Title

Forest regeneration within Earth system models: current process representations and ways forward

Permalink
https://escholarship.org/uc/item/3j76d81f

Journal
New Phytologist, 235(1)

ISSN
0028-646X

Authors

Hanbury-Brown, Adam R
Ward, Rachel E
Kueppers, Lara M

Publication Date
2022-07-01

DOI
10.1111/nph.18131

Tansley Review

Forest regeneration within Earth system models: current process representations and ways forward

Adam R. Hanbury-Brown[1], Rachel E. Ward[1], and Lara M. Kueppers[1, 2]

[1]The Energy and Resources Group, 345 Giannini Hall, University of California, Berkeley, CA 94720, USA

[2]Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA

ORCIDs

Adam R. Hanbury-Brown: 0000-0003-3751-6257

Rachel E. Ward: 0000-0003-2343-1906

Lara M. Kueppers: 0000-0002-8134-3579

Authors for correspondence:

Lara Kueppers

Email: lmkueppers@berkeley.edu

Adam Hanbury-Brown

Email: ahanburybrown@gmail.com

Received: *8 June 2021*

Accepted: *24 February 2022*

Contents

Summary

Earth system models must predict forest responses to global change in order to simulate future global climate, hydrology, and ecosystem dynamics. These models are increasingly adopting vegetation demographic approaches that explicitly represent tree growth, mortality and recruitment, enabling advances in the projection of forest vulnerability and resilience, as well as evaluation with field data. To date, simulation of regeneration processes has received far less attention than simulation of processes that affect growth and mortality in spite of its critical role maintaining forest structure, facilitating turnover in forest composition over space and time, enabling recovery from disturbance, and regulating climate-driven range shifts. Our critical review of regeneration process representations within current Earth system vegetation demographic models reveals the need to improve parameter values and algorithms for reproductive allocation, dispersal, seed survival and germination, environmental filtering in the seedling layer, and tree regeneration strategies adapted to wind, fire, and anthropogenic disturbance regimes. These improvements require synthesis of existing data, specific field data collection protocols, and novel model algorithms compatible with global scale simulations. Vegetation demographic models offer the opportunity to more fully integrate ecological understanding into Earth system prediction; regeneration processes need to be a critical part of the effort.

Key words: forest regeneration, Earth system models, vegetation demographic models, vegetation dynamics, tree recruitment, reproductive allocation

I. Introduction

Forest management, clearing, and abandonment have affected roughly seventy percent of forested area globally (FAO, 2015), while climate change and intensifying natural disturbance regimes are increasingly affecting all forests (Gauthier *et al*., 2015; Abatzoglou & Williams, 2016; Seidl *et al*., 2017; McDowell et al., 2020). Anticipating the individual and combined effects of these pressures on future forests requires understanding not only the resistance and vulnerability of mature trees to global changes, but also the resilience of forest communities through regeneration processes (Lloret *et al*., 2012; Martínez-Vilalta & Lloret, 2016; McDowell *et al*., 2020). Even where land use and climate change pressures are minimal and forests experience relative ecological stability, regeneration processes must compensate for mortality to ensure that forests persist. Regeneration processes, including reproductive allocation, dispersal, seed survival, germination, seedling survival, and growth into adult size classes (see Box 1, Glossary), are known to be sensitive to light, water, temperature, and soil nutrients and can limit where trees regenerate. By acting as a bottleneck to tree establishment, regeneration processes have strong leverage on forest structure, distribution, and composition (Kobe, 1999; Beckage & Clark, 2003; Wright & Calderón, 2006; Engelbrecht *et al.*, 2007; Ibanez *et al.*, 2007; Walck *et al.*, 2011; Conlisk *et al.*, 2018), which ultimately give rise to forest function within the Earth system (Bonan, 2008). Factors that shift resource availability or preferentially select reproductive individuals and their offspring can contribute to compositional turnover within a forest (Engelbrecht *et al.*, 2007; Barlow & Peres, 2008; Comita *et al.*, 2009; Swenson, 2012), and coupled with dispersal, a shift in the geographic distribution of species (Clark *et al.*, 1998). Finally, following landscape-scale disturbances such as fire, harvest, land clearing for agriculture, or hurricanes, the rate of forest recovery or transition to a non-forest state is dictated by regeneration processes (Turner *et al.*, 1999; Chazdon, 2003; Comita *et al.*, 2009; Schoennagel *et al.*, 2009; Martinez-Vilalta *et al*., 2016; Tepley *et al.*, 2017).

Box 1. Glossary

Bioclimatic envelope: Climatic tolerances of a species or plant functional type, generally based on correlations with observed distributions.

Forest: An area ≥ 0.5 hectares with a tree canopy cover of more than 10%, which is not primarily under agriculture or other specific non-forest land use (UN-CBD, 2006).

Patch: In a modeling context, a contiguous area of non-cultivated vegetation (typically 0.1 - 1 ha) with the same disturbance history (e.g. from canopy gap formation or a landscape-scale disturbance event) where vegetation competes for resources. Patches can represent fixed (Smith *et al.*, 2001; Sato *et al.*, 2007) or dynamic fractions (Koven *et al.*, 2020) of their climatic grid cell.

Forest regeneration: The establishment of a new tree cohort (Ponge *et al.*, 1998) or cohorts that reach the post-seedling size classes (≥ 1.35 m in height), which maintain existing forest or replace forest lost during landscape-scale disturbance. We consider regeneration from seed, resprouting, and clonal propagation.

Germination: Emergence of the radicle from the seed coat. This is a critical component of the more observable phenomenon of seedling emergence.

Plant functional type (PFT): A class of plant species that share key traits and life history strategies, which mediate their physiological and demographic responses to environmental conditions (Bonan, 2015). In a modeling context, a PFT can be considered a unique vector of parameter values (typically based on traits) that determine how a PFT interacts with its simulated environment.

Recalcitrant seed: Desiccation-sensitive seed requiring sufficient moisture to remain viable (Walck *et al.,* 2011).

Reproductive allocation: The fraction of available photosynthetic energy (typically modeled as net primary productivity remaining after allocation to tissue turnover and storage) allocated to flowers, seeds, and other reproductive tissues.

Reproductive schedule: Changes in reproductive allocation as a plant increases in size or age (Wenk & Falster, 2015)

Resprouting strategy: The way in which bud location, protection, and resourcing allows trees to regenerate vegetatively, typically after disturbance (Clarke *et al.*, 2013).

Seedling: An individual that has germinated and is < 1.35 m in height.

Tree recruitment: Addition of new trees to the post-seedling size classes (≥ 1.35 m in height) of an existing population or establishment of a new population. Recruitment is one of multiple regeneration processes that contribute to forest regeneration.

The next generation of Earth system models (ESMs) endeavors to dynamically project changes in vegetation function and distribution in response to climate and land use changes to better capture essential biosphere responses and feedbacks. The ESM community is increasingly interested in using vegetation demographic models (VDMs) to represent vegetation dynamics within ESMs (Fisher *et al*., 2018; Bonan, 2019). VDMs comprise a class of Dynamic Global Vegetation Model (DGVM) that tracks multiple size-classes or individuals of the same plant functional type (PFT) and that represents multiple light environments within a single climatic grid cell (Fisher *et al*., 2018). Despite the central role of regeneration processes in forest dynamics, the algorithms for these processes are often missing or overly simplified within VDMs, introducing biases, uncertainties, and compensating errors in the model dynamics (Fisher *et al.*, 2010; Snell *et al.*, 2014; Powell *et al.*, 2018). Parameter sensitivity analyses indicate that some regeneration parameters have leverage on forest function and structure (Fisher *et al.*, 2010; Snell, 2014; Koven *et al.*, 2020), but the lack of a systematic effort to understand model sensitivity to existing parameters and alternative process representations hinders a prioritization of processes for improvement. Additionally, the diverse ways in which VDMs represent forest regeneration have never been critically and comprehensively reviewed, hindering the prioritization of model development goals and supporting research activities.

Based on theoretical and empirical ecological literature, we review the key forest regeneration processes influencing the Earth system (II), critically evaluate how current VDM representations of these processes (III) are likely to succeed or fail in predicting forest responses to global change (IV), and conclude with recommendations for advancing prediction (V). We focus on forests due to their large influence on global biogeochemistry and biogeophysics (Bonan, 2008). Regeneration in non-forest vegetation is largely represented with the same algorithms, but a future review evaluating non-forest regeneration processes is also needed.

II. Forest regeneration processes critical to the Earth system

Forests regulate global terrestrial ecosystem functioning and climate (Bonan, 2008), but only persist where trees are able to recruit after individual mortality and landscape-scale disturbance. Further, regeneration processes help shape forest responses, including range shifts, post-disturbance recovery, and compositional changes, to changing environmental conditions and disturbance regimes both across spatial gradients and over time (Engelbrecht *et al*. 2007; Dupuy & Chazdon, 2008; DeSantis *et al*., 2011; Pausas & Keeley, 2014; Tepley *et al*., 2017). This occurs when abiotic and biotic variables limit one or more aspects of the regeneration niche, which includes conditions necessary for viable seed production, seed dispersal, germination, and seedling growth and survival (Grubb, 1977). The regeneration niche ultimately impacts terrestrial ecosystem functioning by filtering species assemblages, affecting the distribution of functional traits such as maximum growth rates, stature, and drought tolerance, which mediate the cycling of carbon, water, and energy within the Earth system (Bonan, 2019).

Forest regeneration can be conceptualized as a series of environmentally sensitive processes, each acting as a selectively permeable sieve (i.e. potential bottleneck) culminating in tree recruitment (Fig. 1), thereby influencing global forest dynamics. For tree recruitment to occur, seeds must be produced, arrive at a site, remain viable and germinate, survive the vulnerable seedling stage, and grow sufficiently to recruit into the adult population (Box 1, Glossary). Vegetative regeneration via basal or epicormic sprouting can also be critical in some forests. Regeneration bottlenecks may occur at different levels of biological organization, from populations to the entire tree physiognomic class. We are concerned with bottlenecks occurring at the level of 1) the entire tree physiognomic class because they delineate the boundaries between forested and non-forested vegetation, and 2) PFTs because they shape the global distribution of forest types and the distribution of traits within a forest.

The first potential bottleneck to successful forest regeneration is viable seed production, which is largely governed by mature tree productivity and reproductive allocation (RA; Wenk *et al*. 2015). RA and seed production are sensitive to climate (Sirois, 2000; Roland *et al*., 2014) and are believed to contribute to limitation of tree distributions beyond alpine and polar tree line ecotones in some regions (Brown *et al*., 2018; Lantz *et al*., 2019; Anadon-Rosell *et al*., 2020; Körner *et al*., 2021). RA schedules, the changing magnitude of RA over a plant's lifetime, can vary dramatically among PFTs (Wenk *et al*. 2015; Ruger *et al*., 2020), and are adapted to specific disturbance regimes (Wenk & Falster, 2015) such that changing disturbance regimes can render an RA schedule non-viable (Lamont *et al*., 2020). For example, obligate postfire seeders, such as serotinous pines, successfully regenerate in fire-prone forests where fire return intervals remain between maturation age and tree lifespan (Greene *et al*., 1999; Pausas & Keeley, 2014). When fire return intervals deviate from this range, trees face "immaturity risk" (Fig. 2a) or "senescence risk" (Fig. 2b), and other regeneration strategies, such as post-fire resprouting (discussed below) dominate (Pausas & Keeley, 2014).

Seed production depends on temperature and light in species-specific ways (Alfaro-Sanchez *et al*. 2017; Wright & Calderón, 2006; Girardin *et al*., 2016; Gallego Zamorano *et al*., 2018; Detto *et al*., 2018; Hacket-Pain *et al*., 2018) suggesting that as climates change, differential environmental limitations on seed production could mediate functional type turnover. Evidence of environmentally mediated tradeoffs between reproduction and growth (Thomas *et al*., 2011; Berdanier & Clark, 2016; Hacket-Pain *et al.*, 2018) makes the ecophysiology of RA even more important to understand for accurate predictions of global carbon cycle dynamics. In sum, RA and seed production are particularly important for mediating a) climate-driven range shifts where viable seed production contributes to limits on forest distribution, b) functional type turnover within existing forests in response to changing disturbance regimes and possibly climate, and c) carbon available for growth.

After seeds are produced, dispersal limitations affect establishment success if propagules are unable to reach suitable sites (Turnbull, 2000). This particularly influences local species' ability to recolonize sites following landscape-scale disturbance. For example, limited dispersal by non-serotinous conifers after large wildfire is believed to facilitate vegetation transitions from forest to shrubland in the western United States (Tepley *et al*., 2017; Urza & Sibold, 2017; Stevens-Rumann & Morgan, 2019). Similarly, in tropical forests, dispersal limitation slows the rate of forest recovery on abandoned agricultural land (Hooper *et al*. 2005; Chapman & Chapman, 1999). Within mature tropical forests, animal-dispersed trees become more clustered following removal of their seed dispersers and abiotically dispersed species experience higher recruitment rates than sympatric, biotically dispersed species (Wright *et al.*, 2007; Harrison *et al.*, 2013). This evidence suggests that local dispersal limitation affects forest recovery rates after landscape-scale disturbance and forest species composition following defaunation.

On larger spatio-temporal scales, dispersal limitation mediates the geographic range of trees by influencing their migration potential (Clark *et al*., 1998; Caplat *et al*., 2008; Mokany *et al*., 2014; Miller & McGill, 2017). For example, simulations using realistic dispersal assumptions in a temperate deciduous forest suggest that tree migration is incapable of keeping pace with climate change, particularly where anthropogenic land use poses migration barriers (Miller & McGill, 2017). In Australian tropical forests climate change is projected to reduce the seed dispersal services of frugivores, thereby limiting the large scale tree migration potential for multiple species (Mokany *et al*., 2014). Ninety percent of tropical tree species are animal dispersed (Howe & Smallwood, 1982), portending decreasing migration potential and increasing seed and seedling mortality for many trees (de Paula *et al.*, 2018) amid a global wave of “anthropocene defaunation” (Dirzo *et al.*, 2014). These observations highlight the central role of dispersal in mediating tree range shifts.

After dispersal, seeds rely on a variety of traits designed to help them survive, germinate, and emerge under favorable conditions following a disturbance event. Seed size mediates functional type turnover from boreal coniferous to deciduous forests after severe fire because the smaller-seeded deciduous species lack large reserves and a taproot to mitigate moisture stress and benefit from a shallow post-fire organic layer where it is easier to reach stable water supplies (Johnstone & Chapin, 2006; Beck *et al*., 2011). Similarly, where compounding disturbances (such as windthrow followed by fire) kill not just the canopy trees but also the seedling layer in spruce-fir dominated forests, smaller, wind-dispersed seeds such as paper birch (*Betula papyrifera*) and aspen (*Populus tremuloides*) are favored for establishment (Johnstone *et al*., 2016), thereby altering functional composition. These observations demonstrate that seed size and germination are linked to establishment strategies which regulate functional composition in response to disturbance regimes.

After emergence, seedling survival and growth determine tree recruitment. Seedlings are particularly sensitive to abiotic stress, competition, herbivory, and fire (Burns & Honkala, 1990; Bond, 2008; McLaughlin & Zavaleta, 2012) making the seedling stage a critical filter on future forest distribution and functional composition. Climate-driven range shifts beyond current high-elevation range edges in many alpine ecosystems are expected to be limited by seedling survival (where viable seed is not already limiting), due to interactions between increasing temperature and moisture stress (Loranger *et al*., 2016; Conlisk *et al*. 2017; Kueppers *et al*. 2017). For example, experimental warming in the absence of additional summer moisture decreased lodgepole pine seedling establishment rates (Fig. 3a), which may prevent tree population growth above current range limits (Fig. 3b) in the Central Rocky Mountains. More broadly, seedling associations with particular regeneration niches are well correlated with mature tree composition across light, aridity, and nutrient gradients (Engelbrecht *et al*., 2007; Dupuy & Chazdon, 2008; Zalamea *et al.*, 2016), indicating the strong leverage these early life stages have over mature forest composition. Engelbrecht *et al*. (2007) showed that tropical tree seedling drought tolerance (Fig. 4a) explained species turnover across local plant available water and regional precipitation gradients (Fig. 4b). Disturbance regimes also influence forest composition through their impacts on seedling survival. For example, hurricanes alter understory conditions and thereby differential seedling success (Comita *et al.,* 2009). Changes in forest and woodland management that alter fire or grazing regimes dramatically alter forest composition over time through effects on seedling survival (Cierjacks & Hensen, 2004; Brown & Wu, 2005; Boulanger *et al.*, 2015). These examples demonstrate that seedling survival and growth are particularly important for mediating tree ranges and forest composition in response to changing disturbance regimes and not all regeneration relies on seed. Resprouting is common following cyclones, fire, and pest outbreaks (Van Bloem *et al*., 2007; Clarke *et al*., 2013; Pausas & Keeley, 2014; Pausas *et al*., 2015) and strongly influences recovery rates and post-disturbance forest structure and composition (Cooper-Ellis *et al*., 1999; Van Bloem *et al*., 2007; Uriarte *et al*., 2012; Saha *et al*., 2003; DeSantis *et al*., 2011; Kulakowski *et al*. 2013). Resprouting potential is not a binary trait. Bud bank location (underground, epicormic, or apical), protection (e.g. bark) and resourcing mediate resprouting success (Clarke *et al*., 2013; Pausas *et al*., 2015). For example, eucalyptus recovers quickly after fire despite its thin bark because deeply embedded epicormic meristems are well protected (Clarke *et al*., 2013). Many landscape-scale disturbances are predicted to become more frequent and severe in the coming decades (Turner, 2010; Trumbore *et al.*, 2015; Seidl *et al.*, 2017), likely driving changes in the relative abundance of seed and non-seed-based regeneration strategies. Epicormic or apical resprouting by mature trees also limits the short-term loss of ecosystem carbon relative to what occurs with mature tree mortality followed by recruitment from seed (Zeppel *et al.*, 2015; Walden *et al.*, 2019).

III. Forest regeneration in Vegetation Demographic Models

Our review of seven state-of-the-art VDMs reveals diverse approaches to representing forest regeneration processes (Table 1). The VDMs reviewed here predominantly rely on prescribed parameters in place of environmentally sensitive functions, all (except for one) omit at least one key process, and each one emphasizes different aspects of the regeneration niche. We reviewed models that 1) satisfy the definition of a VDM (Fisher *et al.*, 2018) by tracking multiple size classes or individuals of the same PFT that compete with each other under varying light environments (excluding e.g., Haverd *et al*. 2014, Argles *et al*. 2020 and "big leaf" models such as Lawrence *et al.*, 2019), 2) show potential for global scale simulations (excluding gap and forest landscape models), and 3) represent regeneration with a minimal level of complexity to warrant a full review (see bottom of Table S1 for models that represent regeneration with one equation, e.g. Joetzjer *et al.*, 2018). We note where additional versions of each VDM have been published with meaningful updates to the regeneration scheme. We present cohort-based and spatially implicit models first, followed by spatially explicit and individual-based models.

ED2, CLM(ED), and FATES

The Ecosystem Demography model version 2 (ED2; Medvigy *et al*., 2009), the implementation of ED within the Community Land Model (CLM(ED); Fisher *et al*., 2015), and the Functionally Assembled Terrestrial Ecosystem Simulator (FATES; Koven *et al*., 2020; FATES Development Team, 2020), are derived in part from the modeling framework presented in the original ED model formulation (Moorcroft *et al*., 2001). ED-based models simulate a mosaic of spatially implicit forest "patches"

(Box 1, Glossary), which have unique disturbance histories, areas, stand structure and composition, but no notion of spatial adjacency or location within the larger, spatially explicit climatic grid cells used by ESMs (typically 50-500 km resolution) for global simulations, hereafter referred to as "grid cells". Within each patch, trees of the same size and PFT are tracked as a single cohort that is represented by an average individual. Within the patch, stand structure and composition emerge deterministically from size- and PFT-structured competition. With the exception of seed production, which is proportional to productivity, no regeneration processes in ED-based models are sensitive to environmental variation. Reproductive allocation (RA) in all ED-based models occurs in each daily timestep, wherein a prescribed, PFT-specific fraction, F_{repro} (typically 0.1-0.37, of cohort-level carbon for growth and reproduction (C_{g+r}; net after respiration and tissue replacement) is allocated to reproduction (Moorcroft *et al*., 2001).

In ED2, reproductive carbon is subsequently reduced by a seedling mortality parameter, M_s (0.95 month^{-1} by default), which implicitly represents non-seed reproductive carbon and all causes of mortality between reproductive allocation and recruitment into the smallest size class (Longo *et al.*, 2019a; Longo *et al.*, 2019b). Reproductive carbon (numerator in Eqn 1) is converted into a number density (i.e. recruitment rate) of new recruits ($N_{recruit}$) based on the amount of biomass required to make an individual in the smallest size class (Z_0) such that

$$N_{recruit} = \frac{((C_{g+r})(F_{repro}) + S)(1 - M_s)}{Z_0} \qquad \text{(Eqn 1)}$$

where S is additional "seed rain", represented as a prescribed input of reproductive carbon arriving to each grid cell with no specific origin. ED2 includes a PFT-specific minimum size threshold required for RA to occur (regardless of canopy position), but this parameter is not typically differentiated among tree PFTs within a biome. This is also true for CLM(ED) and FATES. Dispersal between patches within a grid cell can be represented by prescribing the fraction of reproductive carbon that can leave its patch of origin..

Unlike ED2, intra-grid cell dispersal is not represented in CLM(ED) or FATES, but they do explicitly represent PFT-specific seed banks (Eqn 2) which gain carbon from RA ($Seed_{in}$) and lose carbon due to seed decay ($Seed_{decay}$) at a prescribed, constant rate for all PFTs (0.51 yr^{-1}; Fisher *et al*., 2015). Subsequently, carbon emerges from the seed bank ($Seed_{germ}$) according to a prescribed germination rate (0.5 yr^{-1}) that is constant across PFTs by default.

$$\frac{\partial Seeds}{\partial t} = Seed_{in} - Seed_{decay} - Seed_{germ} \qquad \text{(Eqn 2)}$$

All carbon emerging from the seed bank is converted to new recruits based on Z_0 (Eqn 1), typically producing recruits between 0.25 and 0.5 cm diameter at breast height (dbh; C. Koven, personal communication, 2018). Aside from M_s in ED2, which has strong leverage on community structure and ecosystem properties (Fisher *et al.*, 2010), seedling survival and growth in ED-based models is only implicitly represented by specifying the size of the smallest size class (Z_0), effectively making recruitment proportional to adult net primary productivity (NPP).

LPJ-GUESS

The Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS) represents the forested landscape as a series of spatially implicit replicate patches where stand structure and composition vary as a function of stochastic disturbance and demographic processes (Smith *et al.*, 2001). LPJ-GUESS is typically considered an individual based model, but for regional- and global-scale simulations it is usually run in “cohort mode” for computational efficiency (Wramneby *et al.*, 2010; Hickler *et al.*, 2012; Smith *et al.*, 2014), using an average individual to represent cohorts in the same way as ED-based models. Therefore, our discussion of LPJ-GUESS pertains to cohort mode. Unlike ED-based models, regeneration occurs on an annual time step and reproduction and establishment only occur within PFT-specific bioclimatic envelopes (BEs) defined by a minimum number of growing degree days for reproduction (GDD_{min}), the maximum mean temperature of the coldest month ($T_c max$*;* a chilling requirement), and the minimum temperature of the coldest month required for establishment and survival ($T_c min$; Hickler *et al*., 2012). Each cohort of climatically eligible PFTs allocates a fixed fraction (0.1 by default; Smith *et al*., 2001) of NPP to reproduction. The main version of LPJ-GUESS does not impose a size threshold for RA, but some versions, LPJ-GUESS-MIGRATION (LPJ-GM) and LPJ-DISP do (Snell, 2014; Lehsten *et al*. 2019).

Dispersal, germination, and seed survival are not explicitly represented in LPJ-GUESS, but LPJ-GM represents germination with a prescribed germination probability parameter and dispersal using dispersal kernels and changes to the sub-grid cell patch structure (Lehsten *et al.*, 2019; discussed in IV). LPJ-GM has an angular dispersion parameter that adjusts the shape of the dispersal kernel to reflect variables such as wind direction. LPJ-DISP also uses dispersal kernels to represent dispersal between grid cells (Snell, 2014) and calculates the probability of inter-patch dispersal (i.e. within a grid cell) as a function of the number of patches already containing the PFT.

In addition to the BEs, PFT-specific limitations on seedling growth and survival are represented by prescribing the minimum amount of photosynthetically active radiation at the forest floor ($\mathrm{PAR_{min}}$) required for establishment (Smith *et al*., 2001) where shade tolerant PFTs have a lower $\mathrm{PAR_{min}}$. The fraction of plant-available water holding capacity (fAWC) in the surface soil layer (Hicker *et al*., 2012) must be above a critical level, but this has only been used in regional, species-specific simulations. The number of new recruits per year is pulled from a Poisson distribution with the expectation, λ, calculated from a PFT-specific observed maximum establishment rate, est_{max}, which is modified by the amount of carbon allocated to reproduction in the prior year (C_{repro}), and f, the potential productivity at the forest floor expressed as a fraction of the maximum (i.e. full sun; Smith *et al*., 2001; Hickler *et al*., 2012, S1; Wramneby *et al*., 2008) such that

$$\lambda = \mu(f)(est_{max})((k_{reprod})(C_{repro}) + k_{bgestab}) \quad \text{(Eqn 3)}$$

$$\mu(f) = e^{\alpha(1-1/f)} \quad \text{(Eqn 4)}$$

where α is a PFT-specific shape parameter that governs the sensitivity of recruitment to reductions in productivity at the forest floor (see Fulton, 1991); light demanding species experience steeper declines in recruitment as light at the forest floor decreases. k_{reprod} and $k_{bgestab}$ are constants scaled such that

(k_{reprod}) (C_{repro}) + $k_{bgestab}$ = 1 at the "highest plausible value of C_{repro}" (Wramneby *et al.*, 2008). $k_{bgestab}$ (i.e. background establishment) ensures that recruitment never goes to zero for any PFT due to lack of productivity and carbon allocated to reproduction. This is functionally similar to the concept of seed rain in the ED-based models. New recruits are 1.2 meters in height and allometry determines the corresponding biomass per new recruit. In summary, recruitment occurs for any given PFT if its BE (GDDs, $T_c$max, $T_c$min) and establishment thresholds (PAR_{min}, fAWC) are satisfied. Recruitment rate is governed by est_{max}, PFT-specific sensitivity to reductions in potential understory productivity, the amount of carbon allocated to reproduction and the background establishment ($k_{bgestab}$) rate.

LM3-PPA

Land Model 3 with the Perfect Plasticity Approximation (LM3-PPA; Weng *et al.*, 2015), is a spatially explicit, individual-based model that can be implemented at global scale (Dunne *et al.*, 2020). It handles recruitment similarly to ED-based models, albeit with a slightly different allocation scheme. Nonstructural carbohydrates (NSC) remaining after allocation to leaf and fine root turnover are allocated to C_{g+r}. Trees in the canopy allocate a fixed fraction of C_{g+r} (0.1) to reproduction, while trees in the understory do not reproduce. Dispersal and seed survival in a seed bank are not explicitly represented, but germination is represented with a prescribed germination probability. New recruits in LM3-PPA emerge as seedlings (Box 1, Glossary) because Z_0 (Eqn 1) is generally set much smaller (0.02-0.1 Kg C per seedling) than in the other models reviewed here (E. Weng, personal communication, 2021). Seedling survival is represented with a prescribed establishment probability and a size-dependent understory mortality function (Weng *et al.*, 2015).

SEIB-DGVM

The Spatially-Explicit Individual-Based Dynamic Global Vegetation Model (SEIB) is a spatially explicit, individual-based VDM (Sato *et al.*, 2007). SEIB simulates 30 m x 30 m forest patches (Box 1, Glossary; patch size can vary depending on the application) where individual trees grow, compete for resources, and die. For large-scale applications, the forest dynamics of one patch is assumed to be representative of the grid cell.

Reproduction in SEIB occurs on an annual time step and only individuals with more than 10 kg of biomass reproduce. Reproductive individuals allocate ten percent of C_{g+r} to reproduction regardless of canopy position. Reproductive carbon flows to the litter pool, which is used to create new recruits (starting dbh of 1 cm). Dispersal, seed survival, and germination are not explicitly represented in the main version of SEIB. However, Sato and Ise (2012) published a version of SEIB (referred to here as "SEIB-2012") that gives any climatically eligible PFT the opportunity to establish in ten percent of neighboring grid cells. They prescribed the frequency at which PFTs could jump between grid cells to match rates of plant migration observed in paleo records.

Similar to LPJ-GUESS, SEIB implicitly models environmental constraints on seedling survival and growth using PFT-specific BEs including PAR_{min}, GDD_{min}, $T_c$min, $T_c$max, a soil moisture threshold, and a number of "dry months" (monthly potential evapotranspiration exceeding monthly precipitation) that must not exceed a maximum (Sato & Ise, 2012). Additionally, the simulation plot is divided into a grid wherein each unoccupied 1 m^2 cell has a prescribed, PFT-specific establishment probability. The establishment probability and the number of unoccupied cells determine recruitment rates. At model initialization, all PFTs can establish in any unoccupied cell, but as the simulation progresses, the number of cells available to each PFT is assumed to be proportional to the PFT's biomass. The amount of reproductive carbon allocated to the litter pool does not influence recruitment rates. In sum, the primary factors influencing PFT-specific recruitment rates in SEIB are BEs, prescribed PFT-specific establishment probabilities in open cells, and the relative biomass of PFTs.

aDGVM

The Adaptive Dynamic Global Vegetation Model (aDGVM) represents individual trees within flexibly-sized (typically 1 ha) forest patches and was originally developed with a focus on tropical vegetation at grassland-savanna-forest biome boundaries (Scheiter & Higgins, 2009). aDGVM2 builds upon aDGVM, but instead of having PFTs with a fixed set of traits defined at model initialization, it allows individuals to assume unique trait combinations during the simulation in response to environmental conditions (Scheiter *et al.*, 2013). aDGVM2 has been run at regional to continental scales on multiple continents (Langan *et al.*, 2017; Martens *et al.*, 2021) and shows promise for global scale simulations.

Trees above an age threshold (default = 10 years) produce seeds at a rate proportional to C_{g+r} on the first day, i, of each month such that

$$Seed_{in} = \frac{C_{g+r,i}}{B_{seed}} \qquad \text{(Eqn 5)}$$

where $Seed_{in}$ is the number of seeds entering the seed bank and B_{seed} is the mass of carbon in one seed. Seed banks (Seeds) are tracked similarly to CLM(ED) / FATES, but seeds can't germinate until the next wet season, which is defined as three consecutive days with soil moisture of the upper soil layer at field capacity. In the wet season, a stochastic proportion of the seed bank, P_{sprout}, is available for germination which helps avoid all seeds germinating at once. Individual seedlings are tracked and the number of new seedlings produced is calculated as

$$N_{seedling} = (Seeds)(P_{sprout})(P_{germ}) \qquad \text{(Eqn 6)}$$

where P_{germ} is a prescribed germination probability. Seedling light availability is reduced by competition with grasses and other trees which affects the probability of seedling mortality via negative carbon balance (Scheiter & Higgins, 2009). Seedling biomass has a high probability of being killed by fire due to a size-dependent "topkill" probability. However, stems are not killed by fire which allows them to resprout. A favorable period without fire can allow seedlings, or resprouts, to grow large enough to escape the fire trap, but the risk of mortality via negative carbon balance increases with successive top kill events.

IV. Strengths and weaknesses of current regeneration algorithms

In this section we review how current model representations are positioned to capture critical forest responses to climate change. In Section V we propose directions forward for parameterizing, evaluating, and improving algorithms. The current representations of regeneration processes pose unique challenges for predicting future forests. All models reviewed here represent RA as an environmentally insensitive process and use uncertain parameters which likely bias estimates of biomass accumulation. Aside from aDGVM's representation of resprouting and moisture-sensitive germination, no other models we reviewed explicitly represent environmentally sensitive seed survival, germination, or resprouting sufficiently to capture vegetation responses to changing disturbance regimes and climate. Existing algorithms for dispersal and seedling growth and survival are not well validated posing uncertainty in their ability to capture tree migration potential, post-disturbance recovery, and demographic bottlenecks at the seedling stage.

Reproductive Allocation and Seed Production

Biases in the representation of RA will bias estimates of biomass accumulation in all forests, particularly following landscape-scale disturbance. RA is represented in most of the models reviewed here (Table 1), but parameter values do not have strong empirical support (Harper *et al*., 1977; Wenk & Falster, 2015; Wenk *et al*., 2017). It is likely that RA is being misrepresented in all models for at least some biomes and PFTs because all tree PFTs are typically given the same RA parameter and maturation size. This uniformity conflicts with empirical evidence showing that RA varies widely across biomes, PFTs, and life stages (Thomas *et al*., 2011; Wenk & Falster, 2015; Visser *et al*., 2016; Ruger *et al*. 2020). This is problematic for capturing forest function because there is a direct tradeoff between RA and growth in most VDMs, with empirical support for this assumption (Thomas *et al*., 2011; Berdanier & Clark, 2016; Hacket-Pain *et al*., 2018).

The lack of empirical basis for current representations of RA schedules or the lack of RA schedules in VDMs poses significant uncertainty for estimating range shifts. A tree's RA schedule guides the timing and magnitude of RA over its lifetime (Wenk & Falster, 2015). In LPJ-GUESS, allocation to reproduction starts as soon as new trees emerge in a patch (B. Poulter, personal communication, 2018), which would enable an immediate and unrealistic positive feedback between recruitment, reproductive output (and a commensurate negative bias in biomass accumulation of young trees), and subsequent colonization of nearby patches or grid cells. This will result in unrealistically fast migration rates as suggested by a parameter sensitivity analysis in LPJ-DISP (Snell, 2014), which is consistent with landscape model experiments which show that the age of reproductive maturity affects the speed with which populations can migrate across a landscape (Miller & McGill, 2017).

VDMs may struggle to capture climate-driven compositional turnover within forests where climate change differentially affects the mass of carbon available for reproduction. This mass influences recruitment rates in ED-based models, LM3-PPA, aDGVM, and LPJ-GUESS which positions them relatively well to predict how future productivity will affect seed production and recruitment rates. For SEIB, recruitment is proportional to adult abundance regardless of productivity. Wherever compositional inertia maintains relative PFT abundances via proportional propagule production this is a safe assumption, but where climate change differentially limits RA and/or production between PFTs, PFT-level propagule pressure may no longer reflect adult abundance. Where this is the case, models that do not represent recruitment rates as a function of reproductive carbon would overestimate compositional inertia and underestimate compositional turnover.

Dispersal

Aside from ED2, intra-grid cell dispersal (dispersal between patches within the same grid cell) is not explicitly represented in any of the main versions of the models discussed here, which hinders their ability to represent disturbance-recovery dynamics. Models that implicitly assume unlimited intra-grid cell dispersal will overestimate tropical forest regeneration on abandoned agricultural land (as in Hooper *et al.*, 2005) and will not accurately capture forest to non-forest vegetation transitions in response to increasing fire size, such as in the western United States (as in Tepley *et al*. 2017; Stevens-Rumann *et al*. 2019). Since mean dispersal distance tends to be short (Clark, 1998), rates of post-disturbance regeneration from seed are dependent on the size, shape, and adjacency of the disturbed patch to remnant patches (Tepley *et al.*, 2017). The spatially implicit patch structure used by most VDMs limits the set of possible representations of patch adjacency necessary for inter-patch (i.e. intra-grid cell) seed dispersal. LPJ-DISP (Snell, 2014) overcomes this constraint by representing the spread of PFTs between patches similar to an infection process where the probability of patch adjacency (and transmission) is a function of the number of patches already containing the PFT. LPJ-GM (Lehsten *et al*., 2019) spatializes the sub-grid cell patch structure of LPJ-GUESS and represents dispersal along transects of adjacent patches. These innovations lay promising groundwork towards more realistically limiting recolonization rates in recently disturbed patches and, because they also can be applied across grid cells, may enable simulation of long distance tree migration.

The lack of explicit inter-grid cell dispersal in most VDMs hinders their ability to predict persistent changes in the geographic distribution of trees (i.e. tree migration). Model testing and theory highlights the importance of considering dispersal kernel shape and maturation age in predicting long distance tree migration rates (Clark *et al*., 2001; Nathan, 2011; Miller & McGill, 2017). Both of these parameters are represented in LPJ-GM (Lehsten *et al*., 2019), which best positions it to predict migration rates. With the exception of LPJ-DISP and LPJ-GM, models reviewed here do not use dispersal kernels to represent how dispersal probability declines with distance from the seed source. The use of discrete simulation points in some models (e.g. Sato *et al.*, 2007) introduces a discontinuous representation of vegetated space between adjacent grid cells (50-500 km resolution), which is incongruous with the standard application of dispersal kernels. Furthermore, the need to parallelize computations across grid cells has generally prohibited the inter-grid cell transfer of seed (LPJ-GM used a message passing interface protocol to overcome this; Clarke *et al.*, 1994; Lehsten *et al.*, 2019). These challenges perhaps explain why it is most common for models to allow PFTs to establish in any grid cell for which they are climatically eligible, implicitly assuming that inter-grid cell dispersal is unlimited. LPJ-GM has the ability to prescribe dispersal resistance on a landscape, but no VDM reviewed here explicitly represents dispersal agents or dispersal mechanism (i.e. animal, wind, water, etc.), which may make it hard to assess how topography, land cover (Miller & McGill, 2017), meteorology, or the abundance of faunal dispersal (Mokany *et al*., 2014) agents facilitate or restrict dispersal depending on the trees' dispersal mechanism (Nathan *et al*., 2008). However, it is still unclear to what extent including dispersal agents is appropriate, necessary, or possible in global scale VDMs.

Seed Survival, Germination, and Resprouting

Current model formulations are not well-positioned to capture how interactions between seed traits, disturbance regimes, and environmental conditions can drive compositional turnover, but this may only be critical in some biomes. For example, in boreal forests, models without seed traits would not capture how establishment of smaller-seeded, broad-leaved deciduous species is limited by access to mineral soil until severe fire removes this constraint (Johnstone & Chapin, 2006; Beck *et al.*, 2011). Prescribed, environmentally insensitive seed bank parameters in ED-based models are based on parameter values for central Europe (Lischke *et al*., 2006; Fisher *et al*., 2015), but are applied to tropical forests by default (Holm *et al.*, 2020; Koven *et al.*, 2020). This makes it difficult to capture how functionally different tropical seeds banks, with recalcitrant seeds, offer limited resilience after agricultural land abandonment (Chapman & Chapman, 1999; Martins & Engel, 2007; Lipoma *et al*., 2020), potentially leading to changes in forest composition.

Resprouting is a critical forest regeneration process that is only explicitly represented in aDGVM, likely biasing predictions of forest recovery and compositional turnover in disturbance prone regions in all other VDMs. Most models prioritize growth to meet allometric targets, which would implicitly enable resprouting of leaf or stem biomass if they have formulations that allow for non-lethal damage. However, no VDM we reviewed explicitly represents different resprouting strategies, such as bud protection and bud location (e.g. from epicormic vs. underground buds; but see Kelley *et al.*, 2014), which can promote clonal reproduction, mediate forest compositional response to disturbance regimes, and affect biomass loss from fire (Del Tredici, 2001; Clarke *et al*., 2013; Pausas & Keeley, 2014; Pausas *et al*., 2015). aDGVM is best positioned to capture resprouting mediated forest responses such as LAI recovery following wildfire in Australian sclerophyll eucalypt forests (Nolan *et al*., 2020), the observed functional turnover between fire-sensitive species and resprouters in response to changing fire frequency (Pausas & Keeley, 2014, DeSantis *et al*. 2011), and the denser, shorter forests created via resprouting after hurricanes (Van Bloem *et al.*, 2006).

Seedling Survival and Growth

The models reviewed here represent differential seedling survival and growth with bioclimatic envelopes (BEs) or by explicitly using a very small (i.e. seedling stage) smallest size class. LPJ-GUESS and SEIB's BEs both include PFT-specific PAR_{min} parameters, positioning them relatively well to capture how changing disturbance regimes are predicted to differentially affect seedling establishment among light demanding and shade tolerant PFTs (Comita *et al*., 2009). BEs required for "establishment" have proven useful for predicting historical PFT distributions and forest succession, but have been critiqued for operating in lieu of physiological understanding of climate tolerances (Fisher *et al*., 2015) and for assuming that historical correlations with climate will predict future plant distributions subject to novel climates and PFT assemblages (Journé *et al*., 2019). It is particularly unclear if these correlations, used to parameterize BEs in LPJ-GUESS and SEIB (based on Sykes *et al*., 1996), would capture seedling-specific survival and growth responses to changing understory conditions. BEs are not well-suited to capture niche differences between seedlings and conspecific adults (Dobrowski *et al.*, 2015) making it difficult to predict tree range contractions around favorable establishment microsites (Mclaughlin & Zavaleta, 2012). Nevertheless, if these uncertainties can be overcome, SEIB and LPJ-GUESS's BEs may capture the role that future temperature and moisture regimes play in changing global forest distribution.

Models that explicitly represent the seedling stage (aDGVM and LM3-PPA) may be better prepared to capture how future understory conditions will impact seedling growth and survival if the associated computational burden remains manageable for global scale simulations. Unlike all other models reviewed here, recruitment in ED-based models is entirely insensitive to conditions at the forest floor where seedling dynamics take place and, in contrast to models that use BEs (e.g. LPJ-GUESS), implicitly assume that establishment processes do not contribute to emergent biogeography. Only mortality following recruitment and differences in the NPP of reproductive cohorts (which has cascading effects on reproductive output) determine biogeography. This limits their ability to capture scenarios where the seedling stage becomes a life history bottleneck (e.g. Mclaughlin & Zavaleta, 2012; Dobrowski *et al.*, 2015). The convention in ED-based models where reproductive carbon is divided by the mass required to produce a new recruit (Z_0; Eqn 1) is a particularly unrealistic abstraction (also shared by LM3-PPA to produce seedlings) that hinders algorithm evaluation against observations of seed and seedling dynamics.

V. Priorities for advancing prediction

Understanding the strengths and weaknesses of current VDMs is a starting point, but diverse research activities are needed to advance the representation of critical regeneration processes.
Unlike prior efforts to review and recommend improvements to the representation of regeneration in gap models (Price *et al.*, 2001), our recommendations are specific to global scale simulations coupled to atmospheric models and therefore account for the unique needs of VDMs to conserve carbon, operate with computational efficiency, and use mechanistic process-based representations. For processes already included in VDMs, research priorities depend on model sensitivity to associated parameters and how well algorithms and parameters have been empirically evaluated. The inclusion of new processes does not guarantee improved prediction and could unnecessarily increase model complexity (Fisher & Koven, 2020). Therefore, our recommendations prioritize the processes identified in section II (Fig. 1) for which there is clear evidence the process influences forest function, composition, or extent at the scale of the ecosystem or PFT in response to global change.

The incorporation of new regeneration processes will inevitably introduce new PFT-specific parameters based on regeneration traits that may not align with the current axes of trait variation used to define PFTs (Ruger *et al*., 2018), potentially necessitating the division of existing PFTs into "subclasses" (Neilson *et al*., 2005). Newer approaches to representing trait diversity, such as flexible individual traits (e.g. LPJmL-FIT, Sakschewski *et al.*, 2016 and aDGVM2, Scheiter *et al.*, 2013) may be required to avoid exponentially expanding the number of fixed PFTs used. To help manage the increasing complexity of VDMs, we recommend employing "modular complexity as a strategy" (Fisher & Koven, 2020), where new algorithms are implemented as distinct modules that can be turned on or off depending on the goal and spatial extent of the simulation.

The trend towards including more mechanistic representations of physiological and demographic processes in VDMs can make them difficult to parameterize and evaluate. Inverse modeling, whereby parameters are calibrated based on their ability to reproduce empirical observations (Hartig *et al.*, 2012), shows promise for parameterizing under-observed processes in vegetation models (e.g. Van Oijen *et al.*, 2005; Hanninen *et al.*, 2019; Chalmandrier *et al.*, 2021). Data syntheses specifically geared towards VDM model development (e.g., Box 2) can constrain parameters and algorithms directly or via inverse modeling approaches (e.g., Box 3), such as Bayesian calibration (e.g. Hartig *et al.*, 2012).

Reproductive Allocation and Seed Production

We recommend prioritizing the evaluation of RA algorithms and parameters because a) reproductive allocation (RA) is represented in nearly all VDMs reviewed here, b) RA affects forest function globally through tradeoffs with growth (Thomas *et al*., 2011; Wenk & Falster, 2015), c) the size or age of reproductive maturity governs forest distribution changes by mediating migration rates (Nathan, 2011; Miller & McGill, 2017), and d) RA algorithms are poorly evaluated and the default parameter values are uncertain (Harper, 1977; Wenk & Falster, 2015).

Like many physiological processes in VDMs, directly observing RA in the field is challenging. Although the most intuitive field measure of RA is the carbon mass of reproductive litter flux (R) as a fraction of NPP, field estimates of R/NPP are rare because of the difficulty of sampling all NPP components. The mass of reproductive litterfall as a fraction of the mass of leaf litter flux (R/L) may offer a more readily available proxy for investment in reproduction (see Box 2). Based on our global synthesis of data from forest sites where R, L, and NPP are reported (Box 2), we find that R/L is a tractable constraint for RA in VDMs. Therefore, we recommend that R/L is used to constrain RA parameters and assess the structure of regeneration algorithms more broadly (Box 3).

Box 2. Global data available to support VDM parameterization of reproductive allocation

Although measures of reproductive output (R) are relatively extensive (e.g. Qiu *et al.*, 2021), it has been unclear whether sufficient data exist to estimate reproductive investment as a fraction of NPP and how these data are distributed across biomes. Because NPP is not often reported at fine spatial or temporal resolution, estimates of R/NPP are spatially and temporarily limited. We propose that using the ratio of carbon allocated to reproductive vs. leaf litterfall, R/L, as a proxy for reproductive investment could aid efforts to constrain RA and characterize its environmental sensitivity.

To assess the viability of R/L as a proxy for R/NPP, we first conducted a global data assessment of forest sites where R/L, R/NPP and R/ANPP (aboveground NPP) are publicly available (Methods S1). We then tested how our proposed proxy, R/L, is correlated with R/NPP (Fig 5a) and R/ANPP (Fig. S1). Finally, we calculated R/L at a larger set of sites where R and L are reported for a minimum of one year (Fig. 5b). See Tables S2, S3 for data.

We found a strong correlation between R/NPP and R/L ($R^2 = 0.87$; Fig 5a) and between R/ANPP and R/L ($R^2 = 0.86$, Fig. S1), suggesting that R/L provides a useful benchmark for reproductive investment and could be used to calibrate RA parameters in VDMs (Box 3). Estimates of R/L from 550 sites ranged from 0 to 3.7 with the lowest variance in tropical rainforests (± 0.012) and the highest variance within woodland/shrubland (± 0.25). The correlation between R/L and R/NPP is driven in part by the correlation between L and NPP ($R^2 = 0.57$, Fig. S2), consistent with prior observations in tropical forests (Malhi *et al.*, 2011).

Figure 5. a) Correlation between observed reproductive litter flux as a fraction of net primary productivity (R/NPP) and leaf litter flux (R/L; all fluxes in g C m^{-2} yr^{-1}) at 61 forest sites (colors indicate Whittaker biomes as in panel b; shaded area shows the 95 % confidence interval), and b) the magnitude of R/L (shown by point size) and the distribution of R/L observations across Whittaker biomes (n = 550); "N sites" shows the approximate number of distinct measures of R/L at mean annual temperature and precipitation coordinates.

Data will be most useful for supporting VDM parameterization and evaluation when they capture a functionally and biogeographically broad sample of individuals identified to species, are generated with standardized field protocols and have concurrent meteorological observations. Long-term forest monitoring plots generally exhibit these characteristics (Anderson-Teixera *et al*. 2015; Metzger *et at.,* 2019; Malhi *et at.,* 2021; Davis *et al.*, 2021), highlighting their utility for advancing the representation of PFT-specific processes as a function of environmental variables in VDMs. For RA specifically, long-term litterfall monitoring at the Center for Tropical Forest Science (CTFS)-ForestGeo network sites are useful because sampling protocols call for species-level identification of reproductive and leaf material. Similar protocols (although not always including species-level identification) have been adopted by GEM, RAINFOR and NEON networks, among others. Although a trend towards open, accessible data is apparent, most of these data are not publicly available, emphasizing the opportunity and need for a large-scale collaborative synthesis among data holders.

Box 3. Evaluating reproductive allocation in FATES at Barro Colorado Island (BCI)

Directly observing RA is challenging, but more observable benchmarks show promise for calibrating RA in vegetation models (Box 2). We used observations of R/L, R/ANPP, and recruitment rates into the 1 cm size class to evaluate a range of potential parameter values for RA in FATES (Methods S1, S2). We ran 7 simulations of a single broad-leaved tropical tree PFT using boundary conditions for BCI (Faybishenko *et al.*, 2018) with RA varying from 0.009 to 0.41, based on a range of empirical estimates for tree species globally (Wenk & Falster, 2015). All other parameters were held constant.

Figure 6. FATES predictions at Barro Colorado Island (circles) of a) reproductive litter flux (R; g C m^{-2} yr^{-1}) as a fraction of leaf litter flux (R/L), b) R as a fraction of above-ground net primary productivity (R/ANPP), and c) recruitment rates into the 1 cm size class. Each model prediction shows the mean across 18 simulation years run with recycled, observed meteorology (2003-2016), initialized with observed stand structure, and using a prescribed value of reproductive allocation (RA; ranging from 0.009-0.41). The standard deviation of inter-annual variation is smaller than the symbol size for most circles. Triangles show

observations (mean and standard deviation among years) at the BCI Forest Dynamics Plot.

The value of RA resulting in the best match to observations of R/L and R/ANPP is 0.15, indicating that the default parameter value (0.1) should be raised slightly at BCI. However, evaluating RA against observations of recruitment rates indicates that RA should be lowered below 0.009. This implies that FATES may allocate too little to reproduction, but insufficiently constrains reproductive carbon between RA and recruitment. This could indicate inaccurate seed bank parameters for decay and germination (i.e. too much seed may be surviving and germinating), and/or larger structural issues related to FATES's implicit assumption that seedling growth and mortality do not constrain carbon available to make new recruits. Due in part to the direct trade-off with growth in FATES, lower reproductive allocation increases growth (data not shown).

Dispersal

We recommend that VDMs incorporate inter- and intra-grid cell dispersal because these processes limit tree migration rates (inter-grid cell scale) and regeneration after landscape-scale disturbance (intra-grid cell scale; Clark *et al*., 2001; Chapman & Chapman, 1999; Nielson *et al*., 2005; Sato & Ise 2012; Stevens-Rumann *et al*., 2019). Dispersal in current VDMs ranges from relatively complex (LPJ-GM), to relatively simple (ED2 and SEIB-2012), but it is still unclear what level of complexity is required. At a minimum we recommend the incorporation of dispersal kernels that represent both short and long-distance dispersal events by incorporating a "long-tailed" dispersal probability kernel (Clark *et al*., 2001). This provides a path towards more mechanistic representations of tree migration without necessarily requiring the explicit representation of dispersal agents and may eliminate the need for "seed rain" which violates the conservation of carbon. Although parameterizing the occurrence of rare, long-distance dispersal events is challenging, it is important because these are believed to explain the pace of historical tree migration (Clark, 1998). Pollen records (Solomon & Kirilenko, 1997; Clark *et al*., 1998; Ordonez & Williams, 2013) can be used to calibrate parameters in dispersal kernels to constrain tree migration more mechanistically than prior VDM experiments that prescribe migration rates without explicit dispersal (e.g. Sato & Ise, 2012). For taxa where pollen records are not available (Nielson *et al*., 2005) observations of tree invasions (e.g. Hunter and Douglas 1984), tree line movement (Kullman *et al*., 2001), and "signatures of range expansion" (Murphy *et al*., 2010) may be used to calibrate dispersal kernel parameters.

The development of intra-grid cell dispersal will enable the simulation of changes in disturbance-recovery dynamics and these algorithms could tractably be evaluated using data from post-disturbance seedling censuses (e.g. Chapman & Chapman, 1999; Nagel & Taylor, 2005; Tepley *et al*., 2017; Young *et al*., 2019) and "seed shadows" (Clark *et al.*, 1999). Evaluation of dispersal algorithms at these smaller spatio-temporal scales is needed to ascertain if relatively simple approaches to intra-grid cell dispersal, such as that in ED2, are sufficient to capture secondary forest development after megafires and agricultural land abandonment, or if the inclusion of PFT-specific dispersal kernels (e.g. LPJ-GM and LPJ-DISP) make meaningful improvements for capturing forest recovery rates after landscape-scale disturbance. Models that don't already represent multiple, coexisting patches per grid cell (e.g. SEIB) will have to do so, perhaps following Lehsten *et al.* (2019).

Seed Survival, Germination, and Resprouting

Resprouting is a well understood process (Clarke *et al*., 2013, Pausas & Keeley, 2014) that mediates forest composition in response to changing disturbance regimes (Saha *et al*., 2003; Beck *et al*. 2011; DeSantis *et al*., 2011) and should be prioritized for biomes and PFTs for which this regeneration strategy is dominant. Forest landscape models (FLMs) represent resprouting probabilities and qualitative fire tolerance scores to capture disturbance-recovery dynamics (Mladenoff, 2004; Scheller *et al*., 2007), but we recommend more mechanistic representations for VDMs. For example, to capture the costs of resprouting, Kelley *et al.* (2014) represented a trade-off between investment in thicker bark (conferring resprouting ability after fire) and seedling establishment. We recommend that the "buds-protection-resources" (BPR) framework (conceptualized by Clarke *et al.*, 2013) is used as the basis for algorithm development where algorithms are designed to include trade-offs within carbon allocation schemes and coordination of PFT-level traits such as bud location and degree of protection (e.g., via bark thickness). This will help VDMs capture PFT-specific, resprouting-mediated carbon stock resilience and LAI recovery in response to changing disturbance regimes where well-protected aerial buds survive fire, resulting in far less biomass loss than where resprouting occurs from basal or underground buds (Clarke *et al*., 2013). TRY, the global database of plant traits (Kattge *et al.*, 2011), includes observations of bark thickness, budbank height distribution, and storage organs for thousands of individuals which makes it a promising resource for parameterizing BPR-based resprouting algorithms in VDMs.

Seed size and longevity can filter functional composition (e.g. Johnstone & Chapin, 2006; Beck *et al.*, 2011; Pausas & Keeley, 2014), but may only be necessary in some contexts. For example, in boreal forests we recommend representing how seed size mediates seedling establishment success after fire because this is important for capturing changes in functional composition across large spatial extents (e.g. Beck *et al.*, 2011). For tropical secondary forests it may be necessary to represent seed longevity and recalcitrance to capture how tropical seed banks offer limited biomass resilience following agricultural land abandonment (Chapman & Chapman, 1999; Martins & Engel, 2007; Lipoma *et al*., 2020).

Seedling Survival and Growth

Seedling survival and growth influence recruitment rates, drive compositional turnover across moisture gradients (Engelbrecht *et al*., 2007) and shape biome boundaries (Bond, 2008; Conlisk *et al*. 2017; Kueppers *et al*. 2017). The VDMs reviewed here implicitly represent one or more aspects of seedling survival and growth with diverse approaches, but none have evaluated their approaches against observations of recruitment rates into the 1 cm size class. Evaluations against recruitment data (e.g. Box 3, Fig. 6c) can help determine if current representations sufficiently capture seedling-specific ecophysiological limits and responses to disturbance regimes, or if current models are allowing recruitment to occur where the seedling stage would actually be limiting.

We used census data to calculate recruitment rates into the 1 cm size class at four sites (Methods S3, Tables S4, S5) and compared these observations to predictions in FATES at Barro Colorado Island (BCI; Box 3, Fig. 6c), Panama. We found evidence of excessive recruitment in FATES (Fig. 6c) which aligns with recent simulations in ED2 showing that excess mortality in the smallest size classes

(1-10 cm dbh) appears to compensate for hyperactive recruitment (Fig. 7). These observations indicate that the pool of carbon available to produce new recruits in ED-based models is generally too large, and that a representation of seedling growth and mortality processes are needed. We recommend algorithm development in ED-based models to introduce PFT-specific, environmentally sensitive recruitment limitations based on observations of seedling ecophysiology or demography. Experiments that track early seedling development under manipulated temperature, moisture, or CO_2 regimes are critical to support this type of algorithm development (e.g. Kobe, 1999; Levy *et al*., 2000; Engelbrecht *et al.,* 2007; Kueppers *et al*., 2017). However, leveraging these observations will require updating the problematic abstraction in ED-based models that all carbon to make new recruits must come from reproductive carbon (Eqn 1). This could be achieved by tracking a seedling layer, subject to environmentally sensitive mortality, that assimilates atmospheric carbon to produce new recruits.

Additional opportunities to evaluate representations of seedling growth and survival depend on each model. For models that represent the seedling stage by having a very small smallest size class, such as LM3-PPA, benchmarks should test if representing seedlings with the same parameters as adults, some of which interact with size to affect demographic rates, fails or succeeds in capturing observed seedling demographic responses to environmental variation. Initial testing in LM3-PPA found that tuning seedling mortality rates was required to reconcile abundance differences between the seedling stage and larger size classes (Weng *et al*., 2015), indicating that (similar to our conclusions about ED-based models) the current formulation is missing seedling-specific constraints to survival. For models that use BEs we recommend testing how contemporary observations of seedling recruitment aligns with BEs that were parameterized with historical, adult distributions. This will ensure that, in addition to circumscribing the requirements of adult trees for growth and survival, BEs are no wider than the set of conditions required for trees to complete their entire life cycle.

VI. Conclusion

Regeneration processes are critical to the future of global forests. To date, even the most advanced vegetation schemes emerging for use within ESMs, VDMs, represent regeneration processes with relative simplicity compared to growth and mortality processes. Current regeneration formulations may allow most models to maintain the distribution of current vegetation, but they will require improvements to capture future functional composition, range shifts, and post-disturbance recovery. Reproductive allocation and maturation parameters lack strong empirical support, which is needed to avoid biases in migration rates and carbon accumulation rates after landscape-scale disturbance. However, our data synthesis finds empirical support for novel RA benchmarks constructed from litterfall data, which can be used to better constrain RA parameters in VDMs. Future work should leverage spatially extensive and long-running litterfall observations to improve current default RA parameters globally. Dispersal is implemented in some VDM versions, but new benchmarking approaches are needed to assess how well current algorithms can capture tree migration and recovery from landscape-scale disturbance. Algorithms for resprouting lag empirical understanding and synthesis, indicating an opportunity for algorithm development to capture how resprouting strategies mediate functional turnover and biomass recovery in response to changing disturbance regimes. It is unclear how the diverse approaches used to represent seedling survival and growth capture seedling-stage recruitment bottlenecks. Benchmarking is needed to determine to what degree VDMs currently capture how this life stage mediates functional turnover in response to changing environmental conditions. More systematic parameter sensitivity analyses (building upon Fisher *et al.*, 2010; Snell, 2014; Koven *et al.,* 2020) are needed across all processes and models to identify which improvements will have the most meaningful impacts. Adding complexity to already complex models doesn't guarantee improved prediction, and can be deleterious if it expands prediction space without adequate observational constraints. We recommend incorporating well-understood processes using "modular complexity as strategy" (Fisher & Koven, 2020) which shows promise for reconciling the need to represent numerous processes mechanistically with the challenges of increasingly complex ESM components. Without addressing the key gaps outlined here, VDMs will not fulfill their mandate.

Acknowledgements

We thank M. Billing, E. Weng, E. Joetzjer, R. Knox, C. Koven, R. Fisher, B. Poulter, T. Powell, and H. Sato, who clarified details of regeneration algorithms within various VDMs, and anonymous reviewers whose comments improved the manuscript. We thank R. Knox and M. Longo for preparing files used in the FATES simulations. This work was supported as part of the Next Generation

Ecosystem Experiments-Tropics (NGEE Tropics), funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award number DE-AC02-05CH11231. A. Hanbury-Brown was also supported in part by the National Science Foundation and the National Aeronautics and Space Administration while conducting this research.

Author contributions

Conceptualization and Investigation A. R. Hanbury-Brown and L. M. Kueppers; Writing-Original Draft, A. R. Hanbury-Brown and L. M. Kueppers; Model simulations, A. R. Hanbury-Brown; Data synthesis, R. E. Ward; Writing-Review & Editing, A. R. Hanbury-Brown, L. M. Kueppers, R. E. Ward.

Data availability

The data that supports the findings of this study are available in the supplementary material of this article.

References

Abatzoglou JT, Williams AP. **2016**. Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences* **113**: 11770–11775.

Alfaro-Sánchez R, Muller-Landau HC, Wright SJ, Camarero JJ. 2017. Growth and reproduction respond differently to climate in three Neotropical tree species. *Oecologia* **184**: 531–541**.**

Anadon-Rosell A, Talavera M, Ninot JM, Carrillo E, Batllori E. **2020**. Seed production and dispersal limit treeline advance in the Pyrenees. *Journal of Vegetation Science* **31**: 981–994.

Anderson-Teixeira KJ, Davies SJ, Bennett AC, Gonzalez-Akre EB, Muller-Landau HC, Joseph Wright S, Abu Salim K, Almeyda Zambrano AM, Alonso A, Baltzer JL, *et al.* 2015. CTFS-ForestGEO: A worldwide network monitoring forests in an era of global change. *Global Change Biology* **21**: 528–549.

Argles APK, Moore JR, Huntingford C, Wiltshire AJ, Harper AB, Jones CD, Cox PM. **2020**. Robust Ecosystem Demography (RED version 1.0): A parsimonious approach to modelling vegetation dynamics in Earth system models. *Geoscientific Model Development* **13**: 4067–4089.

Barlow J, Peres CA. **2008**. Fire-mediated dieback and compositional cascade in an Amazonian forest. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**: 1787–1794.

Beck PSA, Goetz SJ, Mack MC, Alexander HD, Jin Y, Randerson JT, Loranty MM. **2011**. The impacts and implications of an intensifying fire regime on Alaskan boreal forest composition and albedo. *Global Change Biology* **17**: 2853–2866.

Beckage B, Clark JS. **2003**. Seedling survival and growth of three forest tree species: the role of spatial heterogeneity. *Ecology* **84**: 1849–1861.

Berdanier AB, Clark JS. **2016**. Divergent reproductive allocation trade-offs with canopy exposure across tree species in temperate forests. *Ecosphere* **7**: e01313.

Van Bloem SJ, Murphy PG, Lugo AE. **2007**. A link between hurricane-induced tree sprouting, high stem density and short canopy in tropical dry forest. *Tree Physiology* **27**: 475–480.

Bonan GB. **2008**. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science* **320**: 1444–1449.

Bonan G. 2015. *Ecological Climatology*. Cambridge, UK: Cambridge University Press.

Bonan G. **2019**. *Climate Change and Terrestrial Ecosystem Modeling*. New York, USA: Cambridge University Press.

Bond WJ. **2008**. What Limits Trees in C $_4$ Grasslands and Savannas? *Annual Review of Ecology, Evolution, and Systematics* **39**: 641–659.

Boulanger V, Baltzinger C, Said S, Ballon P, Picard J-F, Dupouey J-L. **2015**. Decreasing deer browsing pressure influenced understory vegetation dynamics over 30 years. *Annals of Forest Science* **72**: 367–378.

Brown CD, Dufour-Tremblay G, Jameson RG, Mamet SD, Trant AJ, Walker XJ, Boudreau S, Harper KA, Henry GHR, Hermanutz L, *et al.* 2019. Reproduction as a bottleneck to treeline advance across the circumarctic forest tundra ecotone. *Ecography* **42**: 137–147.

Brown PM, Wu R. **2005**. Climate and disturbance forcing of episodic tree recruitment in a southwestern ponderosa pine landscape. *Ecology* **86**: 3030–3038.

Caplat P, Anand M, Bauch C. **2008**. Interactions between climate change, competition, dispersal, and disturbances in a tree migration model. *Theoretical Ecology* **1**: 209–220.

Chalmandrier L, Hartig F, Laughlin DC, Lischke H, Pichler M, Stouffer DB, Pellissier L. **2021**. Linking functional traits and demography to model species-rich communities. *Nature Communications* **12**: 2724.

Chapman CA, Chapman LJ. **1999**. Forest restoration in abandoned agricultural land: A case study from East Africa. *Conservation Biology* **13**: 1301–1311.

Chazdon RL. **2003**. Tropical forest recovery: Legacies of human impact and natural disturbances. *Perspectives in Plant Ecology, Evolution and Systematics* **6**: 51–71.

Cierjacks A, Hensen I. **2004**. Variation of stand structure and regeneration of Mediterranean holm oak along a grazing intensity gradient. *Plant Ecology* **173**: 215–223.

Clark JS. **1998**. Why Trees Migrate So Fast: Confronting Theory with Dispersal Biology and the Paleorecord. *The American Naturalist* **152**: 204–224.

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

Clark JS, Fastie C, Hurtt G, Jackson STT, Johnson C, King G a. a, Lewis MA, Lynch J, Pacala S, Prentice C, *et al.* 1998. Reid’s paradox of rapid plant migration - Dispersal theory and interpretation of paleoecological records. *Bioscience* **48**: 13–24.

Clark JS, Lewis M, Horvath L. 2001. Invasion by Extremes: Population Spread with Variation in Dispersal and Reproduction. *The American Naturalist* **157**: 537–554.

Clark JS, Andrus R, Aubry-Kientz M, Bergeron Y, Bogdziewicz M, Bragg DC, Brockway D, Cleavitt NL, Cohen S, Courbaud B, et al. 2021. Continent-wide tree fecundity driven by indirect climate effects. *Nature Communications* **12**: 1242.

Clarke PJ, Lawes MJ, Midgley JJ, Lamont BB, Ojeda F, Burrows GE, Enright NJ, Knox KJEE. **2013**. Resprouting as a key functional trait: how buds, protection and resources drive persistence after fire. *New Phytologist* **197**: 19–35.

Clarke L, Glendinning I, Hempel R. **1994**. The MPI Message Passing Interface Standard. In: Decker KM, Rehmann RM, eds. Programming Environments for Massively Parallel Distributed Systems. Basel: Birkhäuser Basel, 213–218.

Comita LS, Uriarte M, Thompson J, Jonckheere I, Canham CD, Zimmerman JK. **2009**. Abiotic and biotic drivers of seedling survival in a hurricane-impacted tropical forest. *Journal of Ecology* **97**: 1346–1359.

Conlisk E, Castanha C, Germino MJ, Veblen TT, Smith JM, Moyes AB, Kueppers LM. **2018**. Seed origin and warming constrain lodgepole pine recruitment, slowing the pace of population range shifts. *Global Change Biology* **24**: 197–211.

Cooper-Ellis S, Foster DR, Carlton G, Lezberg A. **1999**. Forest response to catastrophic wind: Results from an experimental hurricane. *Ecology* **80**: 2683–2696.

de Paula Mateus D, Groeneveld J, Fischer R, Taubert F, Martins VF, Huth A. **2018**. Defaunation impacts on seed survival and its effect on the biomass of future tropical forests. *Oikos* **127**: 1526–1538.

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

DeSantis RD, Hallgren SW, Stahle DW. **2011**. Drought and fire suppression lead to rapid forest composition change in a forest-prairie ecotone. *Forest Ecology and Management* **261**: 1833–1840.

Detto M, Wright SJ, Calderón O, Muller-Landau HC. **2018**. Resource acquisition and reproductive strategies of tropical forest in response to the El Niño–Southern Oscillation. *Nature Communications* **9**: 913.

Dirzo R, Dirzo R, Young HS, Galetti M, Ceballos G. **2014**. Defaunation in the Anthropocene. *Science* **345**: 401–406.

Dobrowski SZ, Swanson AK, Abatzoglou JT, Holden ZA, Safford HD, Schwartz MK, Gavin DG. **2015**. Forest structure and species traits mediate projected recruitment declines in western US tree species. *Global Ecology and Biogeography* **24**: 917–927.

Dunne JP, Horowitz LW, Adcroft AJ, Ginoux P, Held IM, John JG, Krasting JP, Malyshev S, Naik V, Paulot F, *et al.* 2020. The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics. *Journal of Advances in Modeling Earth Systems* **12**: e2019MS002015.

Dupuy JM, Chazdon RL. **2008**. Interacting effects of canopy gap, understory vegetation and leaf litter on tree seedling recruitment and composition in tropical secondary forests. *Forest Ecology and Management* **255**: 3716–3725.

Engelbrecht BMJ, Kursar TA. **2003**. Comparative drought-resistance of seedlings of 28 species of co-occurring tropical woody plants. *Oecologia* **136**: 383–393.

Engelbrecht BMJ, Kursar TA, Tyree MT. **2005**. Drought effects on seedling survival in a tropical moist forest. *Trees* **19**: 312–321.

Engelbrecht BMJ, Comita LS, Condit R, Kursar T, Tyree MT, Turner BL, Hubbell SP. **2007**. Drought sensitivity shapes species distribution patterns in tropical forests. *Nature* **447**: 80–82.

FAO. **2010**. *Global Forest Resources Assessment*. Rome: FAO.

FATES Development Team. **2018**. FATES Technical Documentation. [WWW Document] URL https://fates-docs.readthedocs.io/en/latest/fates_tech_note.html#seed-dynamics-and-recruitment [accessed 18 June 2019].

Faybishenko B; Paton S; Powell T; Knox R; Pastorello G; Varadharajan C; Christianson D; Agarwal D. 2018. QA/QC-ed BCI meteorological drivers. 1.0. NGEE Tropics Data Collection. [dataset] http://dx.doi.org/10.15486/ngt/1423307 [accessed 28 February, 2022].

Fisher RA, Christoffersen O, Longo M, Viskari T, Koven CD, Anderegg WRL, Dietze MC, Farrior CE, Holm JA, Hurtt GC, *et al.* 2018. Vegetation demographics in Earth System Models : A review of progress and priorities. *Global Change Biology* **24**: 35–54.

Fisher RA, Koven CD. **2020**. Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *Journal of Advances in Modeling Earth Systems* **12**: e2018MS001453.

Fisher R, Mcdowell N, Purves D, Moorcroft P, Sitch S, Cox P, Huntingford C, Meir P, Ian Woodward F, Woodward FI, *et al.* 2010. Assessing uncertainties in a second-generation dynamic vegetation model caused by ecological scale limitations. *New Phytologist* **187**: 666–681.

Fisher RA, Muszala S, Verteinstein M, Lawrence P, Xu C, McDowell NG, Knox RG, Koven C, Holm J, Rogers BM, *et al.* 2015. Taking off the training wheels: The properties of a dynamic vegetation model without climate envelopes, CLM4.5(ED). *Geoscientific Model Development* **8**: 3593–3619.

Fulton MR. **1991**. Adult Recruitment as a Function of Juvenile Growth Rate in Size-Structured Plant Populations. *Oikos* **62**: 102–105.

Gauthier S, Bernier P, Kuuluvainen T, Shvidenko AZ, Schepaschenko DG. **2015**. Boreal forest health and global change. *Science* **349**: 819–822.

Girardin CAJ, Malhi Y, Doughty CE, Metcalfe DB, Meir P, del Aguila-Pasquel J, Araujo-Murakami A, da Costa ACL, Silva-Espejo JE, Farfán Amézquita F, *et al.* 2016. Seasonal trends of Amazonian rainforest phenology, net primary productivity, and carbon allocation. *Global Biogeochemical Cycles* **30**: 700–715.

Greene DF, Zasada JC, Sirois L, Kneeshaw D, Morin H, Charron I, Simard MJ. **1999**. A review of the regeneration dynamics of North American boreal forest tree species. *Canadian Journal of Botany/Revue Canadienne de Botanique* **29**: 824–839.

Grubb PJ. **1977**. the Maintenance of Species-Richness in Plant Communities: the Importance of the Regeneration Niche. *Biological Reviews* **52**: 107–145.

Hacket-Pain AJ, Ascoli D, Vacchiano G, Biondi F, Cavin L, Conedera M, Drobyshev I, Liñán ID, Friend AD, Grabner M, *et al.* 2018. Climatically controlled reproduction drives interannual growth variability in a temperate tree species. *Ecology Letters* **21**: 1833–1844.

Hänninen H, Kramer K, Tanino K, Zhang R, Wu J, Fu YH. **2019**. Experiments Are Necessary in Process-Based Tree Phenology Modelling. *Trends in Plant Science* **24**: 199–209.

Harper JL. **1977**. *Population biology of plants*. London, UK: Academic Press.

Harrison RD, Tan S, Plotkin JB, Slik F, Detto M, Brenes T, Itoh A, Davies SJ. **2013**. Consequences of defaunation for a tropical tree community. *Ecology letters* **16**: 687–94.

Hartig F, Dyke J, Hickler T, Higgins SI, O'Hara RB, Scheiter S, Huth A. **2012**. Connecting dynamic vegetation models to data – an inverse perspective. *Journal of Biogeography* **39**: 2240–2252.

Haverd V, Smith B, Nieradzik LP, Briggs PR. **2014**. A stand-alone tree demography and landscape structure module for Earth system models: Integration with inventory data from temperate and boreal forests. *Biogeosciences* **11**: 4039–4055.

Hickler T, Vohland K, Feehan J, Miller PA, Smith B, Costa L, Giesecke T, Fronzek S, Carter TR, Cramer W, *et al.* 2012. Projecting the future distribution of European potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model. *Global Ecology and Biogeography* **21**: 50–63.

Holm JA, Knox RG, Zhu Q, Fisher RA, Koven CD, Nogueira Lima AJ, Riley WJ, Longo M, Negrón-Juárez RI, de Araujo AC, *et al.* 2020. The Central Amazon Biomass Sink Under Current and Future Atmospheric CO2: Predictions From Big-Leaf and Demographic Vegetation Models. *Journal of Geophysical Research: Biogeosciences* **125**: e2019JG005500.

Hooper E, Legendre P, Condit R. **2005**. Barriers to forest regeneration of deforested and abandoned land in Panama. *Journal of Applied Ecology* **42**: 1165–1174.

Hunter GG, Douglas MH. **1984**. Spread of exotic conifers on South Island rangelands. *New Zealand Journal of Forestry* **29**: 78–96.

Ibáñez I, Clark JS, LaDeau S, Hille Ris Lambers J, Ibanez I, Clark JS, LaDeau S, Hille Ris Lambers J. **2007**. Exploiting temporal variability to understand tree recruitment response to climate change. *Ecological Monographs* **77**: 163–177.

Joetzjer E, Maignan F, Chave J, Goll D, Poulter B, Barichivich J, Maréchaux I, Luyssaert S, Guimberteau M, Naudts K, *et al.* 2018. The importance of tree demography and root water uptake for modelling the carbon and water cycles of Amazonia. *Biogeosciences Discussions* **2018**: 1–33.

Johnstone JF, Chapin FS. **2006**. Effects of Soil Burn Severity on Post-Fire Tree Recruitment in Boreal Forest. *Ecosystems* **9**: 14–31.

Johnstone JF, Allen CD, Franklin JF, Frelich LE, Harvey BJ, Higuera PE, Mack MC, Meentemeyer RK, Metz MR, Perry GL, *et al.* 2016. Changing disturbance regimes, ecological memory, and forest resilience. *Frontiers in Ecology and the Environment* **14**: 369–378.

Journé V, Barnagaud J, Bernard C, Crochet P, Morin X. **2020**. Correlative climatic niche models predict real and virtual species distributions equally well. *Ecology* **101**: e02912.

Kattge, J., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J., *et al.* 2011. TRY - a global database of plant traits. *Global Change Biology* **17**:2905-2935.

Kelley DI, Harrison SP, Prentice IC. **2014**. Improved simulation of fire–vegetation interactions in the Land surface Processes and eXchanges dynamic global vegetation model (LPX-Mv1). *Geoscientific Model Development* **7**: 2411–2433.

Kobe RK. **1999**. Light Gradient Partitioning among Tropical Tree Species through Differential Seedling Mortality and Growth. *Ecology* **80**: 187–201.

Körner C. **2021**. Alpine treelines. In: Körner C, ed. Alpine Plant Life: Functional Plant Ecology of High Mountain Ecosystems. Cham: Springer International Publishing, 141–173.

Kueppers LM, Conlisk E, Castanha C, Moyes AB, Germino MJ, de Valpine P, Torn MS, Mitton JB. **2017**. Warming and provenance limit tree recruitment across and beyond the elevation range of subalpine forest. *Global Change Biology* **23**: 2383–2395.

Kulakowski D, Matthews C, Jarvis D, Veblen TT. **2013**. Compounded disturbances in sub-alpine forests in western Colorado favour future dominance by quaking aspen (*Populus tremuloides*). *Journal of Vegetation Science* **24**: 168–176.

Kullman L. **2001**. 20th Century climate warming and tree-limit rise in the Southern Scandes of Sweden. *Ambio* **30**: 72–80.

Lamont BB, Pausas JG, He T, Witkowski ETF, Hanley ME. **2020**. Fire as a Selective Agent for both Serotiny and Nonserotiny Over Space and Time. *Critical Reviews in Plant Sciences* **39**: 140–172.

Langan L, Higgins SI, Scheiter S. **2017**. Climate-biomes, pedo-biomes or pyro-biomes: which world view explains the tropical forest–savanna boundary in South America? *Journal of Biogeography* **44**: 2319–2330.

Lantz TC, Moffat ND, Fraser RH, Walker X. **2019**. Reproductive limitation mediates the response of white spruce (*Picea glauca*) to climate warming across the forest–tundra ecotone. *Arctic Science* **5**: 167–184.

Lawrence DM, Fisher RA, Koven CD, Oleson KW, Swenson SC, Bonan G, Collier N, Ghimire B, van Kampenhout L, Kennedy D, *et al.* **2019**. The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. *Journal of Advances in Modeling Earth Systems* **11**: 4245–4287.

Lehsten V, Mischurow M, Lindström E, Lehsten D, Lischke H, Lindstrom E, Lehsten D, Lischke H. **2019**. LPJ-GM 1.0: simulating migration efficiently in a dynamic vegetation model. *Geoscientific Model Development* **12**: 893–908.

Levy PE, Lucas ME, McKay HM, Escobar-Gutierrez AJ, Rey A. **2000**. Testing a process-based model of tree seedling growth by manipulating [CO_2] and nutrient uptake. *Tree Physiology* **20**: 993–1005.

Lipoma ML, Fortunato V, Enrico L, Díaz S. **2020**. Where does the forest come back from? Soil and litter seed banks and the juvenile bank as sources of vegetation resilience in a semiarid Neotropical forest. *Journal of Vegetation Science* **31**: 1017–1027.

Lischke H, Zimmermann NE, Bolliger J, Rickebusch S, Löffler TJ. **2006**. TreeMig: A forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. *Ecological Modelling* **199**: 409–420.

Lloret F, Escudero A, Iriondo JM, Martínez-Vilalta J, Valladares F. **2012**. Extreme climatic events and vegetation: The role of stabilizing processes. *Global Change Biology* **18**: 797–805.

Longo M, Knox RG, Medvigy DM, Levine NM, Dietze MC, Kim Y, Swann ALS, Zhang K, Rollinson CR, Bras RL, *et al.* 2019a. 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**: 4309–4346.

Longo M, Knox R, Medvigy DM, Levine NM, Dietze M, Swann ALS, Zhang K, Rollinson C, Brugnera M di P e, Scott D, *et al.* 2019b. Ecosystem Demography Model, version 2.2 (ED-2.2). https://doi.org/10.5281/zenodo.3365659 [accessed 28 February, 2022].

Loranger H, Zotz G, Bader MY. **2016**. Early establishment of trees at the alpine treeline: idiosyncratic species responses to temperature-moisture interactions. *AoB Plants* **8**: plw053.

Malhi Y, Doughty C, Galbraith D. **2011**. The allocation of ecosystem net primary productivity in tropical forests. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**: 3225–3245.

Malhi Y, Girardin C, Metcalfe DB, Doughty CE, Aragão LEOC, Rifai SW, Oliveras I, Shenkin A, Aguirre-Gutiérrez J, Dahlsjö CAL, *et al.* 2021. The Global Ecosystems Monitoring network: Monitoring ecosystem productivity and carbon cycling across the tropics. *Biological Conservation* **253**: 108889.

Martínez-Vilalta J, Lloret F. **2016**. Drought-induced vegetation shifts in terrestrial ecosystems: The key role of regeneration dynamics. *Global and Planetary Change* **144**: 94–108.

Martens C, Hickler T, Davis-Reddy C, Engelbrecht F, Higgins SI, von Maltitz GP, Midgley GF, Pfeiffer M, Scheiter S. **2021**. Large uncertainties in future biome changes in Africa call for flexible climate adaptation strategies. *Global Change Biology* **27**: 340–358.

Martins AM, Engel VL. **2007**. Soil seed banks in tropical forest fragments with different disturbance histories in southeastern Brazil. *Ecological Engineering* **31**: 165–174.

McDowell NG, Allen CD, Anderson-Teixeira K, Aukema BH, Bond-Lamberty B, Chini L, Clark JS, Dietze M, Grossiord C, Hanbury-Brown A, *et al.* 2020. Pervasive shifts in forest dynamics in a changing world. *Science* **368**: eaaz9463.

McGill BJ, Enquist BJ, Weiher E, Westoby M. **2006**. Rebuilding community ecology from functional traits. *Trends in Ecology and Evolution* **21**: 178–185.

Mclaughlin BC, Zavaleta ES. **2012**. Predicting species responses to climate change: Demography and climate microrefugia in California valley oak (*Quercus lobata*). *Global Change Biology* **18**: 2301–2312.

Medvigy D, Wofsy SC, Munger JW, Hollinger DY, Moorcroft PR. **2009**. Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. *Journal of Geophysical Research: Biogeosciences* **114**: 1–21.

Metzger S, Ayres E, Durden D, Florian C, Lee R, Lunch C, Luo H, Pingintha-Durden N, Roberti JA, SanClements M, *et al.* 2019. From neon field sites to data portal: A community resource for surface-atmosphere research comes online. *Bulletin of the American Meteorological Society* **100**: 2305–2325.

Miller KM, McGill BJ. **2018**. Land use and life history limit migration capacity of eastern tree species. *Global Ecology and Biogeography* **27**: 57–67.

Mladenoff DJ. **2004**. LANDIS and forest landscape models. *Ecological Modelling* **180**: 7–19.

Mokany K, Prasad S, Westcott DA. **2014**. Loss of frugivore seed dispersal services under climate change. *Nature Communications* **5**: 1–7.

Moorcroft PR, Hurtt GC, Pacala SW. **2001**. A method for scaling vegetation dynamics: The ecosystem demography model (ED). *Ecological Monographs* **71**: 557–586.

Murphy HT, VanDerWal J, Lovett-Doust J. **2010**. Signatures of range expansion and erosion in eastern North American trees. *Ecology Letters* **13**: 1233–1244.

Nagel TA, Taylor AH. **2005**. Fire and Persistence of Montane Chaparral in Mixed Conifer Forest Landscapes in the Northern Sierra Nevada, Lake Tahoe Basin, California, USA. *The Journal of the Torrey Botanical Society* **132**: 442–457.

Nathan R, Horvitz N, He Y, Kuparinen A, Schurr FM, Katul GG. **2011**. Spread of North American wind-dispersed trees in future environments. *Ecology Letters* **14**: 211–219.

Nathan R, Schurr FM, Spiegel O, Steinitz O, Trakhtenbrot A, Tsoar A. **2008**. Mechanisms of long-distance seed dispersal. *Trends in Ecology & Evolution* **23**: 638–647.

Neilson RP, Pitelka LF, Solomon AM, Nathan R, Midgley GF, Fragoso JMV, Lischke H, Thompson K. **2005**. Forecasting regional to global plant migration in response to climate change. *BioScience* **55**: 749–759.

Nolan RH, Rahmani S, Samson SA, Simpson-Southward HM, Boer MM, Bradstock RA. **2020**. Bark attributes determine variation in fire resistance in resprouting tree species. *Forest Ecology and Management* **474**: 118385.

Ordonez A, Williams JW. **2013**. Climatic and biotic velocities for woody taxa distributions over the last 16 000 years in eastern North America. *Ecology Letters* **16**: 773–781.

Pausas JG, Keeley JE. **2014**. Evolutionary ecology of resprouting and seeding in fire-prone ecosystems. *New Phytologist* **204**: 55–65.

Pausas JG, Pausas JG, Pratt RB, Keeley JE, Jacobsen AL, Ramirez AR, Vilagrosa A, Paula S, Kaneakua-pia IN, Davis SD. **2015**. Towards understanding resprouting at the global scale. *New Phytologist* **209**: 1–10.

Ponge JF, André J, Zackrisson O, Bernier N, Nilsson MC, Gallet C. **1998**. The forest regeneration puzzle: Biological mechanisms in humus layer and forest vegetation dynamics. *BioScience* **48**: 523–530.

Powell TL, Koven CD, Johnson DJ, Faybishenko B, Fisher RA, Knox RG, McDowell NG, Condit R, Hubbell SP, Wright SJ, *et al.* 2018. Variation in hydroclimate sustains tropical forest biomass and promotes functional diversity. *New Phytologist* **219**: 932–946.

Price DT, Zimmermann NE, Van Der Meer PJ, Lexer MJ, Leadley P, Jorritsma ITMM, Schaber J, Clark DF, Lasch P, Mcnulty S, *et al.* 2001. Regeneration in gap models: Priority issues for studying forest responses to climate change. *Climatic Change* **51**: 475–508.

Qiu T, Aravena M-C, Andrus R, Ascoli D, Bergeron Y, Berretti R, Bogdziewicz M, Boivin T, Bonal R, Caignard T, *et al.* 2021. Is there tree senescence? The fecundity evidence. *Proceedings of the National Academy of Sciences* **118**: e2106130118.

Roland CA, Schmidt JH, Johnstone JF. **2014**. Climate sensitivity of reproduction in a mast-seeding boreal conifer across its distributional range from lowland to treeline forests. *Oecologia* **174**: 665–677.

Rüger N, Comita LS, Condit R, Purves D, Rosenbaum B, Visser MD, Wright SJ, Wirth C. **2018**. Beyond the fast–slow continuum: demographic dimensions structuring a tropical tree community. *Ecology Letters* **21**: 1075–1084.

Rüger N, Condit R, Dent DH, Dewalt SJ, Hubbell SP, Lichstein JW, Lopez OR, Wirth C, Farrior CE. **2020**. Demographic trade-offs predict tropical forest dynamics. *Science* **368**: 165–168.

Saha S, Howe HF. **2003**. Species composition and fire in a dry deciduous forest. *Ecology* **84**: 3118–3123.

Sakschewski B, Von Bloh W, Boit A, Poorter L, Peña-Claros M, Heinke J, Joshi J, Thonicke K. **2016**. Resilience of Amazon forests emerges from plant trait diversity. *Nature Climate Change* **6**: 1032–1036.

Sato H, Ise T. **2012**. Effect of plant dynamic processes on African vegetation responses to climate change: Analysis using the spatially explicit individual-based dynamic global vegetation model (SEIB-DGVM). *Journal of Geophysical Research: Biogeosciences* **117**: G03017.

Sato H, Itoh A, Kohyama T. **2007**. SEIB-DGVM: A new Dynamic Global Vegetation Model using a spatially explicit individual-based approach. *Ecological Modelling* **200**: 279–307.

Scheiter S, Higgins SI. 2009. Impacts of climate change on the vegetation of Africa: an adaptive dynamic vegetation modelling approach. *Global Change Biology* **15**: 2224–2246**.**

Scheiter S, Langan L, Higgins SI. **2013**. Next-generation dynamic global vegetation models: learning from community ecology. *New Phytologist* **198**: 957–969.

Scheller RM, Domingo JB, Sturtevant BR, Williams JS, Rudy A, Gustafson EJ, Mladenoff DJ. **2007**. Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. *Ecological Modelling* **1**: 409–419.

Schoennagel T, Turner MG, Romme WH. **2009**. The Influence of Fire Interval and Serotiny on Postfire Lodgepole Pine Density in Yellowstone National Park. *America* **84**: 2967–2978.

Seidl R, Thom D, Kautz M, Martin-Benito D, Peltoniemi M, Vacchiano G, Wild J, Ascoli D, Petr M, Honkaniemi J, *et al.* 2017. Forest disturbances under climate change. *Nature Climate Change* **7:** 395–402.

Sirois L. **2000**. Spatiotemporal variation in black spruce cone and seed crops along a boreal forest - Tree line transect. *Canadian Journal of Forest Research* **30**: 900–909.

Smith B, Prentice IC, Sykes MT. **2001**. Representation of vegetation dynamics in the modelling of terrestrial ecosystems: Comparing two contrasting approaches within European climate space. *Global Ecology and Biogeography* **10**: 621–637.

Snell RS. **2014**. Simulating long-distance seed dispersal in a dynamic vegetation model. *Global Ecology and Biogeography* **23**: 89–98.

Snell RS, Huth A, Nabel JEMS, Bocedi G, Travis JMJ, Gravel D, Bugmann H, Gutiérrez AG, Hickler T, Higgins SI, *et al.* 2014. Using dynamic vegetation models to simulate plant range shifts. *Ecography* **37**: 1184–1197.

Solomon AM, Kirilenko AP. **1997**. Climate Change and Terrestrial Biomass: What if Trees do not Migrate? *Global Ecology and Biogeography Letters* **6**: 139–148.

Stevens-Rumann CS, Morgan P. **2019**. Tree regeneration following wildfires in the western US: a review. *Fire Ecology* **15**: 15.

Swenson NG. **2012**. Temporal turnover in the composition of tropical tree communities: functional determinism and phylogenetic stochasticity. **93**: 490–499.

Sykes MT, Prentice IC, Cramer W. **1996**. A bioclimatic model for the potential distributions of north European tree species under present and future climates. *Journal of Biogeography* **23**: 203–233.

Tepley AJ, Thompson JR, Epstein HE, Anderson-Teixeira KJ. **2017**. Vulnerability to forest loss through altered postfire recovery dynamics in a warming climate in the Klamath Mountains. *Global Change Biology* **23**: 4117–4132.

Thomas SC. **2011**. Age-Related Changes in Tree Growth and Functional Biology: The Role of Reproduction. In: Meinzer FC, Lachenbruch B, Dawson TE, eds. Size- and Age-Related Changes in Tree Structure and Function. Dordrecht: Springer Netherlands, 33–64.

Del Tredici P. **2001**. Sprouting in temperate trees: A morphological and ecological review. *Botanical Review* **67**: 121–140.

Trumbore S, Brando P, Hartmann H. **2015**. Forest health and global change. *Science* **349**: 814–8.

Turnbull LA, Crawley MJ, Rees M. **2000**. Are plant populations seed-limited? A review of seed sowing experiments. *Oikos* **88**: 225–238.

Turner MG, Romme WH, Gardner RH. **1999**. Prefire heterogeneity, fire severity, and early postfire plant reestablishment in subalpine forests of Yellowstone National Park, Wyoming. *International Journal of Wildland Fire* **9**: 21.

[UN-CBD] United Nations Convention on Biological Diversity. **2006**. Definitions. [WWW document] URL https://www.cbd.int/forest/definitions.shtml [accessed 13 Oct 2021].

Uriarte M, Clark JS, Zimmerman JK, Comita LS, Forero-Montaña J, Thompson J. **2012**. Multidimensional trade-offs in species responses to disturbance: implications for diversity in a subtropical forest. *Ecology* **93**: 191–205.

Urza AK, Sibold JS. **2017**. Climate and seed availability initiate alternate post-fire trajectories in a lower subalpine forest (F Gilliam, Ed.). *Journal of Vegetation Science* **28**: 43–56.

Burns RM, Honkala BH. **1990**. *Silvics of North America: Volume 1. Conifers*. Washington D.C., USA: United States Department of Agriculture.

Van Oijen M, Rougier J, Smith R. **2005**. Bayesian calibration of process-based forest models: bridging the gap between models and data. *Tree Physiology* **25**: 915–927.

Walck JL, Hidayati SN, Dixon KW, Thompson K, Poschlod P. **2011**. Climate change and plant regeneration from seed. *Global Change Biology* **17**: 2145–2161.

Walden LL, Fontaine JB, Ruthrof KX, Matusick G, Harper RJ, Hardy GESJ. 2019. Carbon consequences of drought differ in forests that resprout. *Global Change Biology* **25**: 1653–1664.

Weng ES, Malyshev S, Lichstein JW, Farrior CE, Dybzinski R, Zhang T, Shevliakova E, Pacala SW. **2015**. Scaling from individual trees to forests in an Earth system modeling framework using a mathematically tractable model of height-structured competition. *Biogeosciences* **12**: 2655–2694.

Wenk EH, Abramowicz K, Westoby M, Falster DS. **2017**. Coordinated shifts in allocation among reproductive tissues across 14 coexisting plant species. *BioRxiv*. doi: 0.1101/141473

Wenk EH, Falster DS. **2015**. Quantifying and understanding reproductive allocation schedules in plants. *Ecology and Evolution* **5**: 5521–5538.

Wramneby A, Smith B, Zaehle S, Sykes MT. **2008**. Parameter uncertainties in the modelling of vegetation dynamics-Effects on tree community structure and ecosystem functioning in European forest biomes. *Ecological Modelling* **216**: 277–290.

Wramneby A, Smith B, Samuelsson P. **2010**. Hot spots of vegetation-climate feedbacks under future greenhouse forcing in Europe. *Journal of Geophysical Research* **115**: D21119.

Wright SJ, Calderón O. **2006**. Seasonal, El Nino and longer term changes in flower and seed production in a moist tropical forest. *Ecology Letters* **9**: 35–44.

Wright SJ, Hernandéz A, Condit R. **2007**. The bushmeat harvest alters seedling banks by favoring lianas, large seeds, and seeds dispersed by bats, birds, and wind. *Biotropica* **39**: 363–371.

Young DJN, Werner CM, Welch KR, Young TP, Safford HD, Latimer AM. **2019**. Post-fire forest regeneration shows limited climate tracking and potential for drought-induced type conversion. *Ecology* **100**: 1–13.

Zalamea P-C, Turner BL, Winter K, Jones FA, Sarmiento C, Dalling JW. **2016**. Seedling growth responses to phosphorus reflect adult distribution patterns of tropical trees. *New Phytologist* **212**: 400–408.

Gallego Zamorano J, Hokkanen T, Lehikoinen A. **2018**. Climate-driven synchrony in seed production of masting deciduous and conifer tree species. *Journal of Plant Ecology* **11**: 180–188.

Zeppel MJB, Harrison SP, Adams HD, Kelley DI, Li G, Tissue DT, Dawson TE, Fensham R, Medlyn BE, Palmer A, et al. 2015. Drought and resprouting plants. *New Phytologist* **206**: 583–589.

Supporting Information

Table S1. The representation of key forest regeneration processes among leading Vegetation Demographic Models.

Table S2. Estimates of ecosystem-level mean annual reproductive litter flux (R), leaf litter flux (L), aboveground net primary productivity (ANPP), NPP, and benchmarks R/L, R/ANPP, and R/NPP for forest sites.

Table S3. Metadata for Table S2.

Methods S1. Synthesis methods for Table S2.

Methods S2. Data preparation for model-data comparison at Barro Colorado Island, Panama.

Methods S3. Synthesis methods for Table S4.

Figure S1. Correlation between R/L and R/ANPP

Figure S2. Correlation between L and NPP

Table S4. Estimates of area-based, species-level recruitment rates into the 1 cm size class for four forest dynamics plots in the CTFS-ForestGeo network.

Table S5. Metadata for Table S4.

Figure 1. Forest regeneration is a series of environmentally sensitive processes, each represented by a sieve. Environmental variables act as filters (depicted as varying sieve mesh sizes, with the small red arrow setting the mesh size) that impose constraints on each process. Each plant functional type (PFT; colored seeds) is uniquely sensitive to environmental constraints, which may limit the number of individuals and PFTs progressing to the next stage. Only the PFTs that make it through all sieves will recruit into post-seedling size classes, which along with growth and mortality, determines the structure, distribution, and functional composition of the mature forest. Thick arrows represent pathways addressed in this review. “NPP” is net primary productivity.

Figure 2. Obligate postfire seeders occur where fire return intervals are between maturation age and plant longevity. Changes to the fire regime risk regeneration failure through “immaturity risk” (a) and “senescence risk” (b). Non-serotinous species (i.e. where seeds are stored in a seed bank) have a wider window of viability because they can wait in the seed bank for suitable conditions even if plants reduce seed production. “SBL” is seed bank longevity. Reproduced from Pausas & Keeley (2014).

Figure 3. Experimental warming (“Heat” treatment) reduces seedling germination and initial survival (mean ± SE) beyond the local elevation limit of lodgepole pine (a), an upper montane-lower subalpine species, preventing population establishment (b). Supplemental summer water (“Heat-water” treatment) improves first-year seedling establishment (a), which, along with greater survival of older seedlings, facilitates population establishment (b). The black line in panel b is the “Control” group; all other line colors match panel a. Remade from Conlisk *et al.* (2017) without error bars on panel b.

Figure 4. Differential seedling survivorship under manipulated drought (a) is well correlated with compositional turnover over an aridity gradient on the isthmus of Panama (b). In panel b, each point is a species. The “Drought sensitivity” axis quantifies survivorship under manipulated drought compared to control conditions (higher values are more drought sensitive). The y-axis is each species’ occurrence probability in dry sites as a fraction of its occurrence probability in relatively wet sites, log transformed (low value indicates preference for wet areas). GAR = *Garcinia intermedia*; POU =

Pouteria unilocularis; SWA = *Swartzia simplex*; VIR = *Virola surinamensis.* Panel a is remade from Engelbrecht *et al*. (2005); panel b is remade from Engelbrecht *et al.* (2007).

Figure 5. a) Correlation between observed reproductive litter flux as a fraction of net primary productivity (R/NPP) and leaf litter flux (R/L; all fluxes in g C m^{-2} yr^{-1}) at 61 forest sites (colors indicate Whittaker biomes as in panel b; shaded area shows the 95 % confidence interval), and b) the magnitude of R/L (shown by point size) and the distribution of R/L observations across Whittaker biomes (n = 550); "N sites" shows the approximate number of distinct measures of R/L at mean annual temperature and precipitation coordinates.

Figure 6. FATES predictions at Barro Colorado Island (circles) of a) reproductive litter flux (R; g C m^{-2} yr^{-1}) as a fraction of leaf litter flux (R/L), b) R as a fraction of above-ground net primary productivity (R/ANPP), and c) recruitment rates into the 1 cm size class. Each model prediction shows the mean across 18 simulation years run with recycled, observed meteorology (2003-2016), initialized with observed stand structure, and using a prescribed value of reproductive allocation (RA; ranging from 0.009-0.41). The standard deviation of inter-annual variation is smaller than the symbol size for most circles. Triangles show observations (mean and standard deviation among years) at the BCI Forest Dynamics Plot.

Figure 7. Simulated mortality rates from ED2-hydro ("Model") are positively biased in the smaller size classes at Barro Colorado Island (BCI), Panama, indicating compensating errors for hyperactive recruitment (Powell *et al.*, 2018). "Obs" show field observations.

Table 1. The representation of key forest regeneration processes among seven leading Vegetation Demographic Models.

Model	CLM(ED) / FATES	ED2	LPJ-GUESS (cohort mode)	LPJ-DISP (version of LPJ-GUESS)	LPJ-GM (version of LPJ-GUESS)	LM3-PPA	SEIB-DGVM	SEIB-DGVM-2012 (version of SEIB-DGVM)	aDGVM/aDGVM2
Reproductive Allocation & Seed Production	Cohorts allocate a **fixed fraction** ($F_{repro,pft}$) **of carbon for growth and reproduction** ($C_{g+r,pft}$) to reproduction ($C_{repro,pft}$). Cohorts above a **size threshold** can allocate additional $C_{g+r,pft}$ to reproduction. This is added to external **seed rain** (SR_{pft}). $C_{repro,pft} = (C_{g+r,pft})(F_{repro,pft}) + SR_{pft}$	Cohorts above a size threshold allocate a **fixed fraction** ($F_{repro,pft}$) **of carbon for growth and reproduction** (C_{g+r}) to reproduction (C_{repro}). This is added to external **seed rain** (SR_{pft}). $C_{repro,pft} = (C_{g+r,pft})(F_{repro,pft}) + SR_{pft}$ More recent code developments in ED2 give the option to have $F_{repro,pft}$ vary as a function of size (see Table S1).	Cohorts allocate **a fixed fraction** ($F_{repro,pft}$) **of** $\mathbf{NPP_{pft}}$ to reproduction ($C_{repro,pft}$). $C_{repro,pft} = (NPP_{pft})(F_{repro,pft})$	Same as LPJ-GUESS	Reproductive allocation is not explicitly represented, but the number of **seeds produced** (S_{pft}) **is** calculated as **a function of LAI and a maximum fecundity parameter**$_{pft}$.	**Trees in the canopy allocate a fixed fraction** ($F_{repro,pft}$) **of carbon for growth and reproduction** (C_{g+r}) to reproduction (C_{repro}). $C_{repro,pft} = (C_{g+r,pft})(F_{repro,pft})$	Trees above a **biomass threshold** allocate **a fixed fraction** ($F_{repro,pft}$) **of carbon for growth and reproduction** ($C_{g+r,pft}$) to reproduction ($C_{repro,pft}$). $C_{repro,pft} = (C_{g+r,pft})(F_{repro,pft})$ Reproductive carbon is sent to litter and does not influence recruitment rates.	Same as SEIB-DGVM	Trees above an **age threshold** produce seeds. The number of seeds produced (S_{pft}) by each tree is a **function of carbon for growth and reproduction** ($\mathbf{C_{g+r,pft}}$) **on the first day of each month**, i (to represent stochasticity of environmentally sensitive seed production), and the mass of carbon in one seed ($B_{seed,pft}$). $C_{repro} = C_{g+r,i,pft} / B_{seed,pft}$

Model	CLM(ED) / FATES	ED2	LPJ-GUESS (cohort mode)	LPJ-DISP (version of LPJ-GUESS)	LPJ-GM (version of LPJ-GUESS)	LM3-PPA	SEIB-DGVM	SEIB-DGVM-2012 (version of SEIB-DGVM)	aDGVM/aDGVM2
Dispersal	C_{repro} mixes evenly among patches within a grid cell. Additional "**seed rain$_{pft}$**" can be added to all patches in all grid cells. Therefore, **inter- and intra-grid cell dispersal is assumed to be non-limiting** despite no transfer of seed mass between grid cells.	A **prescribed fraction** of C_{repro} (i.e. "seed") **may leave its origin patch**. Additional "**seed rain$_{pft}$**" is added to all PFTs in all grid cells. Therefore, **inter-grid cell dispersal is implicitly assumed to be non-limiting** despite no transfer of seed mass between grid cells.	$C_{repro,pft}$ mixes evenly among patches within a grid cell. Therefore **intra-grid cell dispersal is implicitly assumed to be non-limiting**. PFTs can establish in any grid cells in which they are, or become, climatically eligible. Therefore **inter-grid cell dispersal is implicitly assumed to be non-limiting** despite no transfer of seed mass between grid cells.	**Dispersal is explicitly represented. Dispersal kernels$_{pft}$** and the number of patches containing a PFT are used to approximate inter- and intra-grid cell transfer of C_{repro} (see Table S1 for more details).	**Dispersal is explicitly represented** between 1km^2 patches. Dispersal rates between cells are calculated from **dispersal kernels$_{pft}$** and seed production rates of neighboring cells. All seeds share a common seed bank within cells, so dispersal is assumed to be unlimited within grid cells (see Table S1 for more details).	Seeds mix evenly within a grid cell, therefore **intra-grid cell dispersal is implicitly assumed to be non-limiting**. There is **no inter-grid cell dispersal**.	PFTs can establish in any grid cell for which they are or become climatically eligible, therefore, **inter-grid cell dispersal is implicitly assumed to be non-limiting** despite no transfer of seed mass between grid cells. **Intra-grid cell dispersal is implicitly assumed to be non-limiting** because one patch per grid cell is simulated.	Climatically eligible PFTs can establish in 10 % of neighboring grid cells. The frequency at which PFTs can jump between grid cells (i.e. **inter-grid cell dispersal) is tuned to paleo records**. **Intra-grid cell dispersal is implicitly assumed to be non-limiting**.	PFTs can establish in any grid cell, therefore **inter-grid cell dispersal is implicitly assumed to be non-limiting.** **Intra-grid cell dispersal is implicitly assumed to be non-limiting** because one patch per grid cell is simulated (similar to SEIB-DGVM).

Model	CLM(ED) / FATES	ED2	LPJ-GUESS (cohort mode)	LPJ-DISP (version of LPJ-GUESS)	LPJ-GM (version of LPJ-GUESS)	LM3-PPA	SEIB-DGVM	SEIB-DGVM-2012 (version of SEIB-DGVM)	aDGVM/aDGVM2
Seed Survival & Germination	**Seed bank is tracked** $C_{repro,pft}$ becomes the $Seed_{in,pft}$ flux to a seed bank ($Seeds_{pft}$) which is reduced by **prescribed fractions of seeds germinating** ($Seed_{germ,pft}$) **and decaying** ($Seed_{decay,pft}$) $dSeeds_{pft} / dt = Seed_{in,pft} - Seed_{decay,pft} - Seed_{germ,pft}$ There is an upper limit to the germination flux to avoid excessive dominance of a PFT (see Table S1).	**Seed bank is not explicitly represented.**	**Seed bank is not explicitly represented.**	Same as LPJ-GUESS	**Seed bank is tracked.** PFT-specific seed banks grow with inputs from dispersed and produced seeds that stay in their grid cell of origin. Seed banks are reduced by **prescribed fractions of seeds germinating and decaying** (similar to CLM(ED)/FATES).	**Seed bank is not explicitly represented.** $C_{repro,pft}$ is reduced by a **prescribed germination** rate, $P_{germ,pft}$ (see row "Seedling Growth & Survival / Recruitment", column "LM3-PPA").	**Seed bank is not explicitly represented.**	Same as SEIB-DGVM.	**Seed bank is tracked.** As in CLM(ED)/ FATES C_{repro} becomes $Seed_{in,pft}$ to a seed bank ($Seeds_{pft}$), which is reduced by **prescribed fractions of seeds germinating and decaying**. Seed **germination only occurs after 3 consecutive days with soil moisture of the upper soil layer at field capacity.** In aDGVM2 seed banks are represented as community trait pools where genetic crossover (between individuals with the same species label) and mutation can occur, giving rise to new trait combinations.

Model	CLM(ED) / FATES	ED2	LPJ-GUESS (cohort mode)	LPJ-DISP (version of LPJ-GUESS)	LPJ-GM (version of LPJ-GUESS)	LM3-PPA	SEIB-DGVM	SEIB-DGVM-2012 (version of SEIB-DGVM)	aDGVM/aDGVM2
Seedling Growth & Survival / Recruitment	Carbon germinating out of the seed bank ($Seed_{germ,pft}$) becomes carbon to make new recruits ($C_{recruit,pft}$) at a **rate proportional to the amount of carbon required to make 1 new recruit ($Z_{0,pft}$)** $N_{recruit,pft} = C_{recruit,pft} / Z_{0,pft}$	A **prescribed seedling mortality rate** ($M_{seedling}$) reduces $C_{repro,pft}$ such that the amount available for recruitment ($C_{recruit,pft}$) is $C_{recruit,pft} = (C_{repro,pft}) (1 - M_{seedling})$ The number of **new recruits is then calculated as in FATES/CLM(ED).**	**Bioclimatic envelopes** ($GDD_{min,pft}$, $T_c min_{,pft}$, $T_c max_{,pft}$), $PAR_{min,pft}$ and $fAWC_{pft}$ determine establishment eligibility. Recruitment rate is based on **maximum establishment rates** ($est_{max,pft}$), the **potential productivity at the forest floor** (f, a fraction of the maximum), and shade tolerance α_{pft}. $N_{recruit,pft} \sim Pois(ƛ_{pft})$ $ƛ_{pft} = \mu(f_{pft}) (est_{max,pft}) ((k_{reprod})(C_{repro,pft}) + k_{bgestab})$ $\mu(f_{pft}) = exp[\alpha_{pft}(1 - 1 / f^{pft})]$	Same as LPJ-GUESS	**Bioclimatic envelopes**, $PAR_{min,pft}$, and $ƛ_{pft}$ limit recruitment as in LPJ-GUESS. Additionally, a n **establishment probability** ($P_{est,pft}$) reduces recruitment rates. $P_{est,pft}$ is calculated as a **function of the number of seeds in the seed bank** ($Seeds_{pft}$), the prescribed germination probability ($P_{germ,pft}$), and p_x which takes the area of each patch into account such that $P_{est,pft} = (Seeds_{pft}) (p_x) (P_{germ,pft})$	$C_{repro,pft}$ is reduced by a **prescribed establishment probability** ($P_{est,pft}$) such that $C_{seedling,pft} = (C_{repro,pft}) (P_{germ,pft}) (P_{est,pft})$ New individuals emerge as seedlings (i.e., $Z_{0,pft}$ is set very small) **at a rate proportional to the amount of carbon required to make 1 new seedling ($Z_{0,pft}$)** $N_{seedling,pft} = C_{seedling,pft} / Z_{0,pft}$. **Seedling mortality is a function of size**.	**Bioclimatic envelopes** (GDD_{min}, $T_c min$, $T_c max$) and $PAR_{min,pft}$ determine establishment eligibility. Additionally, total precipitation in the current year (in mm) must exceed 20 times the annual mean temperature (in °C; Köppen, 1936). **Recruitment rates are assumed to be a function of space in the understory** (see Table S1) and a **prescribed establishment probability** ($P_{est,pft}$).	Same as SEIB-DGVM	The number of new seedling recruits is determined from **prescribed germination probability** ($P_{germ,pft}$) out of the seed bank ($Seeds_{pft}$), and a **stochastic fraction of seed available to germinate** ($P_{sprout,pft}$): $N_{seedling,pft} = (Seeds_{pft}) (P_{sprout,pft}) (P_{germ,pft})$ **Seedlings explicitly compete with grasses and taller trees,** increasing the chance of mortality. **Seedlings can be disproportionately killed by fire** due to a size-dependent "top kill" probability, but stems are not killed, allowing them to **resprout.**

Model	CLM(ED) / FATES	ED2	LPJ-GUESS (cohort mode)	LPJ-DISP (version of LPJ-GUESS)	LPJ-GM (version of LPJ-GUESS)	LM3-PPA	SEIB-DGVM	SEIB-DGVM-2012 (version of SEIB-DGVM)	aDGVM/aDGVM2
References	Fisher et al. (2015); FATES code (see Table S1 for specific lines of code)	Medvigy et al. (2009); ED2 code (see Table S1 for specific lines of code)	Smith et al. (2001); Fulton (1991); see Table S1 for more details	Snell (2014)	Lehsten et al. (2019)	Weng et al. (2015)	Sato et al. (2007)	Sato & Ise (2012)	Scheiter et al., 2009; Scheiter et al., 2020

; "PFT" indicates plant-functional-type-specific parameters; C_{g+r} = Carbon for growth and reproduction (typically net primary productivity (NPP) remaining after tissue replacement and allocation to storage); GDD_{min} = minimum number of growing degree days for reproduction; $T_c max$ = the maximum mean temperature of the coldest month (a chilling requirement); $T_c min$ = the minimum temperature of the coldest month required for establishment and survival; PAR_{min} = minimum photosynthetically active radiation at the forest floor required for establishment; fAWC = fraction of plant-available water holding capacity in the top soil layer; k_{reprod} and $k_{bgestab}$ are constants (see Supporting Information Table S1)

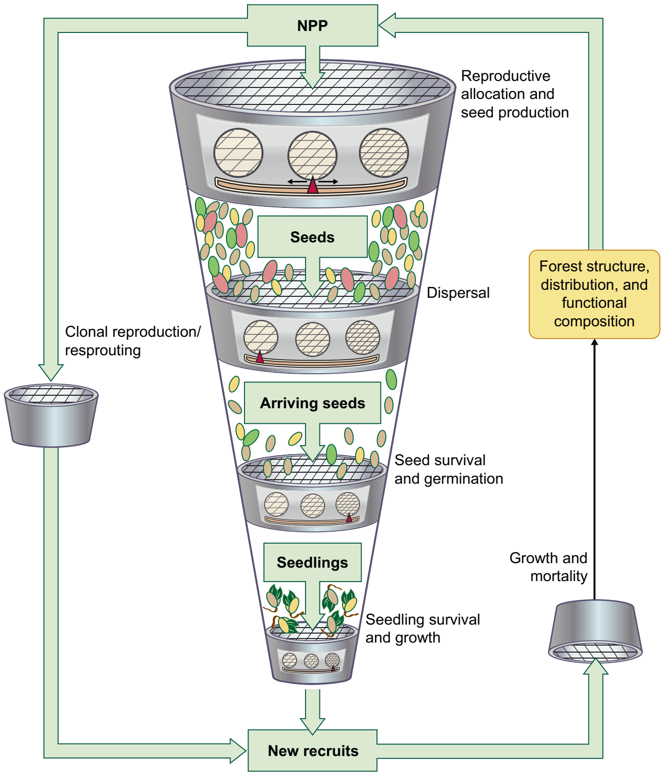

Figure 1

Tansley Review 36908

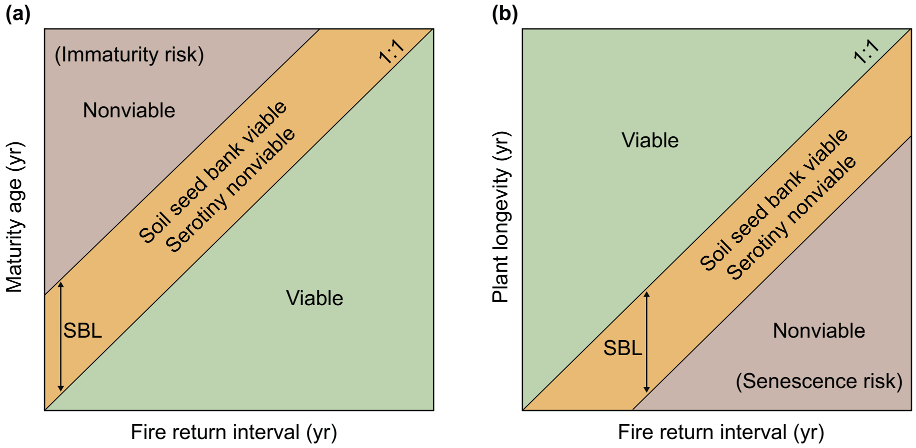

Figure 2

Tansley Review 36908

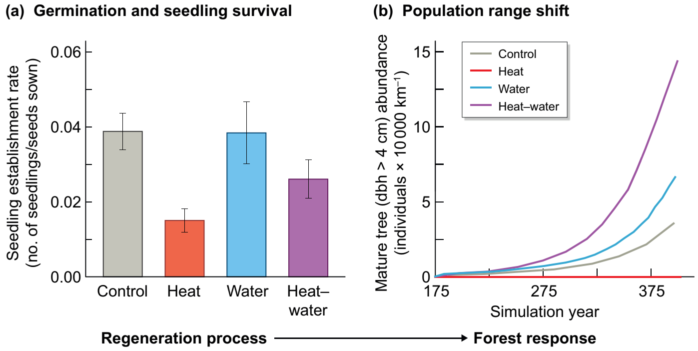

Figure 3

Tansley Review 36908

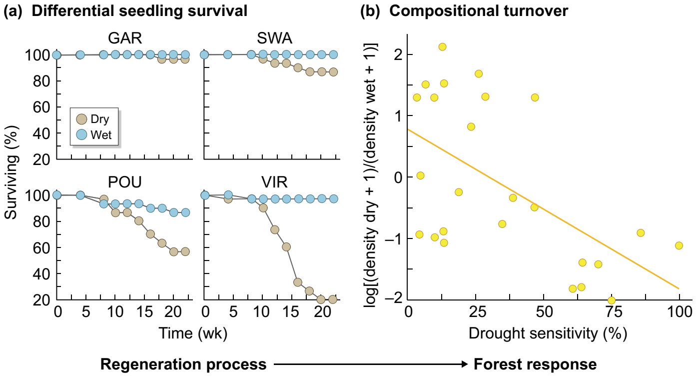

Figure 4

Tansley Review 36908

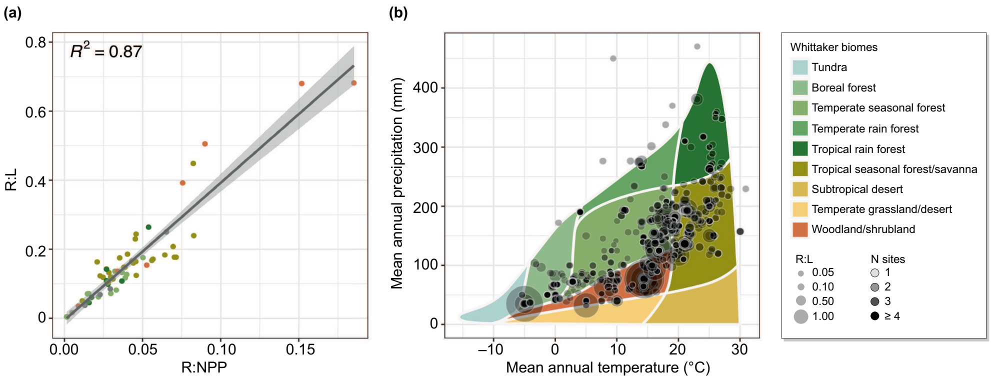

Figure 5

Tansley Review 36908

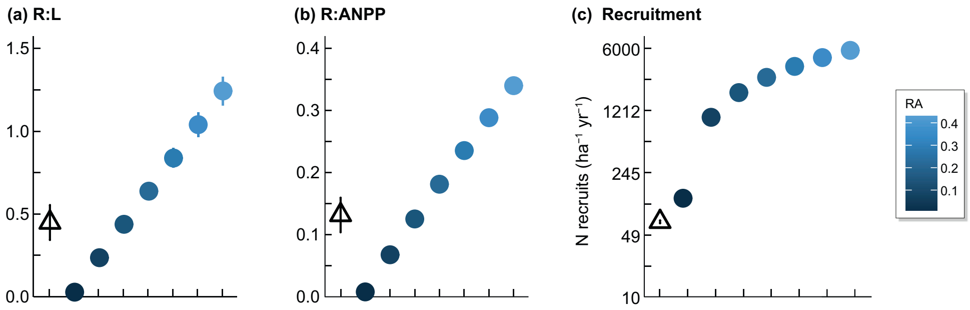

Figure 6

Tansley Review 36908

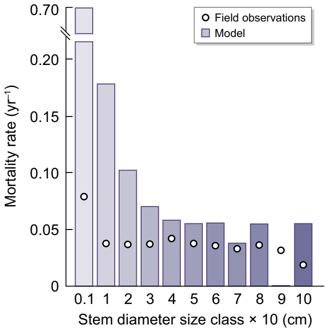

Figure 7

Tansley Review 36908