Ecological Archives M080-020-A1

James S. Clark, David Bell, Chengjin Chu, Benoit Courbaud, Michael Dietze, Michelle Hersh, Janneke HilleRisLambers, Ines Ibáñez, Shannon LaDeau, Sean McMahon, Jessica Metcalf, Jacqueline Mohan, Emily Moran, Luke Pangle, Scott Pearson, Carl Salk, Zehao Shen, Denis Valle, and Peter Wyckoff. 2010. High-dimensional coexistence based on individual variation: a synthesis of evidence. *Ecological Monographs* 80:569–608.

Appendix C0

In this appendix we describe prior distributions, conditional relationships and distribution theory needed for algorithm development, algorithms used for Metropolis within Gibbs, and some issues related to MCMC diagnostics.

PRIOR DISTRIBUTIONS

The analysis includes both informative and non-informative prior distributions. Where possible, we used informative prior distributions that are flat but truncated at limits based on previous study, to maximize transparency, i.e., for clear identification of the contributions of prior vs likelihood. Here we summarize prior distributions and how they were selected to balance information.

Parameters for the logit function of maturation eqn 7b were assigned the prior

$$\boldsymbol{\beta}^{\boldsymbol{\theta}} \sim N\left(\mathbf{b}^{\boldsymbol{\theta}}, \mathbf{V}^{\boldsymbol{\theta}}\right)$$
 A.1

The prior mean values are consistent with knowledge that maturation probability increases with diameter and exposed canopy area. Large variances in \mathbf{V}^{θ} make the contribution from \mathbf{b}^{θ} small. Furthermore, there is a prior on minimum and maximum maturation diameters D_{\min} , D_{\max} , based on values below and above which we have not observed immature and mature status (Table A1). The recognition probability for observing maturation status (eqn 4) has prior density

$$v \sim Be(v_1, v_2) \tag{A.2}$$

where values center the prior density on 0.5 weighted by the number of tree years, $v_1 = v_2 = 0.002n_{IJT}$ (Table A1). The prior mean of 0.5 was based on previous success rates (Clark et al. 2004). The small values for v_1 and v_2 give the prior low weight.

For dioecious species, the female probability has the prior

$$\phi \sim Be(h_1, h_2) \tag{A.3}$$

with $h_1 = h_2 = 4$, having a mean of 0.5. This weak prior had no discernable effects on the fit.

The coefficients associated to the fixed effects in the state-space model in eqn 8 have flat priors bounded by values either having theoretical justification or sufficiently wide to not impact estimates,

$$vec(\mathbf{A}) \sim unif(vec(\mathbf{a}^{(1)}), vec(\mathbf{a}^{(2)}))$$
 A.4

where $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ are matrices having the same dimensions as **A** containing of minimum and maximum values, respectively (Table A1), and $vec(\mathbf{A}) = [A_{11}, \dots, A_{p1}, A_{12}, \dots, A_{p2}]$ is a vector obtained by stacking columns of **A**. We describe prior values for specific elements of **A**, using indexing A_{11} for the first row and column of **A** (eqn 8e):

 (A_{11}, A_{12}) - *intercept:* for all species intercepts we use priors with limits wide enough to have no influence on estimates.

 (A_{21}, A_{22}) - diameter effect on diameter growth increment and fecundity: Priors on the four parameters for diameter effects reflect the facts that increasing diameter increases fecundity until trees become large, and at some point both growth and fecundity could decline. The rationale is as follows. Allometric arguments and empirical evidence suggest that potential fruit yield should scale with canopy width, which, in turn is roughly proportional to diameter. However, this potential should not be realized for trees crowded by neighbors. Priors help us separate the two effects, diameter versus light availability. We allow for the effect of size on potential yield with the constraint $1.5 < A_{22} < 3.5$)(Niklas and Enquist 2003, Clark et al. 1998). Modification of this effect by competition enters in the constraints on exposed canopy area, i.e., the term including λ . Empirical data summarized by Niklas and Enquist (2003) suggest that reproductive organs should be proportional to (stem biomass)^{*a*}, where 0.6 < a < 1.0. Because stem biomass scales as D^3 , this gives a ranges of $1.8 < A_{22} < 3$. Thomas (1995) reports reproductive organs proportional to D^5 , but none of our estimates approached values this high in extensive experiments with the model. We note that such high exponents could be obtained if both immature and mature individuals were included in the fecundity model, but our fecundity model is conditional on mature status.

Diameter can be correlated with diameter growth increment, because tall trees have greater access to light. Because that relationship should be attributed to canopy exposure, rather than to tree diameter directly, we constrain parameter A_{21} to be near zero. This assumption still allows a relationship between growth rate and diameter, but it acknowledges that the relationship operates through the capacity to capture light. From open-grown trees there is no clear evidence for a direct size effect on diameter increment once trees exceed the seedling stage, until they become large. Instead of a direct effect of diameter on growth rate, the correlation between tree size and diameter growth increment is expected to result from the fact that large trees are more likely to have higher light exposure. The limits are $-0.02 < A_{21} < 0.02$. The effect of diameter as trees become large is taken up by the next parameter.

 (A_{31}, A_{32}) - *large diameter effect is negative*: The squared diameter term is included to allow for potential senescence, a decline in physiological function with age (Ryan and Yoder 1997). Tree data sets rarely have sufficient large (potentially old) individuals to estimate these effects, but we can allow that senescence does eventually occur by specifying that this effect only has impact for especially large individuals. These terms are constrained to be negative.

 (A_{41}, A_{42}) - *canopy exposure effect is non-negative*: Light availability increases growth rate and fecundity.

 (A_{51}, A_{52}) - *lag-1 effect of growth rate on growth rate and fecundity*: This effect was constrained to be effectively zero for growth rate (A_{51}) but unconstrained for fecundity (A_{52}) . We wanted to parameterize the effect on fecundity, so it could be used for predictive modeling of potential tradeoffs in time between growth and fecundity.

Because estimates of growth and fecundity balance contributions from the regression (i.e., the diameter and light covariates in eqn 8) and data models for growth rates and seed data (Fig. 3), we used an informative prior on the error covariance matrix Σ to represent a level of variation expected after that taken up by covariates, random effects, and year effects and to assure that covariates were not overwhelmed by noise. The values used for variances on the log growth (cm) and log fecundity (seeds per tree) were 0.05 and 0.5, respectively. The prior density is Wishart,

$$\boldsymbol{\Sigma}^{-1} \sim W\left(\mathbf{V}_{\boldsymbol{\Sigma}}^{-1}, \boldsymbol{n}_{LJT}\right)$$
 A.5

where the weight of the prior is controlled by $n_{IJT} = \sum_{i,j} (T_{ij} - t_{ij})$, the number of tree-years (Table 5), and $\mathbf{V}_{\Sigma} = \text{diag}(0.05, 0.5)$. Through extensive sensitivity analysis, this prior was found

to provide an acceptable balance of data and regression model, contributing to the conditional posterior approximately twice the weight coming from the regression.

Priors on random effects and year effects were weak--we wanted data to dominate these estimates. The prior for the random effects variance is

$$\mathbf{V}_b^{-1} \sim W\!\left(\mathbf{R}_b^{-1}, r_b\right) \tag{A.6}$$

where $r_b = n_{IJ}/100$, $\mathbf{R}_b = \text{diag}(0.2,2)$ and $n_{IJ} = \sum_j I_j$ is the number of trees. r_b was rounded to an integer value and ranged from 3 to 180 for different species.

The prior for fixed year effects is noninformative

$$\mathbf{b}_{r,t} \sim N\left(\begin{bmatrix} 0 & 0 \end{bmatrix}^T, diag(100, 100)\right)$$
A.7

and includes a sum-to-zero constraint (intercepts are included in A), implemented directly in the Gibbs sampler.

Because many individuals are not mature, a separate univariate model applies (eqn 6). The covariates are the same as those listed for the multivariate regression given above with priors being flat and truncated at the same values for adult growth.

Diameter growth increments have a prior for each tree year taken from the posterior from the analysis of Clark et al. (2007b) truncated at two standard deviations, based on the prior belief that true increments should be within this range. Diagnostics showed that posterior estimates from this analysis did not tend to accumulate at these truncation values. The prior for diameter increments is

$$\ln d_{ij,t} \sim N \Big(\ln d_{ij,t}^{(0)}, v_{ij,t} \Big) I \Big(d_{ij,t}^{(1)}, d_{ij,t}^{(2)} \Big)$$
A.8

where $d_{ij,t}^{(0)}$ and $v_{ij,t}$ are the prior mean and variance for log growth rate from the analysis of Clark et al. (2007)(Fig. 4), and $d_{ij,t}^{(1)}$ and $d_{ij,t}^{(2)}$ are one standard deviation below and above the prior mean. In other words, values cannot deviate widely from the posterior estimates from the growth analysis, but there is flexibility for values having weak support.

Priors for fecundity and missing seed data were either noninformative or derived from previous observations. A flat prior was used for fecundity, truncated at the smallest number of seeds observed for a tree and at values much larger than implied by observation of seed densities,

$$f_{ij,t} \sim unif(f^{(1)}, f^{(2)})$$
 A.10

(Table A2). For instance, when defining $f^{(1)}$, we did not expect that a mature individual would produce less seeds than typically contained in a single fruiting structure (e.g., *Pinus, Liriodendron, Liquidambar*). For maximum values, we used the prior estimates for parameter *u* and inverted observations of seed densities from the forest floor (eqn 2), to solve for fecundity, $\mathbf{f}_{j,t} = \left(\mathbf{F}_{j,t}^T \mathbf{F}_{j,t}\right)^{-1} \mathbf{F}_{j,t}^T \left[\left(\mathbf{A}_j^T \mathbf{A}_j\right)^{-1} \mathbf{A}_j^T \mathbf{g}_{j,t} - \mathbf{1}_m cB_j \right]$. For example, *Acer rubrum* seeds have been

observed at average densities of 10^2 seeds m⁻² beneath mature trees but not at average densities of 10^3 seeds m⁻². This inversion was used not only to set limits on fecundities and as a crude initialization for posterior simulation (Appendix). The prior for missing seed data is

 $s_{jk,t} \sim Pois(g_{jk})I(g_{jk}^{(1)},g_{jk}^{(2)})$

where $g_{k,t}$ is average for the trap and year for which data are missing and the upper and lower bounds are limits based on observation for the stand values ever observed for the site.

Prior densities for the seed data model in eqns 1-3, including the dispersal parameter and the seed fraction originating outside the map, were

$$(u,c) \sim N(u_0,v_u)N(c_0,v_c)I(u^{(1)},u^{(2)})I(c^{(1)},c^{(2)})$$
 A.11

where parameter values were chosen to be informative (Table A2).

The monotonicity priors on the parameters in eqn 9 are detailed in Clark et al. (2007, Appendix).

CONDITIONAL RELATIONSHIPS

State-space model

For parameters in the state space submodel (eqn 8), all sampling was direct. The conditional posterior for fixed effect parameters includes the prior from eqn A.4

$$\operatorname{vec}(\mathbf{A})|\mathbf{X},\mathbf{Y},\boldsymbol{\Sigma},\ldots \sim N_{2p}\left(\left(\mathbf{X}^{T}\mathbf{X}\right)^{-1}\mathbf{X}^{T}\mathbf{Z},\left(\boldsymbol{\Sigma}+\mathbf{V}_{b}\right)\otimes\left(\mathbf{X}^{T}\mathbf{X}\right)^{-1}\right)I\left(\operatorname{vec}(\mathbf{a}^{(1)}),\operatorname{vec}(\mathbf{a}^{(2)})\right)$$
 A.12

where p is the number of covariates, X is the stacked matrix of $(T_{ij} - \tau_{ij})$ by p X_{ij} matrices for

each tree, $\mathbf{Z} = [\mathbf{Z}_{11} \mathbf{Z}_{21} \dots]$ is the stacked matrix $\mathbf{Z}_{ij} = \mathbf{Y}_{ij} - \mathbf{1}_{ij} \mathbf{b}_t$, \mathbf{Y}_{ij} is the $(T_{ij} - \tau_{ij})$ by 2 matrix of responses (eqn 8a), $\mathbf{1}_{ij}$ is the length $(T_{ij} - \tau_{ij})$ vector of 1's, and τ_{ij} and T_{ij} are the first and last years during the study in which individual *ij* is imputed to be mature. The truncated multivariate normal is sampled from the conditional univariate truncated normals bounded by $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ (eqn 13).

The error covariance matrix was sampled from an inverse Wishart conditional posterior, incorporating the prior from eqn A.5. The conditional posterior for Σ^{-1} is

$$\boldsymbol{\Sigma}^{-1} \sim W \Biggl(\Biggl[\sum_{i,j,t} Q_{ij,t} \mathbf{z}_{ij,t}^T \mathbf{z}_{ij,t} + n_{IJT} \mathbf{V}_{\boldsymbol{\Sigma}} \Biggr]^{-1}, \sum_{i,j,t} Q_{ij,t} + n_{IJT} \Biggr)$$

where $\mathbf{z}_{ij,t} = \mathbf{y}_{ij,t+1} - \mathbf{x}_{ij,t}\mathbf{A} - \mathbf{b}_{r,t} - \mathbf{b}_{ij}$, n_{LJT} is the total number of tree years. The conditional posterior for the random effects variance is

$$\mathbf{V}_{b}^{-1} \sim W\left(\left[\sum_{i,j}\max_{t} \left(Q_{ij,t}\right)\mathbf{b}_{ij}^{T}\mathbf{b}_{ij} + r_{b}\mathbf{R}_{b}\right]^{-1}, \sum_{i,j,t}\max_{t} \left(Q_{ij,t}\right) + r_{b}\right)$$

the prior coming from eqn A.6. Note that individual *ij* contributes to the conditional posterior only if it is imputed to be mature at some point during the study, in which case $\max_t(Q_{ij,t}) = 1$. The random effects are sampled from

$$\mathbf{b}_{ij} \sim N_2 (\mathbf{V} \mathbf{v}, \mathbf{V}),$$
$$\mathbf{V}^{-1} = \Sigma^{-1} \sum_t Q_{ij,t} + \mathbf{V}_b^{-1}$$

$$\mathbf{v} = \boldsymbol{\Sigma}^{-1} \sum_{t} Q_{ij,t} \left(\mathbf{y}_{ij,t+1}^{T} - \mathbf{A}^{T} \mathbf{x}_{ij,t}^{T} - \mathbf{b}_{r,t} \right).$$

The fixed year effects are sampled from a conditional normal with prior from eqn A.7. The conditional posterior is

$$\mathbf{b}_{t} \sim N_{2} (\mathbf{V} \mathbf{v}, \mathbf{V}),$$
$$\mathbf{V}^{-1} = \Sigma^{-1} \sum_{i,j} Q_{ij,t} + 100^{-1}$$
$$\mathbf{v} = \Sigma^{-1} \sum_{i,j} Q_{ij,t} (\mathbf{y}_{ij,t+1} - \mathbf{x}_{ij,t} \mathbf{A} - \mathbf{b}_{ij}).$$

This draw was followed by subtraction of the mean year effect for both lnd and lnf, imposing a sum-to-zero constraint.

Diameter growth

Diameter growth increments were updated from the conditional posterior that is the product of eqns 6a or 8b for growth, eqns 5 and 9 for survival, and the prior given in eqn A.8. For survival, we use a Gaussian approximation to $\zeta_{ij,t} \approx 1 - \mu_d$, where the μ_d sequence contains probabilities for discrete diameter increment bins. The contribution of diameter is omitted, because its effect on survival probability is small relative to that of growth rate. Then the conditional distribution for bin *k* in the sequence of μ_d estimates is

$$p(\ln d_k | z) = \frac{p(z | \ln d_k) p(\ln d_k)}{\sum_k p(z | \ln d_k) p(\ln d_k)}$$

where $p(z = 1|\ln d_k) = 1 - \mu_k$ and $p(z = 0|\ln d_k) = \mu_k$, z being survival (1) or death (0) in the subsequent year, and $p(\ln d_k)$ the distribution of log growth rates. The conditional means and variances for log growth rates are

$$\mu_{d|z_{ij,t+1}} = E\left(\ln d | z_{ij,t+1}\right) = \sum_{k} \ln d_{k} p\left(\ln d_{k} | z_{ij,t+1}\right)$$

and

$$V_{d|z_{ij,t+1}} = Var(\ln d|z_{ij,t+1}) = \sum_{k} (\ln d_{k})^{2} p(\ln d_{k}|z_{ij,t+1}) - [E(\ln d|z_{ij,t+1})]^{2}$$

There is a conditional mean and variance for z = 0 and z = 1. The log growth rates are sampled from

$$\ln d_{ij,t} \sim N(Vv,V)$$

$$V^{-1} = V_{d|f}^{-1} + v_{ij,t}^{-1} + V_{d|z_{ij,t+1}}^{-1}$$

$$v = \mu_{d|f} V_{d|f}^{-1} + \ln\left(d_{ij,t}^{(0)}\right) v_{ij,t}^{-1} + \mu_{d|z_{ij,t+1}} V_{d|z_{ij,t+1}}^{-1}$$

with conditional means and variances contributed by survival are for z = 0 or z = 1, depending on whether or not the individual survived until the next year. The terms in these conditionals come from the state space model, observations, and survival, respectively.

Survival

Because slow growth is associated with death, the observations of growth rate below a certain threshold are rarely observed. This lack of slow growth observations results from the fact that mortality risk increases sharply at slow growth rates. The survival model includes a sequence of μ_d values (probabilities) bounded by one and zero, declining monotonically with growth rate, and a sequence of μ_D increasing monotonically with tree diameter. In addition to monotonicity, there is an informative prior for values within the sequence μ_D , which was $Be(a_k, b_k)$, where a_k is 0.001 for k = 1, 2, 3 and $a_k = 10$ for k = 4, 5, 6,

$$b_k = a_k \left(\frac{1}{\mu_k^0} - 1\right)$$

 $\mu_k^0 = [0.00001, 0.00002, 0.00003, 0.00004, 0.00005]$

This prior assures essentially zero values for juvenile trees (bins 1, 2, 3), but is non-informative (but monotonically increasing) for large trees. Thus, diameter only affects survival of large trees, but the prior has impact for species lacking large individuals. Although small trees grow slowly and thus are at higher mortality risk, this is a growth effect, not a diameter effect. This informative prior allows us to separate the effect of slow growth from that of large size, which could indicate senescence. The monotonicity constraints on survival model were imposed by a Metropolis step, where proposed values have this characteristic. The full sequence of $\{\mu_d\}, \{\mu_D\}$ are proposed together, subject to the monotonicity constraint, and accepted as a block. To mix over the unknown year of death for the interval-censored data, a year of death was selected at random from the death interval (ranging in width from one to four years), and taken to be the currently imputed year of death.

Dispersal parameters

The parameters *u* and *c* were updated with a Metropolis step, both proposed from a normal distribution truncated at zero.

Fecundity, maturation, gender

Due to their conditional dependence structure, fecundity, maturation, and (for dioecious species) gender are sampled together in a Metropolis step. Here we describe sampling. The basic factoring used for maturation, gender, and fecundity is

$$p(f_{ij,t}, Q_{ij,t}, H_{ij}, s_{j,t} | q_{ij}, h_{ij}, d_{ij,t-1}, D_{ij,t}, \lambda_{ij,t}) = p(s_{j,t} | \{f_{ij,t}, Q_{ij,t}, H_{ij}\}) p(f_{ij,t}, Q_{ij,t}, H_{ij} | q_{ij}, h_{ij}, d_{ij,t-1} D_{ij,t}, \lambda_{ij,t})$$

where q_{ij} and h_{ij} represent the history of observations on individual *ij*, including present (*t*), past (before *t*) and future (after *t*). The first factor on the RHS is the likelihood for seed trap data, indicating that all seed traps on plot *j* in year *t* conditionally depend on every tree *i* on plot *j*. The second factor on the RHS is the probability of being mature (Q = 1), female (H = 1), and having fecundity *f*.

For monoecious species, we use a Metropolis step where maturation status and fecundity are jointly proposed and rejected for all trees in a given stand *j* in a given year *t*. For dioecious species we must further sample gender. Because gender applies to an individual across all years, dioecious species are sampled in a different way and are discussed after monoecious species. The blocking differs between these two data types, which we describe here.

Efficient Gibbs sampling requires blocking of variables to facilitate mixing, which is challenging given the ways in which latent variables are linked with the unknown year in which an individual becomes mature τ_{ij} . These relationships include:

- i) the $Q_{ij,t}$ and $f_{ij,t}$ are inherently linked, by virtue of the fact that non-zero fecundity is defined only for mature individuals, $f_{ij,t} | (Q_{ij,t} = 0) = 0$, and $f_{ij,t} | (Q_{ij,t} = 1) > 0$.
- ii) maturation statuses for an individual over time are mutually dependent according to the one-way transition to maturity in year τ_{ij} .
- iii) gender is considered to be fixed, and
- iv) seed trap data conditionally depend on all trees in the plot in a given year.

In light of the conditional relationships involving status and seed production, the choices for blocking are to 1) sample individually every year for every tree (conditioned on other trees for that year and other years for that individual), 2) sample as a block all individuals within a plot for a given year, and 3) sample as a block all trees and years within a plot. The first option has the advantage that high acceptance rates can be achieved, but is computationally slow, entailing loops over plots, individuals, and years, e.g., a Metropolis step for every tree-year in the data set.

The third option can result in a high rejection rate, each proposal consisting of $\sum_{i=1}^{n_j} (T_{ij} - t_{ij})$ values.

The binary nature of Q and H proposals can make acceptance rates low. Nonetheless, because gender H_{ij} applies to an individual across all years, we use a modification of option 3 for dioecious species. We begin with a description for monoecious species, followed by the description for dioecious species.

*Monoecious species--*We use the second option for monoecious species, blocking on time and modeling each year successively. The factoring is

$$p(f_{j,t},Q_{j,t}|q_{j,t},x_{j,t-1},x_{j,t},Q_{j,t-1},Q_{j,t+1},s_{j,t}) = p(s_{j,t}|f_{j,t},Q_{j,t}) p(Q_{j,t},f_{j,t},|q_{j,t},x_{j,t-1},Q_{j,t-1},Q_{j,t+1},x_{j,t})$$

We propose all values of $\{Q, f\}_{j,t}$ together and accept or reject them as a block. The Markov transition probabilities from *t* to *t*+1 are conditioned on observations of status $q_{j,t}$, and they must be combined with probabilities for fecundities and seed trap data. The transition from immature to mature is a hidden Markov process, but only for tree-years in which status is unknown, which is the case after the last year in which immaturity is certain, and before reproduction has been observed. If the status is known through past observations (if previously observed to be mature, then still mature), a current observation (mature or immature), or future observations (if later known to be immature, then immature now), then status $Q_{ij,t}$ is known. These relationships also apply to imputed statuses. If unknown, status must be modeled as the conditional probability of being in state $Q_{ij,t}$ given $Q_{ij,t-1} = 0$ and $Q_{ij,t+1} = 1$. These probabilities involve the age-specific rates of making the transition from immature to mature states and can be derived from the

cumulative logit probability of being mature given diameter $D_{ij,t}$ and canopy status $\lambda_{ij,t}$. Because blocking is year-by-year, we condition the transition probability on both the foregoing and the following years. Then the trivial probabilities are $p(Q_{ij,t} = 1|Q_{ij,t-1}) = 1$ and

 $p(Q_{ij,t} = 1 | Q_{ij,t+1} = 0) = 0$. For failure to recognize the reproductive state, we need the additional factor

$$p(q_{ij,t} = 0 | Q_{ij,t} = 1) = 1 - v$$

For Gibbs sampling, we need the year-by-year transition probabilities from immature to mature between *t*-1 and *t* given that the transition was made between *t*-1 and *t*+1. Let $\delta_{ij,t}$ be the probability of being in the mature state conditional on states in years *t*-1, *t*+1, and on observations. Ignoring observations for the moment, we have

$$\begin{split} \delta_{ij,t} &= p \Big(Q_{ij,t} = 1 \Big| Q_{ij,t-1} = 0, Q_{ij,t+1} = 1 \Big) = \frac{p \Big(Q_{ij,t+1} = 1 \Big| Q_{ij,t} = 1 \Big) p \Big(Q_{ij,t} = 1 \Big| Q_{ij,t-1} = 0 \Big)}{\sum_{k=0,1} p \Big(Q_{ij,t+1} = 1 \Big| Q_{ij,t} = k \Big) p \Big(Q_{ij,t} = k \Big| Q_{ij,t-1} = 0 \Big)} \\ &= \frac{d \theta_{ij,t} / \Big(1 - \theta_{ij,t-1} \Big) + \Big[1 - d \theta_{ij,t} / \Big(1 - \theta_{ij,t-1} \Big) \Big] \times d \theta_{ij,t+1} / \Big(1 - \theta_{ij,t} \Big)}{d \theta_{ij,t} + \Big(\frac{1 - \theta_{ij,t-1} - d \theta_{ij,t}}{1 - \theta_{ij,t}} \Big) d \theta_{ij,t+1}} \\ &= \frac{d \theta_{ij,t}}{d \theta_{ii,t} + d \theta_{ii,t+1}} \end{split}$$

where

$$d\theta_{ij,t} \approx \left(\frac{d\theta_{ij,t}}{dD_{ij,t}} \times \frac{dD_{ij,t}}{dt}\right) dt \approx \left(\frac{d\theta_{ij,t}}{dD_{ij,t}} \times d_{ij,t}\right) dt$$
$$= \beta_1^{\theta} d_{ij,t} \theta_{ij,t} \left(1 - \theta_{ij,t}\right) dt$$

Because $\lambda_{ij,t}$ changes much slower than $D_{ij,t}$, we do not include it in the chain rule calculation for the derivative. Because *dt* is always equal to 1 yr, we hereafter omit it.

Observations change the transition probabilities. Thus far, the equation for $\delta_{ij,t}$ describes the probability of transition in the absence of an observation. If there is an observation in yr *t* and it is 'uncertain' ($q_{ij,t} = 0$: see Table 7), then the observer did not identify the tree as mature, and the probability becomes

$$\delta_{ij,t} = \Pr(Q_{ij,t} = 1 | Q_{ij,t-1} = 0, Q_{ij,t+1} = 1, q_{ij,t} = 0)$$
$$= \frac{d\theta_{ij,t}(1-v)}{d\theta_{ij,t}(1-v) + d\theta_{ij,t+1}}$$

For the first study year, in the absence of an observation (maturation statuses were not obtained on all individuals the first year of the study), we have

$$\delta_{ij,t} = p\left(Q_{ij,t} = 1 \middle| Q_{ij,t+1} = 1\right) = \frac{\theta_{ij,t}}{\theta_{ij,t} + d\theta_{ij,t+1}}$$

If there was an observation and that observation was $q_{ij,t} = 0$ (Table 4), this becomes

$$=\frac{\theta_{ij,t}(1-v)}{\theta_{ij,t}(1-v)+d\theta_{ij,t+1}}$$

For the last observation year, absent observation,

$$\delta_{ij,T} = p(Q_{ij,T} = 1 | Q_{ij,T-1} = 0) = d\theta_{ij,T} / (1 - \theta_{ij,T})$$

If there was an observation, we have

$$\delta_{ij,T} = d\theta_{ij,T}(1-\nu) / (1-\theta_{ij,T})$$

The Metropolis steps entail a loop over years, at each time step proposing values for Q_t and f_t , with the constraints on Q discussed above and $f_{ij,t} = 0$ for all imputed $Q_{ij,t} = 0$. Those mature at *t*-1 remain mature. Those immature at *t*+1 remain immature. Those of known status retain that status. For those imputed to be immature at *t*-1 and mature at *t*+1 and of unknown status, candidate values come from $Q_{ij,t}^* \sim Bernoulli(0.5)$. For individuals proposed to be mature and previously imputed to be mature, we propose from $\ln f_{ij,t}^* \sim N(\ln f_{ij,t}^{(g)}, 0.1)$, where (g)

denotes the current Gibbs step, prior to updating. For individuals previously imputed to be immature, we propose in three ways, implemented randomly with equal probability. Let **f** be the vector of log fecundity values selected for updating. We draw from $\mathbf{f} \sim N(\mathbf{f}', 0.1 \times \mathbf{I})$ where the

mean vector assumes one of three forms:

- 1) *Previous value*: **f**' is a vector of auxiliary variables having values retained from the most recent iteration of the sampler in which $Q_{ij,t} = 1$.
- 2) *Match process support:* $\mathbf{f}' = \mu_{f|d}$ is the vector of conditional means from the state space model, i.e., the value finding most support from the multivariate regression, based on current parameter values. It is given by eqn 8b.
- 3) *Match likelihood support from seed data:* Conditional on currently imputed values for individuals not being updated on plot *j* at time *t* $\mathbf{e}_{j,t}$, having dispersal matrix $\mathbf{E}_{j,t}$ solve for mean values finding most support from seed data from eqn

2,
$$\mathbf{f}'_{j,t} = \left(\mathbf{F}'_{j,t}\mathbf{F}'_{j,t}\right)^{-1}\mathbf{F}'_{j,t}\left[\left(\mathbf{A}_{j}^{T}\mathbf{A}_{j}\right)^{-1}\mathbf{A}_{j}^{T}\mathbf{g}_{j,t} - \mathbf{1}_{m}cB_{j} - \mathbf{E}_{j,t}\mathbf{e}_{j,t}\right]$$
, where $\mathbf{F}'_{j,t}$ is the

dispersal matrix for individuals to be updated.

The acceptance criterion involves the products from expressions given above,

$$p\left(Q_{j,t}, f_{j,t} \middle| Q_{j,t-1}, Q_{j,t+1}, q_{j,t}, s_{j,t}, \mu_{j,t}^{f|d}, V_{j,t}^{f|d}\right) \propto \prod_{i} \left(1 - \delta_{ij,t}\right)^{1 - Q_{ij,t}} \left[\delta_{ij,t} N\left(\ln f_{ij,t} \middle| \mu_{ij,t}^{f|d}, V_{ij,t}^{f|d}\right)\right]^{Q_{ij,t}} \prod_{k=1} Pois\left(s_{jk,t} \middle| g_{ijk,t}\right)$$

Note that all individuals imputed to be mature have a conditional density associated with the state-space model. The $\delta_{ij,t}$ are different for each individual and year, as discussed above. All trees contribute to the likelihood for the seed data for plot *j* in yr *t* in that by producing seed or not, they affect the parameters of the Poisson sampling distribution for seed. Of course, the set of proposals is accepted or rejected as a block. The sampler is more efficient than it appears, because we propose statuses and fecundities for all plots simultaneously, and accept/reject them on a plot by plot basis, without actually looping over plots. Once states are updated for time *t*, we move to *t*+1. The acceptance rate for typical species is 15 - 50%, when proposals 1 and 2 apply and 8 to 45% when proposal 3 applies. We repeated sampling two times for each iteration of the MCMC.

Dioecious species--For dioecious species gender is unchanging over time, so we evaluate the full history of observations for each tree, but still avoiding loops over individual trees. It is efficient to factor the conditional somewhat differently, taking together all trees on plot *j* over all years,

$$p(f_j, Q_j, H_j | q_j, h_j, X_j, s_j) \propto$$

$$p(s_j | f_j, Q_j, H_j) p(f_j, Q_j, H_j | q_j, h_j, X_j)$$

Because maturation is no longer modeled year-by-year, we require the probability for a history of maturation status, conditioned on observations obtained sporadically over individuals and years.

Let τ_{ij} be the year in which an individual becomes mature, $\tau_{ij}^0 = \max_t (q_{ij,t} = -1)$ be the last year an individual is known to have been immature, and $\tau_{ij}^1 = \min_t (q_{ij,t} = 1)$ be the first year an individual is known to have been mature. Thus, we have the constraint $\tau_{ij}^0 \le \tau_{ij} \le \tau_{ij}^1$. The probability assigned to an individual that became mature in year *t* is,

$$\begin{split} \delta_{ij} &= p \Big(\tau_{ij} = t \Big| \tau_{ij}^0 < \tau_{ij} < \tau_{ij}^1 \Big) = \left. d\theta_{ij,t} \right/ \left(\theta_{ij,\tau_{ij}^1} - \theta_{ij,\tau_{ij}^0} \right) \\ &= \beta_1^\theta d_{ij,t} \frac{\theta_{ij,t} \Big(1 - \theta_{ij,t} \Big)}{\Big(\theta_{ij,\tau_{ij}^1} - \theta_{ij,\tau_{ij}^0} \Big)} \end{split}$$

For individuals imputed to be still immature at the end of the observation period at T_{ij} , the probability is

$$\delta_{ij} = p\left(\tau_{ij} > T_{ij} \middle| \tau_{ij} > \tau_{ij}^{0}\right) = 1 - p\left(\tau_{ij} \le T_{ij} \middle| \tau_{ij} > \tau_{ij}^{0}\right) = \frac{1 - \theta_{T_{ij}}}{1 - \theta_{ij,\tau_{ij}^{0}}}$$

For individuals imputed to be already mature before observations began at t_{ij} , the probability is

$$\delta_{ij} = p \left(\tau_{ij} < t_{ij} \middle| \tau_{ij} < \tau_{ij}^{1} \right) = \underbrace{\theta_{t_{ij}}}_{\theta_{ij,\tau_{ij}^{1}}}$$

We now have a probability for the history of an individual that became mature at time τ_{ij} or remained immature throughout. Our prior specification allows for a minimum and maximum

maturation diameter, in which case θ_{ij,τ_{ij}^0} and θ_{ij,τ_{ij}^1} are the values of θ taken at these prior maturation diameter values.

The probability for maturation is combined with observations of status between those years that established it as immature and mature and for gender. Thus far, we have considered observations that definitively establish maturity or immaturity ($q_{ij,t} = -1$ or 1 in Table 4). For $q_{ij,t} = 0$, status is uncertain. Detection probabilities are $p(q_{ij,t} = 1|Q_{ij,t} = 0) = 0$ and

 $p(q_{ij,t} = 1|Q_{ij,t} = 1) = v$. The individual has unknown gender if the gender is not observed, the observation is uncertain, or if flowers are observed but not identified to sex, and no observation is available from the fruiting season (Table 4). Considering both observations and gender, the probability for individual *ij* becomes

$$p\left(\tau_{ij}, H_{ij} \middle| \tau_{ij}^{0} < \tau_{ij} < \tau_{ij}^{1}, q_{ij} = 0 \right) = \delta_{ij} \left(1 - \nu\right)^{n_{ij}^{\nu}} \left(1 - \phi\right)^{1 - H_{ij}} \phi^{H_{ij}}$$

where $n_{ij}^{\nu} = \sum_{\tau_{ij}}^{\tau_{ij}^{1}} (1 - q_{ij,t})$ is defined to be the number of times mature status was 'undetected' during the interval $(\tau_{ij}, \tau_{ij}^{1})$, i.e., the number of times that an individual imputed to be mature in year τ_{ij} was not identified as such. If gender is known, the third factor disappears,

$$p\left(\tau_{ij} \left| \tau_{ij}^{0} < \tau_{ij} < \tau_{ij}^{1}, q_{ij} = 0 \right) = \delta_{ij} \left(1 - v \right)^{n_{ij}^{v}}.$$

The full reproductive history on all individuals has conditional probability

$$p\left(Q_{j}, H_{j}, f_{j} | q_{j}, h_{j}, s_{j}, \mu_{j}^{f|d}, V_{j}^{f|d}\right) = \\ = \prod_{i} \delta_{ij} (1 - v)^{n_{ij}^{v}} (1 - \phi)^{1 - H_{ij}} \left[\phi \prod_{t} N\left(\ln(f_{ij,t}) | \mu_{ij,t}^{f|d}, V_{ij,t}^{f|d} \right)^{Q_{ij,t}} \right]^{H_{ij}} \\ \times \prod_{t} \prod_{t} Pois(s_{jk,t} | g_{ijk,t})$$

For each individual a maturation diameter is proposed from a uniform distribution

$$\tau_{ij} \sim unif\left(t_{ij}^{0}, t_{ij}^{1}\right)$$
$$t_{ij}^{0} = \max\left(t_{ij}^{\min}, \tau_{ij}^{0}\right)$$
$$t_{ij}^{1} = \min\left(t_{ij}^{\max}, \tau_{ij}^{1}\right)$$

The bounds for minimum and maximum maturation diameters are not sooner than the first year in which *ij* reached the minimum prior diameter for maturation $t_{ij}^{\min} = \max(t | (D_{ij,t} > D_{\min}))$ or it was last known to be immature τ_{ij}^0 and not later than the last year in which *ij* had not yet reached the prior diameter for certain maturation $t_{ij}^{\max} = \min(t|(D_{ij,t} < D_{\max})))$ or it was known to be mature τ_{ij}^1 . There are prior minimum and maximum diameters, which differ among species. The fecundity for an individual proposed to be immature is zero. For individuals currently imputed to be mature, the proposed fecundity is a truncated normal on $(f^{(1)}, f^{(2)})$ centered on the current estimate. The conditional densities are then the product of Poisson seed data, Gaussian fecundity, and the probability associated with maturation in yr *t*. Because the probability of seed data conditionally depends on all trees in the stand in all years, the set containing $\{f_j, Q_j, H_j\}$ is accepted or rejected as a block.

Recognition success is sampled from the conditional posterior

$$Bin\left(\sum_{ij,t} q_{ij,t} \left| \sum_{ij,t} Q_{ij,t}, v \right| Be(v|v_1, v_2) = Be(v|v_1 + \sum_{ij,t} q_{ij,t}, v_2 + \sum_{ij,t} (1 - q_{ij,t}) Q_{ij,t}\right)$$

For female fraction we sample from the conditional posterior

$$Bin\left(\sum_{ij}H_{ij}|n_{IJ},\phi\right)Be\left(\phi|h_1,h_2\right) = Be\left(\phi|h_1 + \sum_{ij}H_{ij},h_2 + \sum_{ij}\left(1 - H_{ij}\right)\right)$$

Prior values are $h_1 = h_2 = 4$, which has a mean of 0.5 and is dominated by the data.

Parameters for the logit function of maturation eqn 1 are sampled with Metropolis step. Conditionally we have

$$\prod_{i,j} \delta(\boldsymbol{\tau}_{ij}) N\left(\boldsymbol{\beta}^{\boldsymbol{\theta}} \middle| \boldsymbol{b}^{\boldsymbol{\theta}}, \boldsymbol{V}^{\boldsymbol{\theta}}\right)$$

where the first product is the probability associated with maturation years, which depend on β^{θ} , and the truncated normal prior. A proposal is generated from a multivariate normal with initial proposal covariance matrix $(\mathbf{M}^{T}\mathbf{M})^{-1}$, where **M** is the design matrix for the logit in eqn 7, $\operatorname{logit}(\theta) = \mathbf{M}\beta^{\theta}$. Sampling is adaptive, with **M** being updated several times in the MCMC from the covariance matrix of previous β^{θ} values.

Parameters for the seed data model are sampled in a single Metropolis step. Conditionally,

$$\prod_{t}\prod_{j}\prod_{k=1}Pois(s_{jk,t}|g_{ijk,t})N(u|u_0,V_u)N(c|c_0,V_c)I(u,c>0)$$

Values are proposed from a truncated normal distribution. In the case of missing data, seed counts were replaced with the currently imputed seed value.

Imputation of missing data involved a Metropolis step with proposals of plus or minus 1 from the current value with probability 0.5. The conditional posterior includes a Poisson prior with a mean density as discussed in Section 4 multiplied by Poisson density for sample jk, t. Proposals were accepted as a block for $s_{j,t}$.

A Metropolis step is used to simultaneously update all of the diameter growth and diameter bins for the non-parametric survival relationship. The growth rates and diameters of all individuals are binned in the sequences μ_d and μ_D are for the all years. For diameter increment there are 31 bins equally spaced with width 0.1 on the log₁₀ scale. For diameter there are six bins, also equally spaced on the log₁₀ scale, with the maximum value chosen to exceed that

largest diameter in the data set. Survival from year *t* to *t*+1 is the event $z_{ij,t} = 1$ and death in the subsequent year is $z_{ij,t} = 0$. At each Gibbs step new sequences of μ_d and μ_D are proposed each being Gaussian and centered on the currently imputed values, but truncated midway between the current values. For diameter increment the proposal density is

$$\mu_d^* \sim N(\mu_d, V) I((\mu_{d-1} - \mu_d) / 2, (\mu_d - \mu_{d+1}) / 2)$$

where V is a small value (0.1 in this case). In other words, if the currently imputed value for $\mu_{d,k}$ was 0.5 and those for bins k-1 and k+1 were 0.6 and 0.48, then the proposal would come from the normal centered at 0.5, truncated at 0.55 and 0.49. This procedure allows for any shape subject to monotonic decline. The proposals for the diameter values are done in the same way, with the constraint being monotonic increase and with an informative prior $Be(a_k, b_k)$. All

values (both growth rate and diameter) are proposed together and accepted as a block.



Fig0A1. The initial 50,0000 MCMC steps for diameter growth (above) and fecundity (below) random effects, 90% thinned. The five individuals were selected at random from *Ulmus alata*, also chosen arbitrarily. Maps and distance relationships are shown in Figs0A2 and A3, respectively.

POSTERIOR SIMULATION

Due to the size of the model, efforts were made to optimize code, written in R. Despite the large number of years across many individuals within multiple plots, the main Gibbs loop contains only three loops over years (including one to update maturation/fecundity, one for missing seed data, and another for dispersal and Poisson parameters), and no loops over

individuals or plots. Data structures that include pointer arrays were used to rebuild (reorder and restack) matrices of state variables based on the changing gender and maturation statuses of trees and tree-years and imputed values for variables.



East

Fig0A2. Random effects for diameter growth (a) and fecundity (b) for *Ulmus* in stand DB showing lack of spatial coherence. Positive and negative values are red and blue, respectively, with symbol size scaled by magnitude of the effect. For fecundity, black dots indicate immature individuals. Differences are plotted against distance in Fig0 A3.

We initialize fecundity using an approximation from equation 2, which has the solution

$$\mathbf{f}_{j,t} = \left(\mathbf{F}_{j,t}^{T}\mathbf{F}_{j,t}\right)^{-1}\mathbf{F}_{j,t}^{T}\left[\left(\mathbf{A}_{j}^{T}\mathbf{A}_{j}\right)^{-1}\mathbf{A}_{j}^{T}\mathbf{g}_{j,t} - \mathbf{1}_{m}cB_{j}\right]$$
A.13

This solution requires the prior mean value for dispersal parameter u, and it is available only if $n_j > m_j$. We initially set all tree years for which an individual is known mature, based on observations (Table 7), or is unknown but exceeds the diameter at which $p(Q_{ij,l}) > 0.5$, or diameter $D_{ij,t} > -\beta_0^{\theta} / \beta_1^{\theta}$, to $Q_{ij,t} = 1$. Prior values are used for $(\beta_0^{\theta}, \beta_1^{\theta})$. If $n_j < m_j$, then we first select the m_j trees closest to seed traps and solved for their fecundities $\mathbf{f}_{j,t}$. If any of the solutions in $\mathbf{f}_{j,t}$ exceeded $f^{(2)}/2$ or were less than $2 \times f^{(1)}$, their values are set to these limits. We then solve for the remaining fecundities $\mathbf{f}_{j,t}$ conditioned on $\mathbf{f}_{j,t}$, with kernel matrix $\mathbf{F}_{j,t}^{"}$,

$$\mathbf{f}_{j,t}^{\prime\prime} = \left(\mathbf{F}_{j,t}^{\prime\prime}\mathbf{F}_{j,t}^{\prime\prime}\right)^{-1}\mathbf{F}_{j,t}^{\prime\prime}\left[\left(\mathbf{A}_{j}^{T}\mathbf{A}_{j}\right)^{-1}\mathbf{A}_{j}^{T}\mathbf{g}_{j,t} - \mathbf{1}_{m}cB_{j} - \mathbf{F}_{j,t}^{\prime}\mathbf{f}_{j,t}^{\prime}\right]$$
A.14

This approach is sensible, because trees closest to seed traps are most strongly constrained by seed counts and thus contribute strongly to dispersal estimates, which contributes in turn to estimates of others. This approach was faster than an Expectation-Maximization approach, but EM would have also work here.

Extensive experiments with different initial conditions were the basis for development of adaptive sampling and initializations discussed above, including thousands of experiments with all species. Like any large model, sensible initialization is required, but experiments with the model, including diagnostics summarized in the text, confirmed that it converges to the posterior.

Although many initializations were used in experiments, production runs came from single long chains. Convergence was achieved with 10,000 iterations for species with moderate numbers of individuals, but required up to 50,000 iterations for trees with many individuals. There are a large number of parameters, not all of which could be sampled efficiently. The slowest updating was obtained for fecundities, one of two Metropolis steps, due to the discrete nature of $Q_{ij,t}$ and H_{ij} , and the blocking over all tree-years within a plot. We selected for updating at random 30% of the trees for a given iteration and embedded 5 such iterations within each Gibbs step. A counter had been inserted into the submodel for monoecious species for development purposes, with results shown in Table A1. These rates are the range that is expected to optimize mixing. No such counter had been included in the submodel for dioecious species.



Fig0A3. Pairwise differences between random individual effects plotted against the distances between individuals. Diameter growth and fecundity are shown in blue and red, respectively.

Execution times for the slowest fecundity submodel in Table A1 are shown for a Dell T5500 workstation (Quad Core Xeon processor E5504 series, 4×2 GHz). CPU time (actual time the processor is engaged), were always < 0.02 s, but clock time was > 3 s for *Acer*, which included 80,930 tree-years. Production runs were completed on the Duke University cluster with

faster processors than used for experiments reported in Table A1, depending on the specific node.

The rapid convergence was achieved due to an informed approach to initialization, the fact that proposals were efficient, and random effects were marginalized. At the initialization stage, we have approximate fecundities for tree-years. There is still substantial mixing required, because we cannot know which individuals are mature and female. Nonetheless, starting with the majority of individuals close to the posterior mean speeds convergence for remaining individuals.

Genus	Acceptance fractions				Execution (s)	
	Auxiliary variable	State-space model	Seed data likelihood	Total	CPU time	Elapsed time
acer	-	-	-	-	0.021	3.562
betula	0.433	0.431	0.294	0.386	0	0.39
carpinus	0.511	0.491	0.328	0.444	0	0.291
carya	0.411	0.408	0.323	0.38	0	0.781
cercis	0.642	0.637	0.268	0.515	0	0.276
cornus	0.314	0.313	0.264	0.297	0	0.949
fagus	0.653	0.652	0.619	0.642	0	0.019
fram	-	-	-	-	0	0.59
list	0.214	0.2	0.119	0.178	0.001	0.517
litu	0.454	0.452	0.292	0.4	0	0.976
nyssa	-	-	-	-	0	0.783
pinus	0.201	0.199	0.134	0.178	0	0.762
quercus	0.261	0.26	0.195	0.238	0	2.172
robinia	0.46	0.46	0.44	0.453	0	0.164
tilia	0.605	0.604	0.357	0.522	0	0.17
tsuga	0.704	0.704	0.59	0.666	0	0.316
ulmus	0.387	0.374	0.148	0.303	0	0.415

Table A1. Acceptance rates for three types of proposals in the Metropolis step and execution time for the fecundity submodel, the rate-limiting step in the algorithm.

Fecundities were sampled efficiently because the proposal strategy included alternately matching proposals with the likelihood for seed data, the state-space model, or an auxiliary variable, being the most recent estimate. Had we not included this proposal, it would have been difficult to accept proposed fecundities for trees previously immature. The burst of new seed

from these individuals would be unlikely, in terms of seed data, unless balanced by a proposed reduction by other trees. This algorithm results in high acceptance rates for maturation status and effective mixing. Alternating proposals in this fashion minimizes the probability that the algorithm becomes trapped in one part of the posterior.

If they cannot be marginalized out, random effects substantially slow convergence. The marginalization in eqn A.12 means that fixed effects do not depend on convergence of random effects. Rapid convergence of fixed effects, in turn, speeds convergence of random effects. Sample chains show this rapid convergence (Fig. A1).

Burnin steps were discarded, followed by >100,000 iterations that were retained for analysis, thinned to reduce storage overhead. We inspected Gibbs chains for all parameters as well as for samples of individual effects (e.g., Fig A1). To help evaluate results we compared priors and posteriors, we considered predictive capacity, in terms of data used to the fit model, and we compared predictive intervals from the model with estimates of latent states that could not be directly observed (see main text).

Analysis of the random effects showed lack of spatial coherence, thus obviating need for a spatial prior. Maps of random effects (Fig. A2) showed no tendency for spatial clustering of these effects. Plots of pairwise differences in random effects summarize the evidence for lack of spatial coherence (Fig. A3).

References

- Clark, J. S., E. Macklin, and L. Wood. 1998. Stages and spatial scales of recruitment limitation in southern Appalachian forests. Ecological Monographs 68:213-235.
- Clark, J.'S., O 0Wolosin, M. Dietze, Kolbanez, UOLaDeau, O 0Welsh, and D0Kloeppel. 2007b. Tree growth inferences and prediction from diameter censuses and ring widths. Ecological Applications 174942-1953.
- Niklas, K. J. and B. J. Enquist. 2003. An allometric model for seed plant reproduction. Evolutionary Ecology 5:79–88.
- Ryan, M.'G. and B.'J. Yoder 1997. Hydraulic limits to tree height and tree growth. Bioscience 47: 235-242.