq. (S2.10)), that a random seed of type m was produced by species

h. The credible intervals are visible for **M** in fig. S8.6a, but too narrow is see for **H** in fig. S8.6b.

As for the simulated example, seed prediction from the fitted model shows the additional uncertainty marginalized from the estimates latent fecundity and maturation states (fig. S8.7).
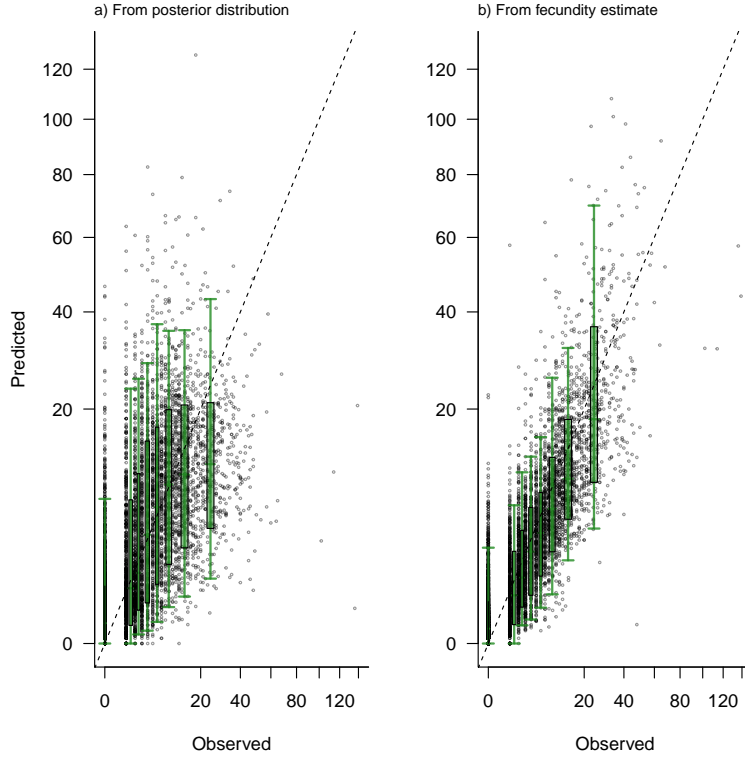

Figure S8.7: Seed prediction from the AR(4) model fitted to *Pinus*. Predictions are from the full posterior distribution (a) and from the fecundity estimates (b).

Random coefficients associated with individual trees are shown in fig. S8.8. The random effects are not centered on zero, because they have been added to the fixed effect, thus allowing more informative species comparisons. The range of variation differs by species and is greatest for the effects of tree diameter, in large part due to the limited representation of small trees in the data for this early successional genus.

Model comparisons and variable selection by DIC, RMSPE, and prediction scores available in the MASTIF package are not the focus here, where the attention centers on description of the basic model. Full MCMC chains are available for convergence diagnostics.
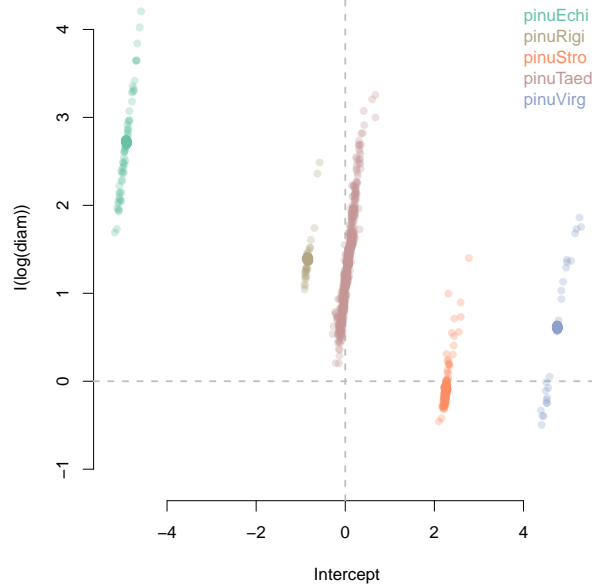
Figure S8.8: Random slopes and intercepts from the coefficient vector $\boldsymbol{\beta}^w$ for the AR(4) model.

S9 POSTERIOR SIMULATION: ALGORITHM NOTES

S9.1 *Joint fecundity and maturation updates*

To impute states, fecundity and maturation status are either proposed jointly (this section) or conditional fecundity of currently imputed mature individuals is updated separately (next section). For joint updates, we factor maturation and conditional fecundity as

$$[\psi^*_{ij,t}, \rho^*_{ij,t}] = [\psi^*_{ij,t}|\rho^*_{ij,t}][\rho^*_{ij,t}] \tag{S9.19}$$

Maturation year is a random walk, centered on the currently imputed maturation year and subject to constraints imposed by observed or currently imputed states (mature individuals cannot become immature, mature individuals remain mature). Possible maturation years range from the last year in which an individual was observed in the immature state to the first year in which that individual was observed in the mature state (section S2.3). If there are no maturation/fecundity observations, then all information on maturation status comes through the seed data. Maturation is proposed and accepted jointly with fecundity for all trees on a plot year, with proposals subject to this one-way constraint on maturation (section S2.3). This blocking

23

is necessitated by the fact that the likelihood for each trap-year conditionally depends on all tree-years for that plot (Clark et al., 2004). Sampling of a plot-year block is summarized this way:

$$[\boldsymbol{\psi}_{j,t}, \boldsymbol{\rho}_{j,t} | \mathbf{Y}_{j,t}, \mathbf{z}_{j,t}, \{\mathbf{c}_{ij,t}\}_{i=1}^{n_j}, \mathbf{M}] \propto P_1 \times P_2 \times P_3 \qquad \text{(S9.20)}$$

where

$$P_1 = \prod_{s=1}^{S_j} \prod_{m=1}^{M} Poi\left(y_{smj,t} | A_{sj} \lambda_{smj,t}(\boldsymbol{\psi}_{j,t}, \boldsymbol{\rho}_{j,t}, \mathbf{u}, \mathbf{M})\right)$$

$$P_2 = \prod_{i=1}^{n_i} [\mathbf{z}_{ij} | \rho_{ij,t}][\rho_{ij,t} | \mathbf{v}'_{ij,t} \boldsymbol{\beta}^v]$$

$$P_3 = \prod_{i=1}^{n_i} [\mathbf{c}_{ij,t} | \psi_{ij,t}][\psi_{ij,t} | \mathbf{x}'_{ij,t} \boldsymbol{\beta}^x, \rho_{ij,t}]$$

(Conditioning on parameters is omitted to improve clarity.) P_2 is given by eq. (S2.7), and P_3 is given by eq. (S2.8).

Two methods are used, depending on the process model, both beginning with proposed $\{\rho^*_{ij,t} | \rho_{j,t-1}, \rho_{ij,t+1}\}_{i=1}^{n_j}$. Only cases where $\rho_{ij,t-1} = 0$ and $\rho_{ij,t+1} = 1$ can change maturation state. If the current maturation year is t', then a new maturation year is sampled from $(t'-1, t'+1)$ with probability 0.5. Fecundity is then proposed for all imputed maturate ij from a normal distribution, with censoring imposed by the proposed $\rho^*_{ij,t}$ (eq. (S2.8)). Here are the two methods:

Method 1:

1. The proposed $\psi^*_{ij,t} | \rho^*_{ij,t}, \psi_{ij,t}$ come from a normal distribution centered on the currently imputed $\psi_{ij,t}$ and censored as in eq. (S2.8).

2. Accept/reject $\boldsymbol{\rho}^*_{j,t}, \boldsymbol{\psi}^*_{j,t}$ as a block with probability $P_1 P_2 P_3$.

Method 2:

1. The proposed $\psi^*_{ij,t} | \rho^*_{ij,t}$ come from the conditional normal distribution obtained from P_3. We limit details to the most complex case of the AR(p) model (section S9.7.5).

2. Accept/reject $\boldsymbol{\rho}^*_{j,t}, \boldsymbol{\psi}^*_{j,t}$ as a block with probability $P_1 \times P_2$.

S9.2 *Hamiltonian updates of conditional fecundity*

Hamiltonian updates accelerate mixing for continuous states, but they cannot be used with discrete states, in this case maturation status. However, mixing of fecundity is accelerated with Hamiltonian updates for currently imputed mature individuals. Each observation is a length-M vector $\mathbf{y}_{sj,t}$. Element m has Poisson intensity $A_{sj}\lambda_{smj,t} = A_{sj}\sum_{i=1}^{n_{j,t}}\mathbf{S}_{[sjt,i]}e^{\psi_{ij,t}}\mathbf{m}_{i[h]}$, where $\mathbf{m}_{i[h],m}$ is the row vector $\mathbf{m}_h$ corresponding to the species of individual i and the column for seed type m. The Hamiltonian can be written as

$$H(\psi_{ij,t}, p) = B(\psi_{ij,t}) + C(p) \tag{S9.21}$$

where $C(p) = \sum_{i=1}^{n_j}\frac{p_i^2}{2m_i}$ is the kinetic energy, taken as a quadratic function of momentum variables m_i, which are tuned to optimize performance (Neal, 2011). The first term incorporates the conditional distribution,

$$B(\psi_{ij,t}) = -\log[\psi_{ij,t}|\mathbf{y}_{j,t}, \mu_{ij,t}, \sigma^2)] \tag{S9.22}$$

$$\propto \sum_{m,s}\left(-y_{smj,t}\log\lambda_{smj,t} + A_{sj}\lambda_{smj,t}\right) + \frac{1}{2\sigma^2}(\psi_{ij,t} - \mu_{ij,t})^2 \tag{S9.23}$$

The gradient is used to direct proposals efficiently:

$$\frac{\partial B}{\partial\psi_{ij,t}} = e^{\psi_{ij,t}}\sum_m\mathbf{m}_{h[i],m}\sum_s\mathbf{S}_{[sjt,i]}\left(-\frac{y_{smj,t}}{\lambda_{smj,t}} + A_{sj}\right) + \frac{1}{\sigma^2}(\psi_{ij,t} - \mu_{ij,t}) \tag{S9.24}$$

Hamiltonian updates are individually slow, but affect larger steps then a Metropolis random walk, especially with large data sets. The two methods are mixed stochastically in the Gibbs sampler.

S9.3 *Fecundity coefficients*

Direct sampling of coefficients in $\boldsymbol{\beta}^x$ and $\boldsymbol{\beta}^v$ is available from Gaussian conditional posterior distributions. Gaussian prior distributions are non-informative. For $\boldsymbol{\beta}^x$ conditional distributions marginalize random effects (see below). The variance σ^2 has an inverse gamma prior – and is sampled directly from the conjugate inverse gamma posterior.

S9.4 *Random individual effects*

Let $\mathbf{w}_{ij,t}$ be a design vector holding all or some of the columns in $\mathbf{x}_{ij,t}$. There is an individual-effects coefficient matrix $\boldsymbol{\beta}^w_{ij}$,

$$\psi_{ij,t} \sim N\left(\mathbf{x}'_{ij,t}\boldsymbol{\beta}^x + \mathbf{w}'_{ij,t}\boldsymbol{\beta}^w_{ij}, \sigma^2\right) \tag{S9.25}$$

The prior distribution includes

$$\boldsymbol{\beta}^w_{ij}|\mathbf{B}_w \sim MVN(\mathbf{0}, \mathbf{B}_w) \tag{S9.26}$$

$$\mathbf{B}_w \sim IW\left(\mathbf{B}^w, df\right) \tag{S9.27}$$

where $df = Q^w + 2$, Q^w is the number of columns in $\mathbf{w}_{ij,t}$, and $\mathbf{B}^w = \mathbf{I}_r$ is a prior diagonal matrix. The conditional posterior matrix is

$$\boldsymbol{\beta}^w_{ij}|\boldsymbol{\beta}^x, \mathbf{B}_w \sim MVN(\mathbf{V}_{ij}\mathbf{v}_{ij}, \mathbf{V}_{ij}) \tag{S9.28}$$

where

$$\mathbf{v}_{ij} = \frac{1}{\sigma^2} \sum_{t \in \{t_i\}} \mathbf{w}_{ijt}(\psi_{ijt} - \mathbf{x}'_{ijt}\boldsymbol{\beta}^x) \tag{S9.29}$$

$$\mathbf{V}_{ij} = \frac{1}{\sigma^2} \sum_{t \in \{t_i\}} \mathbf{w}_{ijt}\mathbf{w}'_{ijt} + \mathbf{B}_w^{-1} \tag{S9.30}$$

The summations are taken over all observation years for an individual i, the set $\{t_i\}$ for which the individual is mature. Here is the conditional for the covariance,

$$\mathbf{B}_w|\{\boldsymbol{\beta}^w_{ij}\} \sim IW\left(\sum \boldsymbol{\beta}^w_{ij}\boldsymbol{\beta}^{w'}_{ij} + df \times \tilde{\mathbf{B}}, \sum_j n_j + df\right) \tag{S9.31}$$

where $\tilde{\mathbf{B}}$ is the prior covariance.

S9.5 *Random groups in the year and AR(p) models*

The year and AR(p) models allow for group random effects on year and lag coefficients, respectively–if groups are defined by the user, they will be treated as random. This is done, because year and lag terms across groups

are highly unbalanced. Plot-species groups can hold different numbers of plots and trees in different years. Plots can be established at different times, have different plot areas, and support very different communities of species. For a given species, abundance across plots may range from zero to high. Within plots, numbers of mature individuals vary across years with recruitment, maturation, and death. Within posterior simulation, their imputed maturation statuses change by tree and year. The sizes of design matrices are thus dynamic.

Given this imbalance, treating groups as random provides the advantage that no arbitrary rules are needed to catch computation errors that would result from plot-years that are at some iterations imputed to have mature trees and other iterations not.

S9.6 *Year effects*

For a single group, years effects are fixed, drawn from the conditional

$$\gamma_t \sim N(V_t v_t, V_t) \tag{S9.32}$$

$$V_t^{-1} = \frac{n_t}{\sigma^2} + 1/\tau^2 \tag{S9.33}$$

$$v_t = \frac{1}{\sigma^2} \sum_i (\psi_{i,t} - \mu_{i,t}) \tag{S9.34}$$

With multiple groups, there are random year effects across groups:

$$\gamma_{g,t} \sim N(V_t v_t, V_t) \tag{S9.35}$$

$$V_{g,t}^{-1} = \frac{n_{g,t}}{\sigma^2} + 1/\tau_t^2 \tag{S9.36}$$

$$v_{g,t} = \frac{1}{\sigma^2} \sum_{i \in g} (\psi_{i,t} - \mu_{i,t} - \gamma_t) \tag{S9.37}$$

Years have a sum-to-zero constraint imposed in Gibbs sampling. The intercept for a given year is the overall intercept plus the year effect for that year. The variance for random effects is

$$\tau_t^2 \sim IG\left(2 + \frac{G_t}{2}, 1 + \frac{1}{2} \sum_g \gamma_{g,t}^2\right) \tag{S9.38}$$

where G_t is the number of groups available in year t, i.e., those having mature individuals in that year.

S9.7   *AR(*p*) model*

The AR(p) model allows for the dependence of the current states of ψ_t on p previous states. The process is homogeneous in time, because the lag coefficients α_p are constant. I start with a few words on structure.

S9.7.1   *Imputed past, predicted future*

AR models handle the early years in different ways. There is no AR(p) estimate for years $t \in \{1, \ldots, p\}$. One of the more common ways to deal with these years is to simply condition on them. This seems like a big price to pay. Because MASTIF is a state-space model, and we are imputing fecundity and maturation anyway, it makes sense to impute fecundity/maturation for years $t - p, \ldots, t - 1$.

So while we are imputing the past, it makes sense to predict the future. Conditionally, fecundity in the final year T_i depends on the future, up to year $T_i + p$. To accommodate past and future, MASTIF imputes backward p years from the first observation and predicts forward p years beyond the last observed year.

A consistent treatment would appear to demand that AR effects be restricted to individuals that have been mature for the past $t - p$ years. Note that this would not be a concern if maturation state was known. We adopt this rule, so lag effect estimates are not biased downward by the inclusion of trees that might have immature during one or more of years $t - p, \ldots, t - 1$. Although fecundity is imputed for all years, including before observations began, sampling of coefficients for fixed effects and lag effects is restricted to years in which trees were observed and mature.

S9.7.2   *AR(p) model structure*

To avoid further notation, the description that follows applies only to tree-years for which the mature state extends back to $t - p$ years. Also to simplify notation we omit the subscript j. Note that multiple plots j might fall within a group g.

Conditionally, the model for an individual i in group $g \in \{1, \ldots, G\}$ can be written as

$$\psi_{ig,t}|\mu_{ig,t}, \boldsymbol{\alpha}, \boldsymbol{\alpha}_{g[i]}, \tilde{\boldsymbol{\psi}}_{ig,t} \sim N\left(m_{ig,t}, \sigma^2\right) \tag{S9.39}$$

where

$$m_{ig,t} = \mu_{ig,t} + \sum_{l=1}^{p} (\alpha_l + \alpha_{g[i]l})\psi_{ig,t-l} \tag{S9.40}$$

$$= \mu_{ig,t} + (\boldsymbol{\alpha} + \boldsymbol{\alpha}_{g[i]})' \tilde{\boldsymbol{\psi}}_{ig,t} \tag{S9.41}$$

$\mu_{ig,t} = \mathbf{x}'_{ig,t}\boldsymbol{\beta}^x$ is the fixed effect, $\tilde{\boldsymbol{\psi}}_{ig,t} = (\psi_{ig,t-1}, \ldots, \psi_{ig,t-p})'$ is the vector of lagged fecundities for (ig, t), $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)'$ is the vector of fixed effects for lag $l = 1, \ldots, p$, and $\boldsymbol{\alpha}_g = (\alpha_{g1}, \ldots, \alpha_{gp})$ is the random effect for group g with prior distribution

$$\boldsymbol{\alpha}_g \sim MVN(\mathbf{0}, \mathbf{A}_\alpha) \tag{S9.42}$$

$$\mathbf{A}_\alpha \sim IW\left(\tilde{\mathbf{A}}_\alpha, df\right) \tag{S9.43}$$

To facilitate sampling, the fecundity values are organized into a vector $\boldsymbol{\psi} = \{\psi_{ig,t}|i = 1, \ldots, n, g = 1, \ldots, G, t = 1, \ldots, T_i\}$ and a corresponding matrix of p lag terms. For example, a vector with these subscripts

$$\boldsymbol{\psi} = (\psi_{i,g,t}, \psi_{i,g,t+1}, \ldots, \psi_{i,g,T_i}, \psi_{i+1,g,t}, \ldots,) \tag{S9.44}$$

has the lag matrix with matching rows and p columns,

$$\tilde{\boldsymbol{\Psi}} = \begin{pmatrix} \psi_{i,g,t-1} & \cdots & \psi_{i,g,t-p} \\ \psi_{i,g,t} & \cdots & \psi_{i,g,t+1-p} \\ \vdots & \vdots & \vdots \\ \psi_{i,g,T_i-1} & \cdots & \psi_{i,g,T_i-p} \\ \psi_{i+1,g,t} & \cdots & \psi_{i+1,g,t+1-p} \\ \vdots & \vdots & \vdots \end{pmatrix} \tag{S9.45}$$

S9.7.3 *Sample fixed effects*

To sample fixed effects, I move a few terms to the left,

$$\mathbf{m} = \tilde{\boldsymbol{\Psi}}\boldsymbol{\alpha} \tag{S9.46}$$

where $\mathbf{m}$ has elements $\psi_{ig,t} - \mu_{ig,t} - \boldsymbol{\alpha}'_{g[i]}\tilde{\boldsymbol{\psi}}_{ig,t}$, and, again, $\boldsymbol{\alpha}_{g[i]}$ indicates the vector of lags for the group to which individual i belongs. The conditional posterior matrix for fixed effects is

$$\boldsymbol{\alpha}|\{\boldsymbol{\alpha}_g\} \sim MVN(\mathbf{Vv}, \mathbf{V}) \tag{S9.47}$$

$$\mathbf{v} = \sigma^{-2}\tilde{\boldsymbol{\Psi}}'\mathbf{m} \tag{S9.48}$$

$$\mathbf{V}^{-1} = \sigma^{-2}\tilde{\boldsymbol{\Psi}}'\tilde{\boldsymbol{\Psi}} + 0.001 \times \mathbf{I}_p \tag{S9.49}$$

S9.7.4 *Random group effects*

For random effects, make a slight change in the mean vector and write

$$\boldsymbol{\alpha}_g \sim MVN(\mathbf{V}_g\mathbf{v}_g, \mathbf{V}_g) \tag{S9.50}$$

$$\mathbf{v}_g = \frac{1}{\sigma^2}\sum_{i,t}\tilde{\boldsymbol{\psi}}'_{ig,t}m_{ig,t} \tag{S9.51}$$

$$\mathbf{V}_g = \frac{1}{\sigma^2}\sum_{i,t}\tilde{\boldsymbol{\psi}}_{ig,t}\tilde{\boldsymbol{\psi}}'_{ig,t} + \mathbf{A}_\alpha^{-1} \tag{S9.52}$$

where $m_{ig,t} = \psi_{ig,t} - \mu_{ig,t} - \boldsymbol{\alpha}'\tilde{\boldsymbol{\psi}}_{ig,t}$. The summations are taken over all observation years in which individual i has been in the mature state for the previous p years, for all individuals in group g. Here is the conditional for the $p \times p$ covariance matrix,

$$\mathbf{A}_\alpha|\{\boldsymbol{\alpha}_g\} \sim IW\left(\sum_{g=1}^{G}\boldsymbol{\alpha}_g\boldsymbol{\alpha}'_g + df \times \mathbf{I}_p, G + df\right) \tag{S9.53}$$

S9.7.5 *Latent states*

Latent states in the AR(p) model are sampled by proposing from the conditional posterior for the fecundity/maturation submodel $\psi_t, z_t|\psi_{\{-t\}}, z_{t-1}, z_{t+1}$, where $\psi_{\{-t\}}$ is the set of all fecundity values except t, and accepting from the likelihood (Method 2 in section S9.1). To reduce clutter, we now omit subscripts ijg. If there are random groups in the model, then everything below is handled at the group level, with lag coefficient α_l being replaced with $\alpha_l + \alpha_{gl}$ for group g.

To isolate the terms in ψ_t the AR(p) model can be written as

$$\psi_t \sim N\left(m_t, \sigma^2\right) \tag{S9.54}$$

where

$$m_t = \mu_t + \sum_{l=1}^{p} \alpha_l \psi_{t-l} \tag{S9.55}$$

The exponent of the conditional distribution $\psi_t | \psi_{\{-t\}}$ can be factored this way:

$$\frac{1}{\sigma^2}\left[(\psi_t - m_t)^2 + \sum_{k=1}^{p}\left(n_{t,k} - \alpha_k \psi_t\right)^2\right] \tag{S9.56}$$

where

$$n_{t,k} = \psi_{t+k} - \mu_{t+k} - \sum_{l=1}^{p} \alpha_l \psi_{t+k-l} I(l \neq k) \tag{S9.57}$$

To sample latent states, we propose from

$$\psi_t \sim N\left(V v_t, V\right) \tag{S9.58}$$

where

$$v_t = \frac{1}{\sigma^2}\left(m_t + \sum_{k=1}^{p} n_{t,k}\alpha_k\right) \tag{S9.59}$$

$$V^{-1} = \frac{1}{\sigma^2}\left(1 + \sum_{k=1}^{p} \alpha_k^2\right) \tag{S9.60}$$

Proposals are accepted as a block for each plot-year in the data set, based on the likelihood for seed data (see Method 2).

S9.8 *Other parameters*

The error variance σ^2 is sampled from the conditional inverse gamma posterior distribution.

If there are no random groups, the dispersal parameter u is sampled with Metropolis, with an adaptive proposal variance and truncated normal prior distribution,

$$[u] \propto L \times N(u|u_0, U_0)I(u_{min} < u < u_{max}) \quad (S9.61)$$

where the likelihood L is $\prod_{r,s,j,t} Poi\left(y_{rsj,t}|A_{sj}\lambda_{smj,t}(u, \boldsymbol{\psi}_{j,t}, \boldsymbol{\rho}_{j,t}, \mathbf{m})\right)$, u_0 and U_0 are the prior mean and variance dispersal parameters, and (u_{min}, u_{max}) is an interval selected with prior knowledge.

If there are random groups, then there is an additional stage for the global mean. The previous distribution applies to u_g for group g,

$$[u_g] \propto L \times N(u_g|u, U) \quad (S9.62)$$

The u_g are proposed and accepted as a block. The global mean and variance have conditional distributions:

$$u|u_1, \ldots, u_G, u_0, U, U_0 \sim N(Vv, V) \quad (S9.63)$$

$$V^{-1} = \frac{G}{U} + \frac{1}{U_0} \quad (S9.64)$$

$$v = \frac{1}{U}\sum_g u_g + \frac{u_0}{U_0} \quad (S9.65)$$

$$U|u_1, \ldots, u_G, u \sim IG\left(2 + \frac{G}{2}, 1 + \frac{1}{2}\sum_g (u_g - u)^2\right) \quad (S9.66)$$

References

Clark, J., S. LaDeau, and I. Ibanez. 2004. Fecundity of trees and the colonization-competition hypothesis. Ecological Monographs **74**:415–442.

Clark, J., E. Macklin, and L. Wood. 1998. Stages and spatial scales of recruitment limitation in southern Appalachian forests. Ecological Monographs **68**:213–235.

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

Ladeau, S. L., and J. S. Clark. 2006. Elevated CO2 and tree fecundity: the role of tree size, interannual variability, and population heterogeneity. Global Change Biology **12**:822–833. URL `http://dx.doi.org/10.1111/j.1365-2486.2006.01137.x`.

Muller-Landau, H. C., S. J. Wright, O. Calderon, R. Condit, and S. P. Hubbell. 2008. Interspecific variation in primary seed dispersal in a tropical forest. Journal of Ecology **96**:653–667. URL `<GotoISI>://WOS:000256635800010`.

Neal, R., 2011. MCMC Using Hamiltonian Dynamics, Book section 5, pages 1–49 . CRC Press, New York, New York.

West, M., and J. Harrison. 1997. Bayesian Forecasting and Dynamic Models. 2nd edition. Springer.