A minimal model of a forest after neutral theory

Matthew D. Schwartz
Department of Physics, Harvard University

James P. O'Dwyer
Department of Plant Biology, University of Illinois at Urbana-Champaign

Abstract

Community models aim to shed light on the composition and dynamics of ecological systems; properties like which species are rare, which are common, and how their populations change over time. Neutral theory is an ambitious example of this type of model. Treating individuals alike, neutral theory predicts the overall shape of rarity and abundance. But it doesn't tell us which species will gain or lose over time, and its predictions for timescales in forest communities underestimate the pace of change. Coupling this model to environmental variation, and incorporating species' life-history differences, have been proposed independently as remedies for these failings. But neither has made predictions for each species' dynamics within a community, or its long-run abundance. Here we develop a new, minimal model, joining these two processes with immigration from outside. This minimal model successfully predicts the dynamics and abundances for species in the Barro Colorado forest dynamics plot, in Panama. Species are characterized by their measured life histories; how fast they grow, survive and reproduce. We also show that environmental variation over time couples in a surprisingly simple way to the community, as a single common effect on growth rates, shared across all species. These two features allow the model to predict species dynamics, while regional commonness and rarity, alongside a simple competition-colonization trade-off, predict relative abundances in the plot. Our model could a priori have been very complex; but what we read from the censuses is much simpler than the ingredients of the model would have allowed, and yet predicts dynamics and abundances well. Given appropriate data, our approach can be applied across any forest. Whether a comparably simple model performs as well at different locations remains an open question.

Keywords: Barro Colorado Island; environmental stochasticity; forest dynamics; immigration; life history; matrix population model; neutral theory; species abundance distribution; stage structure; tropical forest

1 Introduction

A model of a forest community is asked two things above all. One is which species will gain or lose trees over the coming decades, the forecast that repeated censuses are taken to inform. The other is which species the forest holds in the long run and at what abundance, the oldest question of community ecology (Preston, 1948). With them come the stand's census-to-census fluctuations and its distribution of stems over sizes. Three kinds of answer have been offered. Niche theory explains where each species is common by its match to soil, water and light (Engelbrecht et al., 2007; Condit et al., 2013), but seldom in numbers for a whole community. Demographic models project a population, or a few functional types standing for many species, from measured rates of growth, survival and reproduction (Lefkovitch, 1965; Caswell, 2001; Moorcroft et al., 2001). Stochastic community models, descending from the lottery model (Chesson and Warner, 1981), predict the

distribution of abundance across species from births, deaths and recruitment into limited space. Neutral theory is their tractable extreme: every individual, whatever its species, has the same odds of dying, reproducing and arriving from the surrounding region (Hubbell, 1979, 1997, 2001). It fits species-abundance distributions across many kinds of community with two or three parameters (Volkov et al., 2003; Rosindell et al., 2011), has a closed-form stationary state and sampling formula (McKane et al., 2000, 2004; Etienne, 2005), and by that parsimony became the standard null model against which richer models are judged (McGill, 2003; Chave, 2004).

Its limits are of two kinds, one structural, since a species-symmetric model cannot name a species, the others contingent, predictions inconsistent with the repeated censuses of the 50-ha plot on Barro Colorado Island (BCI), Panama (Condit et al., 2012, 2017). There, species composition changes several times faster than demographic drift allows, and the excess over drift grows with a species' abundance (Chisholm et al., 2014; Kalyuzhny et al., 2015; Fung et al., 2016, Schwartz & O'Dwyer, companion manuscript). Two remedies for this excess have been proposed independently. The first, environmental variance, reproduces the observed rate of change at BCI in otherwise neutral models in which each species responds separately to a fluctuating environment (Chisholm et al., 2014; Kalyuzhny et al., 2015; Fung et al., 2016). Other non-neutral mechanisms favor whichever species is currently rare (Chesson, 2000). One of the oldest such mechanisms proposed for tropical trees is the pressure of natural enemies near a species' own adults (Janzen, 1970; Connell, 1971), a density dependence since measured at BCI and across plots worldwide (Comita et al., 2010; Kalyuzhny et al., 2023; Hülsmann et al., 2024). A whole-community model, however, needs each mechanism quantified for every species at once, and so far enemy effects, niche differences and species-specific environmental responses have been characterized a few species at a time or by one estimated parameter per species.

Life-history variation, the second remedy, is measured directly and for whole communities. Species differ in how fast they grow, how long they live and when they reproduce, in ways that partly determine which species gains stems in a decade. Repeated censuses measure these rates (survival, growth and recruitment by size) for every species (Condit et al., 2017; Jops et al., 2025). Demographic noise acting on stage-structured populations is by itself not enough. With each species' life history in it, that noise raises the amplitude of abundance change above the unstructured neutral value and toward what the censuses show, without reaching it (Jops et al., 2025). More to the point, it is silent on direction: it says how far a species is likely to move and not which way; and the model that established it left immigration out by design, so it makes no statement about which species the forest holds in the long run or at what abundance. A forecast needs both. The direction of each species' change must then come from its measured rates, and its long-run abundance from regional immigration, as in neutral theory, where locally common species are largely the regionally common ones.

Our minimal model joins two bodies of work that have mostly run in parallel, and the join turns out to predict more than the sum of its parts with a model simpler than either part would lead one to expect. It combines stage-structured demography (Lefkovitch, 1965; Caswell, 2001), regional immigration and a shared fluctuating environment under one carrying capacity for all censused stems, each dead stem being replaced by a seedling drawn at random from those of all species (Fig. 1). Immigrant seedlings arrive at a fixed rate in proportion to each species' share of the surrounding forest, computed with the plot itself left out so that it does not seed its own external pool. Two forms of the immigrant stream are run in parallel. In the first, the naive one, arrivals are weighted by regional abundance alone, and a species' local life history then decides how many standing trees each arrival yields. In the second, arrivals are weighted by regional abundance and against that local yield exactly, a competition-colonization trade-off (Levins and Culver, 1971; Hastings, 1980; Tilman, 1994) in its simplest form.

The environmental part of such a model could easily be its most heavily parametrized. Once populations have stage structure, environmental variation can have it too, a good or bad year acting on seedling establishment, sapling growth or adult survival, and on each species differently. That is more realistic than an unstructured response but has more parameters than even one response per species. Stage structure could also do the opposite: one factor shared by all species and acting on one life stage moves species with different life histories by different amounts, because each has a different fraction of its stems passing through that stage at the time, so that a single shared effect, with an amplitude and a memory, might do what unstructured models need a response per species to do. Not all of that freedom need be present in a real forest, or needed for a forecast, and one of our questions is whether allowing for stage structure reveals a hidden simplicity or inflates the model, which we answer by admitting a factor for every species on every transition and estimating its form from the forest's own transition data.

Here we ask one minimal model for both predictions, parametrizing it for 84 BCI species, those with published projection matrices (Jops et al., 2025) that gained or lost at least one tree of 10 cm dbh over the censuses: each species' change in abundance one and two censuses ahead, the test a species-symmetric theory cannot take, and the long-run relative abundances by name, with that state's abundance distribution, size structure and census-to-census fluctuations (Chisholm et al., 2014; Kalyuzhny et al., 2015). Demanding both from one small set of shared constants keeps the model simple enough for its residuals to mean something: where it misses, the discrepancy points to a mechanism. We find that each species' measured life history forecasts the direction and relative size of its change a decade ahead, better than models without life-history information, and that the long-run rank order of species follows regional abundance. We estimate the form of the environmental noise, whether all species share it and which rates it multiplies, from the recorded stem transitions. What a repeated census allows is to read the form rather than assume it, and at Barro Colorado the reading is unexpectedly simple: the effect of the environmental noise amounts to one number per census interval, the same for every species, acting on growth, the promotion of stems from one size class to the next, with an amplitude and a memory. Our model differs from the community model of Jops et al. (2025) in adding regional immigration, simulating all species together under a cap on censused stems, and testing each named species' abundance and decadal change.

2 Material and methods

2.1 The minimal model

We represent each species s as a stage-structured population (Lefkovitch, 1965; Caswell, 2001), a vector $n_s(t)$ of stem counts in a seedling class below the census threshold and in diameter classes above it. We project this vector one census interval Δt ahead by the species' measured matrix,

$$n_s(t+\Delta t) \;=\; A_s\, n_s(t), \qquad A_s \;=\; U_s + F_s, \tag{1}$$

where U_s contains survival and movement among classes and F_s seed production. We realize the projection stem by stem, with multinomial draws for survival and movement and Poisson draws for seed production (Fig. 1). A stem here is interpreted as one individual, counted once at its largest stem, which is how census tables count. The rates in each matrix A_s were measured in the standing community and so already include part of the effect of competition.

The one capacity in the model stands for competition for space and resources. We cap the censused stems of all species together at their observed number, averaged over censuses. When a

censused stem dies its place is filled by a seedling drawn at random, whatever its species, from that interval's candidates, those whose own transitions take them out of the seedling class (Chesson and Warner, 1981). If candidates outnumber vacancies this lottery admits as many as fit, the others remain seedlings, and no standing stem is removed. The seedling class therefore has no cap, and its size is a prediction. We leave seedlings outside the cap because smaller stems use fewer resources, so that a cap treating a large tree and a seedling as equivalent would be biologically unreasonable. By the same reasoning the capped stems could be weighted by basal area, or by some other scaling of resource use with size, a generalization we explore in Supporting Information S4, where the predicted composition proves insensitive to it. As a simple approximation to the variation of resource use with size, then, we cap the number of individuals in the census and leave seedlings outside it.

Immigration enters as a fixed arrival rate, M seeds per interval in expectation (Section 2.3), divided among species by weightings w_s that sum to one, the seedling class of species s receiving a Poisson number of immigrant seeds with expectation Mw_s each interval. An imposed stream of this kind, the form neutral theory itself uses, is what keeps every species present in the long run: with no supply from outside, a mixture of species each with growth rate exactly one drifts until one species is left. The weightings derive from regional relative abundance p_s, each species' fraction of reproductive-size stems in the region's census plots, the focal plot excluded. A species the regional sample happens to miss is entered at the smallest weighting any sampled species receives, reflecting what we take to be its rarity in the regional pool.

Environmental variation enters as the year effect, one random factor g_t per census interval. In general form it may multiply any transition of any species' matrix: survival, and so mortality, in any class; the promotion of a stem from one diameter class to the next at any size; and, through the fecundity row, seed production and the establishment of seedlings. It may be common to all species or drawn separately for each, with or without memory between intervals, and its form, like the mean rates, is estimated from each forest's transition data. With T the set of transitions multiplied, the same in every species, the year effect acts in interval t as

$$\begin{gathered}[A_s]_{ij} \;\to\; g_t\,[A_s]_{ij}, \qquad (i,j)\in T, \\ \ln g_t \text{ Gaussian}, \qquad \mathbb{E}[g_t]=1, \qquad \mathrm{sd}(\ln g_t)=\sigma_g, \qquad \mathrm{corr}(\ln g_t,\, \ln g_{t+\Delta t})=\phi=e^{-\Delta t/\tau_g},\end{gathered} \tag{2}$$

with unpromoted stems staying in their class and all other entries unchanged. We estimate T and the amplitude σ_g from the transition data at Barro Colorado and set the memory τ_g by assumption (Sections 2.4 and 2.5). There T proves to be the growth (promotion) rates of the three sapling classes, out of classes 2, 3 and 4, so that the shared effect acts on the growth of small trees and not on mortality, whose common variation is weaker and out of step with it, nor on seed production and establishment.

2.2 Stationary state and local gain

For a tractable analytical result we first remove the environmental variation, setting $g_t = 1$ in Eq. (2), and replace each stem-by-stem draw by its expectation, a mean-field approximation. The stationary state of these deterministic dynamics has a closed form (Supporting Information S2); the stochastic simulations restore both the demographic noise and the environmental variation. Competition then enters through a single number common to all species, the fraction d of candidate recruits left in the seedling class (class 1) for lack of a vacancy. We write $U_{d,s}$ for U_s with its class-1 promotion entries multiplied by $1-d$ and the removed probability returned to the class-1 diagonal, and e_1 for the unit vector on class 1. The stationary stage vector of species s is

$$\bar{n}_s = M\, w_s \left(I - U_{d,s} - F_s\right)^{-1} e_1, \tag{3}$$

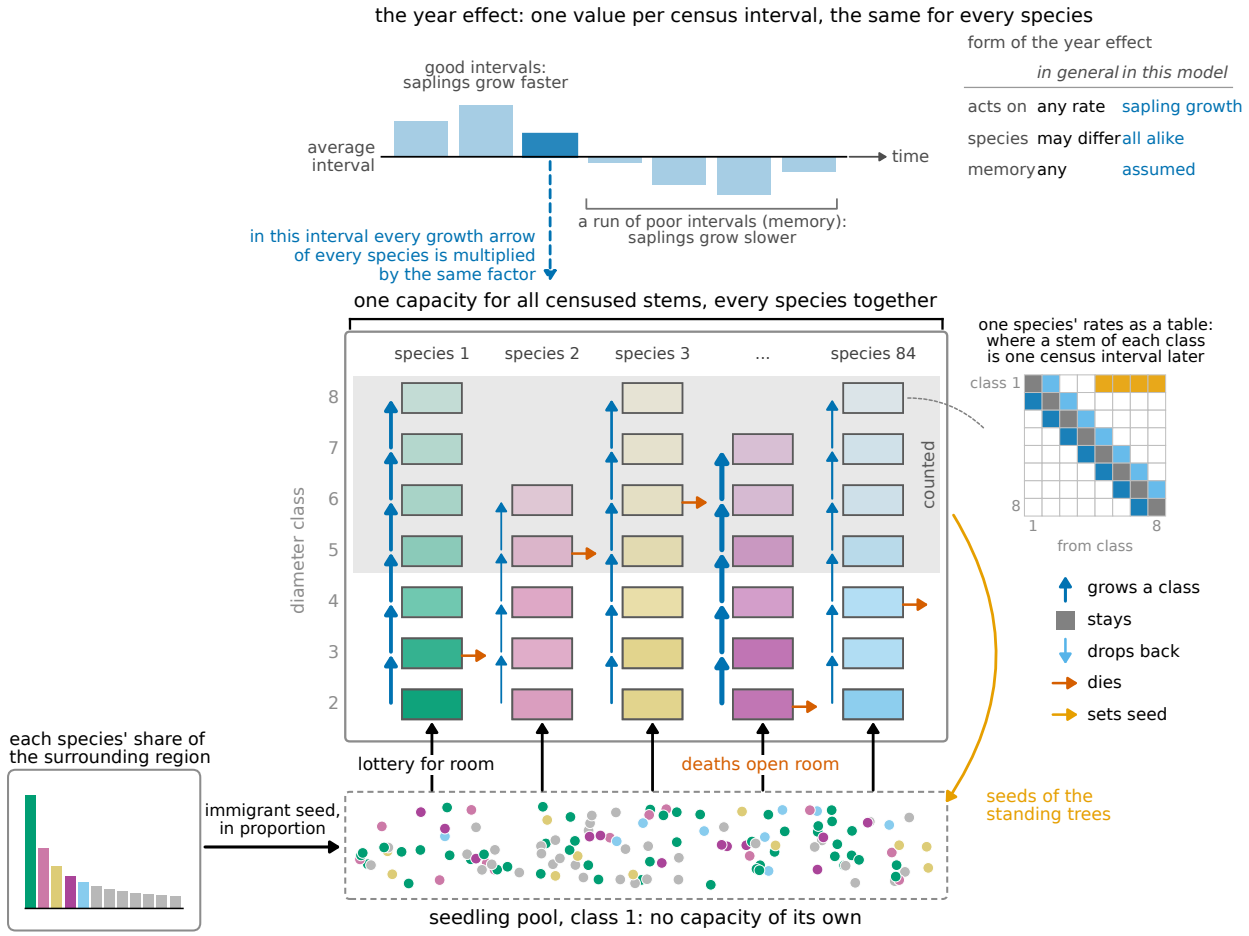

Figure 1: The minimal model. The censused stems of every species stand under one capacity, shared by all species together (diameter classes 2 to 8; classes 5 to 8, shaded and labeled 'counted', are the large trees). Each species is a ladder of diameter classes that its stems climb at the species' own measured rates, a stem staying, advancing, dropping back or dying in each census interval; five of the 84 species are drawn. Deaths open room, and seedlings enter by a lottery for that room, irrespective of species. The seedling pool (class 1) has no capacity of its own and is fed by the seeds of standing trees and by immigrant seed arriving in proportion to each species' share of the surrounding region (lower left). The year effect (top) is one value per census interval, good and poor intervals coming in runs; within an interval it multiplies the growth arrows of every species by the same factor. The key (top right) contrasts the general form such a factor could take with its form in the model, which the Barro Colorado transition data select (Material and methods 2.5: sapling growth, all species alike) except for the memory, which is assumed. Right: one species' rates as the table the model uses (A_s of Eq. (1)). Occupancies, interval values, arrow weights and regional shares are illustrative, and no value in the drawing is particular to one forest.

with d fixed by the capacity condition, that the censused stand hold its observed number of stems, and therefore an output of the model (Table 1).

Because every projection matrix has growth rate one, Eq. (3) factorizes (Supporting Information S2): above the seedling class every species stands in its own stable stage structure. The stationary number of large trees (the four largest diameter classes) of species s is

$$\bar{c}_s = M\, w_s\, G_s, \qquad G_s = \frac{(1-d)\,\Omega_s}{d\,\mu_s}, \tag{4}$$

with μ_s its seedling mortality per interval and Ω_s the number of large trees per seedling in its stable structure. The local gain G_s, the number of large trees standing per arriving seed, summarizes the whole life history: every species falls short of replacement by about $d\mu_s$ per interval, and its abundance is its immigration divided by that shortfall, projected into the large-tree classes. Under the first scheme, the regional pool ($w_s = p_s$), the large-tree composition is proportional to $p_s\Omega_s/\mu_s$, independent of M and d. Under the second, the compensated scheme, we offset the local advantage of species with large gains by also weighting arrivals inversely to the gain, $w_s \propto p_s\mu_s/\Omega_s \propto p_s/G_s$, a competition-colonization trade-off (Levins and Culver, 1971; Hastings, 1980; Tilman, 1994). By Eq. (4) the composition is then p_s itself, whatever d and M.

With equal gains the composition would follow regional abundance, and without immigration drift would leave one species, life history acting only through demographic variance (Jops and O'Dwyer, 2023). With immigration and unequal gains, that inequality is an induced fitness difference (Jops et al., 2025) acting only through competition for recruitment from the seedling class. The advantage of the steadier life history in the two-species problem of Turkia and Shnerb (2025) acts here through G_s rather than beside it: the abundance standing per immigrant is a lineage's expected lifetime, which at growth rate one varies inversely with demographic variance, and G_s is that quantity for a structured life history, so that across species the gain falls as demographic variance rises. Equations (3) and (4) are mean-field results, not exact expectations of the stochastic model, whose lottery is nonlinear, but the simulations with full demographic noise reproduce them: no species drifts to fixation, long-run abundances match the closed form (under the regional pool the median species within 0.6%, the large-tree total within 0.5%), and the small departures show no trend with demographic variance (Supporting Information S2), so that demographic noise scatters the composition about the mean-field state without displacing it or favoring the steadier life histories further. With the year effect on, the simulated composition stays at the closed form's and only the fluctuations about it change (Supporting Information S2).

2.3 Barro Colorado data and model constants

On the BCI plot every free-standing stem of at least 1 cm diameter at breast height (dbh) is recensused every five years (Condit et al., 2012, 2017, 2019), and we use the censuses of 1982–2015 (seven intervals, the first lasting three years, 33 years in all). Of the 90 species with published stage-structured matrices (Jops et al., 2025; Jops and O'Dwyer, 2024), we follow the 84 that recruited or lost one or more trees of at least 10 cm dbh over these censuses, and keep the matrices' eight stages (Table 1). Other stage structures, from four classes to a continuum, lead to the same qualitative conclusions (Supporting Information S1). The matrices act as five-year operators, matching the census interval, and their published normalization sets the long-run growth rate of every species to exactly one. Fecundity is not observed directly. Jops et al. (2025) inferred it from the diameter class at which trees of each species typically reach reproductive maturity and a seed output rising as the square of diameter, and found their conclusions insensitive to the exponent.

The two predictions we ask for, each species' change a decade ahead and its long-run abundance, are measured on the species' large trees, stems of at least 79.4 mm dbh (classes 5–8). This is the published class edge nearest the 10-cm threshold at which trees are conventionally counted and plots compared (Condit et al., 2002); all-stem results are in Supporting Information S7. Four community-level constants complete the model, with no free parameter per species (Table 1): the capacity (the mean count of censused stems of the modeled species over the censuses of 1982–2010, 200,721), the arrival rate, and the amplitude and memory of the year effect, which alone acts on both timescales.

We estimate the regional shares p_s from the network of census plots across Panama (Condit et al., 2022, 2016), excluding BCI and the six plots on or beside the island. The nine species absent from the remaining plots take the smallest sampled share. The arrival rate is the product of two numbers, neither of which needs to be precise, because no censused prediction depends on M and the seedling density (Section 3.4) responds to it only through the immigrant share of the seed input: an immigrant fraction of recruits $\hat{m} \approx 0.026$, the value at which Hubbell's model best reproduces the census-to-census turnover of the plot's trees of at least 10 cm dbh, and a nominal whole-stand seed production B of 2.80 million seeds per interval, the order the matrices' fecundity rows give for a stand of this size, not a measured seed rain. Their product is $M = \hat{m}B = 73{,}729$ seeds per interval ($\hat{m}$ unrounded), about 15% of the seedling input at stationarity under the regional pool, the rest being seed of the standing trees, which the matrices fix; varying M widely in simulations leaves the predicted composition unchanged (Supporting Information S2). The class-1 rates come from a separate census of free-standing woody stems at least 20 cm tall and below 1 cm dbh in 19,313 one-square-meter plots (Comita et al., 2023a,b). We compare their mean density over the fourteen censuses of 2001–2018 with the model's seedling pool in Section 3.4.

2.4 Year-to-year variation in the transition rates

For every species and interval the censuses count the stems of each diameter class promoted to the next, the deaths, and the recruits across 10 mm dbh. These counts fix each species' mean projection matrix and, taken interval by interval, show how the rates vary about their means, which the year effect must represent (Fig. 2). Sapling growth, the promotion rates out of the three smallest censused classes, varies in common across species, high from 1982 to 1990, low from 1990 to 2000 and near its mean since. One shared component accounts for 71% of its systematic between-interval variance. Promotions among large trees follow that component with a loading of 0.86 (90% CI 0.75–0.97). Recruitment across 10 mm dbh per standing stem, itself a growth transition out of the seedling pool, follows the common variation with a larger amplitude. Species respond in the same direction: of the 175 sapling-promotion series of individual species, 99% move with the common component estimated from the other species.

Mortality has a common component weaker than that of sapling growth and not synchronized with it, both below and above the large-tree threshold. Yet within a species small-stem mortality varies about as much as growth (median species' standard deviation 0.19 on the log scale, versus 0.19–0.27 for the sapling promotions), and this variation is largely particular to the species. We measure it (Supporting Information S3) and, by the criterion of the next subsection, leave it out of the model. A year effect particular to each species has no persistent direction and no synchrony across species, so that to first order it leaves each species' mean matrix, and hence the decadal forecast and the long-run composition, unchanged, and partly cancels in the large-tree total. A single environmental effect on growth shared by every species can, by contrast, produce large differences between species in how their abundances move, because it acts through stage structures that differ, whereas coupling each species' mortality separately to unmeasured environmental variables would

Table 1: **Constants of the model: the rates are estimated from individual stems' growth, survival and recruitment transitions, the capacity is the census total of stems and the immigrant fraction comes from the turnover of stems between censuses, and the memory of the year effect is assumed; none is fitted to any species' abundance, to changes in abundance or to the variability of any count.** Values for the 84 species at Barro Colorado, with the meaning of each, its source (published per species, set once for the whole community from a census quantity, or a modeling choice) and the quantity from which it is set. Classes 2–8 of the matrices are diameter classes whose lower edges nearly double from class to class, starting at 10 mm, with class 5 beginning at 79.4 mm. Class 1, the seedling class, holds stems below 10 mm, whose rates were estimated from the annual seedling census. The amplitude and set of transitions of the year effect are estimated, and its memory set, in Material and methods 2.5 (alternative forms in Supporting Information S3).

quantity	symbol	meaning	source and value
projection matrices	U_s, F_s	five-year survival and growth (U_s) and fecundity (F_s) among eight stages; long-run growth rate one	published per species: fits to the BCI censuses of 1982–2015 and the seedling census (Jops et al., 2025)
species followed	s	the 84 species that have a published matrix and that recruited or lost one or more trees of at least 10 cm dbh over the censuses used	the plot's censuses; the other six species, none of which had a tree of at least 10 cm dbh, are listed in Supporting Information S1
large trees		stems of at least 79.4 mm dbh (classes 5–8), the published class edge nearest the 10-cm threshold at which plots are conventionally compared (Condit et al., 2002)	the abundance compared throughout; all-stem results in Supporting Information S7
capacity		censused stems of those species, classes 2–8, all species together	set once: the observed censused-stem total, mean over the censuses of 1982–2010, 200,721
regional shares	p_s	share of reproductive-size stems in the regional plots; a species the sample misses takes the smallest share	set once: the Panama plot network with BCI and its six neighbors excluded (Condit et al., 2022)
immigrant fraction	$\hat{m}$	fraction of recruits that are immigrants from outside the plot	set once: the value at which Hubbell's model best reproduces the census-to-census turnover of BCI trees of at least 10 cm dbh, 0.026; affects only the seedling pool
arrival rate	M	immigrant seeds per interval, $\hat{m}B$ with B a nominal whole-stand seed production of 2.80 million seeds per interval	73,729; no censused prediction depends on it; it supplies about 15% of the seedling input under the regional pool and can move the predicted seedling pool by no more than about that share
year effect, amplitude	σ_g	standard deviation of $\ln g_t$ (g_t has mean one)	estimated from the transition data: the component of the three sapling promotion rates common to all species has a standard deviation of 0.22 in log rate over the seven intervals, which a year effect sampled at the census dates reproduces at $\sigma_g = 0.32$ (90% CI 0.17–0.84)
year effect, memory	τ_g	correlation time of g_t	not estimable from the censuses (Material and methods 2.5); assumed: 18 years ($\phi = 0.76$ per five-year step), with 5–40 years carried (Supporting Information S3 and S8)
year effect, transitions	T	the set of transitions g_t multiplies; at BCI the three sapling promotions, out of classes 2, 3 and 4, with one amplitude	estimated from the transition data: one component common to all species, on growth, and no factor on mortality (Fig. 2; Supporting Information S3, where simulations with a shared factor on each transition in turn lead to the same choice)
immigration weighting	w_s	regional share p_s (regional pool), or p_s/G_s with G_s the local gain (compensated scheme); both are reported	modeling choice; the two extremes (Material and methods 2.2)
derived	d	fraction of candidate recruits not admitted to the censused stand	output: 0.153 under the regional pool
derived		large-tree total (classes 5–8)	output, compared in Results with the observed seven-census mean of about 21,700 (21,000–22,200 census by census)
derived		seedling pool below 10 mm	output, compared with the seedling census in Results

make the model conceptually and practically complex, for a gain to be settled by test rather than assumed.

Each species' own mortality variation, which we estimate net of the sampling noise of finite stem numbers, does add to its census-to-census movement though not to its direction, the long-run order or the total's movement, and in Section 3.3 we add it at its measured size to the simulations as a generalization of the minimal model.

Every statement above also holds on 1985–2015 (Supporting Information S3); the first interval, 1982–1985, when small stems were censused by a different method, shows the largest common deviations in growth (Fig. 2).

2.5 Selecting the environmental terms and setting their constants

The transition data offer more environmental terms than a predictive model should keep, year effects common to all species and particular to each, on growth, mortality and recruitment, all measurable. We choose among them by asking which are large enough to change a prediction rather than which can be detected, and then set the two constants of the term kept, its amplitude and memory. In penalized comparisons of year-effect structures fitted to the 4,228 transition counts by species, interval and rate (4.0 million stem-intervals; Supporting Information S3), every candidate term is detectable and the Bayesian information criterion retains nearly all. What the model *should* include is decided differently. A term enters the model if it is large enough to move a species' decadal change, the long-run composition or the stand's census-to-census movement, and is otherwise left out with its measured size reported. By that test we include the interval effects on growth common to all species and leave out the species' deviations from them, which account for 10% of the systematic between-interval variance in sapling promotion (Supporting Information S3).

We therefore let g_t of Eq. (2) multiply the three sapling promotions of every species (the set T) with one amplitude, leaving promotions among large trees at their mean rates, since these do not change the large-tree total. Simulations with a shared factor on each transition in turn support the choice (Supporting Information S3). On seed production or establishment (the usual assumption) the variation is absorbed within the sapling classes, each holding a stem for 13 to 25 years, comparable to the memory τ_g assumed below, so that T omits establishment despite its shared variation. On survival the variation reaches the large trees only by changing the total or the order of species. Only on sapling promotion does large-tree recruitment vary as observed with the total and the order intact.

Simulated at the census dates and sample sizes, the year effect reproduces the spread of the pooled sapling path in Fig. 2 at $\sigma_g = 0.32$ (90% CI 0.17–0.84), or 0.27 without the first interval. The seven serially correlated intervals cannot constrain the shared factor's memory, and nothing else in the census does. Species' relative performance does persist, the rank correlation of their net gains in trees ≥ 10 cm dbh between consecutive intervals being 0.33 across 255 species and decaying over two to three decades (Supporting Information S3), but a factor common to all species moves them together and produces no such persistence in records simulated from the model, whatever its correlation time; the persistence belongs to species-specific variation the model leaves out. We therefore take $\tau_g = 18$ years ($\phi = 0.76$ per five-year step) as an assumption and carry the range 5 to 40 years, over which the amplitude estimated from the transitions moves from 0.23 to 0.43.

2.6 Forecast and stationarity tests

Because almost no seedling reaches the large-tree classes within two intervals (Supporting Information S5), we project a species' large-tree count N_s one or two censuses ahead from its matrix A_s

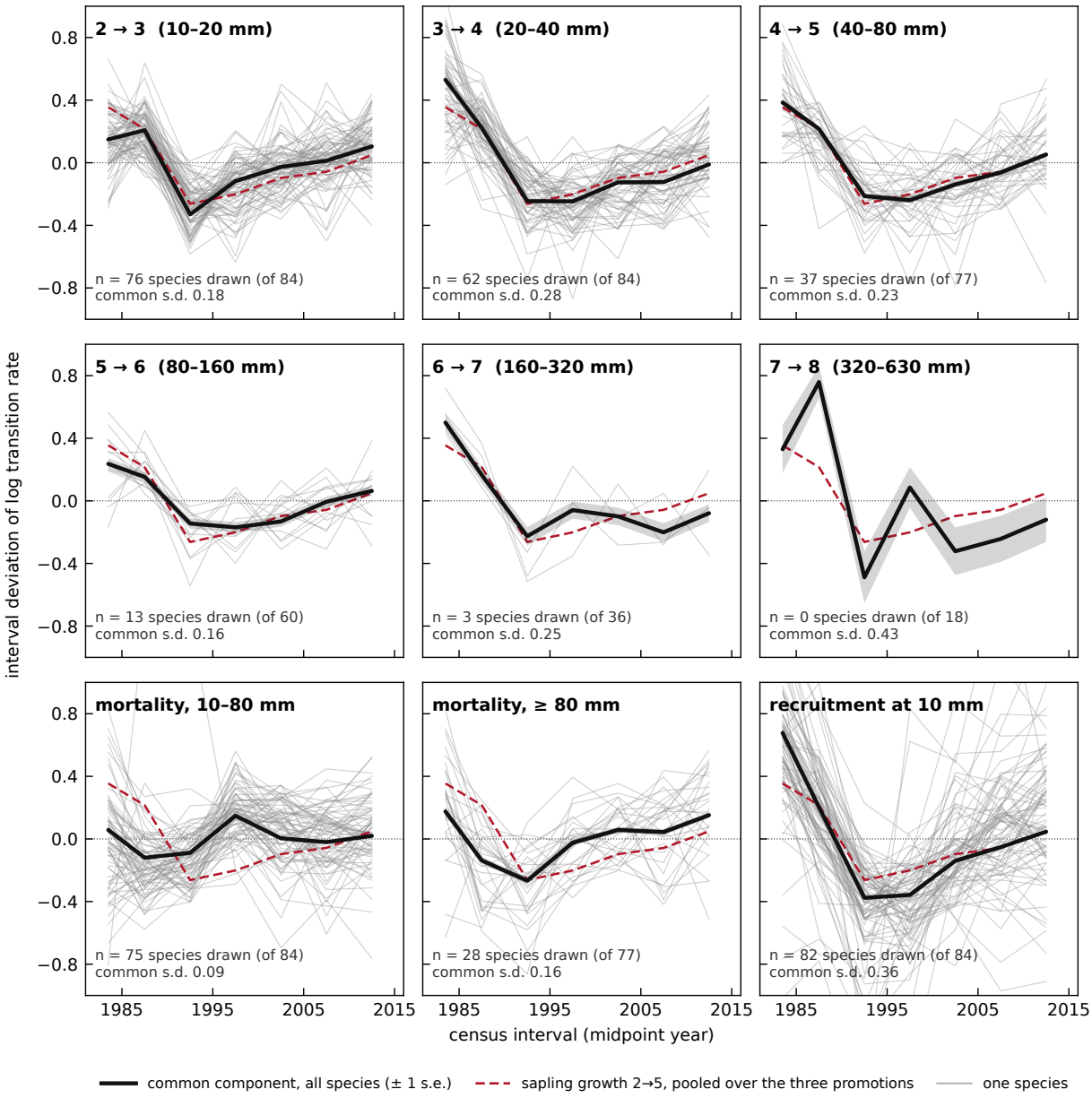

Figure 2: Between-census variation of nine transition rates for the 84 modeled species at Barro Colorado, 1982–2015 (census data; species-level statistics in Supporting Information S3). Top row, sapling promotions (classes 2 → 3, 3 → 4, 4 → 5); middle row, promotions among large trees (5 → 6, 6 → 7, 7 → 8); bottom row, mortality below and above 79.4 mm dbh, and recruitment across 10 mm per standing stem. Gray lines, single species' log-rate deviations from their own seven-interval means, for species with an expected count of at least twenty events per interval (n in each panel; none for 7 → 8). Bold line and band, the component common to all species with its standard error; numbers, the standard deviation of that component. Dashed line, the three sapling promotions pooled into one rate (the pooled sapling path, standard deviation 0.22), from which σ_g is estimated. One vertical scale throughout. Sapling growth varies between intervals in common across species, promotions among large trees follow it, and the common component of mortality is weaker and not synchronized with growth.

and observed stage vector $n_s(t)$ alone:

$$\Delta_5 N_s(t) = \sum_{k=5}^{8} \left[A_s\, n_s(t)\right]_k - N_s(t), \qquad \Delta_{10} N_s(t) = \sum_{k=5}^{8} \left[A_s^2\, n_s(t)\right]_k - N_s(t), \qquad N_s(t) = \sum_{k=5}^{8} n_{s,k}(t). \tag{5}$$

Applying the matrices of Jops et al. (2025) at the first five censuses (420 species-start pairs, one species at one starting census) tests only consistency, because the matrices were estimated from every interval. For the forecast we re-tabulated the rates the same way from the first four censuses and applied them at 1995, 2000 and 2005 (252 pairs). This is the position of most forest plots, which have only a few censuses: the test asks what could have been predicted then, the whole series serving instead to resolve each species' life history more finely (Supporting Information S5). The comparator available uses only the counts: each species' trend, its mean change per interval outside the predicted window (over 1982–1995 for the forecast), continued; Supporting Information S5 adds other predictors without demography.

For the projections and the trend we compute the fraction of nonzero observed changes predicted in the right direction and the Spearman correlation over pairs (90% CI by bootstrap over species). A reversal is an observed change opposite in sign to the species' own net change from the first census to the start; we count how many a predictor anticipates among pairs where neither change is zero, in sample at every start census after the first, for the published matrices and for each species' trend up to that start, which cannot reverse (out of sample in Supporting Information S5). We also test the observed changes in N_s against neutral drift under Hubbell's model fitted to this plot, by the rank correlation across species between the change in large trees after 1990 and the demographic momentum measured by then (Caswell, 2001), the eventual change a species' size structure implies under its early rates (Supporting Information S6).

The stationary composition under either immigration scheme follows from Eqs. (3) and (4), and we compare predicted large-tree abundances with census means, since the censuses are serially correlated. We measure agreement in rank order by Spearman correlation, in absolute abundance by the identity-line R^2 on log abundance (negative when the observed mean alone fits better) and in the abundance distribution by Kolmogorov distance. For fluctuations we compare model and census values of the median absolute five-year relative change of the large-tree total and the between-census coefficient of variation of large-tree recruitment. Within ten abundance bands we compare the mean squared five-year change in a species' large trees. The tolerances for these comparisons, a factor of two by abundance band and a factor of three on the seedling total, were set in advance (Supporting Information S8).

3 Results

3.1 Decade-ahead change of each species

Within the census series (in sample), the matrices of Jops et al. (2025) project changes in each species' large trees whose rank correlation with the observed changes is 0.81 (90% CI 0.75–0.87) one census ahead and 0.86 two ahead (Fig. 3a). Out of sample, rates from the 1982–1995 censuses alone are applied at 1995 and the next two censuses (Fig. 3b): the forecast one could have made in 1995, and the kind available at the many plots censused only a few times. These forecasts have a rank correlation of 0.71 (90% CI 0.60–0.80) with the observed changes one census ahead and 0.68 two ahead, and their sign is right in 74% of the pairs whose count changed. Matrices from the whole series do better chiefly because more censuses resolve each species' life history more finely, not because they have seen the outcome (Supporting Information S5): the better a life history

is measured, the better the forecast. Within a species, forecast and observed deviations of each start from the species' mean change are positively correlated as well (0.47), so that the forecasts also time each species' changes. They also anticipate changes of direction: of the 128 species-start pairs one census ahead, at every start after the first, whose observed change reverses the species' net change since 1982, the published matrices (in sample) anticipate 78, 61% (90% CI 52–69%), the misses presumably reflecting rates changing within the decade (Supporting Information S5). In absolute numbers these forecasts run high (+1,729 large trees in sum against -517 observed one census ahead), the early rates coming from the intervals of fastest sapling growth. What makes these forecasts is each species' own life history, its measured rates, acting on the standing variation in its stage structure that past environmental variation has left. The forecast thus tests those rates and that structure alone: run through the full model, with reproduction, immigration and the capacity acting, the predictions rank almost identically (rank correlation 0.999; Supporting Information S5), so the mechanisms of long-run abundance are tested only in Section 3.2.

None of the model's predecessors, neutral, demographic or environmental (Jops et al., 2025; Chisholm et al., 2014; Kalyuzhny et al., 2015; Fung et al., 2016), can make such a forecast. Under neutral drift among equivalent species, Hubbell's model at the plot's own fitted parameters, nothing measured by 1990 could rank the later changes. At Barro Colorado alone that null, neutral drift among equivalent species, is rejected with more than 99.9% confidence: demographic momentum measured in the 1980s ranks the 75 species' changes in large trees from 1990 to 2015 at Spearman +0.67, $P < 10^{-5}$, which none of 600,000 simulated neutral forests and reassignments of the scores reaches (Supporting Information S6). A neutral model whose species share one set of size-dependent rates is rejected as firmly (Supporting Information S6). The simplest alternative that does give a direction uses no biology, each species' own past trend continued. Out of sample it ranks the later changes at 0.54 and 0.51, below the forecasts by +0.17 (90% CI +0.09 to +0.27) one census ahead and by a positive margin in every other division of the censuses (Supporting Information S5); it can never anticipate a reversal, and it amounts to a net growth rate per species that does not persist between decades, which is why our model keeps every species at growth rate one and lets immigration set absolute abundances.

3.2 Long-run composition under the two immigration schemes

Under the regional pool (arrivals weighted by regional relative abundance p_s), stationary large-tree abundance is proportional to $p_s G_s$ by Eq. (4). A forest need not be near this state, which our model reaches only after about 12,000 model years, and 33 years of censuses can compare the present composition with it but not establish that the forest has arrived. Plots subject to large disturbances make the point directly: in the Luquillo forest of Puerto Rico hurricanes reset the stand's structure and shift its composition at intervals far shorter than any approach to a stationary state (Zimmerman et al., 1994; Uriarte et al., 2019). At Barro Colorado the observed order of species already broadly matches the stationary order under the regional pool (Spearman correlation 0.67, 95% CI 0.51–0.79; Fig. 4a). But the regional shares p_s alone correlate with observed abundance at 0.72 (95% CI 0.56–0.83), so that the local dynamics do not improve the predicted order. In absolute numbers the regional pool distributes the stems wrongly, between species and therefore between size classes, with identity-line $R^2 = -0.693$ on log abundance, a large-tree total 2.6 times that observed and 16 species below one stem (5 observed). Recruitment without regard to species favors the life histories whose seedlings most often become large trees, the gain G_s varying by 0.84 in $\log_{10}$ units (standard deviation) over the 79 species reaching that size. That dependence is the pool's one prediction beyond the regional shares, and the census rejects it: abundance relative to regional share should rise with the gain, at log-log slope one, and instead falls (slope −0.20, 90% CI

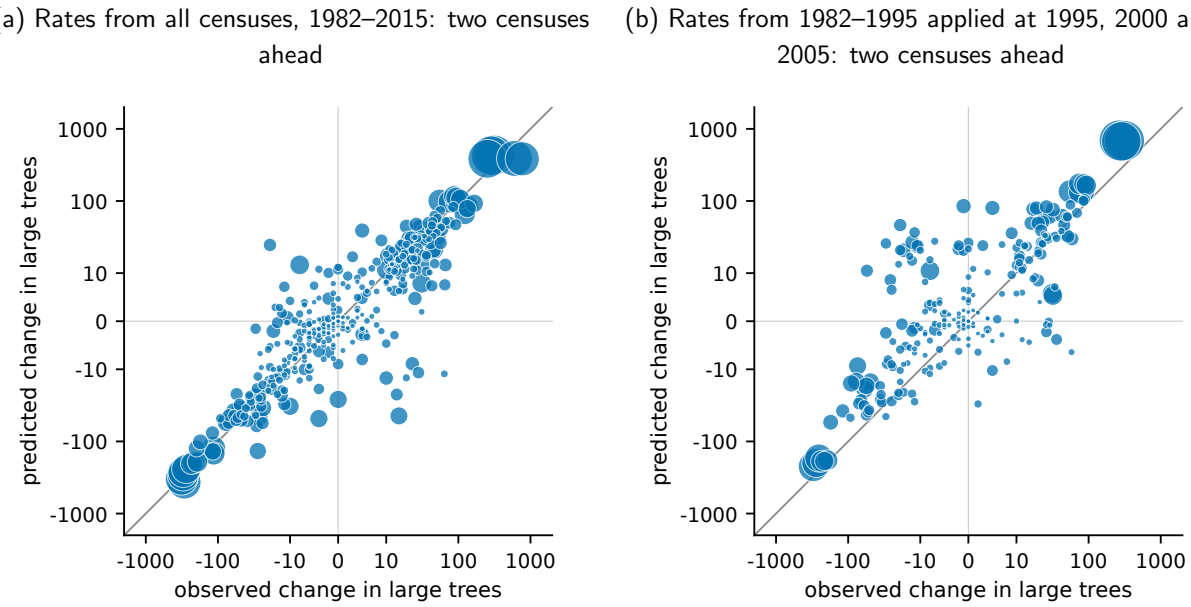

Figure 3: Predicted versus observed change in each species' large trees (stems $\geq$ 79.4 mm dbh, classes 5–8) two censuses ahead at Barro Colorado, one point per species-start pair (one species at one starting census). Symbol area scales with the species' large-tree abundance at the start, the line is the identity, and the axes are logarithmic away from zero. (a) Projection matrices from all censuses of 1982–2015, applied at the first five censuses (420 pairs), a test of consistency because the matrices come from those censuses. (b) Matrices from the censuses of 1982–1995 alone, applied at 1995, 2000 and 2005 (252 pairs), a forecast because none of the predicted intervals enters the rate estimates; in absolute numbers these forecasts are biased upward (Supporting Information S5 and S8). Predictions are each species' matrix applied twice to its observed stage vector (Eq. (5)); observations are census counts. Rank correlations 0.86 (a) and 0.68 (b); sign, error, the trend comparator and one-census panels are in Supporting Information S5.

-0.32 to -0.07). A large local gain does not make a species common on this plot today, although the species now gaining share are those the gain favors, moving toward the pool's composition at a pace of centuries (Supporting Information S7). The pool fails, then, largely because local life history biases the competition it feeds, through the gain G_s.

By Eq. (4), equalizing the gains across species through fecundity and weighting arrivals against the gain ($w_s \propto p_s/G_s$) both give the same large-tree composition, p_s itself. The second correction, the compensated scheme, takes one factor per species from its published matrix and no new parameter. It is a phenomenological repair, what a competition-colonization trade-off would produce, not a measured property of arriving seed. It does better than the alternatives on every count (Fig. 4b, c): it keeps the rank order (rank correlation 0.727), brings absolute abundances close to the census ($R^2 = +0.495$ against -0.693 for the regional pool) and the large-tree total to 1.25 times the observed. As a species abundance distribution, where the regional pool's stationary stand is too steep (Kolmogorov distance 0.290), of the two weightings it is the closer to the census (0.127). It is also closer than a dispersal-limited neutral model (Hubbell, 2001; Volkov et al., 2003) fitted, for this comparison only, to the abundance distribution of the same 79 species (0.18). Since the compensated scheme's rank order is that of the regional shares whatever the local demography, its own tests are the stand's fluctuations (Section 3.3) and the seedling density (Section 3.4). Under

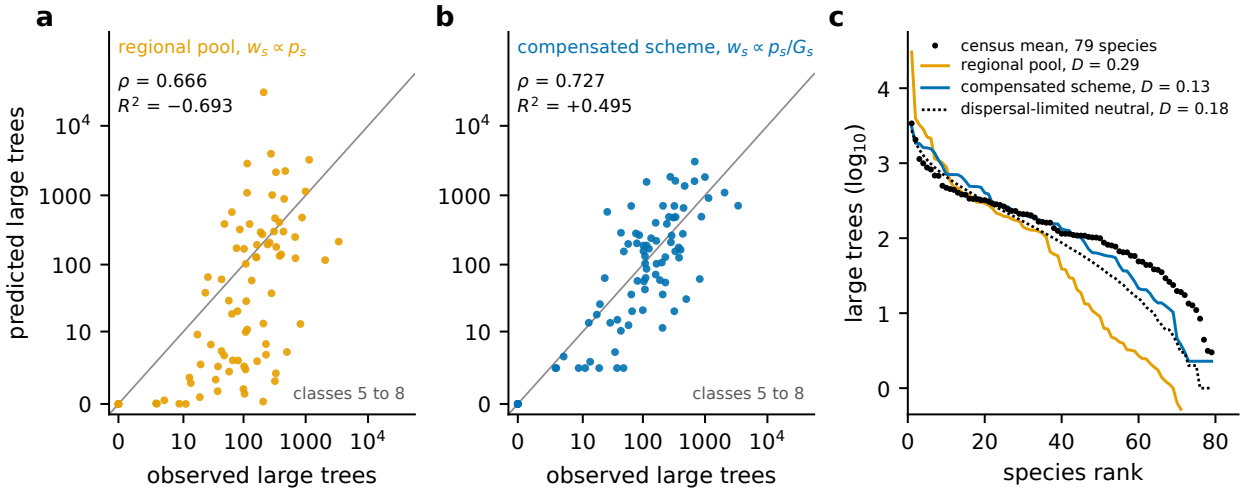

Figure 4: Stationary large-tree abundances (stems of at least 79.4 mm dbh, classes 5–8; Material and methods 2.2, year effect switched off) under the two immigration schemes, compared with the 1982–2015 Barro Colorado censuses. (a) Regional pool, arrivals weighted by regional share p_s: predicted versus observed mean abundance for the 84 species (log axes, identity line); Spearman $\rho = 0.666$, identity-line $R^2 = -0.693$ on log abundance. (b) Compensated scheme, arrivals weighted by p_s/G_s, inversely to the local gain G_s (a competition-colonization trade-off): $\rho = 0.727$, $R^2 = +0.495$. (c) Ranked $\log_{10}$ abundance of the 79 species reaching the large-tree classes: census (points), both schemes (lines), Hubbell's dispersal-limited neutral model fitted, with its two parameters free, to the census abundance distribution of the same species (dotted; $\theta = 21$, $m = 0.0085$). The neutral fit is separate from the $\hat{m}$ of Table 1 and uses neither the regional shares nor species' identities. D, Kolmogorov distance from the census. Intermediate weightings and stems by size class are in Supporting Information S4 and S7.

either scheme the absolute abundances, unlike the rank order, hold only for species of exactly equal long-run fitness, which the published normalization imposes and the censuses cannot certify: with every transition probability redrawn at its census sample size and the matrices left unnormalized, the predicted dominant usually changes and the identity-line R^2 turns negative, while the median rank correlation stays at 0.64 and 0.70 (Supporting Information S7). The total, the R^2, the named dominant and the abundance distribution are therefore consequences of the model at equal fitness; the rank order and the decadal forecasts are its tested predictions. The two schemes read this forest differently, regional supply and local life history jointly setting abundance under the pool, arrivals offsetting the local gain under compensation. On the standing large trees the census favors compensation wherever the two differ; against it stand the seedling total below and the direction of recent change, which follows the gain. We therefore adopt the compensated scheme as the stationary reference, to be tested by the composition of arriving seed, keep the naive pool beside it, and report the two extremes, no compensation and exact compensation.

3.3 Fluctuations of the whole stand

Without the year effect our model is too steady: under the regional pool the five-year changes of the modeled large-tree total are smaller than observed by a factor of 12. With the year effect, one factor common to all species, estimated from the sapling transitions and not from any abundance series, brings the stand's movement to about what the censuses show. Recruitment into the large-

tree classes then varies between censuses with a coefficient of variation of 0.21 (measured 0.18), and the total's median five-year relative change under the regional pool is 0.016 (observed 0.0127). The composition remains that of Eq. (4), and the mean squared five-year change of species in the top two of ten abundance bands is within a factor of two of that observed (Supporting Information S8). A factor with nothing species-specific in it reproduces so much species-level fluctuation because the same good or poor interval for sapling growth moves each life history by a different amount; the rest of an abundant species' movement is largely its own steady trend, which its matrix carries. The factor falls short in two places. First, species with 16 to 255 large trees move 2.4 to 3.3 times more in the census than in the model. Second, under the compensated scheme, whose total otherwise fluctuates like the pool's, the stationary stand is 32% lower than without the year effect. This thinning may be an artifact of confining the year effect to sapling growth, or it may be real: if the observed stand is what competition under environmental noise leaves standing, the capacity behind it is larger than the observed count, and the size of the community is itself an outcome of competition and the year effect rather than a datum. We have not re-set the capacity.

Each species' own mortality variation, added at its measured size as a generalization of the minimal model (Section 2.4), improves the species-level agreement somewhat and changes little else. Every species fluctuates more, abundant species most, and the observed mean squared change then falls within the simulated range in 7 bands under either scheme (5 and 6 without the term, and 6 with each species' mean survival held at its measured value, which a mean-one factor on the hazard otherwise raises slightly). But the shared growth variation still moves the total 3.6 to 4.3 times as much as the mortality term alone, and the composition is unchanged; under the compensated scheme the term also offsets the thinning, but only through that slight rise in survival (Supporting Information S3). At intermediate abundance the census's mean squared change is still 2.1 to 2.7 times the model's, leaving species-specific deviations in sapling growth as the term to test next.

3.4 The seedling pool

Among our predictions only the stationary seedling pool involves the arrival rate, and only weakly. Standing seedlings are the seed input times a seedling's expected lifetime, and most of the input is seed of the standing trees, fixed by the capacity and the matrices, immigrant seed supplying close to the fraction d of it (Supporting Information S2). Hence $M = \hat{m}B$, whose factor B is set and not measured (Section 2.3), can move the densities below by about that fraction at most, and their ratio between the schemes hardly at all. Under the regional pool we find 2.5 seedlings per m^2, compared with 1.9–2.8 for the modeled species in the seedling census (Comita et al., 2023b). This density is 1.10 times the census mean (0.89–1.35 census by census) and within a factor of three, whereas the compensated scheme gives 4.8 times the mean, beyond that factor. Because the compensated scheme favors species whose seeds yield few large trees, the stand stays at capacity only with a larger seed input and every species closer to replacement ($d = 0.026$ of candidate recruits left waiting instead of 0.153), each arriving seed then leaving $1/(d\mu_s)$ seedlings standing (Supporting Information S2), and a lower arrival rate does not remove the excess, because d falls with it. The seedling total thus tells against the compensated scheme's stationary stand, at the one stage the model treats most simply. Species by species the model predicts large-tree composition far better than small-stem composition: both schemes order the seedling layer poorly and the saplings little better (rank correlation 0.23 to 0.45 within classes 2 to 4 under the compensated scheme, whose class totals are nevertheless right; Supporting Information S7 and S9). This failure of composition, not of level, enters with the recruits: the observed seedling composition resembles neither the region's nor the plot's large-tree composition, and may reflect conspecific density dependence in

seedling survival (Comita et al., 2010), which our model omits and which is the next term to test.

4 Discussion

Community ecology has wanted a model that says, for a real forest and by name, which species will gain or lose trees over the coming decades and which the forest holds in the long run, and the two traditions that might have supplied one grew up apart: stochastic community models, neutral theory first among them, predict abundance distributions but no species' direction, while stage-structured demography measures every species' rates but says nothing about the long run. Our minimal model closes that gap with little more than the two already contain: each Barro Colorado species keeps its measured rates, all species compete under one capacity, and immigrant seed arrives by regional abundance weighted against each species' local advantage, a competition-colonization trade-off (Levins and Culver, 1971; Hastings, 1980; Tilman, 1994). One environmental effect on sapling growth, shared by every species and estimated from the same censuses, supplies the variation between years. So built, the model's measured rates predict what its predecessors could not, each named species' change in large trees a decade ahead, and predict it well. In the long run it recovers the census order of species' large trees, and at exactly equal long-run fitness their absolute abundances and abundance distribution too, where arrivals by regional abundance alone recover only the order. Its clear misses lie below 79.4 mm: sapling and seedling composition under either scheme, and under compensation a seedling layer several times too dense.

The transition data, not a modeling choice, placed the year effect on sapling growth, the transition in which the censuses show variation shared across species. In the model growth moves no stem into or out of the censused stand, which the capacity holds; the factor changes large-tree numbers because the last sapling promotion is the entry into the large-tree classes, so that a good year takes more saplings of every species into those classes and a poor year fewer, and because a stem that has crossed then stands for decades, each year's pulse of entrants becomes a departure that outlasts it. What drives the factor is not identified, though seven intervals constrain it. The common component was high through 1982–1990, the first of those intervals spanning the El Niño drought that killed many trees (Condit et al., 1995), lowest in 1990–1995, and has recovered since, as diameter growth did plot-wide (Condit et al., 2017); no steadily changing driver such as rising CO_2 or warming gives a fall and recovery. Canopy opening, which droughts and storms synchronize across a plot (Condit et al., 1995; Rutishauser et al., 2020; Araujo et al., 2021), would act on sapling growth just where the transitions place the factor and so mimic environmental noise, but the light-demanding species follow the factor no more strongly than the shade-tolerant. Of the climatic drivers, light, temperature and rainfall (Graham et al., 2003; Feeley et al., 2007; Clark et al., 2010; Dong et al., 2012), no series we examined tracks the factor beyond chance, the drier intervals having if anything the faster growth and the sunnier dry seasons the slower, and none has a memory near the two decades we assume (Supporting Information S3). Two features stay unexplained under any driver: every species responds in the same direction, and both the factor's runs of good and poor intervals and the species' own departures from their mean rates persist for a decade or two, the latter a persistence the model does not reproduce (Material and methods 2.5). Mortality variation, by contrast, is largely particular to each species, the kind of response earlier models assumed (Chisholm et al., 2014; Kalyuzhny et al., 2015; Fung et al., 2016); added at its measured size it brings the abundant species' movement close to that observed and changes little else (Section 3.3). Measured life histories and a single common effect on growth acting through them thus account for most of the faster-than-drift change that motivated the model, and species-specific couplings add little for their cost.

Beside the two kinds of model the Introduction contrasted, stochastic community models and species demography, which ours joins, stands a third kind, built from physiology to predict something else. Ecosystem-demography and cohort models (Moorcroft et al., 2001; Medvigy et al., 2009; Longo et al., 2019b,a) scale physiology, through parameter-rich functional types, to biomass, fluxes and their response to climate, and are benchmarked at this plot (Fisher et al., 2015, 2018; Koven et al., 2020); with species entered as types they are not built to say which named species gain trees next decade or which the plot holds in the long run, although the nearest of them to ours (Rüger et al., 2020) predicts this forest's composition through succession from five functional groups. Species-resolved models like ours have no physiology and say nothing directly about fluxes. The division need not stand: given every species' stems by size, biomass and fluxes could follow from allometry and physiology added on top. The shared growth variation is a stand-wide signal that weather-driven models ought to reproduce. These results also speak, with data, to an old question: whether a plant community is an integrated whole (Clements, 1916) or a set of species that merely share a place, each answering the environment on its own (Gleason, 1926), the individualistic view that the gradient analyses of the 1950s upheld (Curtis and McIntosh, 1951; Whittaker, 1956). At Barro Colorado there is something shared by all species in their response to the environment, coherent even in direction, while the idiosyncrasy in how species' abundances then move is mediated by their life histories, which differ widely. The community responds to a single, coherent environmental signal, but the signal is rendered invisible by life-history diversity. Individuality is preserved, and indeed generated, by demography.

Our parametrization is particular to Barro Colorado, and so are the simplifying features the transition data supported, one year effect common to species and acting on growth and a seedling class with no limit of its own, together with the memory assumed for that effect; the construction is not, and what limits its transfer is data. Data will be the obstacle at most forests, and there is more than one route around it. A new plot needs two censuses of tagged stems for the projection matrices, several intervals for the year effect, and a regional sample for immigration. Several dozen large plots already have such censuses (Davies et al., 2021), and where series are short, inventories pooled across sites can supply survival rates and repeated airborne surveys growth (Araujo et al., 2021). Demographic rates inferred from population-genomic samples, which have already predicted community dynamics in a temperate forest (O'Dwyer et al., 2025), could provide other parts of the matrices.

What remains open follows from where the results come from, one forest and rates re-derived from its early censuses. First is the driver of the shared growth variation, and with it the memory, which nothing in the censuses fixes; annual growth set against monthly climate at this and other Panamanian plots may identify the one and bound the other. Second is each species' own growth variation, testable as its mortality variation was. Third is the size of a species' decadal change, which the early rates over-predict with intervals still too narrow (Supporting Information S5 and S8), and fourth the composition of the layers below 79.4 mm, where density-dependent seedling survival is the term to test (Section 3.4). The Barro Colorado results nevertheless suggest that a forest census already contains most of what such forecasts need: four censuses of tagged stems were enough to rank the species by their change in large trees over the following two decades, a regional sample to predict which species are common, and one number per interval, estimated from the same stems, to reproduce how the whole stand moves. Later censuses here, and the same construction elsewhere, will test all three.

Acknowledgements

We thank R. Condit, S. P. Hubbell, R. B. Foster, R. Pérez, S. Aguilar and S. Lao for making the Barro Colorado Island (BCI) and Panama plot censuses public, L. S. Comita and her coauthors for releasing the seedling census, and K. Jops for archiving the demographic matrices. The BCI forest dynamics research project was made possible by National Science Foundation grants to Stephen P. Hubbell: DEB-0640386, DEB-0425651, DEB-0346488, DEB-0129874, DEB-00753102, DEB-9909347, DEB-9615226, DEB-9615226, DEB-9405933, DEB-9221033, DEB-9100058, DEB-8906869, DEB-8605042, DEB-8206992, DEB-7922197, support from the Forest Global Earth Observatory (ForestGEO), the Smithsonian Tropical Research Institute, the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, the Small World Institute Fund, and numerous private individuals, and through the hard work of over 100 people from 10 countries over the past three decades. The plot project is part of the Forest Global Earth Observatory (ForestGEO), a global network of large-scale demographic tree plots. We acknowledge the support of the Forest Global Earth Observatory of the Smithsonian Tropical Research Institute for the BCI plot.

Author contributions

[Draft for the authors.] M.D.S. and J.P.O'D. conceived the study, decided its scientific questions and tests, and wrote the paper. The modelling, computation and data analysis were carried out with an AI research assistant (Claude, Anthropic) working under M.D.S.'s direction: it wrote and ran the simulation and analysis code, kept the record of every analysis, and drafted text and figures for the authors' revision, each test being specified and registered before it was computed and its outcome checked by the authors, who take responsibility for all results and their interpretation. [Computational resources: to be completed; the CPU-hours of the registered runs are recorded in the run ledgers, and an energy estimate is to follow.]

Data accessibility

All data analyzed here are public, and we used them under the licenses stated with each source. The 1982–2015 censuses of the 50-ha BCI plot are the complete-plot release archived on Dryad (Condit et al., 2019, doi:10.15146/5xcp-0d46, CC0); the plot and its census methods are described by Condit (1998) and Hubbell et al. (1999). The seedling and small-sapling census of the same plot is the Dryad release of Comita et al. (2023a, doi:10.5061/dryad.05qfttf85, CC0), described by Comita et al. (2023b). We took each species' stage-structured projection matrix from the Zenodo archive of Jops and O'Dwyer (2024, doi:10.5281/zenodo.14563651, CC BY 4.0) that accompanies Jops et al. (2025). We built the regional species pool from the Panama plot-network tables of Condit et al. (2016, Smithsonian, doi:10.5479/data.stri.2016.0622, CC BY 4.0); the network's current archive is on Dryad (Condit et al., 2022, doi:10.15146/mdpr-pm59, CC0). Species' light responses, functional traits and the climate series, which we use only for exploratory comparisons with the year effect (Supporting Information S3), come from further public datasets listed with their DOIs and licenses in Supporting Information S10. All of our code, the stochastic simulation engine, the closed-form solver, the estimation of the year effect and every analysis script behind the numbers, tables and figures reported here and in the Supporting Information, is held with its run records in the project repository and will be archived with a README and a DOI [repository and DOI to be added before submission]; it will be available to reviewers at submission, together with the derived tables from which all results can be reproduced. The archive also contains the projection matrices that

we estimated from the 1982–1995 censuses alone for the out-of-sample forecasts, and the code that computes those matrices from the census data.

References

Araujo, R. F., Grubinger, S., Celes, C. H. S., Negrón-Juárez, R. I., Garcia, M., Dandois, J. P., and Muller-Landau, H. C. (2021). Strong temporal variation in treefall and branchfall rates in a tropical forest is related to extreme rainfall: results from 5 years of monthly drone data for a 50 ha plot. *Biogeosciences*, 18(24):6517–6531.

Caswell, H. (2001). *Matrix Population Models: Construction, Analysis, and Interpretation*. Sinauer Associates, Sunderland, MA, second edition.

Chave, J. (2004). Neutral theory and community ecology. *Ecology Letters*, 7(3):241–253.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, 31(1):343–366.

Chesson, P. L. and Warner, R. R. (1981). Environmental variability promotes coexistence in lottery competitive systems. *The American Naturalist*, 117(6):923–943.

Chisholm, R. A., Condit, R., Rahman, K. A., Baker, P. J., Bunyavejchewin, S., Chen, Y.-Y., Chuyong, G., Dattaraja, H. S., Davies, S., Ewango, C. E. N., Gunatilleke, C. V. S., Nimal Gunatilleke, I. A. U., Hubbell, S., Kenfack, D., Kiratiprayoon, S., Lin, Y., Makana, J.-R., Pongpattananurak, N., Pulla, S., Punchi-Manage, R., Sukumar, R., Su, S.-H., Sun, I.-F., Suresh, H. S., Tan, S., Thomas, D., and Yap, S. (2014). Temporal variability of forest communities: empirical estimates of population change in 4000 tree species. *Ecology Letters*, 17(7):855–865.

Clark, D. B., Clark, D. A., and Oberbauer, S. F. (2010). Annual wood production in a tropical rain forest in NE Costa Rica linked to climatic variation but not to increasing CO_2. *Global Change Biology*, 16(2):747–759.

Clements, F. E. (1916). *Plant Succession: An Analysis of the Development of Vegetation*. Number 242 in Carnegie Institution of Washington Publication. Carnegie Institution of Washington, Washington, DC.

Comita, L., Aguilar, S., Hubbell, S., and Pérez, R. (2023a). Long-term seedling and small sapling census data from the Barro Colorado Island 50 ha Forest Dynamics Plot, Panama. Dryad, dataset.

Comita, L. S., Aguilar, S., Hubbell, S. P., and Pérez, R. (2023b). Long-term seedling and small sapling census data from the Barro Colorado Island 50 ha Forest Dynamics Plot, Panama. *Ecology*, 104(9):e4140.

Comita, L. S., Muller-Landau, H. C., Aguilar, S., and Hubbell, S. P. (2010). Asymmetric density dependence shapes species abundances in a tropical tree community. *Science*, 329(5989):330–332.

Condit, R. (1998). *Tropical Forest Census Plots: Methods and Results from Barro Colorado Island, Panama and a Comparison with Other Plots*. Springer, Berlin, Heidelberg.

Condit, R., Chisholm, R. A., and Hubbell, S. P. (2012). Thirty years of forest census at barro colorado and the importance of immigration in maintaining diversity. *PLoS ONE*, 7(11):e49826.

Condit, R., Engelbrecht, B. M. J., Pino, D., Pérez, R., and Turner, B. L. (2013). Species distributions in response to individual soil nutrients and seasonal drought across a community of tropical trees. *Proceedings of the National Academy of Sciences*, 110(13):5064–5068.

Condit, R., Hubbell, S. P., and Foster, R. B. (1995). Mortality rates of 205 neotropical tree and shrub species and the impact of a severe drought. *Ecological Monographs*, 65(4):419–439.

Condit, R., Pérez, R., Aguilar, S., and Lao, S. (2016). Data from tree censuses and inventories in Panama. Smithsonian Institution, dataset (plot tables of January 2013).

Condit, R., Pérez, R., Aguilar, S., and Lao, S. (2022). Census data from 65 tree plots in Panama, 1994–2015. Dryad dataset.

Condit, R., Pérez, R., Aguilar, S., Lao, S., Foster, R., and Hubbell, S. (2019). Complete data from the barro colorado 50-ha plot: 423617 trees, 35 years. Dryad dataset, doi:10.15146/5xcp-0d46.

Condit, R., Pérez, R., Lao, S., Aguilar, S., and Hubbell, S. P. (2017). Demographic trends and climate over 35 years in the Barro Colorado 50 ha plot. *Forest Ecosystems*, 4:17.

Condit, R., Pitman, N., Leigh, Jr., E. G., Chave, J., Terborgh, J., Foster, R. B., Núñez, P., Aguilar, S., Valencia, R., Villa, G., Muller-Landau, H. C., Losos, E., and Hubbell, S. P. (2002). Beta-diversity in tropical forest trees. *Science*, 295(5555):666–669.

Connell, J. H. (1971). On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. In den Boer, P. J. and Gradwell, G. R., editors, *Dynamics of Populations: Proceedings of the Advanced Study Institute on Dynamics of Numbers in Populations, Oosterbeek, 1970*, pages 298–312. Pudoc, Wageningen.

Curtis, J. T. and McIntosh, R. P. (1951). An upland forest continuum in the prairie-forest border region of Wisconsin. *Ecology*, 32(3):476–496.

Davies, S. J., Abiem, I., Abu Salim, K., Aguilar, S., Allen, D., Alonso, A., Anderson-Teixeira, K., et al. (2021). ForestGEO: Understanding forest diversity and dynamics through a global observatory network. *Biological Conservation*, 253:108907.

Dong, S. X., Davies, S. J., Ashton, P. S., Bunyavejchewin, S., Supardi, M. N. N., Kassim, A. R., Tan, S., and Moorcroft, P. R. (2012). Variability in solar radiation and temperature explains observed patterns and trends in tree growth rates across four tropical forests. *Proceedings of the Royal Society B: Biological Sciences*, 279(1744):3923–3931.

Engelbrecht, B. M. J., Comita, L. S., Condit, R., Kursar, T. A., Tyree, M. T., Turner, B. L., and Hubbell, S. P. (2007). Drought sensitivity shapes species distribution patterns in tropical forests. *Nature*, 447(7140):80–82.

Etienne, R. S. (2005). A new sampling formula for neutral biodiversity. *Ecology Letters*, 8(3):253–260.

Feeley, K. J., Wright, S. J., Nur Supardi, M. N., Kassim, A. R., and Davies, S. J. (2007). Decelerating growth in tropical forest trees. *Ecology Letters*, 10(6):461–469.

Fisher, R. A., Koven, C. D., Anderegg, W. R. L., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., Holm, J. A., Hurtt, G. C., Knox, R. G., Lawrence, P. J., Lichstein, J. W., Longo, M., Matheny, A. M., Medvigy, D., Muller-Landau, H. C., Powell, T. L., Serbin, S. P., Sato, H.,

Shuman, J. K., Smith, B., Trugman, A. T., Viskari, T., Verbeeck, H., Weng, E., Xu, C., Xu, X., Zhang, T., and Moorcroft, P. R. (2018). Vegetation demographics in Earth System Models: a review of progress and priorities. *Global Change Biology*, 24(1):35–54.

Fisher, R. A., Muszala, S., Verteinstein, M., Lawrence, P., Xu, C., McDowell, N. G., Knox, R. G., Koven, C., Holm, J., Rogers, B. M., Spessa, A., Lawrence, D., and Bonan, G. (2015). Taking off the training wheels: the properties of a dynamic vegetation model without climate envelopes, CLM4.5(ED). *Geoscientific Model Development*, 8(11):3593–3619.

Fung, T., O'Dwyer, J. P., Rahman, K. A., Fletcher, C. D., and Chisholm, R. A. (2016). Reproducing static and dynamic biodiversity patterns in tropical forests: the critical role of environmental variance. *Ecology*, 97(5):1207–1217.

Gleason, H. A. (1926). The individualistic concept of the plant association. *Bulletin of the Torrey Botanical Club*, 53(1):7–26.

Graham, E. A., Mulkey, S. S., Kitajima, K., Phillips, N. G., and Wright, S. J. (2003). Cloud cover limits net CO_2 uptake and growth of a rainforest tree during tropical rainy seasons. *Proceedings of the National Academy of Sciences*, 100(2):572–576.

Hastings, A. (1980). Disturbance, coexistence, history, and competition for space. *Theoretical Population Biology*, 18(3):363–373.

Hubbell, S. P. (1979). Tree dispersion, abundance, and diversity in a tropical dry forest. *Science*, 203(4387):1299–1309.

Hubbell, S. P. (1997). A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs*, 16(Suppl.):S9–S21.

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*, volume 32 of *Monographs in Population Biology*. Princeton University Press, Princeton, NJ.

Hubbell, S. P., Foster, R. B., O'Brien, S. T., Harms, K. E., Condit, R., Wechsler, B., Wright, S. J., and de Lao, S. L. (1999). Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science*, 283(5401):554–557.

Hülsmann, L., Chisholm, R. A., Comita, L., Visser, M. D., de Souza Leite, M., Aguilar, S., Anderson-Teixeira, K. J., Bourg, N. A., Brockelman, W. Y., Bunyavejchewin, S., Castaño, N., Chang-Yang, C.-H., Chuyong, G. B., Clay, K., Davies, S. J., Duque, A., Ediriweera, S., Ewango, C., Gilbert, G. S., Holík, J., Howe, R. W., Hubbell, S. P., Itoh, A., Johnson, D. J., Kenfack, D., Král, K., Larson, A. J., Lutz, J. A., Makana, J.-R., Malhi, Y., McMahon, S. M., McShea, W. J., Mohamad, M., Nasardin, M., Nathalang, A., Norden, N., Oliveira, A. A., Parmigiani, R., Perez, R., Phillips, R. P., Pongpattananurak, N., Sun, I.-F., Swanson, M. E., Tan, S., Thomas, D., Thompson, J., Uriarte, M., Wolf, A. T., Yao, T. L., Zimmerman, J. K., Zuleta, D., and Hartig, F. (2024). Latitudinal patterns in stabilizing density dependence of forest communities. *Nature*, 627(8004):564–571.

Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *The American Naturalist*, 104(940):501–528.

Jops, K., Dalling, J. W., and O'Dwyer, J. P. (2025). Life history is a key driver of temporal fluctuations in tropical tree abundances. *Proceedings of the National Academy of Sciences*, 122(4):e2422348122.

Jops, K. and O'Dwyer, J. (2024). Space-limited matrix community model - Barro Colorado Island tree species. Zenodo, dataset.

Jops, K. and O'Dwyer, J. P. (2023). Life history complementarity and the maintenance of biodiversity. *Nature*, 618(7967):986–991.

Kalyuzhny, M., Kadmon, R., and Shnerb, N. M. (2015). A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. *Ecology Letters*, 18(6):572–580.

Kalyuzhny, M., Lake, J. K., Wright, S. J., and Ostling, A. M. (2023). Pervasive within-species spatial repulsion among adult tropical trees. *Science*, 381(6657):563–568.

Koven, C. D., Knox, R. G., Fisher, R. A., Chambers, J. Q., Christoffersen, B. O., Davies, S. J., Detto, M., Dietze, M. C., Faybishenko, B., Holm, J., Huang, M., Kovenock, M., Kueppers, L. M., Lemieux, G., Massoud, E., McDowell, N. G., Muller-Landau, H. C., Needham, J. F., Norby, R. J., Powell, T., Rogers, A., Serbin, S. P., Shuman, J. K., Swann, A. L. S., Varadharajan, C., Walker, A. P., Wright, S. J., and Xu, C. (2020). Benchmarking and parameter sensitivity of physiological and vegetation dynamics using the Functionally Assembled Terrestrial Ecosystem Simulator (FATES) at Barro Colorado Island, Panama. *Biogeosciences*, 17(11):3017–3044.

Lefkovitch, L. P. (1965). The study of population growth in organisms grouped by stages. *Biometrics*, 21(1):1–18.

Levins, R. and Culver, D. (1971). Regional coexistence of species and competition between rare species. *Proceedings of the National Academy of Sciences*, 68(6):1246–1248.

Longo, M., Knox, R. G., Levine, N. M., Swann, A. L. S., Medvigy, D. M., Dietze, M. C., Kim, Y., Zhang, K., Bonal, D., Burban, B., Camargo, P. B., Hayek, M. N., Saleska, S. R., da Silva, R., Bras, R. L., Wofsy, S. C., and Moorcroft, P. R. (2019a). The biophysics, ecology, and biogeochemistry of functionally diverse, vertically and horizontally heterogeneous ecosystems: the Ecosystem Demography model, version 2.2 – Part 2: Model evaluation for tropical South America. *Geoscientific Model Development*, 12(10):4347–4374.

Longo, M., Knox, R. G., Medvigy, D. M., Levine, N. M., Dietze, M. C., Kim, Y., Swann, A. L. S., Zhang, K., Rollinson, C. R., Bras, R. L., Wofsy, S. C., and Moorcroft, P. R. (2019b). The biophysics, ecology, and biogeochemistry of functionally diverse, vertically and horizontally heterogeneous ecosystems: the Ecosystem Demography model, version 2.2 – Part 1: Model description. *Geoscientific Model Development*, 12(10):4309–4346.

McGill, B. J. (2003). A test of the unified neutral theory of biodiversity. *Nature*, 422(6934):881–885.

McKane, A., Alonso, D., and Solé, R. V. (2000). Mean-field stochastic theory for species-rich assembled communities. *Physical Review E*, 62(6):8466–8484.

McKane, A. J., Alonso, D., and Solé, R. V. (2004). Analytic solution of Hubbell's model of local community dynamics. *Theoretical Population Biology*, 65(1):67–73.

Medvigy, D., Wofsy, S. C., Munger, J. W., Hollinger, D. Y., and Moorcroft, P. R. (2009). Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. *Journal of Geophysical Research: Biogeosciences*, 114(G1).

Moorcroft, P. R., Hurtt, G. C., and Pacala, S. W. (2001). A method for scaling vegetation dynamics: the ecosystem demography model (ED). *Ecological Monographs*, 71(4):557–586.

O'Dwyer, J. P., Lutz, J. A., Schappe, T., Alegre, D., Black, A. N., Grünwald, N. J., and Jones, F. A. (2025). Genomic demography predicts community dynamics in a temperate montane forest. *Science*, 389(6766):eadu6396.

Preston, F. W. (1948). The commonness, and rarity, of species. *Ecology*, 29(3):254–283.

Rosindell, J., Hubbell, S. P., and Etienne, R. S. (2011). The Unified Neutral Theory of Biodiversity and Biogeography at Age Ten. *Trends in Ecology & Evolution*, 26(7):340–348.

Rüger, N., Condit, R., Dent, D. H., DeWalt, S. J., Hubbell, S. P., Lichstein, J. W., Lopez, O. R., Wirth, C., and Farrior, C. E. (2020). Demographic trade-offs predict tropical forest dynamics. *Science*, 368(6487):165–168.

Rutishauser, E., Wright, S. J., Condit, R., Hubbell, S. P., Davies, S. J., and Muller-Landau, H. C. (2020). Testing for changes in biomass dynamics in large-scale forest datasets. *Global Change Biology*, 26(3):1485–1498.

Tilman, D. (1994). Competition and biodiversity in spatially structured habitats. *Ecology*, 75(1):2–16.

Turkia, O. and Shnerb, N. M. (2025). Asymmetric variability: the impact of uneven stochasticity on competitive dynamics. arXiv:2501.03682 [q-bio.PE].

Uriarte, M., Thompson, J., and Zimmerman, J. K. (2019). Hurricane María tripled stem breaks and doubled tree mortality relative to other major storms. *Nature Communications*, 10:1362.

Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2003). Neutral theory and relative species abundance in ecology. *Nature*, 424(6952):1035–1037.

Whittaker, R. H. (1956). Vegetation of the Great Smoky Mountains. *Ecological Monographs*, 26(1):1–80.

Zimmerman, J. K., Everham, III, E. M., Waide, R. B., Lodge, D. J., Taylor, C. M., and Brokaw, N. V. L. (1994). Responses of tree species to hurricane winds in subtropical wet forest in Puerto Rico: implications for tropical tree life histories. *Journal of Ecology*, 82(4):911–922.