UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Predicting the future of forests under global change: the critical role of regeneration processes

Permalink

https://escholarship.org/uc/item/1g92859f

Author

Hanbury-Brown, Adam

Publication Date

2022

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Predicting the future of forests under global change: the critical role of regeneration processes

By

Adam R Hanbury-Brown

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Energy and Resources

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Lara Kueppers, Chair

Professor Jeffrey Chambers

Dr. Charles Koven

Spring 2022

Abstract

Predicting the future of forests under global change: the critical role of regeneration processes

by

Adam R Hanbury-Brown

Doctor of Philosophy in Energy and Resources

University of California, Berkeley

Professor Lara Kueppers, Chair

Climate change, land use, and intensifying natural disturbance regimes are affecting forests globally. Earth system models (ESMs) must predict forest responses to these global changes to simulate future climate, hydrology, and ecosystem dynamics. The Earth system modeling community is increasingly using vegetation demographic models (VDMs) to simulate vegetation dynamics in ESMs. VDMs explicitly represent tree growth, mortality and recruitment, enabling advances in the projection of forest vulnerability and resilience, as well as evaluation with field data. Predicting future forests requires understanding if, where, and how forest communities will be resilient to global change through regeneration processes such as allocation to reproduction, dispersal, and seedling survival. Simulation of regeneration processes has received far less attention than simulation of processes that affect mature tree growth and mortality in spite of its critical role in maintaining forest structure, facilitating turnover in forest composition over space and time, enabling recovery from disturbance, and regulating climate-driven range shifts. The role of forest regeneration processes is particularly important in California where intensifying fire regimes and climate change are reducing conifer seedling recruitment and causing concern that vegetation regime shifts will push conifer forests towards more shrub-dominated ecosystems.

This dissertation advances empirical understanding and modeling infrastructure for predicting how forest regeneration processes will mediate changes in forest distribution, structure, and functional composition (i.e. “forest responses”) in response to global change. It is structured around three papers, two of which were previously published (Chapters 1 and 2). Chapter 1 critically reviews how forest regeneration processes are currently represented in VDMs and makes recommendations for advancing model algorithms and parameterizations. This critical review finds that regeneration processes are not currently represented sufficiently to capture how regeneration will mediate changes in future forest distribution, structure, and functional composition. There is a need to improve parameter values and algorithms for reproductive allocation, dispersal, environmental filtering in the seedling layer, and tree regeneration strategies adapted to wind, fire, and anthropogenic disturbance regimes. These improvements will require synthesis of existing data, specific field data collection protocols, such as long-term forest

monitoring plots that identify individuals to species, and novel model algorithms compatible with global scale simulations.

Chapter 2 draws upon the insights and conclusions from Chapter 1 to develop a new modeling scheme, the Tree Recruitment Scheme (TRS), for simulating environmentally sensitive tree recruitment in VDMs. We evaluate the TRS by predicting tree recruitment for four tropical tree functional types under varying meteorology and canopy structure at Barro Colorado Island, Panama. By comparing TRS predictions to those of a current VDM, quantitative observations, and ecological expectations we find that it improves the magnitude and rank order of recruitment rates among four tropical tree plant functional types (PFTs). It also captures recruitment limitations in response to variable understory light, soil moisture, and changing precipitation regimes. These results indicate that adopting this framework will improve VDM capacity to predict PFT-specific tree recruitment in response to climate change, thereby improving predictions of future forest distribution, composition, and function. We implement the TRS in a leading VDM, the Functionally Assembled Terrestrial Ecosystem Simulator (FATES). Technical documentation for this implementation is provided in Appendix A.

Chapter 3 uses remote sensing and machine learning to analyze how changing climate and fire regimes in California are affecting long-term, post-fire conifer forest regeneration.. We quantify decadal-scale variation in post-fire conifer canopy recovery in 44 burned areas throughout California's Sierra Nevada, Klamath Mountains, and Interior Coast Ranges and use generalized additive models to identify key drivers of recovery outcomes. The work in this chapter shows that post-fire conifer canopy recovery is reduced after higher severity fire, in areas that re-burned within 32 years, and when mean annual precipitation in the first three years after fire falls below 1,000 mm yr^{-1}. This extends the spatio-temporal scope of prior work focused on shorter term post-fire seedling regeneration by providing insight into what drives more persistent changes in forest structure. It also provides a valuable dataset for benchmarking predictions of forest responses to changing fire regimes and supports prior hypotheses that increasingly intense fire regimes will negatively affect post-fire conifer regeneration, likely leading to an increase in the area dominated by shrubs.

Table of Contents

Acknowledgments....iii
Introduction....1
Background and motivation....1
Dissertation goals and structure....2
Chapter 1: Forest regeneration within Earth system models: current process representations and ways forward....4
Summary....4
I. Introduction....5
II. Forest regeneration processes critical to the Earth system....7
III. Forest regeneration in Vegetation Demographic Models....13
ED2, CLM(ED), and FATES....19
LPJ-GUESS....20
LM3-PPA....21
SEIB-DGVM....21
aDGVM....22
IV. Strengths and weaknesses of current regeneration algorithms....23
Reproductive Allocation and Seed Production....23
Dispersal....24
Seed Survival, Germination, and Resprouting....25
Seedling Survival and Growth....26
V. Priorities for advancing prediction....27
Reproductive Allocation and Seed Production....28
Dispersal....30
Seed Survival, Germination, and Resprouting....31
Seedling Survival and Growth....32
VI. Conclusion....33
Supplementary Information....35
Chapter 2: Simulating environmentally sensitive tree recruitment in vegetation demographic models....43
Summary....43
Introduction....43
Materials and Methods....45
Model description....45
Simulations at Barro Colorado Island, Panama....52
Parameterization for BCI....54
Results....57
Benchmarking recruitment at BCI....57
Sensitivity to variable light and soil moisture....58
Recruitment predictions under ENSO and WET precipitation scenarios....59
Parameter sensitivity....61
Discussion....62
Benchmarking recruitment at BCI....63

Sensitivity to variable light and moisture conditions....63
Responses to ENSO and WET precipitation scenarios....64
Compatibility with VDMs....65
Limitations and future work....65
Conclusion....67
Supplementary Information....69
Chapter 3: Drivers of post-fire conifer canopy recovery in California's mixed conifer forests....79
Summary....79
Introduction....79
Methods....82
Study Area....82
Quantifying percent conifer canopy cover....84
Field validation....85
Quantifying post-fire conifer canopy recovery....85
Predictor variables....86
Generalized additive models to predict conifer canopy recovery....88
Results....89
Predicting conifer canopy cover....89
Conifer canopy recovery....90
Discussion....94
Performance of the RF algorithm....94
The long-term effect of post-fire precipitation....95
Shrub competition as a likely mechanism mediating the influence of post-fire precipitation....96
Changing fire severity, temperature, and the potential for vegetation transitions....96
Post-fire planting and seed availability....98
Management implications....99
Conclusion....99
Supplementary Information....101
Conclusion....106
References....108
Appendices....134
Appendix A: Technical documentation for the implementation of the Tree Recruitment Scheme in the Functionally Assembled Terrestrial Ecosystem Simulator (FATES-TRS) 134

Acknowledgments

Thank you to my friends, family, mentors, and coworkers who helped make this PhD possible. Thank you to my advisor, Lara, who spent countless hours helping me hone my thinking and skills as a scientist. Thank you to my parents who helped me cultivate the curiosity, self discipline, and confidence needed for this kind of endeavor. Thank you to the love of my life, Micah, who supported me in every way imaginable through the ups and downs of the PhD process.

Introduction

Background and motivation

Climate change, intensifying natural disturbance regimes and land use are altering forest distribution, structure, and functional composition (Gauthier *et al.*, 2015; Abatzoglou & Williams, 2016; Seidl *et al.*, 2017; McDowell *et al.*, 2020). These forest characteristics affect fluxes of carbon, water, and energy between the land surface and atmosphere (Bonan, 2008). Therefore, predicting the future distribution, structure, and functional composition of forests is necessary for predicting future global climate. At the regional scale of ecosystem management (Lackey, 1998) future forests will influence local hydrology, erosion rates (Lancaster & Baas, 1998; Flint *et al.*, 2019) and determine the type and quality of available wildlife habitat (Verner & Boss, 1980). Ecosystem managers can no longer make decisions or craft management plans under the assumption that forests are in steady state with their climate, nor can they reliably use past ecosystem states and processes to predict future forest distributions or function (Lewis & Maslin, 2015). For example, forest managers in California must prepare for an increased prevalence of drought (Diffenbaugh *et al.*, 2015) and larger, more intense fires (Steel *et al.*, 2018). Predicting how forests will respond to these global change variables is necessary to plan strategies for cultivating resilient forests, building climate-smart hydrological infrastructure, and conserving biodiversity.

A recent global review of drought-induced vegetation transitions shows that forests will only persist where they can regenerate (Martínez-Vilalta & Lloret, 2016). Regeneration processes, including allocation to reproduction, dispersal, germination, and seedling growth and survival are highly sensitive to light, water, temperature, and soil nutrients (Kobe, 1999a; Engelbrecht *et al.*, 2007a; Walck *et al.*, 2011; Conlisk *et al.*, 2017). The sensitive early life stages of a tree act as a bottleneck to mature tree establishment, thereby influencing future forest structure, distribution, and functional composition. For example, conifer seeds must disperse into high severity burn scars and compete with more drought-tolerant shrubs if they are to persist under intensifying fire regimes in California (Tepley *et al.*, 2017; Young *et al.*, 2019). Tropical tree seedlings must disperse, survive, and grow on abandoned, degraded pastureland to facilitate secondary forest development in large areas of the tropics (Hooper *et al.*, 2005; Chapman *et al.*, 2020; Sansevero *et al.*, 2020). Predicting the future of forests under global change requires a focus on regeneration.

The presence of nonstationarity in ecological systems (Rollinson et al., 2021) means that mechanistic, process-based models are best positioned to predict future forest states and function. Vegetation demographic models (VDMs) comprise a class of dynamic global vegetation models (DGVM) that seek to mechanistically represent demographic processes such as forest regeneration (Fisher *et al.*, 2018). The mandate of VDMs is to predict global vegetation dynamics within Earth system models (ESMs), thereby improving the prediction of land surface fluxes and climate feedbacks. In contrast to the sophisticated algorithms VDMs use to predict tree growth and mortality (Fisher *et al.*, 2018), their representation of forest regeneration has received much less attention (McDowell *et al.*, 2020). Prior to the work presented in this

dissertation, VDM representations of regeneration had not been critically reviewed or tested, creating uncertainty around how model representations of regeneration should be advanced.

VDM algorithm development requires an understanding of which processes should be prioritized and what types of data are required to support algorithm development and evaluation. Existing empirical analyses of forest regeneration are often conducted on scales of biological organization and geographic extent that are too narrow to support the development of VDMs algorithms. VDMs are designed for large-scale simulations of plant functional types (PFTs) and therefore require the identification of relationships between environmental variables and plant physiology or demography that are generalizable across broad geographic areas and groups of species. This requires new empirical approaches specifically designed with VDM algorithm development in mind. Filling these research gaps is essential for helping VDMs fulfill their mandate of representing global vegetation dynamics in ESMs. It will also help them become more useful to land management decision-making and planning at regional scales.

Dissertation goals and structure

This dissertation argues that forest regeneration processes will mediate changes in forest distribution, structure, and functional composition (i.e. “forest responses”) in response to global change. The overarching goal is to advance empirical understanding and modeling infrastructure for predicting how forest regeneration processes will mediate these forest responses. It accomplishes this by critically reviewing current representations of forest regeneration processes in VDMs (Chapter 1), presenting a novel modeling scheme for representing forest regeneration in VDMs (Chapter 2) and by advancing empirical understanding of how long-term conifer canopy recovery in California is affected by intensifying fire regimes and a warming climate (Chapter 3).

Chapter 1 first presents a conceptual model for how forest regeneration processes ultimately affect forest distribution, structure, and functional composition at the global scale. It then critically reviews how forest regeneration processes are currently represented in VDMs, analyzes how current representations of regeneration position VDMs to succeed or fail in capturing future tree migration, disturbance recovery dynamics, and compositional turnover. It concludes with recommendations for how to prioritize algorithm development and data collection activities to improve prediction.

Chapter 2 draws upon the insights and conclusions from Chapter 1 to develop a new modeling scheme, the Tree Recruitment Scheme (TRS), for simulating environmentally sensitive tree recruitment in VDMs. The TRS is evaluated by predicting tree recruitment for four tropical tree PFTs under varying meteorology and canopy structure at Barro Colorado Island (BCI), Panama and comparing predictions to those of a current VDM, quantitative observations, and ecological expectations. The TRS is then implemented in a leading VDM, the Functionally Assembled Terrestrial Ecosystem Simulator (FATES). Technical documentation and initial results of this implementation areare available in Appendix A.

Chapter 3 uses remote sensing and machine learning to analyze how changing climate and fire regimes are affecting long-term, post-fire conifer forest regeneration in California's mixed conifer forests (MCFs). This work expands the spatio-temporal scope of prior research to provide new insights on how global change variables affect long-term, conifer forest regeneration outcomes. By focusing on decadal-scale conifer canopy recovery instead of shorter-term seedling regeneration this analysis is positioned to provide new insight into how global change will influence future forest structure and function. The scale of this analysis is also compatible with VDM algorithm development and evaluation which prepares it well to evaluate how current VDMs capture post-fire conifer recovery dynamics in California. It provides an empirical starting point for identifying which processes should be prioritized for development in VDMs that seek to predict the future of California's MCFs. The dissertation concludes by summarizing key findings from Chapters 1-3, discussing the implications of these findings for predicting forest responses to global change and identifying future research directions.

Chapter 1: Forest regeneration within Earth system models: current process representations and ways forward

The work presented in this chapter was previously published in collaboration with the following co-authors.

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

Published in *New Phytologist* (2022): https://doi.org/10.1111/nph.18131

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.

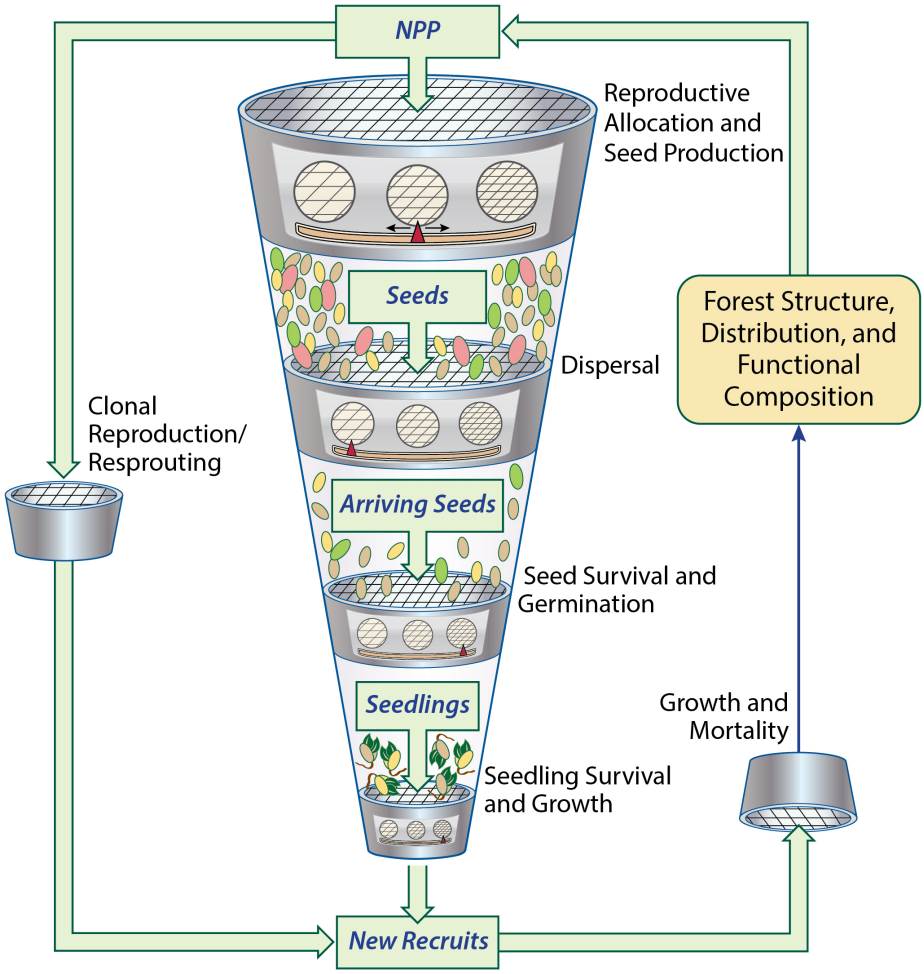


Figure 1. Forest regeneration is a series of environmentally sensitive processes, each represented by a sieve. Environmental variables act as filters (depicted as varying sieve mesh sizes, with the small red arrow setting the mesh size) that impose constraints on each process. Each plant functional type (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.

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).

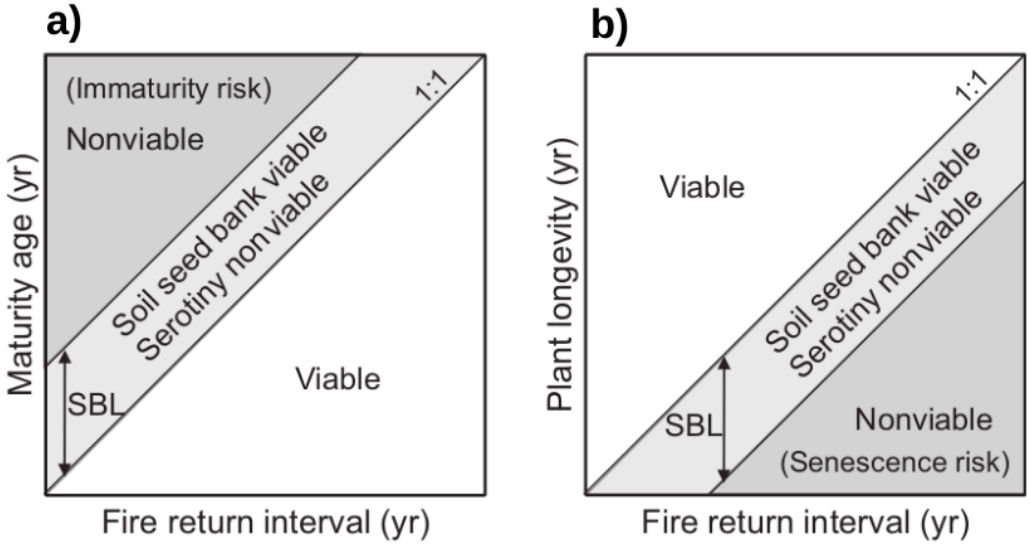


Figure 2. Obligate postfire seeders occur where fire return intervals are between maturation age and plant longevity. Changes to the fire regime risk regeneration failure through "immaturity risk" (a) and "senescence risk" (b). Non-serotinous species (i.e. where seeds are stored in a seed bank) have 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).

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 environmental conditions.

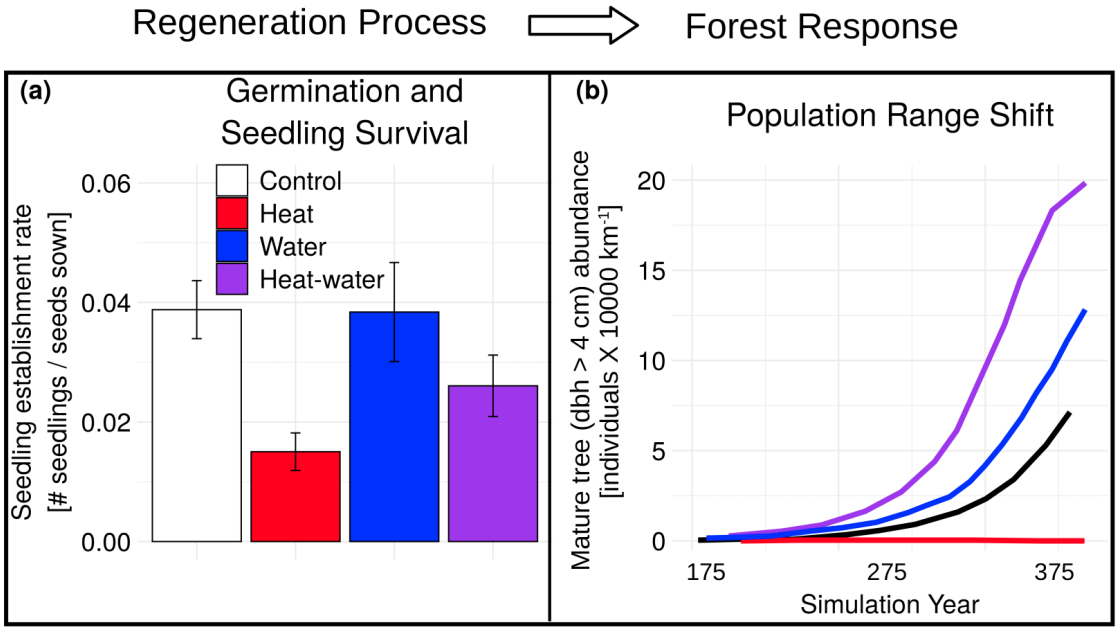


Figure 3. Experimental warming ("Heat" treatment) reduces seedling germination and initial survival (mean ± SE) beyond the local elevation limit of lodgepole pine (a), an upper montane-lower subalpine species, preventing population establishment (b). Supplemental summer water ("Heat-water" treatment) 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.

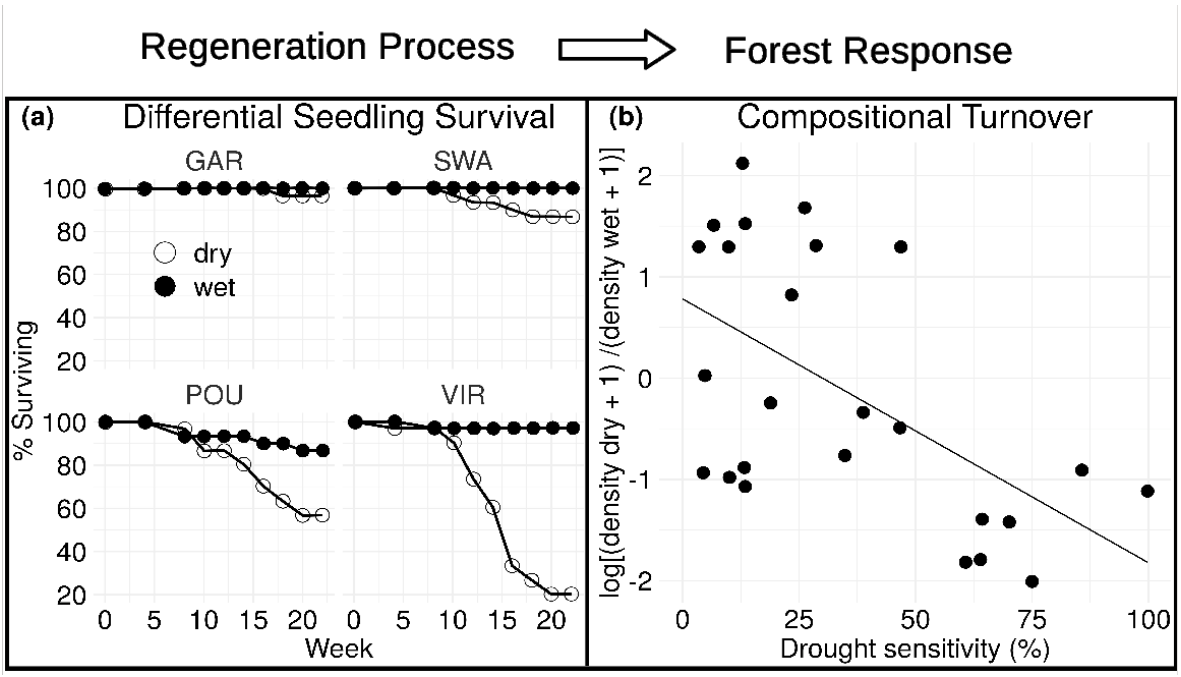


Figure 4. Differential seedling survivorship under manipulated drought (a) is well correlated with compositional turnover over an aridity gradient on the isthmus of Panama (b). In panel b, each point is a species. The "Drought sensitivity" axis quantifies survivorship under manipulated drought compared 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 & Kursar (2005); panel b is remade from Engelbrecht *et al.* (2007).

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.

Table 1. The representation of key forest regeneration processes among seven leading Vegetation Demographic Models; "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 Table S1)

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 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 ($C_{g+r,pft}$) 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 equation below).	**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** ($GDD_{min,pft}$, $T_cmin_{,pft}$, $T_cmax_{,pft}$), $PAR_{min,pft}$ 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}. $N_{recruit,pft} \sim Pois(\lambda_{pft})$ $\lambda_{pft} = \mu(f_{pft}) (est_{max,pft}) ((k_{reprod}) (C_{repro,pft}) + k_{bgestab})$ $\mu(f_{pft}) = exp[\alpha_{pft}(1 - 1 / f^{pft})]$	Same as LPJ-GUESS	**Bioclimatic envelopes**, $PAR_{min,pft}$, and λ_{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 $P_{est,pft} = (Seeds_{pft}) (p_x) (P_{germ,pft})$	$C_{repro,pft}$ 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** (see Fig. 1a in Weng *et al.*, 2015)	**Bioclimatic envelopes** (GDD_{min}, T_cmin, T_cmax) 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

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.

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

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 (PAR_{min}) required for establishment (Smith *et al*., 2001) where shade tolerant PFTs have a lower 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

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

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

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

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).

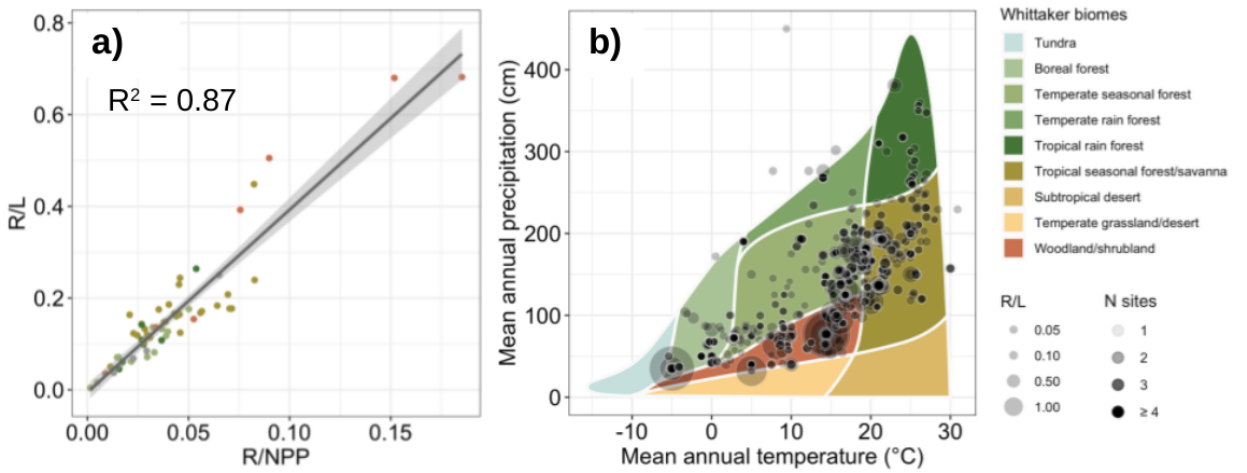


Figure 5. a) Correlation between observed reproductive litter flux as a fraction of net primary productivity (R/NPP) and leaf litter flux (R/L; all fluxes in g C m^{-2} yr^{-1}) at 61 forest sites (colors indicate Whittaker biomes as in panel b; shaded area shows the 95 % confidence interval), and b) 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.

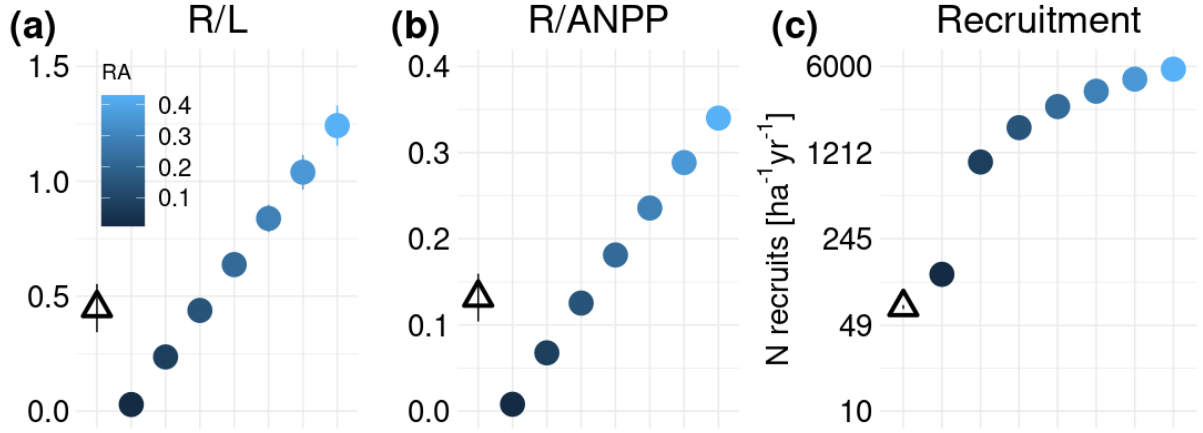


Figure 6. FATES predictions at Barro Colorado Island (circles) of a) reproductive litter flux (R; g C m^{-2} yr^{-1}) as a fraction of leaf litter flux (R/L), b) R as a fraction of above-ground net primary productivity (R/ANPP), and c) recruitment rates into the 1 cm size class. Each model prediction 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.

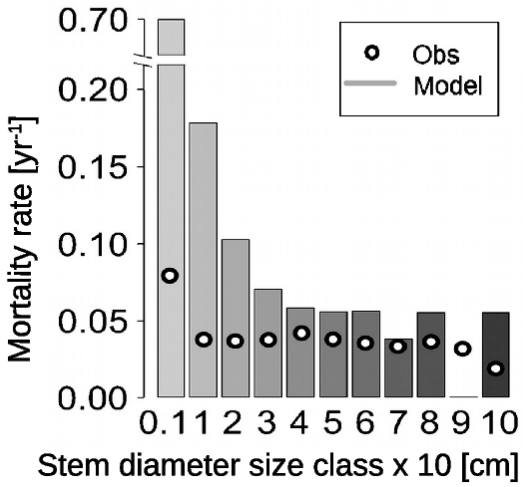


Figure 7. Simulated mortality rates from ED2-hydro ("Model") are positively biased in the smaller size classes at Barro Colorado Island (BCI), Panama, indicating compensating errors for hyperactive recruitment (Powell *et al.*, 2018). "Obs" show field observations.

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.

Acknowledgments
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. Table S1, S2 and S4 are publicly available at *New Phytologist*, doi: https://doi.org/10.1111/nph.18131

Supplementary Information

Table S3. Metadata for Table S2; “NA” = not applicable.

COLUMN NAME	UNITS	DESCRIPTION
Site	NA	Site name
Subsite	NA	Unique id, identifies a subsite within main sampling site
R	g C $m^{-2}year^{-1}$	Flux of reproductive litterfall material
L	g C $m^{-2}year^{-1}$	Flux of leaf litterfall material
ANPP	g C $m^{-2}year^{-1}$	Estimate of aboveground net primary productivity
NPP	g C $m^{-2}year^{-1}$	Estimate of total net primary productivity
Units	-	Units for R, L, ANPP, and NPP
Database_or_data_paper_units	-	Units used by the source database or data paper. Where units are not explicitly masses of carbon, they are masses of dry biomass.
R/L	-	Reproductive litterfall flux divided by leaf litterfall flux (R/L)
R/ANPP	-	Reproductive litterfall flux divided by ANPP (R/ANPP)
R/NPP	-	Reproductive litterfall flux divided by NPP (R/NPP)
Litterfall sampling interval	YYYY.MM-YYYY.MM	Start and end year of reproductive litterfall and leaf litterfall sampling events, if reported in original data
Stem census sampling interval	YYYY-YYYY	Start and end year of stem census interval associated with NPP estimate, if reported in original data.
Country	-	Country
Lat	Decimal degrees	Reported approximate latitude (‘-’ = south of equator; ‘+’ = north of equator)
Lon	Decimal degrees	Reported approximate longitude (‘-’ = west of prime meridian; ‘+’ = east of prime meridian)
MAT	degrees C	Mean annual temperature, in degrees Celsius
MAP	mm $year^{-1}$	Mean total annual precipitation, in mm/year
MASL	meters above	Mean elevation

	sea level	
Whittaker biome	-	Whittaker biome, as defined in Figure 5b of the main text, Whittaker biome is NA if site MAT and MAP fall outside of the limits delineated in Fig. 5b.
References	-	List of references which provide sources of original data and information regarding litterfall sampling, census events, MAT, MAP
Database_or_data_paper	-	Database or data paper used to aggregate data; original data sources are given in “References” field

Methods S1. Data synthesis methods for 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 S2 includes estimates of annual ecosystem-level reproductive litter flux (R; g C m^{-2} yr^{-1}), leaf litter flux (L; g C m^{-2} yr^{-1}), aboveground net primary productivity (ANPP), and NPP for 61 subsites. An additional 499 subsites (n = 550) include estimates of just R, L, and R/L. Each row reports a distinct set of estimates for one sampling interval at one subsite (see Table S3 for all associated field descriptions).

The spatial scope is global and includes tropical, temperate, and boreal forest biomes. The geographical location (approximate latitude, longitude, country) and climate variables (mean annual temperature, mean annual precipitation) for each subsite are also reported. These data cover a climate gradient with a range of mean annual precipitation from 323 to 5302 mm and mean annual temperature from -5.4 to 30.9 degrees C. Across all sites, the benchmark R/L ranged from 0 to 3.7; for sites reporting NPP, R/NPP ranged from 0.001 to 0.185.

Subsites reporting reproductive litterfall, leaf litterfall, ANPP and NPP (all in g C m^{-2} yr^{-1}) were found by querying Web of Science, Google Scholar, Scielo, and publicly available ecological monitoring network metadata available between 1966 and 2021 using the following query terms: "(Litterfall OR litter fall OR litter-fall) AND (fruit OR flower OR seed OR reproductive) AND (leaf OR foliage OR needle) AND (NPP OR "net primary productivity"). This produced ~3,800 publications which were used as a starting point to find published estimates of NPP, ANPP, reproductive flux, and leaf litter fluxes. Of these, ~400 publications reference separating reproductive material from leaf material during sampling, and ~6% report reproductive flux as a unique NPP component. These data were used to calculate the benchmarks R/L, R/ANPP, and R/NPP and to plot the correlation between R/L and R/NPP, shown in Figure 5a in the main text.

Sites reporting only reproductive and leaf litter fluxes were identified by searching for publicly available litterfall databases and data papers. Flux estimates are reported in units of mass of carbon (see "Units" in Table S3); data originally reported in units of dry biomass were converted to carbon assuming dry biomass is 50% carbon by weight. Data from all subsites (n = 550) were used to create Figure 5b in the main text.

The data table (Table S2) and Figure 5 in the main text were created using R 4.0.3 and the tidyverse (Wickham et al., 2019), ggpubr (Alboukadel, 2020) and plotbiome (v0.0.0.9001; Valentin & Levin, 2021) packages.

Methods S2 Model-data comparison at Barro Colorado Island, Panama.

To evaluate model predictions at the Barro Colorado Island (BCI) Forest Dynamics Plot (see Box 3, Fig. 6 in main text), we used litterfall observations from January 2014-December 2018 (Wright and Cornejo, 1990) to estimate mean annual reproductive litter flux (R; g m^{-2} yr^{-1}), mean annual leaf litter flux (L, g m^{-2} yr^{-1}), and R as a fraction of L (R/L). We also used census data from the BCI Forest Dynamics Plot (Hubbell *et al.*, 2015) between 2005 and 2015 to calculate per area recruitment rates (indv. m^{-2} yr^{-1}) for each species into the 1 cm size class. We used recruitment estimation methods that account for unobserved mortality of new recruits between census intervals (see Kohyama *et al.*, 2018, eqn 11). Ecosystem-level tree recruitment rates were calculated by summing all species-level recruitment rates.

For consistency with model predictions, only recruitment and litter fall data from "canopy trees" were included in our analysis, which we define according to prior work (Powell *et al.*, 2018) as any tree with the potential to reach a stature 20 cm dbh. Palms (i.e. members of the family *Arecaceae*) were excluded from all estimates and understory specialists were additionally excluded from estimates of recruitment rates. BCI litterfall data identifies all material to species and categorizes reproductive material by botanical part. Therefore, we used species-specific seed-to-fruit and capsule-to-fruit ratios to correct estimates of R for frugivory. To correct L for herbivory we assumed that leaves lost 9.70% leaf area prior to observation to correct for leaf herbivory (herbivory correction methods provided by J. Wright, personal communication). R/ANPP was calculated using the most recent ANPP estimate available from BCI (Condit *et al.*, 2013).

Model predictions were made with the Functionally Assembled Terrestrial Ecosystem Simulator (FATES; tag: sci.1.47.0_api.17.0.0; Koven *et al.*, 2020) initialized with observed stand structure taken from the tree census ending in 2016 at the BCI Forest Dynamics Plot (Hubbell *et al.*, 2015). Each model prediction for R/L, R/ANPP, and recruitment (i.e. circles in main text Fig. 6) is the mean and SD over 25 simulation years driven with observed, recycled meteorology at BCI (2003-2016; Faybishenko *et al.*, 2018) and one of 7 unique values for reproductive allocation: 0.009, 0.079, 0.149, 0.220, 0.290, 0.360 0.430. This range is based on observed variation of RA in trees (Wenk & Falster, 2015).

Figure S1. Correlation between observed reproductive litter flux as a fraction of above-ground NPP (R/ANPP) and as fraction of leaf litter flux (R/L; g C m^{-2} yr^{-1}) at 61 forest sites. The shaded area represents the 95% confidence interval.

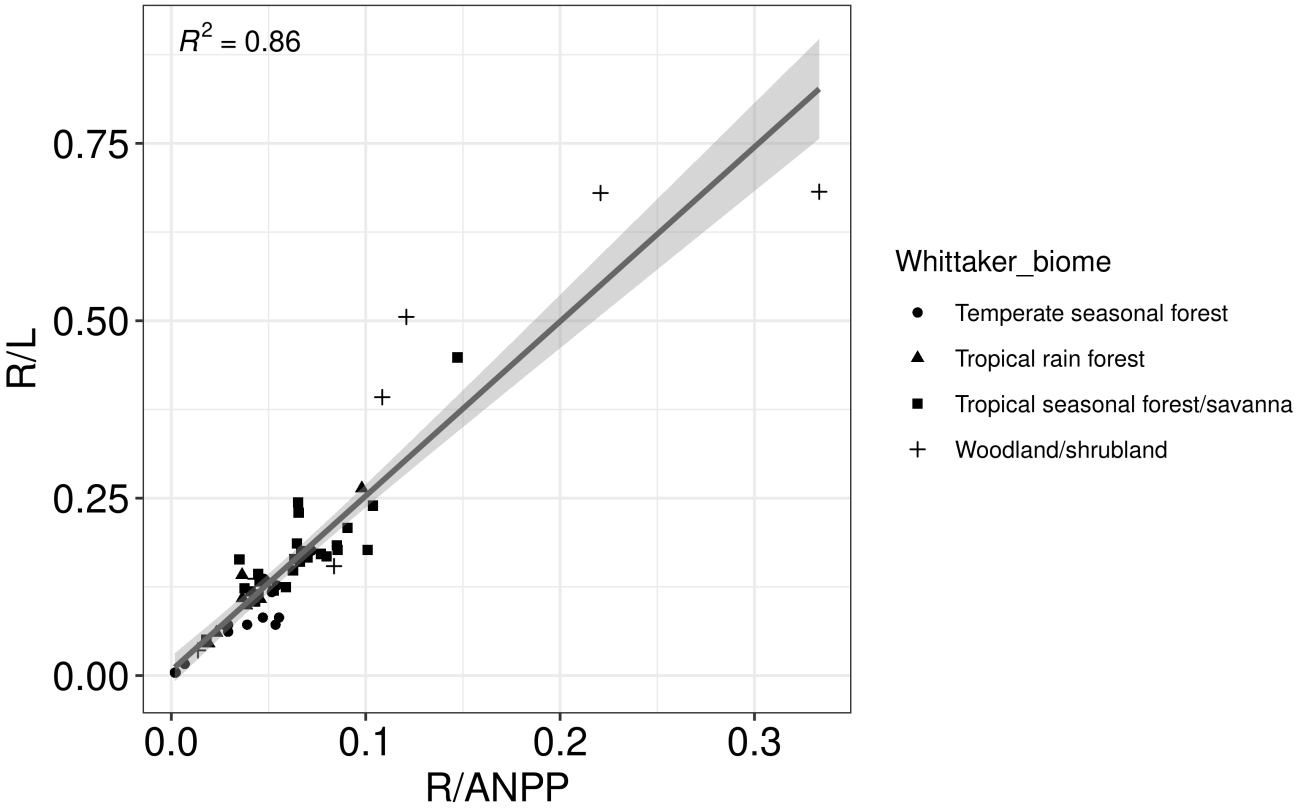
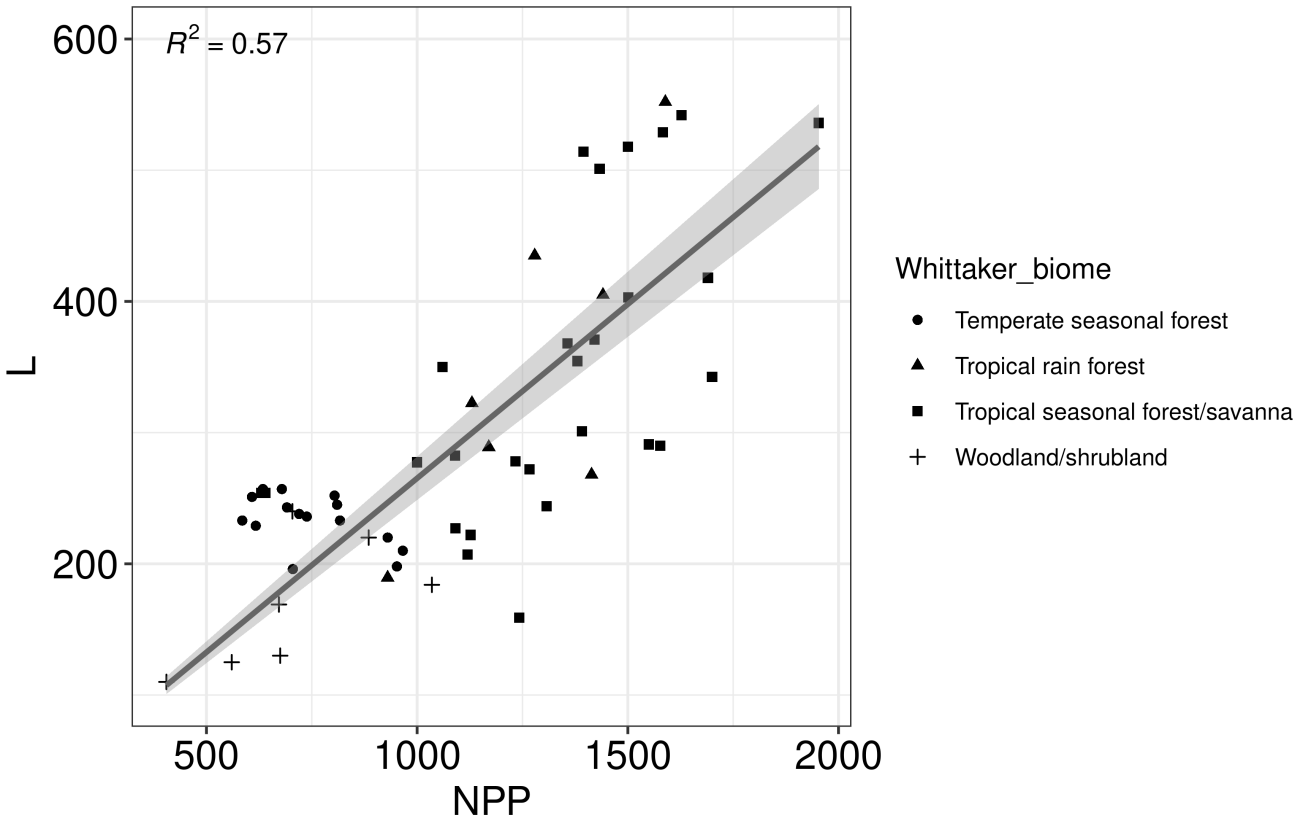


Figure S2. Correlation between observed leaf litter flux (g C m^{-2} yr^{-1}; L) and NPP (g C m^{-2} yr^{-1}) at 61 forest sites. The shaded area represents the 95% confidence interval.



Methods S3. Data synthesis methods for 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.

Species-level recruitment rates (individuals m^{-2} yr^{-1}) into the 1 cm size class were calculated from CTFS-ForestGeo census data at four sites using methods that account for unobserved mortality of new recruits between census intervals (see Kohyama *et al.*, 2018, eqn 11). For consistency with model predictions, only "canopy trees" are included, which we define according to prior work (Powell *et al.*, 2018) whereby any tree with the potential to reach a stature 20 cm dbh is included in the analysis. Palms (i.e. members of the family *Arecaceae*) were excluded from estimates.

Tree census data used to produce this data set came from a) the Barro Colorado Island (BCI) Forest Dynamics Plot (Hubbell *et al.*, 2015), census observations between 1981-2016; b) the Luquillo Forest Dynamics Plot (Zimmerman *et al.*, 2010), census observations between 1990-2016; c) the Smithsonian Conservation Biology Institute (SCBI) Forest Dynamics Plot (Bourg *et al.*, 2013), census observations between 2008-2018; and d) the Smithsonian Environmental Research Center (SERC) Forest Dynamics Plot, census observations between 2008-2014 (McMahon & Parker, 2021). These sites cover tropical rain forest (BCI and Luquillo) and temperate seasonal forest (SCBI and SERC) Whittaker biomes. All data were accessed via the ForestGeo data portal (http://ctfs.si.edu/datarequest/index.php/main/plotdata) on September 26th, 2021 and filtered to main stems only (i.e. excluding resprouts).

Table S5. Metadata for 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. “NA” = not applicable.

COLUMN NAME	UNITS	DESCRIPTION
sp	NA	ForestGeo 6-letter species code
cen_interval	NA	Census interval
tStart	yyyy-mm-dd	Start date of the census interval
tEnd	yyyy-mm-dd	End date of the census interval
tCenInt	years	Length of census interval
Nt0	individuals	The number of living individuals at the start of the census interval
Nt1	individuals	The number of living individuals at the end of the census interval
S_t1	individuals	The number of survivors across the census interval
M_t1	ind. $m^{-2}yr^{-1}$	The per-area mortality rate
R_t1	ind. $m^{-2}yr^{-1}$	The per-area recruitment rate into the 1 cm size class
site	NA	The ForestGeo forest dynamics plot site code: “bci” = Barro Colorado Island, Panama; “luq” = Luquillo, Puerto Rico; “scbi” = Smithsonian Conservation Biology Institute, USA; “serc” = Smithsonian Environmental Research Center, USA
latin	NA	The Latin binomial species name

Chapter 2: Simulating environmentally sensitive tree recruitment in vegetation demographic models

The work presented in this Chapter builds upon insights from the critical review of current VDMs, presented in Chapter 1, to develop an novel scheme for representing environmentally sensitive tree recruitment in vegetation demographic models. It was previously published in collaboration with the following co-authors.

Adam R. Hanbury-Brown[1], Thomas L. Powell[2,3], Helene C. Muller-Landau[4] , S. Joseph Wright[4], 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
[3]Department of Earth and Environmental Systems, The University of the South, 735 University Ave, Sewanee, TN 37383, USA
[4]Smithsonian Tropical Research Institute, Apartado 0843–03092, Balboa, Republic of Panama

Published in *New Phytologist* (2022): https://doi.org/10.1111/nph.18059

Summary

Vegetation demographic models (VDMs) endeavor to predict how global forests will respond to climate change. This requires simulating which trees, if any, are able to recruit under changing environmental conditions. We present a new recruitment scheme for VDMs in which functional-type-specific recruitment rates are sensitive to light, soil moisture, and the productivity of reproductive trees. We evaluate the scheme by predicting tree recruitment for four tropical tree functional types under varying meteorology and canopy structure at Barro Colorado Island, Panama. We compare predictions to those of a current VDM, quantitative observations, and ecological expectations. We find that the scheme improves the magnitude and rank order of recruitment rates among functional types and captures recruitment limitations in response to variable understory light, soil moisture, and changing precipitation regimes. Our results indicate that adopting this framework will improve VDM capacity to predict functional-type-specific tree recruitment in response to climate change, thereby improving predictions of future forest distribution, composition, and function.

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

Introduction

Tree recruitment, the rate at which trees grow into the smallest size class tracked by observations, affects global terrestrial ecosystem functioning by determining the rate of regrowth and the future vegetation composition after disturbance (Chazdon, 2003; Johnstone *et al.*, 2016; Martínez-Vilalta & Lloret, 2016). It is the outcome of a dynamic set of environmentally sensitive

processes including seed production, dispersal, and seedling establishment (Hubbell *et al.*, 1999; Chazdon, 2003; Wright & Calderón, 2006; Wright *et al.*, 2007; Jabot *et al.*, 2008; Markl *et al.*, 2012; Hacket-Pain *et al.*, 2018). Changing climate and disturbance regimes affect recruitment through their influence on the regeneration niche: the set of environmental conditions needed for plants to produce viable seed, establish as seedlings, and recruit (Grubb, 1977). When an established population's regeneration niche contracts, it results in a change in forest composition or distribution (Engelbrecht *et al.*, 2007; Poorter, 2007; Jabot *et al.*, 2008; Bond, 2008; Valdez *et al.*, 2019). Multiple lines of evidence indicate that climate change and land use are already affecting forest regeneration globally (Chazdon, 2003; Kueppers *et al.*, 2017; Serra-Diaz *et al.*, 2018; Valdez *et al.*, 2019; Sansevero *et al.*, 2020). For example, more severe droughts are linked to declines in post-fire tree recruitment (Stevens-Rumann *et al.*, 2017; Tepley *et al.*, 2017), and increasing hurricane intensity is predicted to change forest composition through differential seedling survival (Comita *et al.*, 2009). Limitations to tree recruitment that change biome boundaries (e.g., Bond, 2008, Sansevero *et al.*, 2020) or the functional composition of a forest (e.g., Johnstone *et al.*, 2006) affect ecosystem resilience and function through structural and physiological traits (Poulter *et al.*, 2011; Zhang *et al.*, 2018; Bonan, 2019), making their prediction essential for forecasting terrestrial biosphere function in the Earth system.

There is growing interest in using vegetation demographic models (VDMs) to represent vegetation dynamics within Earth System Models (ESMs; Bonan, 2015; Fisher *et al.*, 2018). VDMs are "a special class of DGVM, which include representation/tracking of multiple size-classes or individuals of the same PFT, which can encounter multiple light environments within a single climatic grid cell" (Fisher *et al.*, 2018). In contrast to the sophisticated algorithms VDMs use to predict growth and mortality, their representations of recruitment lack a sufficiently process-based approach (McDowell *et al.*, 2020; Hanbury-Brown *et al.*, 2022). Gap models, forest landscape models and mechanistic species distribution models have successfully represented key regeneration processes influencing recruitment such as seed production, dispersal, germination, and seed decay in stand- and landscape-scale simulations (Mladenoff, 2004; Lischke *et al.*, 2006; Lischke & Loffler, 2006; Scheller *et al.*, 2007; Holm *et al.*, 2012; Mok *et al.*, 2012; Larocque *et al.*, 2016), but their modeling approaches are generally less suitable for large scale ESM-coupled simulations because they are computationally expensive and do not conserve carbon. To operate within these constraints many VDMs represent "cohorts" of trees, which belong to the same PFT and size class, and are tracked as pools of carbon occupying spatially implicit forest patches (e.g. Medvigy *et al.*, 2009; Fisher *et al.*, 2015). Each cohort's associated number density of stems is calculated based on representative tree diameter and allometry.

VDMs such as LPJ-GUESS and SEIB-DGVM rely on parameter tuning and rough proxies for understory light and space to calculate recruitment rates and use bioclimatic envelopes to predict PFT distributions (Smith *et al.*, 2001; Sato *et al.*, 2007; Sato, 2009). Bioclimatic envelopes rely on the tenuous assumption that historical correlations between species distributions and climate metrics are sufficient to predict future distributions under novel climates and species assemblages (Pearson & Dawson, 2003; Journé *et al.*, 2020). VDMs based on the Ecosystem Demography (ED) concept (Moorcroft *et al.*, 2001), have dropped bioclimatic envelopes in favor

of allowing biogeography to emerge from ecophysiology and competition, but no mechanistic constraints have replaced bioclimatic envelopes for limiting recruitment. Instead, a fixed fraction of net positive carbon production is allocated to a reproductive carbon pool from which new individuals emerge at a rate that is a constant fraction of this pool (Moorcroft *et al.*, 2001; Medvigy *et al.*, 2009; Fisher *et al.*, 2018). This captures the effect of adult productivity on seed production, but environmental conditions in the seedling layer do not affect recruitment in these models. This limits their ability to capture how climate change will affect the regeneration niche, future recruitment limitations, forest distribution, and functional composition. The need to improve predictions is particularly critical for tropical forests which make up the "least certain major component of the global carbon budget" (Mitchard, 2018).

Here we present a new Tree Recruitment Scheme (TRS) for VDMs that more mechanistically constrains recruitment rates based on carbon production from reproductively mature trees, light at the forest floor, and soil moisture in the simulated rooting zone of seedlings. We seek to capture the size dependence of reproductive output and the light- and moisture-dependence of seedling emergence, survival, and the transition out of the seedling pool. By incorporating these environmentally sensitive processes we hope to capture more realistic recruitment responses to varying light, soil moisture, and changing precipitation patterns. We evaluate the TRS by simulating tree recruitment in a seasonally dry tropical forest under observed meteorological conditions, a synthetic El Niño (Powell *et al.*, 2017), wetter-than-observed, and drier-than-observed precipitation scenarios. We compare the scheme's predictions of recruitment rates (at the 1 cm dbh size class) to predictions from the Ecosystem Demography model version 2 (Medvigy *et al.*, 2009), forest demographic data, and ecological expectations and conclude by discussing how the TRS is positioned to improve VDM predictions of forest distribution, composition, and function under global change.

Materials and Methods

Model description

VDMs represent the forested landscape as a mosaic of spatially implicit patches varying in time since the last disturbance. This creates a patchwork of heterogenous biotic and abiotic conditions under which cohorts of trees recruit, grow, compete, and die (Fisher *et al.*, 2018). Trees within cohorts are all the same size and PFT, but each patch can contain multiple cohorts of different sizes and PFTs. We designed the TRS, currently implemented in *R* (R Core Team, 2020), to operate within each of these forest patches where it predicts PFT-specific, environmentally sensitive tree recruitment rates. It was developed primarily from studies at BCI, with an initial focus on tropical tree PFTs that vary along axes of drought and shade tolerance (Fig. 2), but is designed to be extensible to all tropical forests. Our primary goal is to mechanistically constrain recruitment. In each daily timestep the TRS receives cohort-level carbon for growth and reproduction (C_{g+r}; net after tissue turnover and allocation to storage) from its host VDM. Regeneration processes, described in detail below, move dynamic fractions of C_{g+r} through a seedbank and seedling pool (Fig. 1) which are tracked in units of carbon. Carbon emerging out of the seedling pool each day is passed back to the VDM and can be converted into a number

density of new recruits. Carbon in seeds or seedlings that die or that is allocated to non-seed reproductive biomass, moves to a reproductive litter pool (also passed back to the VDM), thereby conserving carbon.

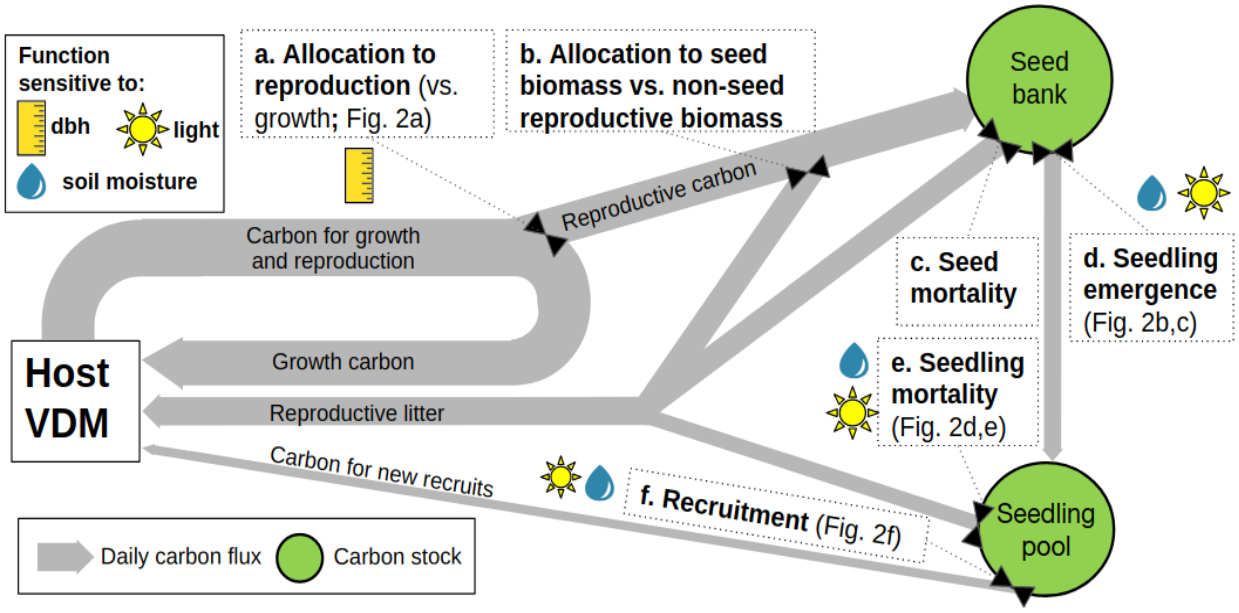


Figure 1. The TRS receives carbon for growth and reproduction from the host VDM. Daily regeneration processes (depicted with hour glasses) transfer reproductive carbon through seed bank and seedling carbon pools (depicted as circles). Processes are sensitive to diameter at breast height (DBH) or environmental conditions (see inset key). The host model's reproductive litter pool receives non-seed reproductive carbon, dead seeds, and dead seedlings. Carbon for new recruits can be passed back to the host VDM in units of carbon or as a number density of new recruits. Parenthetical references to Figure 2 show relationships between regeneration processes and DBH and environmental conditions.

Allocation to reproduction

The probability a tree is reproductive increases sigmoidally with size within species (Visser *et al.*, 2016). Past models assume reproductive allocation is insensitive to size (Smith *et al.*, 2001) or invariant above a fixed size threshold (Sato *et al.*, 2007; Medvigy *et al.*, 2009; Fisher *et al.*, 2015). In contrast, the TRS allocates a dynamic fraction of cohort-level C_{g+r} to reproduction based on the cohort's size and reproductive allocation (RA) function (Eqn 1, Fig. 1a, Fig. 2a). Each mature cohort in the host VDM can contribute to recruitment via the TRS if they are in positive carbon balance. The effective fraction of cohort-level C_{g+r} allocated to reproduction, $F_{E,repro}$ (Eqn 2), is calculated based on a sigmoidal relationship relating the cohort's current dbh to the probability of being reproductive (P_{repro}; Eqn 1). We assume that all reproductive individuals allocate to reproduction at a constant, PFT-specific rate, F_{repro} (see Table 1 for all TRS parameters), which is modified by P_{repro} to calculate $F_{E,repro}$ (Eqn 2)

$$P_{repro} = \frac{e^{(a_{RA}(dbh)+b_{RA})}}{1 + e^{(a_{RA}(dbh)+b_{RA})}} \quad \text{(Eqn 1)}$$

$$F_{E,repro} = (P_{repro})(F_{repro}) \quad \text{(Eqn 2)}$$

where a_{RA} and b_{RA} are PFT-specific parameters describing the shape of the sigmoidal curve (Fig. 2a). This functional form is consistent with empirical data (Visser *et al.*, 2016; Minor & Kobe, 2019; Andrus *et al.*, 2020). The TRS subsequently multiplies $F_{E,repro}$ by C_{g+r} to get reproductive carbon per cohort.

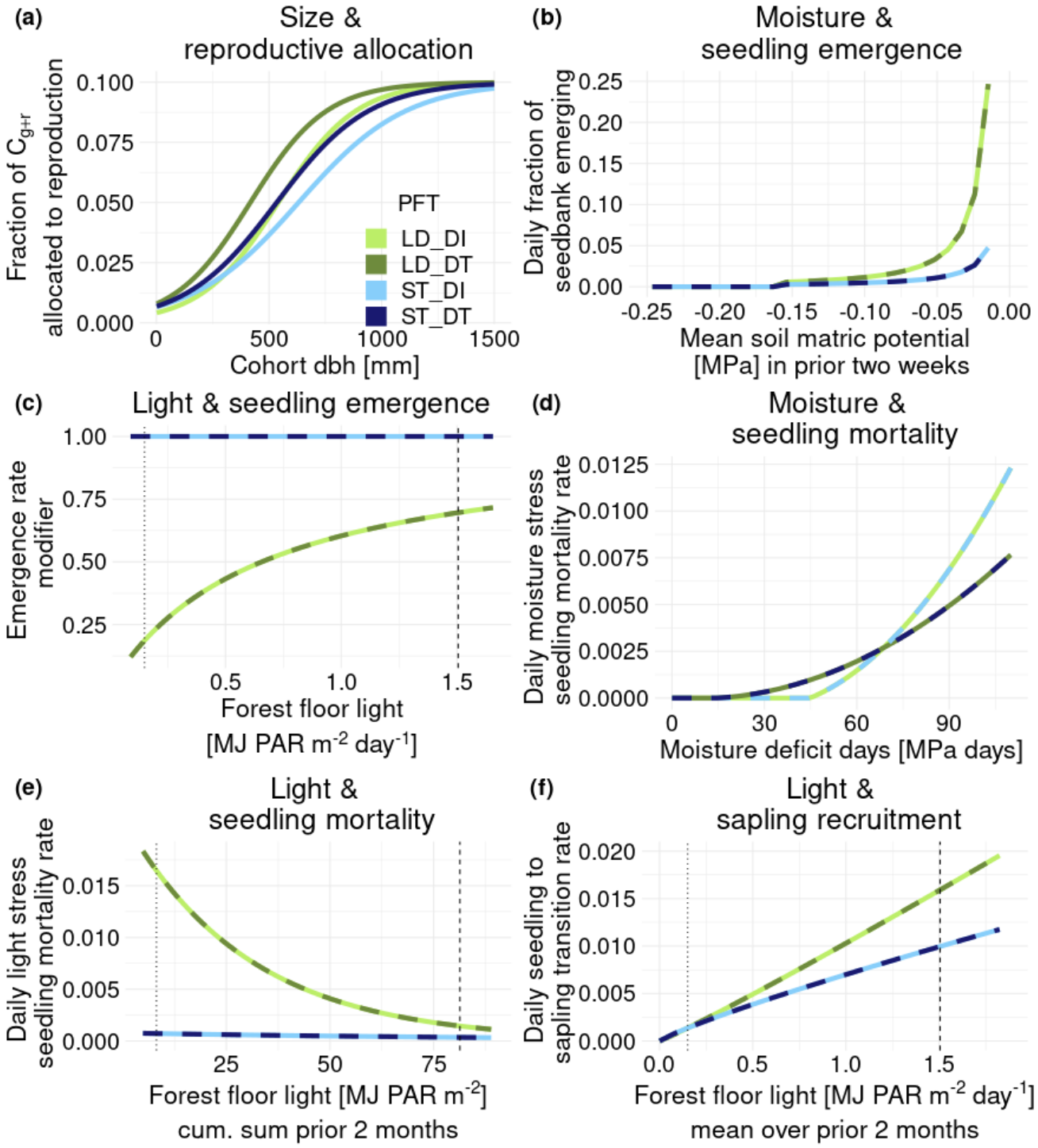


Figure 2. Regeneration processes represented by the Tree Recruitment Scheme are sensitive to size (a), soil moisture (b,d), and light (c,e,f,). Functions are parameterized for four tropical tree PFTs at Barro Colorado Island, Panama: light demanding, drought intolerant (LD-DI); light demanding, drought tolerant (LD-DT); shade tolerant, drought intolerant (ST-DI); and shade

tolerant, drought tolerant (ST-DT). Note that moisture deficit days (panel d) are not equivalent between DT and DI PFTs because they accumulate according to PFT-specific values of Ψ_{crit}. In panels c, e, and f, the dotted vertical lines indicate mean understory light conditions on BCI (2% top of canopy radiation) and the dashed vertical lines indicate conditions in a small-medium sized light gap (20% top of canopy radiation). C_{g+r} = carbon for growth and reproduction; dbh = diameter at breast height; PAR = photosynthetically active radiation.

Allocation to seed vs. non-seed reproductive biomass and seed mortality
In nature, only a subset of the carbon allocated to reproduction becomes seeds, with the rest going to flowers, fruit flesh, capsules, etc. (Wenk *et al.*, 2017), but previous models assume that all reproductive carbon becomes seed (Fisher *et al.*, 2015). The TRS partitions each cohort's reproductive carbon into seed carbon and non-seed reproductive carbon (e.g., flowers, fruit flesh, and capsules) based on a prescribed, PFT-specific fraction of reproductive carbon that is seed, F_{seed} (Fig. 1b). Available seed carbon moves to a seed bank each day and non-seed reproductive carbon moves to a reproductive litter pool (Fig. 1b). Seeds in the seed bank die at a PFT-specific, constant rate, S_{mort}, which represents all modes of seed mortality including predation and decay.

Seedling emergence
Seedling emergence is sensitive to soil moisture (Garwood, 1983; Atondo-Bueno *et al.*, 2016; Ruiz Talonia *et al.*, 2017; Foster *et al.*, 2020) and light (Pearson *et al.*, 2002), but prior regeneration schemes in VDMs either do not represent seedling emergence (Sato *et al.*, 2007; Medvigy *et al.*, 2009) or represent it as an environmentally insensitive constant (Fisher *et al.*, 2015). In the TRS, by contrast, emergence depends on both soil moisture and light.

Light-dependence of germination is captured on day i in a Michaelis-Menten rate modifier [0,1]

$$f(PAR_i) = \frac{PAR_i}{PAR_i + PAR_{crit}} \qquad \text{(Eqn 3)}$$

dependent on PAR_i, the photosynthetically active radiation at the seedling layer on day i, and PAR_{crit}, a PFT-specific threshold governing the shape of the germination response to light (Fig. 2c). Most tropical pioneer species exhibit an increase in germination probability with light, whereas shade-tolerants are insensitive to light (captured by $PAR_{crit} = 0$).

If soil moisture is above a critical threshold, Ψ_{emerg}, the seedling emergence rate on day i, $F_{emerg,i}$, is dynamically calculated based on the mean soil matric potential (SMP) in the top 0-10 cm over a rolling window of days, W_{emerg}, prior to i (Fig. 1d, Fig. 2b). The moisture response parameter, b_{emerg}, modifies the mean seedling emergence coefficient (a_{emerg}) in response to variation in SMP such that

$$F_{emerg,i} = \begin{cases} 0 & SMP_i < \psi_{emerg} \\ f(PAR_i)(a_{emerg}) \left(\frac{\sum_{j=i-W_{emerg}}^{i} (1/-SMP_j)}{W_{emerg}} \right)^{b_{emerg}} & SMP_i \geq \psi_{emerg}. \end{cases} \qquad \text{(Eqn 4)}$$

Eqn 4 produces pulses of seedling emergence in response to seasonal and interannual precipitation events, and stalls seedling emergence under relatively dry conditions.

This formulation captures observed patterns of variation in seedling emergence in relation to fluctuations in soil moisture (Garwood, 1983; Foster *et al.*, 2020) and spatial variation in understory light levels (Vazquez-Yanes *et al.*, 1990; Pearson *et al.*, 2002). At BCI there is a pulse of seedling emergence at the onset of the wet season with the earliest emerging species responding within ~2 weeks of wet season precipitation (Garwood, 1983). The observed seasonal recruitment pulse at BCI is more pronounced for light demanding (LD, "pioneer") species than for shade tolerant (ST) species (see Fig. 7 in Garwood, 1983), which is represented with a higher value for b_{emerg}.

Moisture and light-sensitive seedling survival

Seedling survival decreases at low soil moisture and low light, affecting forest composition across environmental gradients (Kobe, 1999; Engelbrecht *et al.*, 2007), but this dynamic is missing in previous models. The TRS seeks to capture this with a PFT-specific moisture stress threshold, Ψ_{crit}, below which the seedling pool starts to accumulate moisture deficit days (MDD) similar to the concept of growing degree days. The MDD value on day i is summed within a rolling window of days, W_Ψ, prior to i such that

$$MDD_i = \sum_{j=i-W_\psi}^{i} \begin{cases} 0 & \psi_j \geq \psi_{crit} \\ |\psi_j| - |\psi_{crit}| & \psi_j < \psi_{crit}. \end{cases} \quad \text{(Eqn 5)}$$

This formulation simultaneously captures the magnitude and duration of moisture stress. We used observations of seedling wilting points from a prior manipulative drought experiment (Engelbrecht & Kursar, 2003; Engelbrecht *et al.*, 2007) to explore the relationship between moisture deficit day accumulation and mortality. We found that observed drought-induced mortality was 0 up to a critical accumulation of MDD, MDD_{crit}, at which point a convex quadratic relationship best explained drought-induced seedling mortality as a function of MDD (Fig. 1e, Fig. 2d, see SI Methods S1 and Fig. S1 for more details). The mortality rate from moisture stress (M_Ψ) on day i is therefore

$$M_{\psi,i} = \begin{cases} 0 & MDD_i < MDD_{crit} \\ a_\psi MDD_i^2 + b_\psi MDD_i + c_\psi & MDD_i \geq MDD_{crit} \end{cases} . \quad \text{(Eqn 6)}$$

Seedlings also die from insufficient light, which we refer to as light stress. The light stress mortality rate, M_L, on day i is a function of PAR at the seedling layer, $L_{seedling}$, accumulating within a rolling window of days, W_L, prior to i (Fig. 1e, Fig. 2e). Two PFT-specific parameters determine the shape of the negative exponential relationship

$$M_{L,i} = e^{a_{ML}\left(\sum_{j=i-W_L}^{i} L_{seedling,j}\right)+b_{ML}} \quad \text{(Eqn 7)}$$

where a_{ML} is a PFT-specific light response parameter and b_{ML} is the intercept. We based this function on an analysis by Kobe (1999) who tested four functional forms and found that the negative exponential best described light stress mortality for two shade tolerant (ST) and one

light demanding (LD) species that were transplanted into varied light environments. This function is supported by observations that seedling mortality generally increases as light decreases (Kitajima, 1994; Poorter, 1999; Bloor & Grubb, 2003; Comita & Hubbell, 2009). Tolerance to low light conditions varies considerably across species depending on life history strategy (Kitajima, 1994; Wright *et al.*, 2010), which can be captured with PFT-specific values of a_{ML} (Table 1). A background seedling mortality rate, $M_{background}$, represents other seedling mortality (e.g. herbivory, pathogens, tree fall, etc.). Total seedling mortality is the sum of moisture-dependent, light-dependent and background mortality.

Recruitment

The rate of transition from seedling to sapling increases with understory light (Brokaw, 1985; Rüger *et al.*, 2009), but prior models do not capture this sensitivity (Sato *et al.*, 2007; Fisher *et al.*, 2015) or use proxies for absolute light (Smith *et al.*, 2001). In contrast, recruitment in the TRS is represented with a dynamic seedling to sapling transition rate (TR) which is the fraction of total carbon in the seedling pool, $C_{seedling}$, that recruits each day (Fig. 1f). The TR on day i is calculated as a power function of mean PAR at the seedling layer within a rolling window of days, W_L, prior to i (Fig. 2f). If SMP on day i, Ψ_i, is drier than Ψ_{crit} the transition rate goes to zero such that

$$TR_i = \begin{cases} 0 & \psi_i < \psi_{crit} \\ a_{TR}\left(\frac{\sum_{j=i-W_L}^{i} PAR_j}{W_L}\right)^{b_{TR}} & \psi_i \geq \psi_{crit} \end{cases} \quad \text{(Eqn 8)}$$

where a_{TR} is a coefficient derived from the mean transition rate at observed mean understory PAR (see SI Methods S1) and b_{TR} is the light response modifier. The light response modifier produces accelerating (LD PFTs) or decelerating (ST PFTs) responses to light (Fig. 2f) depending on if b_{TR} is greater or less than 1. Of a variety of functional forms tested at BCI, a power function with species-specific light response modifiers best explained observed variation in recruitment rates under spatially heterogenous patch-level light (Rüger *et al.*, 2009). This formulation is more broadly supported by the growth-mortality functional trade-off axis where LD species can take advantage of higher light conditions through faster relative growth rates (Wright *et al.*, 2010). Carbon transitioning out of the seedling layer can be converted to a number density of new recruits based on the amount of carbon required to form an individual in the smallest size class tracked by the VDM, Z_0, such that the number of new recruits predicted on day i, R_i, is

$$R_i = \frac{(TR_i)(C_{seedling,i})}{Z_0}. \quad \text{(Eqn 9)}$$

Eqn 9 is very similar to how VDMs currently convert reproductive carbon into new recruits, but the key difference is that the TRS only makes this conversion after the functions presented here

(Eqns 1–8; Fig. 2) have more mechanistically constrained the amount of carbon available for recruitment.

Simulations at Barro Colorado Island, Panama

We ran the TRS at the 50-ha forest dynamics plot (FDP) on BCI in central Panama (9.151°N, 79.855°W). BCI receives an average of 2662 ± 479 (SD) mm of precipitation per year and experiences a four-month dry season (<100 mm of precipitation per month). All living trees in the FDP > 1 cm in dbh have been censused every 5 years since 1985 (Condit, 1998), which provides the opportunity to benchmark recruitment rates into the 1 cm size class.

We used monthly model output from the Ecosystem Demography model version 2 with hydrodynamics (hereafter referred to as ED2; Medvigy *et al.*, 2009; Powell *et al*., 2018) to run the TRS at BCI. In prior work ED2 was initialized from bare ground at BCI and run with recycled 2008-2014 observed meteorology (i.e. "BASE") until predictions of above ground biomass (AGB) reached dynamic equilibrium after 700 years (see Powell *et al.* (2018), Fig. 3; referred to as simulation year 0 in the simulations presented here). After ED2's spin-up period, its predictions of forest demography were benchmarked against observations of aboveground biomass, size-dependent basal area and tree mortality from the FDP and a series of hydroclimate scenarios were run (discussed below). The TRS requires approximately 4 years of its own spin up before the seed bank and seedling pool come into dynamic equilibrium with mature tree productivity and environmental conditions, so we began TRS evaluations 705 years after ED2 was initialized from bare ground. ED2's output provided the TRS with top of canopy (TOC) solar radiation (W m^{-2}), SMP (we used 6 cm below the surface), and each cohort's dbh and number density. ED2's history of C_{g+r} was not saved, so we used its history of carbon allocated to reproduction, along with knowledge of its relatively simple RA scheme, to back-calculate C_{g+r}. Similar to an offline "one-way-coupled" model setup (Forrest *et al.*, 2020), the TRS's predictions of recruitment were not passed back to ED2.

To test the performance of the TRS, we simulated 15 years of BASE meteorology in a patch with 2% of the TOC light to match average conditions in the understory of the BCI FDP (Rüger *et al.*, 2009), converting solar irradiance to PAR using a conversion rate of 0.45 (García-Rodríguez *et al.*, 2020). We compared predictions of recruitment rates into the 1 cm size from ED2 and the TRS to recruitment from census observations at the BCI FDP between 2005 and 2015 (Condit *et al.*, 2019) because these census intervals overlapped with the observed meteorology used for the ED2 simulations. Following Powell *et al*. (2018), we calculated recruitment only for species that can reach a stature ≥ 20 cm dbh and used expert knowledge (see SI Table S1) to exclude any remaining understory specialists. The FDP census intervals are ~5 years, so we estimated recruitment rates accounting for mortality undetected before individuals' first census following established methods (see Eqn 11 in Kohyama *et al.*, 2018). To compare the TRS's predictions of recruitment to ED2's, we emulated ED2's current recruitment subroutine in *R* which provided more flexibility in making comparisons among scenarios and allowed us to use a value for Z_0 (carbon required to build a new recruit) that matched FDP observations. ED2 allocates a fixed

fraction of C_{g+r} to reproduction (F_{repro}; 0.1) and a seedling mortality parameter ($M_{seedling}$; 0.09 day^{-1}) further reduces the carbon available for recruitment.

To evaluate predictions of PFT-specific recruitment across a range of understory light environments, we ran the TRS with 20 years of BASE meteorology in 20 different patches where light at the seedling layer varied between 1 and 100% of TOC solar radiation. To evaluate PFT-specific recruitment responses to variable soil moisture we ran 20-yr simulations in 20 different patches where seasonal patch-level SMP varied from observed soil moisture to unrealistically dry conditions (mean annual SMP ranged from -0.25 to -2.5 MPa), but where light was constant (2% of the TOC), and all patches experienced BASE meteorology. Therefore, soil moisture was decoupled from light conditions in these simulations to demonstrate recruitment responses to soil moisture gradients that may arise from topographic variation within a site.

In addition to BASE meteorology we ran the TRS with ED2 output from three previously published hydroclimate scenarios: 1) a synthetic "El Niño" time series which includes two exceptionally strong droughts, based on observed precipitation during the 1982/83 El Niño at BCI (Powell *et al.*, 2017; Powell *et al.*, 2018 Note S2), within 30 years (the "SYN-ENSO" scenario), 2) a "WET" scenario where precipitation increased 30% compared to BASE precipitation, and 3) a "DRY-DS" scenario where dry season (January−April) precipitation was reduced 75% compared to BASE (Powell *et al.*, 2018). For each of these scenarios, we ran the TRS for 30 years and recorded annual PFT-specific recruitment rates. Each precipitation scenario changed soil moisture (Fig. S3) and mature tree productivity (Fig. 3), thereby changing C_{g+r} used by the TRS. ED2 predicts that the ST PFTs have a greater share of AGB than the LD PFTs at BCI (Powell *et al.*, 2018), but the relative share of total forest NPP is more variable among the PFTs over time, and the LD-DT PFT often accounts for the greatest share of NPP (Fig. 3).

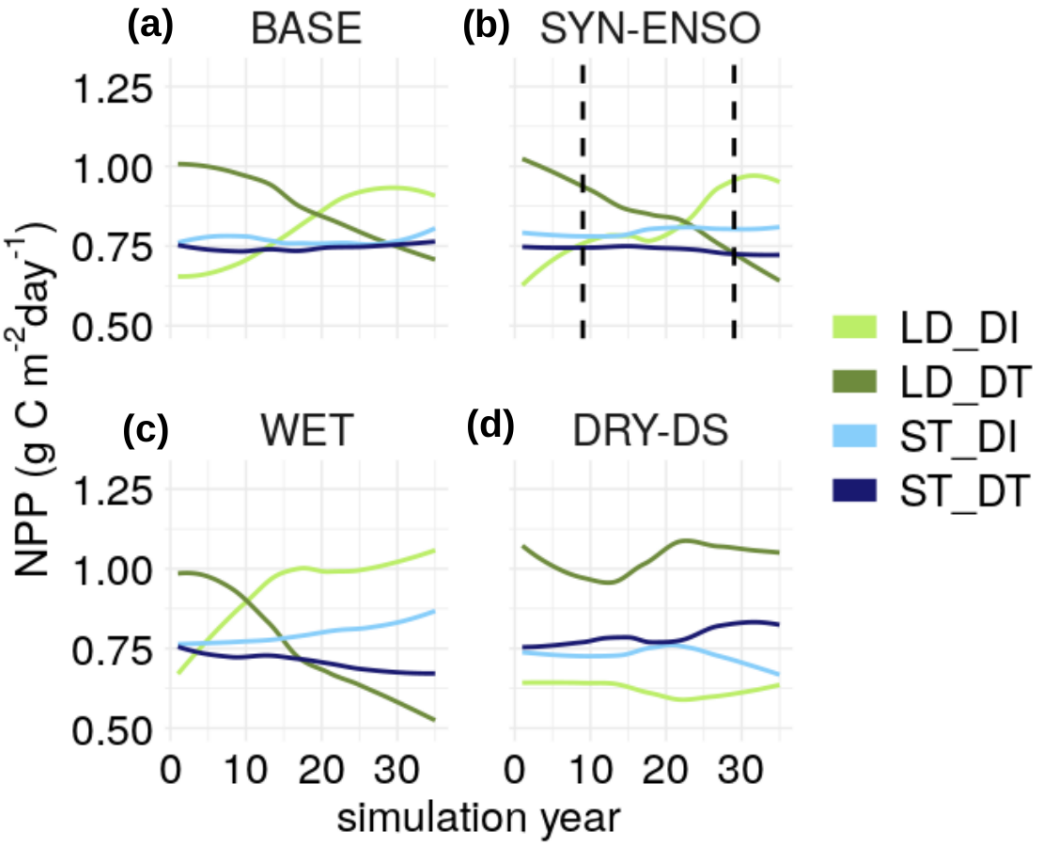


Figure 3. Net primary productivity (NPP) predicted by ED2 under four precipitation scenarios used to run the TRS. BASE = recycled 2008-2014 observed meteorology (a); SYN-ENSO = two exceptionally strong El Niño events within 30 years (b); WET = 30% increase in precipitation compared to BASE (c); DRY-DS = dry season (January−April) precipitation reduced by 75% compared to BASE (d). Dashed lines indicate El Niño events. Note: NPP in the BASE scenario reached dynamic equilibrium (Fig. S4) despite the directional trends apparent over this relatively short 30-yr time period. Lines have a LOESS smoother for easier interpretation of rank order and trends among PFTs. Published estimates of observed PFT-level NPP are not available, but the predicted mean NPP over this period, summed over the four PFTs (2.8 g C m^{-2} day^{-1}), is 15% lower than empirical estimates of total ecosystem-level NPP (3.3 g C m^{-2} day^{-1}; Martínez Cano *et al.*, 2020, Running *et al.*, 2015). LD-DI = light demanding, drought intolerant; LD-DT = light demanding, drought tolerant; ST-DI = shade tolerant, drought intolerant; ST-DI = shade tolerant, drought tolerant.

Parameterization for BCI

We parameterized the TRS with the same four tropical tree PFTs used in Powell *et al.* (2018) which differ along axes of shade tolerance and drought tolerance: 1) light demanding, drought intolerant (LD-DI), 2) light demanding, drought tolerant (LD-DT), 3) shade tolerant, drought intolerant (ST-DI), and 4) shade tolerant, drought tolerant (ST-DT). To calculate PFT-level

recruitment benchmarks from species-level observations we used a wood density threshold whereby species above 0.49 g cm^{-3} were assigned to the ST PFT and species below this threshold were assigned to the LD PFT (Powell *et al.* 2018). We categorized species as DT or DI based on three observational datasets including a manipulative drought experiment (Engelbrecht & Kursar, 2003; Engelbrecht *et al.*, 2007) and occurrence probabilities across aridity gradients at the site scale (Harms *et al.*, 2001) and regional scale (Condit *et al.*, 2013; SI Methods S2, Table S2). All parameter values used for these simulations are shown in Table 1. We tested model sensitivity to all parameters by increasing each parameter value by 10% above the default values and evaluating the corresponding change in predicted recruitment rates.

Table 1. Tree Recruitment Scheme default parameterization for Barro Colorado Island (BCI), Panama grouped by regeneration processes shown in Fig. 1; see SI Methods S1 for parameter derivations; PFT = plant functional type; C_{g+r} = carbon for growth and reproduction; RA = reproductive allocation; dbh = diameter at breast height; PAR = photosynthetically active radiation; MDD = Moisture deficit days; LD-DI = light demanding, drought intolerant; LD-DT = light demanding, drought tolerant; ST-DI = shade tolerant, drought intolerant; ST-DI = shade tolerant, drought tolerant.

Name	Value	Units	Description	Derivation / Source
Allocation to reproduction				
F_{repro}	0.1 (all PFTs)	-	Fraction of C_{g+r} allocated to reproduction	Smith *et al.* (2001), Fisher *et al.* (2015) based on Harper (1977)
a_{RA}	LD-DI: 0.0058 LD-DT: 0.0059 ST-DI: 0.0042 ST-DT: 0.0049	Δ RA $[\Delta$ dbh$]^{-1}$	Governs RA as function of dbh (logit function coefficient)	Logistic regression fit to observations by Wright *et al.* (2015), Visser *et al.* (2016)
b_{RA}	LD-DI: -3.1380 LD-DT: -2.4607 ST-DI: -2.6518 ST-DT: -2.6171	-	Governs RA as function of dbh (intercept in logit function)	Logistic regression fit to observations by Wright *et al.* (2015), Visser *et al.* (2016)
Allocation to seed vs. non-seed reproductive biomass and seed mortality				
F_{seed}	0.5 (tuned to 0.24) (all PFTs)	-	Fraction of reproductive C that is seed	Based on range cited by Wenk *et al.* (2017)
S_{mort}	0.0014 (all PFTs)	day^{-1}	Seed mortality rate	Fisher *et al.* (2015) based on Lischke *et al.* (2006)
Seedling emergence				
a_{emerg}	0.0003	day^{-1}	Coefficient for seedling emergence rate	Calculated from Pearson *et al.* (2002)
b_{emerg}	LD-DI: 1.6 LD-DT: 1.6 ST-DI: 1.2 ST-DT: 1.2	-	Seedling emergence sensitivity to soil moisture	Calibrated to observations of seasonal seedling emergence (Garwood, 1983)
W_{emerg}	14 (all PFTs)	days	Time window for emergence response to soil moisture	Observations of seasonal seedling emergence (Garwood, 1983)
Ψ_{emerg}	-0.15745	MPa	Soil moisture	This study, see SI Methods S1

			required for emergence	
PAR_{crit}	0.656	MJ m^{-2} day^{-1}	Critical PAR level for light-sensitive germination	Based on observations of mean irradiance (PAR) in small gaps ($25m^2$) at BCI (Pearson *et al.*, 2002)
Moisture and light-sensitive seedling survival				
$M_{background}$	LD-DI: 0.17 LD-DT: 0.18 ST-DI: 0.19 ST-DT: 0.11	yr^{-1}	Background seedling mortality rate	Calculated from seedling censuses done at BCI (2003-2012); Johnson *et al.* (2017)
Ψ_{crit}	DI: -0.176 DT: -0.252	MPa	Seedling moisture stress threshold	Based on observations from Engelbrecht & Kursar (2003), Engelbrecht *et al.* (2005)
MDD_{crit}	DI: 46 DT: 14	-MPa days	Moisture deficit day threshold for seedling mortality	Based on observations from Engelbrecht & Kursar (2003), Engelbrecht *et al.* (2005)
a_{Ψ}	DI: 1.04E-16 DT: 4.07E-17	-	Moisture-based mortality coefficient	Based on observations from Engelbrecht & Kursar (2003), Engelbrecht *et al.* (2005)
b_{Ψ}	DI:-5.5E-10 DT:-6.4E-11	-	Moisture-based mortality coefficient	Based on observations from Engelbrecht & Kursar (2003), Engelbrecht *et al.* (2005)
c_{Ψ}	DI:3.5E-04 DT:1.3E-05	-	Moisture-based mortality coefficient	Based on observations from Engelbrecht & Kursar (2003), Engelbrecht *et al.* (2005)
W_{Ψ}	126 (all PFTs)	days	Rolling window for MDD	Based on observations from Engelbrecht & Kursar (2003), Engelbrecht *et al.* (2005)
a_{ML}	LD:-0.033 ST:-0.00990	-	Light-based mortality coefficient	Based on observations by Kobe (1999)
b_{ML}	LD:-3.84 ST:-7.15	-	Light-based mortality coefficient	Based on observations by Kobe (1999)
W_L	64 (all PFTs)	days	Rolling window for seedling light response	Based on observations by Augspurger (1984)
Recruitment				
a_{TR}	LD:0.010 ST:0.007	-	Seedling to sapling transition rate coefficient	Derived from parameters used in CLM(ED) and FATES; Fisher *et al.* (2015)
b_{TR}	LD: 1.0653 ST: 0.8615	-	Recruitment light response parameter	Based on Ruger *et al.* (2009)
Convert carbon to number density of recruits				
Z_0	160 (~1 cm dbh; all PFTs)	g C	C per new recruit	Based on value used in ED2 (Powell *et al.* 2018)

Results

Benchmarking recruitment at BCI

Compared to predictions from the default ED2 formulation, the new TRS improves predictions of recruitment magnitude and rank order across all PFTs using default parameters (i.e. no parameter tuning; Fig. 4). ED2 incorrectly predicts that recruitment is dominated by the LD-DT PFT while the TRS correctly predicts that the ST-DT PFT dominates recruitment. Manually adjusting parameters that control the amount of reproductive carbon available for recruitment can further improve biases in the magnitude of recruitment rates for both models compared to their default parameter set. For example, increasing ED2's seedling mortality parameter from 0.09 to 0.986 day^{-1} and reducing F_{seed} in the TRS by half across all PFTs improves predictions of recruitment magnitude in both models (tuned values are used for subsequent results). However, parameter adjustments do not improve ED2 predictions of the rank order of recruitment rates, nor do they address its environmental insensitivity (discussed below). Although the TRS correctly predicts that ST recruitment is greater than LD recruitment it incorrectly predicts that LD-DT recruitment is greater than LD-DI recruitment.

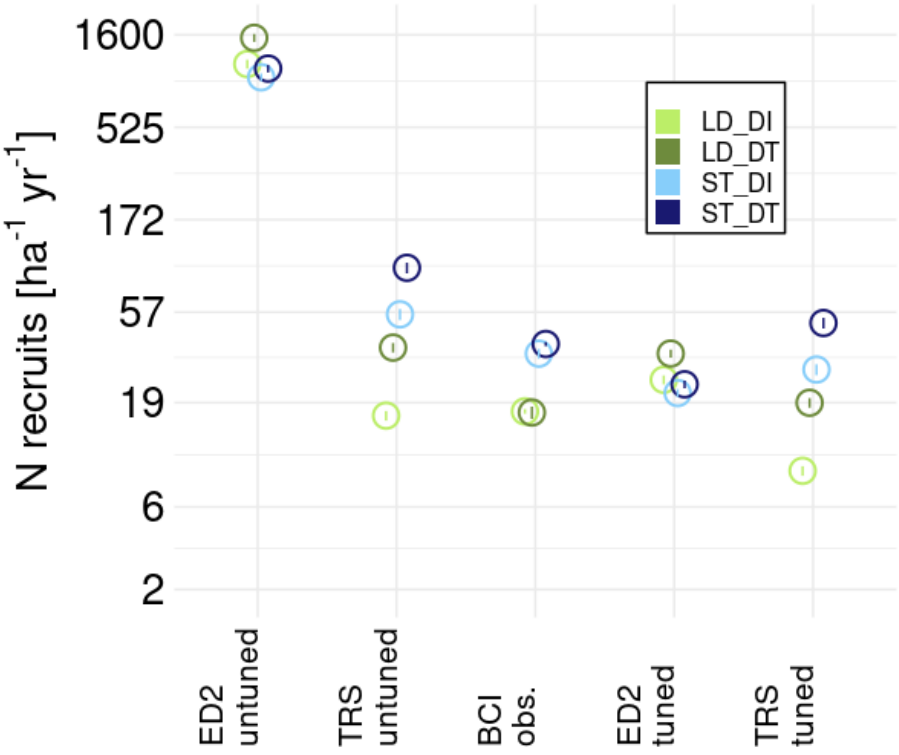


Figure 4. Observed mean annual tree recruitment (at 1-cm dbh) for four plant functional types (PFTs) at BCI (center) compared with model predictions under 20 yr of observed meteorology (2008–2014 recycled) using default and tuned parameter values for the Tree Recruitment Scheme (TRS) and Ecosystem Demography model v.2 with hydrodynamics (ED2). 'TRS tuned' and

'ED2 tuned' are predictions after reducing the TRS's F seed parameter by half and increasing ED2's M seedling parameter from 0.094 to 0.986. 'BCI obs.' are PFT-level mean annual observed recruitment rates averaged from the two 5-yr census intervals between 2005 and 2015 in the BCI Forest Dynamics Plot.

Sensitivity to variable light and soil moisture

Recruitment in ED2 is insensitive to understory light (Fig. 5b). In contrast, the TRS predicts PFT-specific, light-sensitive recruitment responses to varying understory light (Fig. 5a). All PFTs in the TRS show recruitment increasing with light and this variation is strongest for the LD PFTs. Even under high light the LD-DI PFT recruits at a relatively low rate because it comprises a smaller share of total forest NPP under BASE meteorology (Fig. 3a), demonstrating the TRS's dual sensitivity to seedling layer conditions and adult productivity.

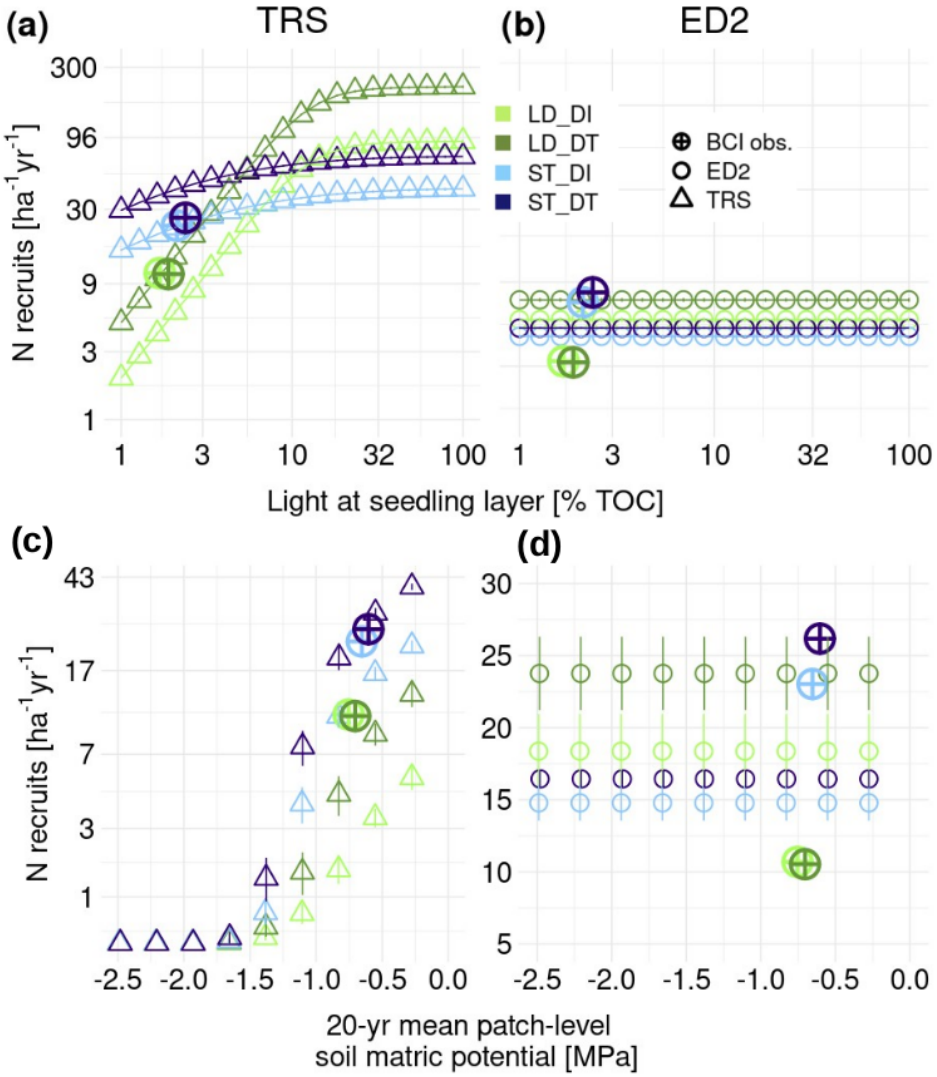


Figure 5. Predictions of tree recruitment (at 1 cm dbh) across a range of idealized patch-level light (a,b) and soil moisture (c,d) conditions with local observed meteorology (2008-2014) at

Barro Colorado Island. Observed PFT-specific mean annual recruitment rates ("BCI obs.") are shown for reference and were calculated from 2005–2010 and 2010–2015 census intervals. Observed means are plotted at light levels equal to the mean understory light level across all patches in the Forest Dynamics Plot (Rüger *et al*., 2009) (a,b) and soil moisture equal to the mean measured at the BCI Lutz catchment between 2008 and 2014 (Paton, 2019) (c,d). Error bars show the interannual variation in recruitment within each patch (SD); in many cases these are smaller than the symbol size. TOC = top of canopy.

Among low light patches (2% TOC) that vary in soil moisture (Fig. 5c,d), the TRS predicts PFT-specific, moisture-sensitive recruitment responses (Fig. 5c). All PFTs show complete recruitment failure when the 20-yr mean SMP reaches -2 MPa (Fig. 5c). Recruitment drops faster for the more vulnerable DI PFTs. When soil moisture is high the ST-DI PFT recruits better than the LD-DT PFT despite a much lower share of NPP, reflecting its ability to recruit in low light conditions when moisture is not limiting. However, when SMP drops below -1.0 MPa, soil moisture limits recruitment more than light and all PFTs show minimal recruitment.

Recruitment predictions under ENSO and WET precipitation scenarios

Predictions of recruitment (at 1 cm dbh) responses to ENSO and increasing precipitation (i.e. WET scenario) differ between the TRS and ED2. ED2 and the TRS both show a sharp reduction (~30%) in recruitment for all PFTs across two synthetic El Niño events, but the TRS slightly buffers this reduction for DT PFTs (Fig. 6a,b). A closer look at recruitment dynamics during an El Niño year (Fig. 6c,d) shows that ED2 and the TRS differ in the duration of recruitment failure for DT and DI PFTs. In ED2, drought stress in the adult cohorts causes recruitment to decline in proportion to NPP and then gradually ramp back up as NPP becomes positive at the end of the El Niño event. Recruitment in the TRS only fails when SMP drops below PFT-specific Ψ_{crit} and seedlings persist (but don't recruit) through the dry season, allowing them to take advantage of wetter conditions with a strong pulse of recruitment at the end of the El Niño dry season (Fig. 6c). This aligns with observations of sapling recovery after the strong 1982/83 El Niño (Condit *et al.*, 2017). The TRS's environmentally sensitive seedling pool changed the transient recruitment response to El Niño.

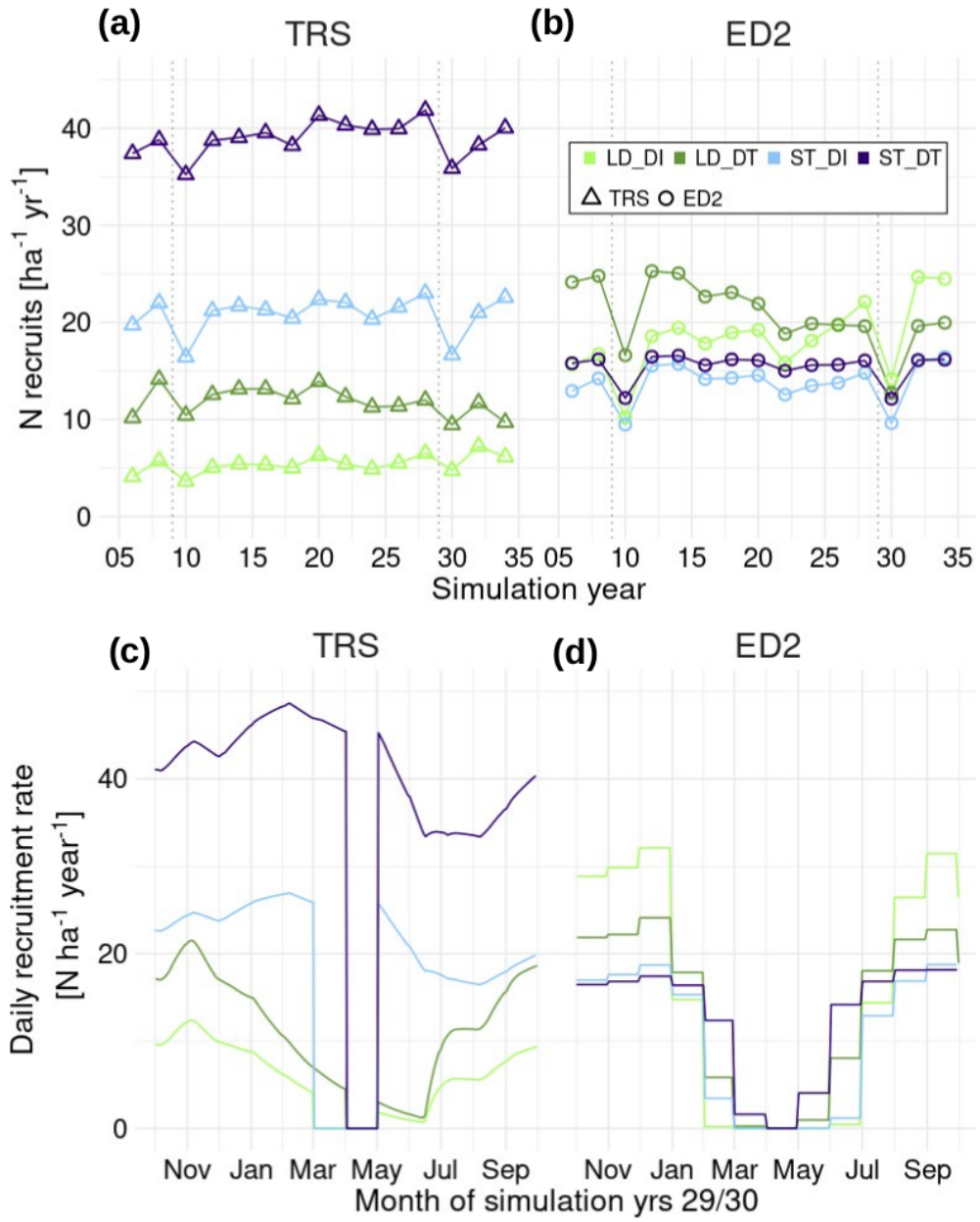


Figure 6. Predictions of tree recruitment (at 1 cm dbh) in a low light patch, 2% top of canopy (TOC) solar radiation, under 30 yrs of local observed meteorology with a strong synthetic El Niño event every 20 years (SYN-ENSO) (a,b). Dotted lines indicate the El Niño years. Seasonal predictions of tree recruitment in a patch with 2% TOC light across the El Niño in simulation years 29/30 are shown in panels c and d. Note that the ED2 output has a monthly timestep. Soil matric potential is below -2 MPa during March, April, and May (see Fig. S5 for more details).

Under the WET scenario the LD-DI PFT starts to dominate total forest NPP (Fig. 3), which results in different recruitment responses between ED2 and the TRS. ED2 predicts a

corresponding, immediate increase in LD-DI recruitment regardless of the light environment (Fig. 7b,d), reflecting its insensitivity to patch-level light. In contrast, the TRS does not allow the LD-DI PFT to dominate recruitment in low light patches (Fig. 7a), because the light component of its regeneration niche is not met despite its increasing share of NPP (Fig. 3). However, under high light the TRS predicts a significant increase in LD-DI recruitment (Fig. 7c), similar to ED2. Under the DRY-DS scenario both models predicted lower recruitment rates, but differ in the timing of recruitment declines (see SI Note 1 and Fig. S2 for more details on the DRY-DS precipitation scenario).

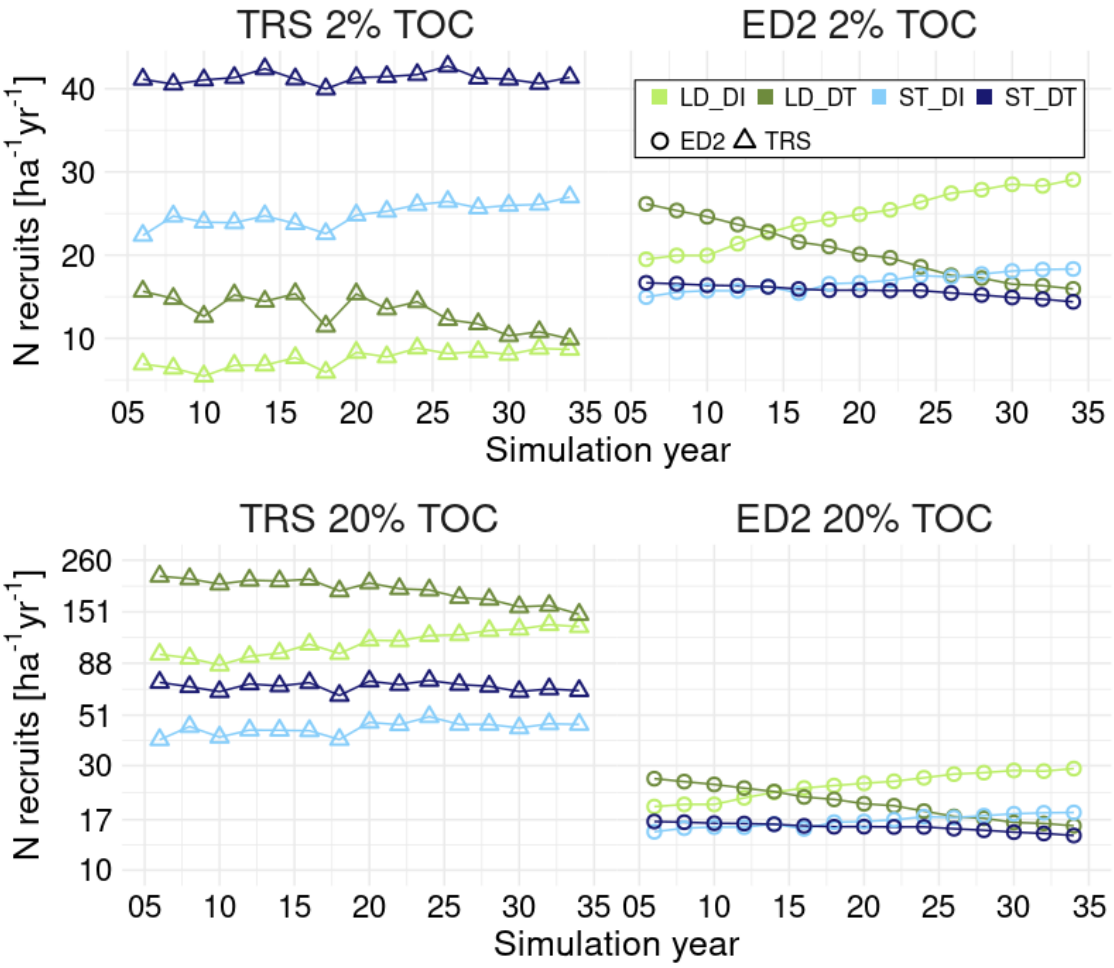


Figure 7. Predictions of tree recruitment (at 1 cm dbh) under 30 yrs of the WET scenario, a 30% increase in precipitation compared to baseline. Predictions are shown under 2% top of canopy (TOC) light (a,b) and 20% TOC light (c,d).

Parameter sensitivity

The light response parameter for seedling mortality (b_{ML}) has the greatest leverage on recruitment outcomes (Fig. 8), which is expected given that light is a key limiting resource in mature tropical forest understories. The remaining top four parameters with the most leverage on recruitment are part of the upstream reproductive allocation scheme (Fig. 8, Table 1). Parameters governing

moisture stress mortality, emergence, and transition rates do not show leverage because the parameter sensitivity simulations were run under BASE meteorology where soil moisture does not fall below the observed wilting thresholds (Engelbrecht *et al.*, 2007) required to trigger moisture-based mortality responses.

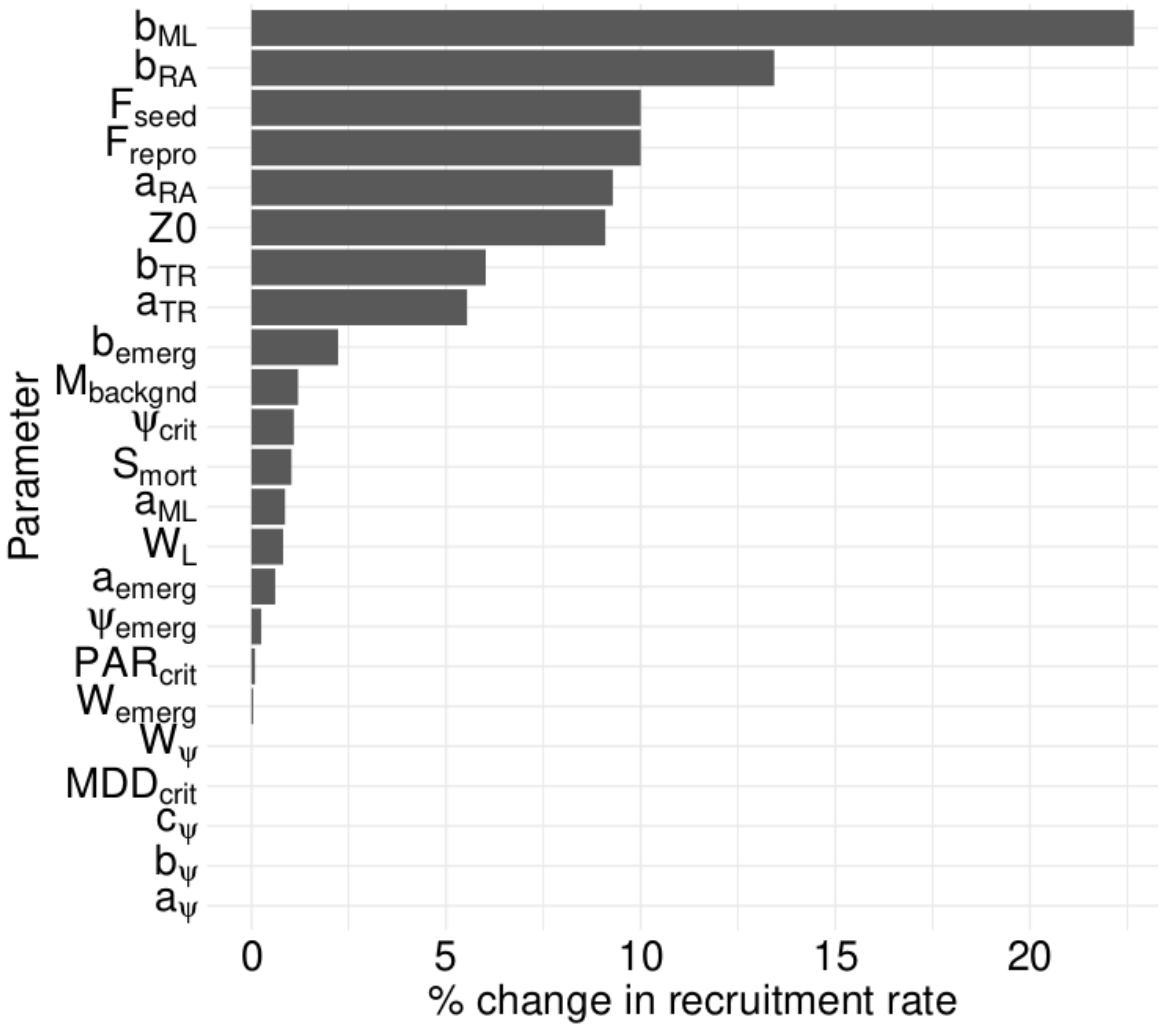


Figure 8. Recruitment sensitivity to a 10% change in parameter value (see Table 1 for parameter descriptions).

Discussion

The TRS predicts tree recruitment as a function of understory light, soil moisture, and productivity of reproductively mature cohorts. We evaluated the TRS by parameterizing it for a seasonally dry tropical forest and running simulations under observed meteorology and canopy structure, variable patch-level understory light and soil moisture, and with three altered precipitation scenarios (ENSO, wetter-than-observed, and drier-than-observed). The TRS improves upon ED2 predictions by capturing recruitment sensitivity to light and soil moisture, and allowing more realistic recruitment responses to environmental heterogeneity without the computational cost of simulating seedling cohorts or individuals explicitly.

Benchmarking recruitment at BCI

We attribute the TRS's improved predictions of recruitment magnitude and PFT rank order (Fig. 4) to PFT-specific, light-sensitive germination, seedling mortality from light stress and light-sensitive seedling to sapling transition rates. Despite its dominant share of total forest NPP, the LD-DT PFT experienced seedling mortality rates almost 13 times greater than the ST PFTs in BCI's understory, resulting in recruitment predictions more consistent with quantitative observations at the BCI FDP (Fig. 4). The TRS's prediction that recruitment is dominated by ST PFTs is consistent with ecological expectations in a mature, closed canopy forest where shade tolerant PFTs should be favored during community assembly (Comita & Hubbell, 2009; Lebrija-Trejos *et al.*, 2010; Wright *et al.*, 2010). In contrast to observations, the TRS predicts that LD-DT recruitment is twice as high as LD-DI (Fig. 4) which could be due to erroneous species assignments to DT and DI PFTs or because we did not simulate observed spatial heterogeneity in patch-level soil moisture at BCI, such as occurs along hillslopes (Becker *et al.*, 1988), or due to soil moisture and NPP biases predicted by ED2.

Sensitivity to variable light and moisture conditions

Light sensitive recruitment in the TRS (Fig. 5) is consistent with prior experimental evidence and ecological expectations showing that LD species recruit at higher densities under brighter conditions, such as in light gaps or forest fragments (Brokaw, 1985; Dupuy & Chazdon, 2008; d'Oliveira & Ribas, 2011). These predictions primarily emerge within the TRS from light-sensitive seedling mortality and seedling-to-sapling transition rates, but the size and productivity of adult cohorts are also influential. For example, the LD-DT PFT is particularly dominant among recruits at high light because of its large share of total forest NPP under BASE meteorology (Fig. 3). PFT-specific responses to increasing light are important for representing gap phase dynamics (Brokaw, 1985) and because disturbance rates are increasing with climate change (Turner, 2010; Seidl *et al.*, 2017), likely chronically increasing understory light. Therefore, incorporating the TRS into VDMs will improve predictions of PFT-specific light responses that mediate functional turnover.

The TRS captured PFT-specific, moisture-sensitive recruitment limitations (Fig. 5), which we attribute to the inclusion of moisture-sensitive seedling emergence, seedling mortality, and seedling to sapling transition rates. The DT PFTs maintained very low recruitment at -2 MPa of SMP due to a more negative (i.e. dry) Ψ_{crit} value, allowing them to accumulate less moisture deficit days than the DI PFTs under dry conditions. Capturing moisture-sensitive recruitment limitations is important for predicting PFT distributions because soil moisture varies dramatically throughout large parts of the tropics and influences differential seedling survival and mature forest composition (Engelbrecht *et al.*, 2007; Condit *et al.*, 2013). By more accurately reflecting the early life-stage at which moisture-based environmental filtering is believed to take place (Engelbrecht *et al.*, 2007), the TRS will enable VDMs to more mechanistically predict functional turnover across topographic and regional moisture gradients and in response to changing precipitation regimes (Martínez-Vilalta & Lloret, 2016).

The TRS's representation of moisture-sensitive recruitment is more consistent with ecological expectations and observations (Engelbrecht *et al.*, 2007) compared to ED2, which assumes that

recruitment rates are insensitive to SMP (Fig. 5). Enabling PFT-specific, moisture-sensitive recruitment rates is helpful for VDMs that already have a sophisticated representation of mature plant hydraulics because the understory is typically more humid than the canopy (Fetcher *et al.*, 1985), which can help seedlings survive drought in some contexts (Gómez-Aparicio *et al.*, 2008; Andivia *et al.*, 2018). Conversely, their shallower root systems can leave them more vulnerable to wide fluctuations in soil moisture occurring near the soil surface (Brum *et al.*, 2018). Representing these aspects of the regeneration niche will help VDMs predict forest composition in response to simultaneous but potentially different hydrological conditions experienced by the canopy and the seedling layer.

Responses to ENSO and WET precipitation scenarios

The TRS's more complete representation of the regeneration niche creates novel predictions of how recruitment of 1-cm dbh saplings will respond to varying meteorological scenarios. Unlike ED2, the TRS allowed DT sapling recruitment to continue into March (Fig. 6c) until SMP dropped below the DT Ψ_{crit} value, thereby capturing PFT-specific responses to seedling layer conditions under El Niño. It also captures a large pulse of recruitment from a persisting seedling pool when the El Niño dry season ends (Fig. 6c). The TRS's ability to capture PFT-specific recruitment responses to El Niño is consistent with observations showing that some seedlings persist through severe El Niños (Engelbrecht *et al.*, 2002), some suffer increased mortality (Gilbert *et al.*, 2001), and that recruitment responses vary by PFT (Slik, 2004). The TRS's temporal decoupling of adult productivity from seedling dynamics allows understory light and soil moisture to drive the transient recruitment response instead of remaining directly proportional to adult NPP as occurs in ED2 (Fig. 6d). Decoupling recruitment from adult NPP by including a seed bank and a seedling pool reproduced observed time lags between seed production and seedling recruitment (Garwood, 1983; Wright & Muller-Landau, 2005) as well as time lags between seedling emergence and recruitment into the 1 cm size class (Chang-Yang *et al.*, 2021). However, the residence time of surviving carbon in the seedling pool, 737 days, may still be shorter than the time it takes for many seedlings in closed canopy tropical forests to recruit (Chang-Yang *et al.*, 2021). Many tropical forest trees lose their leaves in response to El Niño (Detto *et al.*, 2018) letting more light into the understory, which is observed to simultaneously increase understory recruitment and mortality rates depending on functional traits such as wood density (Slik, 2004). The TRS is better positioned to predict these PFT-specific recruitment responses to changing understory conditions as El Niño events become more frequent and severe (Haszpra *et al.*, 2020).

Recruitment predictions under the WET scenario demonstrate the TRS's dual sensitivity to changes in adult productivity and seedling layer conditions. As precipitation increases under the WET scenario the LD-DI PFT dominates total forest NPP (Fig. 3), but the TRS only allows this to translate into recruitment dominance under higher light (Fig. 7c). Conversely, under low light, the TRS only allows recruitment to increase slightly in response to increased LD-DI propagule pressure; low light stops it from dominating recruitment (Fig. 7a). This is consistent with theory and observations that recruitment can be limited by both propagule pressure and environmental filtering during establishment (Jabot *et al.*, 2008). The TRS's dual sensitivity to adult productivity and seedling layer conditions also means that it will provide the host VDM with a

seedling layer composition more reflective of forest composition and understory conditions in any given timestep. This is significant because when a disturbance occurs, pulses of recruitment arise from the existing seedling layer, thereby influencing the composition of trees recruiting into the canopy (Brokaw, 1985).

Compatibility with VDMs

The TRS is positioned to improve VDMs by making the representation of tree recruitment more mechanistic while maintaining a computationally efficient approach. It tracks pools of carbon instead of individuals or cohorts, regeneration processes are predicted as a function of variables that are already tracked by VDMs, and carbon is conserved. The TRS can be run at any site where a host VDM can run, which at a minimum requires local meteorological data and a PFT parameter set. Predictions of recruitment can be evaluated most readily at long term forest dynamics plots with repeated census events that track all trees down to the 1 cm dbh size class (e.g. the CTFS-ForetsGeo plot network; Davies *et al.*, 2021), but once the TRS is coupled to a host VDM it will be possible to evaluate recruitment into larger size classes. In sum, the TRS's formulation is designed to be compatible with existing VDM model architecture and observational data.

The TRS minimizes the introduction of new, sensitive parameters that are hard to empirically constrain. Four out of the five most sensitive TRS parameters are part of the size dependent reproductive allocation (RA) function (Fig. 8, Eqn 1), which is expected because RA determines the amount of carbon flowing to all subsequent processes (Fig. 1). F_{repro} already exists in VDMs and it is possible to use existing litter fall data to empirically constrain it in tropical forests (Hanbury-Brown *et al.*, 2022). F_{seed} has not been quantified at the ecosystem or PFT level, but is measurable from field observations (Wenk *et al.*, 2017), highlighting the need and opportunity to quantify it at VDM-relevant scales. The probabilistic relationship between size and RA (governed by a_{RA} and b_{RA}) can be derived from logistic regression applied to readily available observations of dbh and reproductive status (e.g. Visser *et al*., 2016). The light response parameter for seedling mortality (b_{ML}) can be derived from observations of seedling mortality under experimentally manipulated light (e.g. Augspurger, 1984; Kobe, 1999; Balderrama & Chazdon, 2005). With the exception of a_{TR} (discussed below), the remaining parameters are observable and have less leverage on recruitment, but would still need to be quantified or synthesized to run the TRS in extra-tropical biomes or with new PFTs.

Limitations and future work

Evaluating the TRS in a one-way coupled configuration (i.e. offline from a host VDM) allowed us to evaluate the TRS's behavior in a reduced complexity environment without idiosyncratic host model feedbacks that are hard to diagnose. However, until the TRS is fully integrated into a VDM we can't assess how feedbacks between recruitment and adult demographics will influence predictions of future forest composition. We avoided simulations longer than 30 years for this reason. Secondly, the TRS only includes the first order processes limiting tree recruitment within tropical forest patches. We did not address dispersal because it is an inter-patch process that should be implemented directly in VDMs in a way that is congruent with how each model abstracts spatial processes. The representation of PFT-specific inter-patch dispersal limitation (which could be based on traits such as seed size) will likely alter the recruitment rates presented

here because more dispersal limited PFTs will be less likely to reach new high resource patches. The TRS also does not address vegetative propagation and post-disturbance resprouting, known to be an important regeneration strategy (Dietze & Clark, 2008; Clarke *et al.*, 2013), because this process should be embedded within each VDM's storage allocation scheme. Despite these limitations, testing and presenting the TRS as an offline module is consistent with recent calls from the Earth system modeling community for "modular complexity as a strategy" whereby new functionality is organized as modules to mitigate the intractability of increasingly complex ESM components (Fisher & Koven, 2020).

Following the need to balance process fidelity with manageable complexity and computational cost (Fisher & Koven, 2020), we focused on a first order set of processes that we hypothesize is required to capture future recruitment limitations in tropical forests under climate change. However, the TRS omits processes that may require additional consideration. Pathogen attack (Spear *et al.*, 2015) and herbivory (Weissflog *et al.*, 2018) are both known to play roles iparameters regulating seedling population dynamics at the species level and susceptibility to these causes of mortality may covary with functional traits (Coley & Barone, 1996; Spear & Broders, 2021). We used a constant seed mortality rate because we lacked the data to parameterize PFT-specific environmentally sensitive seed mortality in tropical forests and the parameter had relatively little leverage on recruitment rates (Fig. 8). Nevertheless, seed mortality rates likely vary among stands, biomes, and PFTs by more than the 10% value used in our parameter sensitivity experiment (Notman & Gorchov, 2001; Obroucheva *et al.*, 2016), indicating the need for further evaluation. Nitrogen availability is believed to influence the functional composition of recruits during secondary forest development (Batterman *et al.*, 2013) indicating that nutrient limitations on recruitment may be important to represent. There is evidence that seed production varies with interannual variation in solar irradiance, precipitation, and stand structure (Wright & Calderón, 2006; O'Brien *et al.*, 2018; Hacket-Pain *et al.*, 2018; Detto *et al.*, 2018; Minor & Kobe, 2019; Andrus *et al.*, 2020) and a representation of these sensitivities may help facilitate the prediction of masting in future models (Vacchiano *et al.*, 2018), but observations of how RA varies with these variables is still missing for most tropical tree PFTs, so we have not included environmentally sensitive RA here. Currently, all seedlings in the TRS access water at the same depth, but seedling rooting depth is a critical functional trait mediating seedling physiological responses to drought (Brum *et al.*, 2018), making this a key area for future data collection, synthesis, and algorithm development. Additionally, manipulative experiments analyzing how vapor pressure deficit and temperature extremes interact with soil moisture (e.g. Will *et al.*, 2013) to affect seedling mortality would improve upon existing drought experiments (e.g. Engelbrecht & Kursar, 2003; Engelbrecht *et al.*, 2007) and may facilitate more robust algorithms for moisture stress mortality.

We did not implement a carbon assimilation scheme for the seedling layer. Instead, following the current ED-based convention (Moorcroft *et al.*, 2001; Fisher *et al.*, 2015) all carbon used to produce new recruits must come from the stock of C_{g+r}. This is done to avoid resolving photosynthesis for many small cohorts of seedlings, but the corresponding loss of process fidelity means that the fraction of seedling carbon that becomes new recruits at mean understory light (calculated using a_{TR}) is not comparable with observations of seedling to sapling (>= 1 cm dbh)

transition probabilities. Future work is required to assess the complexity-fidelity tradeoffs associated with implementing a simple carbon assimilation scheme in the seedling pool to allow for easier model-data intercomparison of this transition. Additional functions, such as temperature-sensitive seedling emergence for extra-tropical forests, could easily be added to the TRS, making the framework presented here (Fig. 1) extensible globally.

Conclusion

The TRS provides a more mechanistic constraint on the amount of carbon available for recruitment within VDMs, thereby improving predictions of recruitment compared to a current VDM. Representing tree recruitment as the outcome of critical regeneration processes sensitive to carbon production, light, and soil moisture, enabled predictions of recruitment in a tropical forest that are more consistent with prior observations and ecological expectations. The core infrastructure of the scheme is simple and versatile, and the scheme's parameter set is designed to be minimal and observable. Parameter estimation may require additional empirical synthesis and/or new observations to constrain outside of well-studied ecosystems. Through its improved representation of the regeneration niche, the TRS is well positioned to advance predictions of future tree recruitment under changing climate and disturbance regimes. This is essential to predicting future forest composition, distribution, and function.

Acknowledgments

A. R. Hanbury-Brown, T. L. Powell, and L. M. Kueppers were 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. We used the ForestGEO benchmarking driver, developed by Dan Johnson and Ryan Knox, to calculate forest demographic rates used for benchmarking the simulations. Seedling census data supplied by Liza Comita and Dan Johnson were used to calculate background seedling mortality rates. The BCI tree censuses were made possible by National Science Foundation grants to Stephen P. Hubbell: DEB-0640386, DEB-0425651, DEB-0346488, DEB-0129874, DEB-00753102, DEB-9909347, DEB-9615226, DEB-9615226, DEB-9405933, DEB-9221033, DEB-9100058, DEB-8906869, DEB-8605042, DEB-8206992, DEB-7922197, and by support from the Forest Global Earth Observatory, the Smithsonian Tropical Research Institute, the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, the Small World Institute Fund, and numerous private individuals, and through the hard work of over 100 people from 10 countries over the past three decades. The plot project is part the Forest Global Earth Observatory (ForestGEO), a global network of large-scale demographic tree plots.

Author Contributions

Conceptualization and Investigation A. R. Hanbury-Brown and L. M. Kueppers; Resources and Data Curation, T. L. Powell, H. C. Muller-Landau, S. J. Wright; Writing-Original Draft, A R. Hanbury-Brown; Writing – Review & Editing, A. R. Hanbury-Brown, L. M. Kueppers, T. L. Powell, H. C. Muller-Landau, S. J. Wright; Software, A. R. Hanbury-Brown.

Data Availability

The Barro Colorado Island Tree Reproduction Dataset is openly available through the Smithsonian Tropical Research Institute at https://doi.org/10.5479/si.data.201511251100 (Wright *et al.*, 2015)

The Barro Colorado Island seedling census data is openly available through Dryad at https://datadryad.org/stash/dataset/doi:10.5061/dryad.fm654 (Visser *et al.*, 2016)

The data used to classify species as canopy trees, midstory specialists, and understory specialists is in Table S1 available from *New Phytologist* at https://doi.org/10.1111/nph.18059.

The species assignments to plant functional types are available in Table S2 available from *New Phytologist* at https://doi.org/10.1111/nph.18059.

The Tree Recruitment Scheme source code and model data output is openly available on github at https://github.com/adamhb/regeneration_submodel

The data used to run the Tree Recruitment Scheme is freely available on Zenodo at https://doi.org/10.5281/zenodo.5498285

Supplementary Information

Methods S1: Derivation of parameters for BCI

See Table 1 in the main text for an overview of all parameter values, definitions, and sources.

Allocation to reproduction
F_{repro} is the fraction of carbon for growth and reproduction (C_{g+r}) allocated to reproduction. We used default values for F_{repro} that are currently used in VDMs (Smith et al., 2001; Sato et al., 2007; Fisher et al., 2015). a_{RA} and b_{RA} govern the shape of the relationship between dbh (mm) and the probability that an individual is reproductive. The Barro Colorado Island (BCI) Tree Reproductive Dataset (Wright et al., 2015; Visser et al., 2016) includes observations of reproductive status for 60 species at the BCI Forest Dynamics Plot (FDP) in 1995, 1996, 1997, 2005, 2006 and 2011 – 2015. We used logistic regression to predict the probability of reproductive status as a function of dbh (mm) for light demanding, drought intolerant (LD-DI), light demanding, drought tolerant (LD-DT), shade tolerant, drought intolerant (ST-DI) and shade tolerant, drought tolerant (ST-DT) plant functional types (PFTs). For all PFTs we used observations of canopy trees (n = 34 species, excluding midstory and understory specialists) which were classified using expert knowledge (see Table S1) based on the minimum size at which the species becomes reproductive. The canopy trees in this dataset comprise > 75 % of canopy tree basal area at the BCI FDP for all PFTs except the LD-DT PFT.

Allocation to seed vs. non-seed reproductive biomass and seed mortality
F_{seed} is the fraction of reproductive carbon that is seed. We used a default value of 0.5 for F_{seed} for all PFTs because it is roughly in the middle of the range cited by Wenk et al. (2017), but this assumption requires further empirical analysis to constrain. For all simulations done here, except the "untuned" benchmarking simulation (Fig. 4 in the main text), we tuned this parameter to be 0.24 to better match the observations of recruitment rates at BCI. S_{mort} is the environmentally insensitive seed mortality rate (yr^{-1}) and is derived from the values used in current VDMs (Smith et al., 2001; Sato et al., 2007; Fisher et al., 2015).

Seedling emergence
a_{emerg} is a coefficient in the seedling emergence function (see Eqn 4 in main text). Pearson et al., (2002) conducted a 60-day germination experiment with seeds collected from BCI. All seeds in this experiment had a nonlimiting moisture supply. We calculated the mean daily germination rate across all species in Pearson *et al.*'s (2002, Fig. 2) "light" treatment. We used this value in combination with Eqn 4 to backcalculate the value of a_{emerg} with moisture equal to the mean at BCI (excluding the dry season). The b_{emerg} parameter is the seedling emergence sensitivity to soil moisture (see Eqn 4 in main text). For b_{emerg} we tried a range of values between 0.6 and 1.6 and used values of 1.6 and 1.2 for the LD and ST PFTs respectively because these values best reproduced observations of seedling emergence at BCI where the pioneer species (i.e. LD species on BCI) show a more pronounced pulse of emergence at the onset of the wet season (see Fig. 7 in Garwood, 1983).

Ψ_{emerg} is the minimum soil matric potential (SMP) required for emergence. To capture observations that seedlings do not emerge in the dry season at BCI (Garwood, 1983), we used the minimum wet season SMP simulated by ED2-hydro under observed meterology (2008-2014) as the value for Ψ_{emerg}. This stops emergence during BCI's dry season.

PAR_{crit} is the critical level of photosynthetically active radiation (PAR) determining the shape of the Michaelis-Menten germination response to diminishing light (see Eqn 3 in main text). Below PAR_{crit} germination rates quickly drop to 0 (see Fig. 2c). Pearson et al. (2002) measured solar irradiance in small gaps (25 m^2) on BCI and found the mean PAR to be ~3 mol m^{-2} day^{-1} which is roughly equivalent to 0.656 MJ m^{-2} day^{-1}. Given that photoblastic germinators (i.e. LD species on BCI) are generally found in canopy gaps, we assume this threshold to function as a suitable value for PAR_{crit}.

Moisture and light-sensitive seedling survival
Ψ_{crit} is the seedling moisture stress threshold at which a PFT starts accumulating moisture deficit days (MDDs; see Eqn 5 in main text). Engelbrecht & Kursar (2003) conducted a manipulative drought experiment to monitor seedling wilting and survivorship under decreasing soil moisture for 28 species of trees on BCI. We used observations of seedling wilting stages from Engelbrecht & Kursar (2003) and Engelbrecht et al. (2005) to determine Ψ_{crit} for DI and DT PFTs. W_Ψ is the moving time window over which MDDs are calculated. We used 126 days for W_Ψ because it matches the time span of Engelbrecht and Kursar's (2003) drought experiment.

$a_{\Psi,}$ $b_{\Psi,}$ c_Ψ are coefficients in the quadratic equation that predicts moisture-based seedling mortality as a function of moisture deficit days (MDD). To determine these parameter values for BCI we first used PFT-level values of Ψ_{crit} (see above) to calculate MDD across Engelbrecht & Kursar's (2003) drought experiment for drought tolerant (DT) and drought intolerant (DI) PFTs. Second, we used Engelbrecht & Kursar's (2003) observations of seedling survivorship in manipulated "drought plots" vs. control plots to calculate the mean daily seedling mortality rate attributable to moisture stress in each census interval. We performed quadratic regression to find the parameters ($a_{\Psi,}$ $b_{\Psi,}$ c_Ψ) that predict moisture-based seedling mortality as a function of accumulated MDD for drought tolerant and drought intolerant PFTs (SI Fig. S1). We excluded census intervals towards the end of the experiment when rain started to rehydrate the plots.

a_{ML} and b_{ML} are parameters for the negative exponential light stress seedling mortality function (Eqn 7 in main text). To derive parameter values for a_{ML} and b_{ML} we used Kobe's (1999) first year light-based seedling mortality models for *Cecropia obtusifolia* and *Trophis racemosa* as a starting point. Kobe's (1999) models predict mortality as a function of relative irradiance at La Selva, Costa Rica. Our model predicts mortality as a function of absolute light (MJ PAR m^{-2} day^{-1}) so we used TOC solar radiation data taken during Kobe's experiment at the La Selva field station (Clark, 2017) along with Kobe's first year mortality models to derive parameters for Eqn 7. Although Kobe (1999) found that mortality plateaued for *C. obtusifolia* above a certain light threshold, mortality under high light was attributed to peccary herbivory, not light stress mortality, so we omitted this plateau in Eqn 7. W_L is the rolling time window over which incident light is calculated for the seedling mortality response. We used a default value of 64 days for W_L

because this was the approximate mean number of days required to see tropical seedling mortality responses to different light levels (see Fig. 1 in Augspurger (1984) and Fig. 1 in Myers & Kitajima (2007).

$M_{background}$ is the background seedling mortality rate which represents all other forms of seedling mortality (aside from light and moisture stress). We calculated $M_{background}$ using the total seedling mortality rate (calculated from prior seedling census data; see Johnson et al., 2017) and then substracted the simulated moisture and light-based seedling mortality rates predicted by the Tree Recruitment Scheme (TRS) under BASE meteorology.

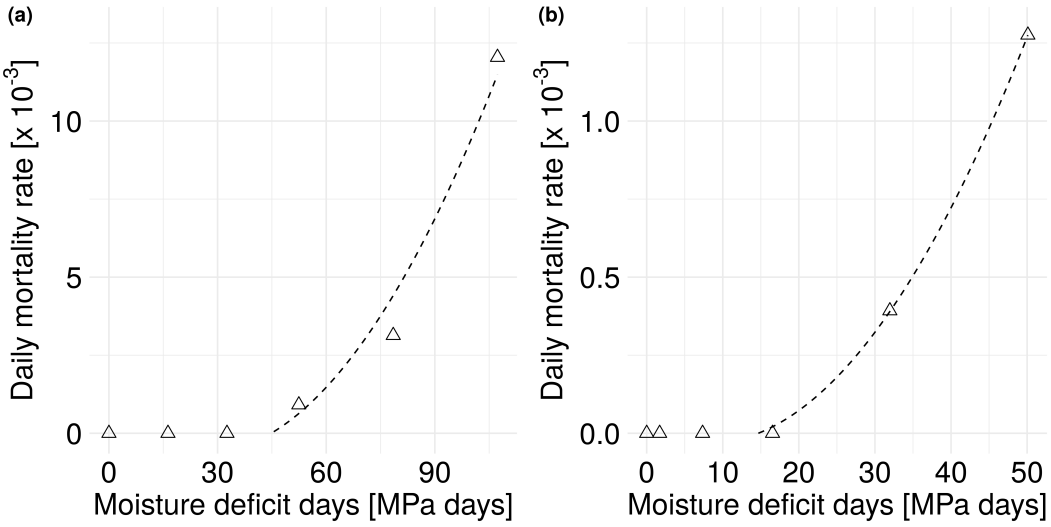
Recruitment

b_{TR} governs the recruitment response to light at the seedling layer. PFT-specific values for b_{TR} were based on the process model component of a Bayesian hierarchical statistical framework applied to observations of spatial variation in recruitment rates on BCI (Rüger et al., 2009). Ruger's model included a light response parameter where values greater than 1 indicated an acceleratnig relationship between increasing light and recruitment, values less than 1 indicated a decelerating relationship, and values less than 0 indicated a negative affect of increasing light on recruitment. We averaged the species-level values (n = 263 species) for this light response parameter to derive PFT-specific valus of b_{TR} for LD and ST PFTs at BCI.

a_{TR} is a coefficient in the seedling to sapling transition rate power function (Eqn 8). Its value was back-calculated using Eqn 8 assuming the mean transition rate was equal to the current value used in VDMs (Fisher *et al.*, 2015) under mean understory light levels at BCI. Note that this parameter is not equivalent to a transition probability because individual seedlings are not tracked.

We used a value of 160 g for Z_0 in both ED2 and the TRS because this is the amount of carbon that ED2 requires to make a new recruit with a dbh of 1 cm, which matches the smallest size class in the FDP census data.

Figure S1: The relationship between observed seedling mortality rates and moisture deficit days (MDDs) based on prior observations made by Engelbrecht & Kursar (2003) and (Engelbrecht et al., 2005). Dashed lines show the quadratic relationship used by the Tree Recruitment Scheme for drought intolerant, DI (a), and drought tolerant, DT (b), plant functional types (PFT). Each PFT accumulates MDDs at different rates based on PFT-specific values of Ψ_{crit} (see Methods S1) so the MDD axes are not directly comparable between panels.



Methods S2. Assigning species to drought tolerant and drought intolerant PFTs

To ensure that as many species as possible could be assigned to drought tolerant and drought intolerant PFTs we used three prior studies, that together, covered 95% of the species in our analysis. First, we used drought sensitivity scores derived from manipulative seedling drought experiments (Engelbrecht & Kursar, 2003; Engelbrecht et al., 2005, 2007) to assign species to PFTs. Engelbrecht *et al.* (2007) calculated seedling drought sensitivity (DS) for 48 species as

$$DS = \frac{S_{irrigated} - S_{dry}}{S_{irrigated}}$$

where $S_{irrigated}$ is survival in an irrigated control plot and S_{dry} is survival in a drought treatment. Species with drought sensitivity scores above the 33rd percentile of their sample of 48 species were assigned to the DI PFT, and species below the 66th percentile were assigned to the DT PFT. For species not included in the seedling drought experiments, we used prior analyses of species habitat associations at BCI (Harms et al., 2001). Species for which stem abundance was previously determined to be statistically positively associated with the relatively dry high plateau on BCI (Harms *et al.*, 2001) were assigned to the drought tolerant PFT, while species negatively associated with the high plateau were assigned to the drought intolerant PFT. For species that were not included in the prior two studies, we used species level occurrence probabilities across the Panamanian aridity gradient (Condit *et al.*, 2013) to calculate their relative moisture optima (dry season precipitation). Species with moisture optima below the 33rd percentile of all species included in the analysis (Condit *et al.*, 2013) were assigned to the DT PFT and species with moisture optima above the 66th percentile were assigned to the DI PFT.

Note S1. The DRY-DS precipitation scenario

Under the drier dry season (DRY-DS) scenario (75 % reduction in dry season rainfall), both ED2 and the TRS showed an increase in the magnitude of interannual fluctuations in recruitment, but these fluctuations are not in sync with each other, and they appear to happen for different reasons (see Fig. S2). The TRS predicts local recruitment minimums towards the end of three-year dry periods and local maximums towards the end of three-year wet periods, demonstrating time lags that emerge from the explicit representation of a seed bank and seedling pool. ED2 has no time lags between dry years and recruitment, instead predicting local maximums in wet years and local minimums in dry years.

The DRY-DS scenario had more years where SMP dropped below Ψ_{crit} for both DI and DT PFTs. This increased the magnitude of interannual recruitment fluctuations as each PFT's regeneration niche became seriously affected in some years, but not others. Capturing these dynamics has the potential to more realistically represent population changes as a transient process (Hastings et al., 2018), rather than an abrupt threshold such as when a bioclimatic envelope is suddenly transgressed by chronic change. The TRS's ability to capture recruitment fluctuations in response to interannual climatic variation may even facilitate coexistence within VDMs through storage effects (Chesson & Warner, 1981) which are believed to maintain coexistence when reproductively mature individuals can persist long enough to take advantage of a transient regeneration niche.

Figure S2. Predictions of tree recruitment in a low light patch (2 % top of canopy radiation) under 30 yrs of the DRY-DS precipitation scenario where the dry season (January–April) has 75% less precipitation compared to baseline. Dotted lines indicate dry years where mean annual SMP was below -0.5 MPa.

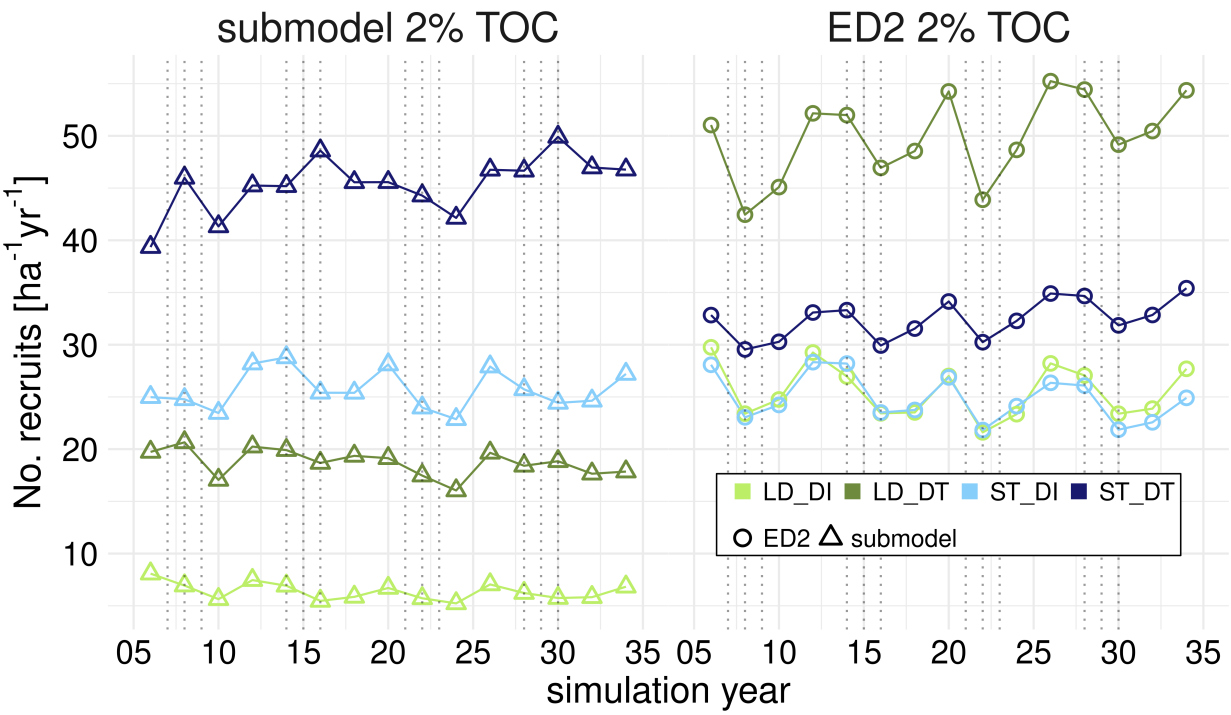


Figure S3. ED2 predictions of soil matric potential at 6 cm soil depth under BASE, WET, SYN-ENSO, and DRY-DS precipitation scenarios.

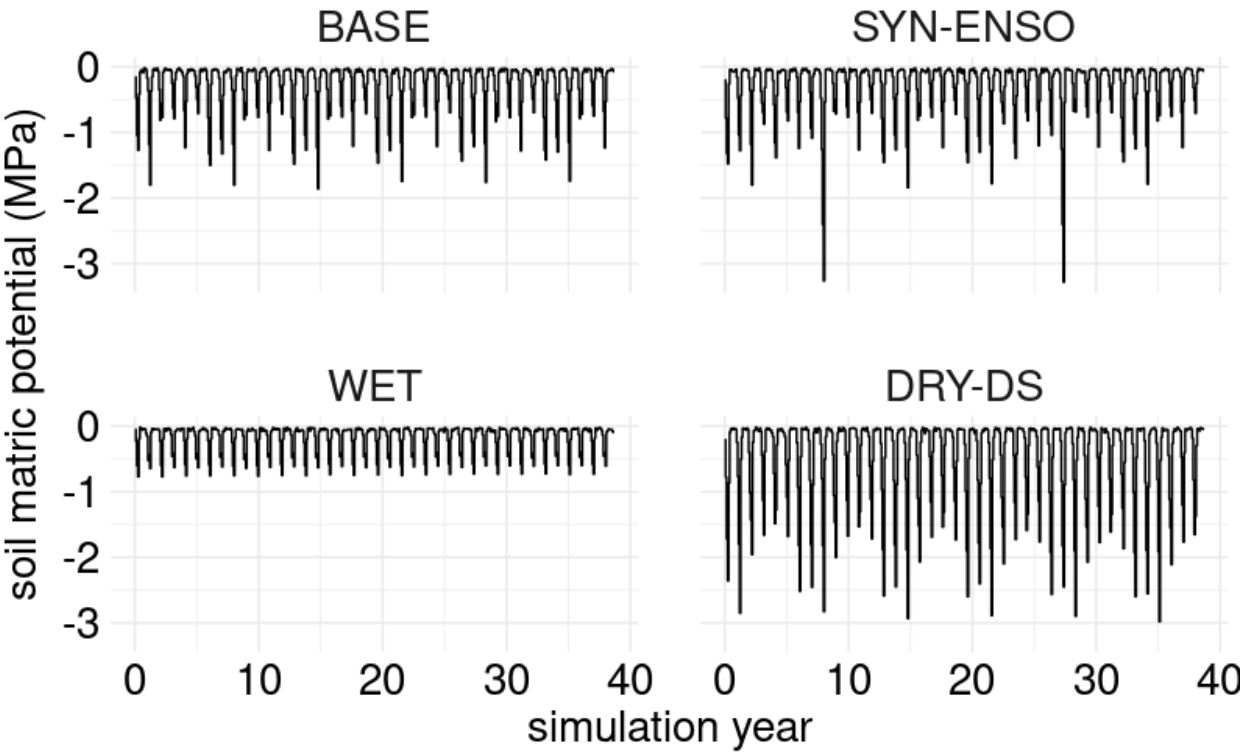


Figure S4. Longterm ED2 predictions of PFT-specific NPP under BASE, WET, SYN-ENSO, and DRY-DS precipitation scenarios.

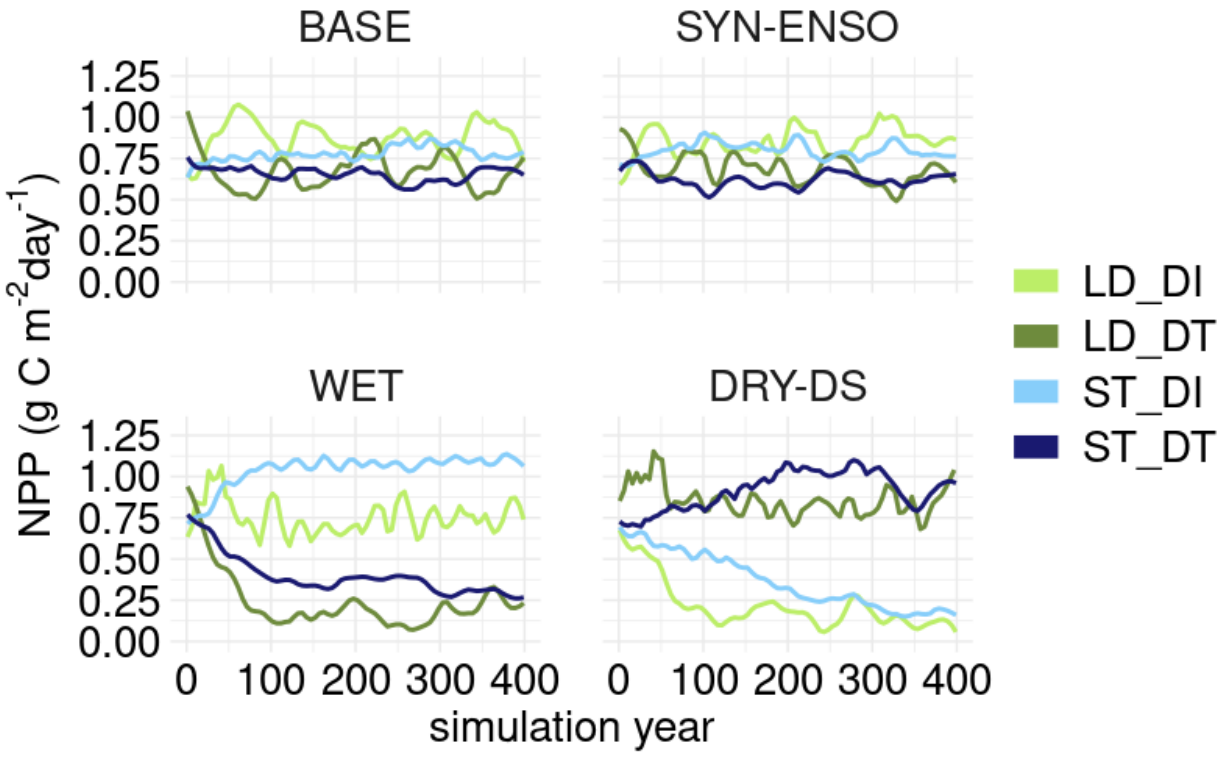
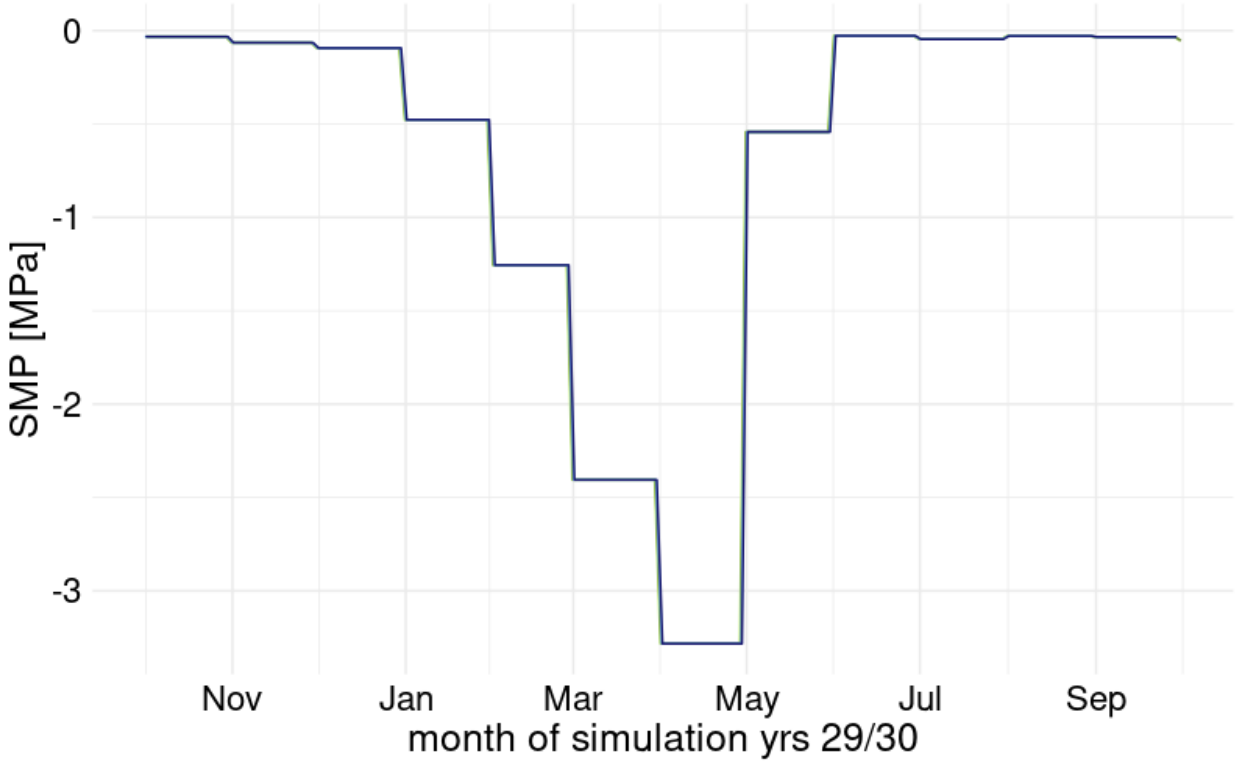


Figure S5. ED2 predictions of soil matric potential at 6 cm soil depth during a simulated El Niño drought at Barro Colorado Island.



Chapter 3: Drivers of post-fire conifer canopy recovery in California's mixed conifer forests

Chapter 1 presented a critical review of current VDMs and Chapter 2 presented a new scheme for representing forest regeneration in VDMs. Developing this modeling infrastructure requires an empirical understanding of how environmental variables and regeneration processes ultimately affect the distribution, structure, and functional composition of forests at broad scales. The work presented in this chapter quantifies the affects of fire characteristics, climate variables, and site conditions on decadal-scale, post-fire conifer canopy recovery throughout 19,698 hectares of California's mixed conifer forests. This extends shorter-term studies of post-fire seedling regeneration to provide new insights into how global change variables affect decadal-scale changes in conifer canopy structure. This analysis is well-positioned to support VDM model development by providing data for benchmarking and parameterization and developing insights that will help prioritize which processes to represent in VDMs that seek to predict the future of California's forests.

Summary

Intensifying fire regimes and climate change are reducing post-fire conifer seedling regeneration in the western United States. However, it remains uncertain if the observed short-term declines in post-fire seedling regeneration will translate into longer-term declines in conifer canopy cover. We quantified decadal-scale variation in post-fire conifer canopy recovery in California's mixed conifer forests (MCFs) and used generalizedgeneralized additive models to identify key drivers of recovery outcomes. We found that post-fire conifer canopy recovery is reduced after higher severity fire, when re-burns occur during the first 30 years after fire, and when mean annual precipitation in the first three years after fire falls below 1,000 mm yr^{-1}. Our results support prior hypotheses that increasingly intense fire regimes will negatively affect post-fire conifer regeneration, likely leading to an increase in the area dominated by shrubs. Despite these broader trends, our results also suggest that topographic micro-climate refugia and periodic pulses of recruitment during particularly wet years may continue to support MCF stands near current range boundaries where natural seed sources or re-planting occurs.

Introduction

Wildfire regimes and climate interact to shape the structure, composition, and function of terrestrial vegetation (Brown & Wu, 2005; Alencar et al., 2015; Airey et al., 2016; Coop et al., 2019). Recent forest management and climate change have intensified fire regimes in California (Stevens *et al.*, 2017; Steel *et al.*, 2018), creating uncertainty around the future distribution of California's forests. This is of particular concern in California's mixed conifer forests (MCFs) which are distributed across semi-arid regions near the boundaries of non-forested vegetation types (Evans *et al.*, 2011). The future distribution of California's conifer forests will regulate energy and carbon fluxes (Bonan 2008), influence regional hydrology and erosion rates (Lancaster & Baas, 1998; Flint *et al.*, 2019), and determine the type and quality of available

wildlife habitat (Verner & Boss, 1980). Therefore, understanding how California's conifer forests will respond to changing climate-fire interactions is an essential component of developing a forward-looking, state-wide land management plan (California Forest Management Task Force, 2021).

Prior to European settlement, California's MCFs where highly spatially heterogenous in species composition, forest structure, and fuel loads (Evans *et al.*, 2011). Euro-American forest management, including fire exclusion and logging, have resulted in higher fuel loads and more homogenous forest structure (Evans *et al.*, 2011) which are associated with larger, more intense fires (Miller & Thode, 2007). Common species in California's MCFs include ponderosa pine (*Pinus ponderosa*), sugar pine (*Pinus lambertiana*), white fir (*Abies concolor*), Douglas-fir (*Pseudotsuga menziesii*), incense cedar (*Calocedrus decurrens*), black oak (*Quercus velutina*), canyon live oak (*Quercus chrysolepis*) and shrubs in the *Arctostaphylos* and *Ceanothus* genera (Plamboeck et al., 2008; Evans et al., 2011). Theory and modeling suggest that more intense fire regimes, coupled with "climatic disequilibrium" (Svenning & Sandel, 2013), is triggering vegetation regime shifts (Folke *et al.*, 2004) whereby California's MCFs are converted to pyrogenic shrublands which store less carbon and perpetuate increased fire activity (Serra-Diaz *et al.*, 2018; Tepley *et al.*, 2018; Stevens-Rumann & Morgan, 2019). Predicting if, where, and when these fire-induced vegetation transitions may occur is predicated on understanding how fire characteristics and post-fire environmental conditions influence the long-term regeneration dynamics of conifers, hardwoods, and shrubs.

Recent field observations in California's MCFs have focused on identifying the drivers of short-term, post-fire (typically 0-5 years after fire) seedling (< 1.37 m in height) regeneration (Shive *et al.*, 2018; Stewart *et al.*, 2020). Conifer seedling regeneration is higher after lower severity fire, on flatter slopes, in areas with lower historical climatic water deficit (CWD), higher post-fire precipitation, in areas closer to potential seed sources, and in areas of less shrub competition (Shive *et al.*, 2018; Young *et al.*, 2019; Rodman *et al.*, 2020; Stewart *et al.*, 2020). Competition from shrubs in large burn scars can increase conifer seedling mortality rates (Tubbesing *et al.*, 2020) and the species composition of competing shrubs influences seedling emergence through the shrub layer 15-20 years after fire (Tubbesing *et al.*, 2022). These observations are consistent with evidence from the Klamath Mountains showing that shrub competition influences conifer dominance during post-fire regeneration trajectories (Tepley *et al.*, 2017). Our ability to predict how changing climate and fire regimes will affect the future functional composition of MCFs hinges on understanding to what extent the factors affecting seedling regeneration and shrub-seedling competition also affect longer-term conifer canopy recovery outcomes. Focusing on short-term seedling recovery may not provide sufficient insight into the extent to which changing fire regimes and climate are altering the competitive balance between shrubs and conifer seedlings during the recovery period.

Prior field studies on conifer regeneration in the western United States indicate declining forest resilience to fire (Stevens-Rumann *et al.*, 2018; Young *et al.*, 2019; Rodman *et al.*, 2020; Stewart *et al.*, 2020; Littlefield *et al.*, 2020), but have not produced a dataset that encompasses a large enough spatio-temporal extent to derive strong insights into the relationships between fire

characteristics, climate and decadal-scale changes in the dominant vegetation cover during the recovery period. A recent review of 49 field studies covering 150 fires in the western US found that most studies used non-standardized field protocols, different recovery timelines, and varying metrics for climate variables, seed availability, and fire characteristics (Stevens-Rumann & Morgan, 2019). The majority of these studies sampled less than 600 m^2 of post-fire terrain to draw conclusions, which is problematic given that post-fire forest recovery is highly spatially variable and dependent on spatial context (Turner, 2010).

Remote sensing (RS) offers a synoptic approach for large spatio-temporal extents. For example, a remote sensing study using post-fire normalized differenced vegetation index (NDVI) found that drought hampers vegetation recovery in lower montane forests in California's Sierra Nevada (Meng *et al.*, 2015). A more recent study spanning Canada's boreal forests used normalized burn ratio (NBR) recovery to conclude that post-fire forest recovery rates have increased since the 1980s (Frazier *et al.*, 2018). Prior RS studies of post-fire conifer forest recovery have used vegetation indices (Meng *et al.*, 2015; Frazier *et al.*, 2018) rather than metrics of conifer canopy cover or functional composition. Although NDVI and NBR can detect the return of photosynthetic activity to a burn scar, these indices tend to saturate in mesic sites (Asner, 1998) and do not provide strong insight into what type of vegetation is returning (e.g. shrub vs. conifer). Machine learning approaches, such as the random forest (RF) algorithm (Schonlau & Zou, 2020), provide effective tools to derive more ecologically meaningful insights from satellite imagery (Lary *et al.*, 2016), but have not previously been used to study post-fire conifer recovery in California's MCFs. Existing vegetation cover products, such as the National Land Cover Dataset (Dewitz, 2021), are produced using decision trees and supervised classification of remote sensing observations (Jin *et al.*, 2019), but existing products are not optimized to distinguish between shrubs and conifers in California's highly structurally and floristically diverse MCFs and do not exist at high-enough temporal resolution (i.e. annual) to sufficiently quantify recovery trajectories. In sum, neither the existing vegetation maps, nor the previously used vegetation indices, are well-equipped to quantify decadal-scale post-fire conifer recovery in California's MCFs.

Here we train a RF algorithm to predict conifer canopy cover as a function of Landsat satellite observations of surface reflectance and use it to quantify variation in long-term, post-fire conifer canopy recovery in California's MCFs. Using prior research on post-fire conifer seedling regeneration as a starting point, we use generalized additive models to test which combination of fire characteristics, historic climate variables, post-fire weather variables, and site conditions best explains variation in conifer canopy recovery 15-32 years after fire. We discuss the probable mechanisms underlying divergent recovery outcomes and implications for the future of California's MCFs under intensifying fire regimes and changing climate.

Methods

Study Area

Our study area included MCFs (Evans *et al.*, 2011) within California's Sierra Nevada Mountains, Cascade Ranges, Klamath Mountains, and northern Interior Coast Ranges (Fig. 1). Pixels were included in our analysis if they met the following criteria: a) were previously classified by LANDFIRE as having the biophysical potential to be MCF (Rollins, 2009), b) had a pre-fire conifer foliage projective cover (Ediriweera et al., 2013; FPC) > 0.3, c) burned once at medium or high severity, as quantified by the Monitoring Trends in Burn Severity (MTBS; Eidenshink et al., 2007) dataset between January 1985 and December 2004 (i.e. the "fire period"), d) did not burn again at high severity between January 2005 and December 2019 (i.e. the "recovery period"). Pixels that re-burned at low or medium severity during the recovery period were included. These criteria produced 44 clusters of pixels (i.e. "burned areas"; Fig. 1) resulting from 37 fires burning during 12 years between 1987 and 2004. This study area is large enough to identify generalizable relationships between post-fire recovery and fire characteristics, climate drivers, and site conditions, but ecologically constrained enough (just MCFs) to allow for the interpretation of results within a manageable inference space. The temporal extent (1987-2019) allowed for the evaluation of long-term regeneration trajectories, 15-32 years long, over an area of 19,698 ha.

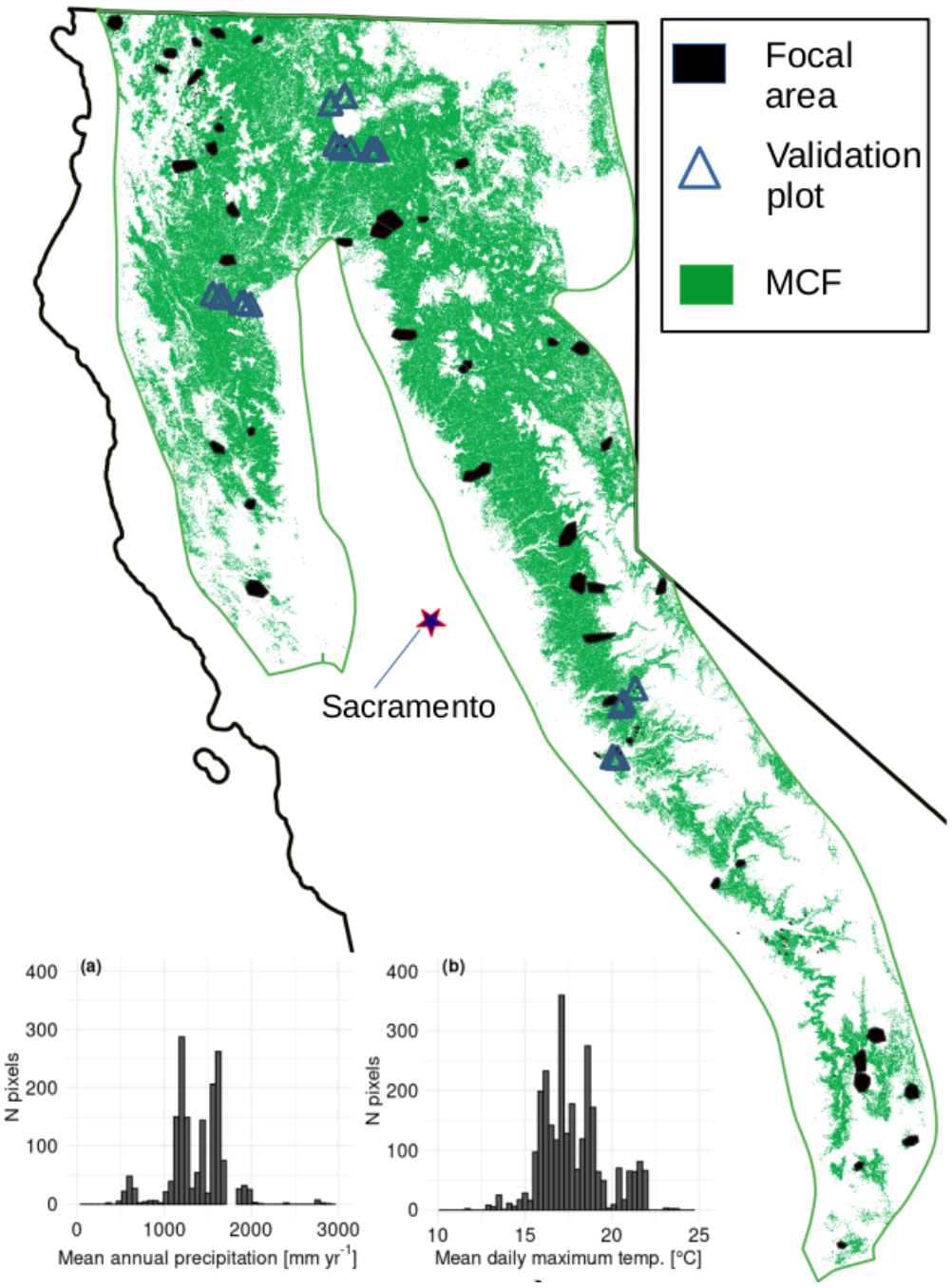


Figure 1. The study area comprised pixels within 44 burned areas (black polygons) of California's Sierra Nevada Mountains, Cascade Ranges, Klamath Mountains and Interior Coast Ranges with the biophysical potential to be mixed conifer forest (MCF; shaded green). Burned areas delineate clusters of pixels that burned once at medium or high severity between the beginning of 1985 and the end of 2004. Blue triangles indicate the locations of 36 validation

plots used to evaluate the RF algorithm for its ability to predict percent conifer canopy cover. The histograms show the distribution of historical (1980-2010) mean annual precipitation (Daly *et al.*, 2015; a) and daily maximum temperature (Daly *et al.*, 2015) in the burned areas (b).

Quantifying percent conifer canopy cover

We used high resolution imagery available from Google Earth (i.e. "training imagery") to manually estimate the FPC of conifer trees (i.e. "conifer canopy cover") in 108 training areas distributed throughout the study area. FPC is a measure of the fractional ground area covered by the vertical projection of a vegetation canopy (Ediriweera *et al.*, 2013) such that a value of 1 would correspond to a fully closed forest canopy. Google Earth imagery is derived from a variety of sources including cameras mounted on satellites or aircraft. The resolution generally varies between 0.15 and 5 m per pixel depending on the specific image source (Yu & Gong, 2012; see Fig. S1 for an example image at one of our training areas). The training image dates we used varied depending on the training area, but were all between May 2014 and June 2021. All training images were accessed on Google Earth in January 2022. Following best practices, the training areas were delineated to represent "areas of homogeneous or uniformly mixed land cover" (Soto-Berelov *et al.*, 2015) and ranged in size from 0.15 – 105 ha. Many training areas with little conifer cover were purposefully selected to included dense shrub or grass cover to help the algorithm distinguish between conifer canopies and other vegetation types. We quantified conifer canopy cover using the point intercept method (Elzinga *et al.*, 2001) where the presence or absence of conifer tree canopies was recorded at each point on a systematic grid overlaying the training image (e.g. Fig. S1). The number of sample points per training area ranged from 42 to 206, depending on training area size, corresponding to roughly two points per ha. Conifer canopy cover in the training areas ranged from 0-0.81.

The conifer canopy cover training data were used to train a RF algorithm (Ho, 1995) to predict conifer canopy cover as a function of Landsat 5-8 Tier 1 surface reflectance values in the red, green, blue, near infrared (NIR), and short-wave infrared (SWIR1 and SWIR2) bands. The set of possible Landsat scenes for each training area was first filtered to include scenes with < 15% cloud cover. Within each training area we chose Landsat imagery with an image date as close as possible to the training imagery (never more than 6 months apart). To account for variable illumination angles in mountainous, forested terrain we applied a previously used topographic correction algorithm to all Landsat data (S. A. Soenen *et al.*, 2005; Poortinga *et al.*, 2019) prior to training the RF algorithm. To avoid the influence of winter snow cover, haze, and other atmospheric anomalies obscuring the vegetation's spectral signature we used a "quality mosaic" approach (Li *et al.*, 2019) where the "greenest" (i.e. highest NDVI) pixel values between June 30th and August 15th were used as RF RF training data. All Landsat images were radiometrically harmonized across the Operational Land Imager and Thematic Mapper + sensors following prior methods (Roy *et al.*, 2016) before applying the RF algorithm to the full Landsat time series. These data were processed in Google Earth Engine (Gorelick *et al.*, 2017).

Field validation

In 2021, we collected independent field observations of conifer canopy cover at 36 validation plots distributed throughout our study area (Fig. 1). Validation plot locations were selected using a random stratified sampling scheme designed to capture geographic variation and a wide spectrum of conifer canopy cover (0-0.73). Each plot a) had the biophysical potential to be MCF, b) was on public land, c) was < 300 m to a road, and d) did not intersect a paved road. All plots were large enough to fully contain a 3 x 3 window of 30 m pixels such that the mean prediction of 9 contiguous Landsat pixels could be compared with field observations.

We used two different methods to quantify conifer canopy cover in the field. For 10 validation plots we used the standard line intercept field method (Etchberger & Krausman, 1997) where the FPC of conifers, broadleaved trees, shrubs, and graminoids/forbes was estimated along four systematically placed 90-m transects. Along each transect we recorded where the vertical projection of vegetation cover intersected the transect tape. All four transects originated at the plot centroid with a heading in each of the four cardinal directions. For 26 validation plots we flew drone missions in pre-programed, systematic grids to capture high resolution (1.5 cm^2 ground area per pixel) aerial images of the vegetation canopy. All images were taken between 100-150 feet above ground level at a nadir angle using a 12 Megapixel camera onboard the DJI Mavic Mini drone (e.g. Fig. S4). We quantified conifer canopy cover in each validation plot using the standard point-intercept method (Elzinga *et al.*, 2001) by recording the presence or absence of conifer tree crowns at the center of each aerial image. This yielded ~200 points per validation plot. To assess the performance of our RF algorithm we predicted conifer canopy cover at each of the 36 validation plots and calculated the RMSE and adjusted R^2 of predictions.

Quantifying post-fire conifer canopy recovery

We quantified post-fire conifer canopy recovery by adapting the relative recovery index (RRI) previously used to quantify short-term, post-fire spectral vegetation recovery in conifer forests (Frazier *et al.*, 2018). We calculated RRI in year i after fire (RRI_{F+i}; Eqn 3) as a function of the absolute recovery of conifer canopy cover (ARI_{F+i}; Eqn 2) normalized by the disturbance size (Δ_C; Eqn 1). C_{F+i} is conifer canopy cover i years after fire and C_{F-i} is conifer canopy cover i years before fire such that

$$\Delta_C = \frac{C_{F-2} + C_{F-1}}{2} - \min\left[C_{F+1}, C_{F+2}, C_{F+3}\right] \qquad \text{(Equation 1)}$$

$$ARI_{F+i} = C_{F+i} - \min\left[C_{F+1}, C_{F+2}, C_{F+3}\right] \qquad \text{(Equation 2)}$$

$$RRI_{F+i} = \frac{ARI_{F+i}}{\Delta_C}. \qquad \text{(Equation 3)}$$

We defined the initial post-fire canopy cover (i.e. the post-disturbance state prior to recovery) as the minimum of the first three years after fire (the right terms in Eqns 1 and 2) because prior work has found that tree cover continues to decline for multiple years after fire with a corresponding decline of NDVI (Meng *et al.*, 2015). Our calculation of RRI has the advantage of normalizing recovery by disturbance size (Δ_C) which facilitates the comparison of recovery among pixels of varying pre-fire conifer cover and fire intensity.

Predictor variables

Based on prior research (Shive *et al.*, 2018; Young *et al.*, 2019; Rodman *et al.*, 2020; Stewart *et al.*, 2020), we hypothesized that RRI would be a function of topography, fire history, seed availability, historical climate, and post-fire weather. We assessed multiple candidate variables within each of these five predictor categories (see Table 1 for descriptions of all predictor variables analyzed) for their ability to explain variation in RRI. All variables were re-projected from their native coordinate reference system to California Albers (EPSG: 3310) at 270 m resolution and standardized prior to analysis. Some variables that were categorical or boolean at their native 30 m resolution (e.g. burn severity or "PostFirePlanting") became continuous when averaged to 270 m resolution.

Table 1. Candidate predictor variables for the generalized additive models used to predict conifer canopy recovery.

Variable name	**Definition and Derivation**	**Native resolution (m)**	**Reference**
Topography			
HLI	Heat load index integrates slope, aspect, and latitude for a continuous measure of terrain-driven solar heat load.	10	(Theobald *et al.*, 2015)
Adj-northness	Adjusted northness = cos(aspect) · sin(slope) such that northerly facing aspects score closer to 1, southerly aspects closer to -1 and flatter areas closer to zero regardless of aspect.	30	(Farr *et al.*, 2007)
Adj-fold-northness	Adjusted, folded northness = cos(aspect + 45) · sin(slope) such that the lowest northness values (-1) are oriented toward the southwest instead of south. Prior work indicates this is a more ecologically relevant metric than Adj-northness because southwest-facing aspects receive maximum solar irradiance during the hottest part of the day.	30	(McCune & Keon, 2002; Farr *et al.*, 2007)
TPI	Topographic position index is a continuous measure of a pixel's elevation relative to the	10	(Theobald *et al.*, 2015)

	mean elevation of a surrounding pixel window such that valleys are negative and ridges are positive.		
Fire history			
BurnSev	Burn severity is a categorical score of fire damage to vegetation (2 = low, 3 = medium, 4 = high) from the Monitoring Trends in Burn Severity program. Categorical scores are based on threshold values for the differenced Normalized Burn Ratio (dNBR).	30	(Eidenshink *et al.*, 2007)
MediumReburn	A boolean variable indicating if a pixel re-burned at medium severity during the recovery period (2004-2019).	30	(Eidenshink *et al.*, 2007)
LowReburn	A boolean variable indicating if a pixel re-burned at low severity during the recovery period (2004-2019).	30	(Eidenshink *et al.*, 2007)
Seed availability			
PostFireSAP	The seed availability proxy is calculated as the mean conifer canopy cover within a 400 m radius of each pixel. Post-fire seedling density drops dramatically above 400 m from remnant seed trees (Donato *et al.*, 2009). The PostFireSAP is taken as the mean SAP of the first three years after fire. SAP was predicted using RF predictions of canopy cover.	30	(Donato *et al.*, 2009; Tepley *et al.*, 2017; Shive *et al.*, 2018)
PostFirePlanting	A boolean variable indicating if a pixel was planted by the USFS within the first 6 years after fire.	30	(United States Forest Service, 2022)
Historical climate			
CWD	Mean historical (1980-2010) climate water deficit (mm) calculated using the basin characterization model, which accounts for topography, soil texture, and meteorology.	270	(Flint *et al.*, 2013)
AET	Mean historical (1980-2010) actual evapotranspiration (mm) calculated using the basin characterization model, which accounts for topography, soil texture, and meteorology.	270	(Flint *et al.*, 2013)
PPT	Historical (1980-2010) mean annual precipitation (mm).	4,000	(Daly *et al.*, 2015)
Tmax	Historical (1980-2010) mean daily maximum temperature for the whole year.	4,000	(Daly *et al.*, 2015)
Post-fire weather			
PPT_Yr1-3	Mean annual precipitation (mm) in the first 3 years after fire.	4,000	(Daly *et al.*, 2015)

PPTmin_Yr1-3	Minimum annual precipitation (mm) in the first 3 years after fire.	4,000	(Daly *et al.*, 2015)
PPT_Yr1-3_anomaly	Mean annual precipitation (mm) anomaly (Z-score) compared to the historical mean (1980-2010) in the first 3 years after fire.	4,000	(Daly *et al.*, 2015)
Tmax_Yr1-3	Mean daily maximum temperature (Tmax) in the first 3 years after fire.	4,000	(Daly *et al.*, 2015)
PPT_Yr1-3_T-adjusted	Temperature-adjusted mean annual precipitation (mm) in the first 3 years after fire. Calculated as PPT_Yr1-3 / Tmax_Yr1-3	4,000	(Daly *et al.*, 2015)
Other			
Recovery time	Time since fire (yr)	NA	NA

Generalized additive models to predict conifer canopy recovery

Following prior modeling approaches to investigate post-fire conifer seedling regeneration (Shive *et al.*, 2018; Stewart *et al.*, 2020) we used generalized additive models (GAMs; Wood, 2006) to predict long-term post-fire conifer canopy recovery. To maximize recovery time, we fit models to predict RRI (Eqn 3) at the end of the recovery period (RRI_{2019}) resulting in a range of recovery times from 15-32 years long. We used a spatially rarefied subset (n = 684) of all eligible pixels in our study area to mitigate residual spatial autocorrelation during model fitting. We chose 270 m^2 as the unit of analysis because this size a) is big enough to accommodate the open spatial structure of many conifer forests which may not have a tree crown in every 30 m pixel, b) is small enough to capture variation in key topographic variables such as slope, aspect, heat load index (HLI), and topographic position index (TPI), and c) is the same spatial resolution as the basin characterization model (Flint *et al.*, 2013) used for historical climate water deficit (CWD) and actual evapotranspiration (AET).

Following the methods used in Shive *et al*. (2018), we fit candidate models using backward step-wise selection within a constrained hypothesis space. More specifically, our approach assumed that top performing models would include at least one variable from each of the five predictor categories (topography, fire history, seed availability, historical climate, post-fire weather; Table 1) previously found to predict post-fire conifer seedling presence (Shive *et al.*, 2018; Rodman *et al.*, 2020; Stewart *et al.*, 2020). Initial models were formulated to include the same variables as in previous work plus uncorrelated novel predictors such as the fraction of each pixel that re-burned at medium severity (mediumReburn) and the fraction of each pixel that was planted by the USFS within six years of fire (PostFirePlanting). Following prior approaches (Shive *et al.*, 2018), highly correlated variables from the same predictor category, such as HLI and adjusted, folded northness (Adj-fold-northness) or mean post-fire precipitation in the first three years (PPT_Yr1-3) and post-fire precipitation anomaly (PPT_Yr1-3_anomaly), were tested one at a time while holding the rest of the model formula constant to avoid multicollinearity. This allowed us to compare the relative performance of models using alternate correlated variables

from each predictor category. Non-significant terms were dropped one at a time and model performance was ranked based on AIC. Model residuals were checked for homogeneity of variance, independence, and spatial autocorrelation following standard protocols (Wood, 2006). All models required the inclusion of an interaction term between latitude and longitude to mitigate spatial autocorrelation of model residuals.

Results

Predicting conifer canopy cover

The RF algorithm predicted conifer canopy cover in the training areas with a root mean squared error of 0.07 (R^2 = 0.95; Fig. S5). It predicted conifer canopy cover at 36 validation plots with a RMSE of 0.11 (adj. R^2 = 0.78; Fig. 1), providing sufficient accuracy for the purpose of this analysis, but not enough accuracy to distinguish between sites with no conifer canopy cover versus sites with very little conifer canopy cover (i.e. conifer FPC < 0.1). Unlike vegetation indices such as NBR and NDVI, which predict post-fire recovery in areas of vigorously regenerating shrub cover (Hope *et al.*, 2007), the RF algorithm correctly predicted low conifer canopy cover (FPC < 0.1) in shrub-dominated plots where conifer cover was confirmed to be low or absent (Fig. 2).

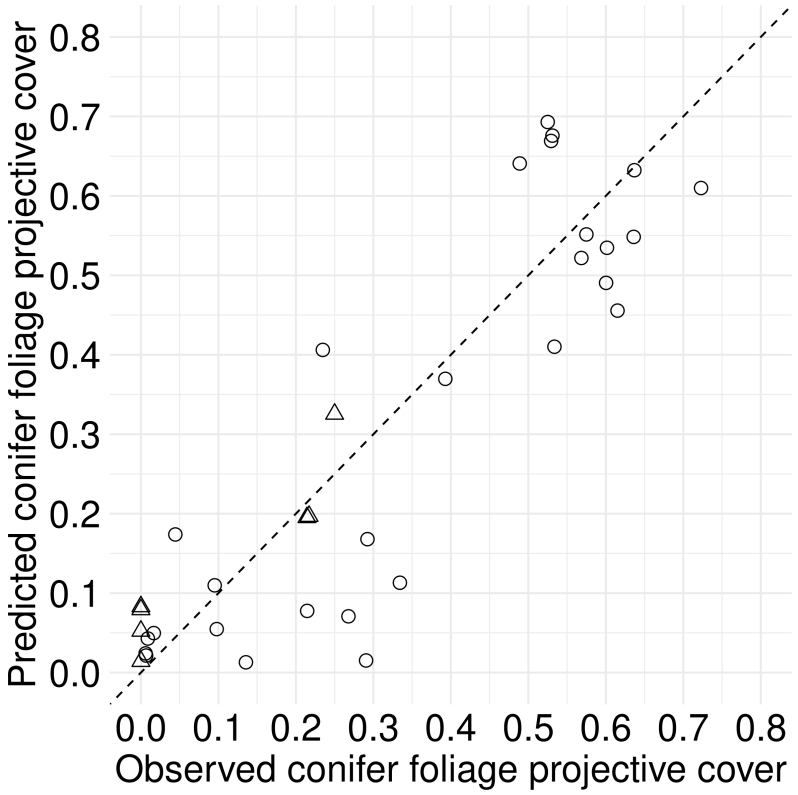


Figure 2. Conifer foliage projective cover predicted by the random forest algorithm (Y-axis) plotted against field observations (X-axis) at 36 validation plots distributed throughout the study area. Triangles indicate shrub-dominated validation plots where shrub foliage projective cover is > 0.5. The dashed line (y = x) represents perfect agreement between predictions and field observations.

Conifer canopy recovery

We found significant variation in the shape and ultimate outcome of individual conifer canopy recovery trajectories (Fig. 3). Nineteen percent of pixels fully recovered to their pre-fire canopy cover (> 0.9 RRI) within 32 years of fire. For pixels that recovered 50-100% of their pre-fire canopy cover ($RRI_{2019} > 0.5$), recovery is generally non-linear with time (Fig. 2b,d). There is minimal recovery in the first 10-15 years and then it accelerates (Fig. 3b,d). Pixels that re-burned at medium severity during the recovery period (2004-2019) showed less recovery 25 years after fire (mean RRI = 0.11) than those that re-burned at low severity (mean RRI = 0.65; Fig. 3a,c; t = 32.958, df = 336.01, p < 2.2e-16). Pixels that originally burned at medium severity (bluer trajectories in Fig. 3) recovered more in the first 15 years of recovery than pixels that burned at high severity (t = 4.8716, df = 272.81, p-value = 1.879e-06; Fig. S6). Surprisingly, pixels that were planted within 6 years of fire (Fig. 3c,d) showed little difference in recovery from pixels that were not planted (Fig. 3a,b).

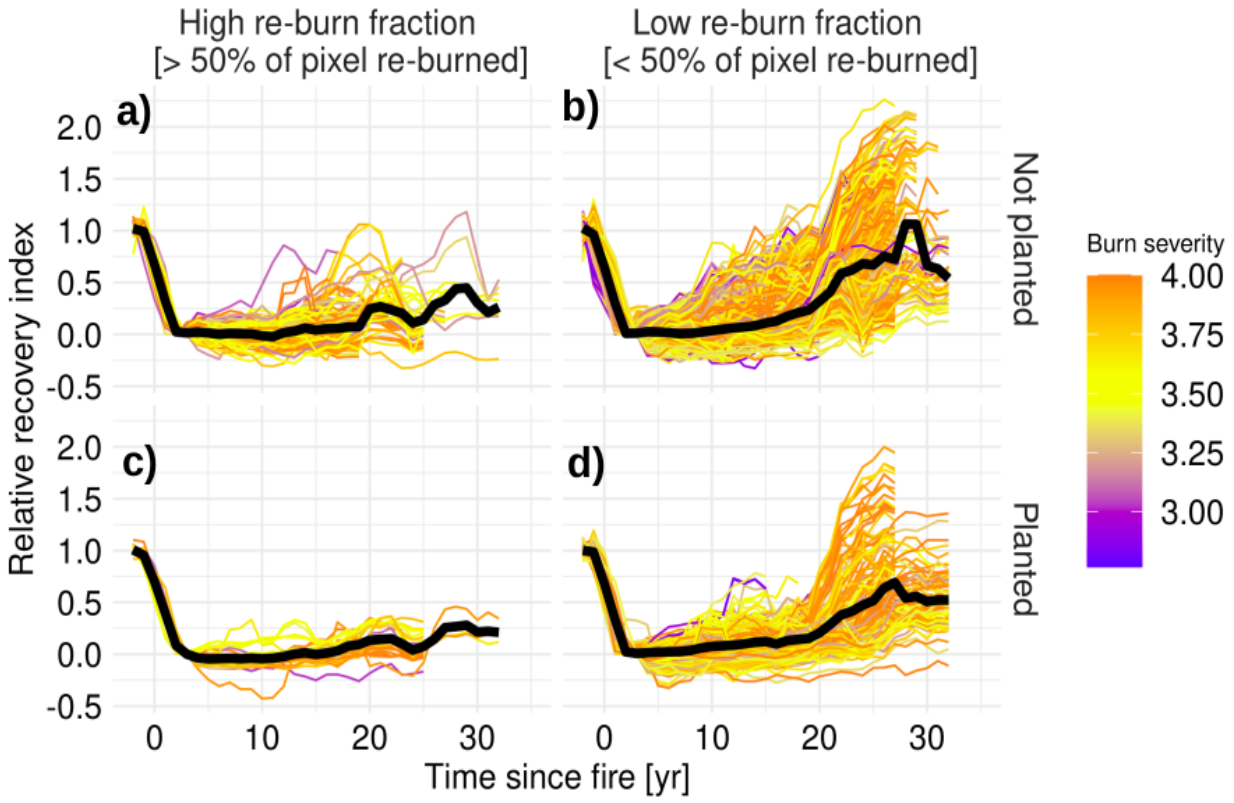


Figure 3. Regeneration trajectories for all pixels (n = 2,529; 18,436 ha) in the study area. A relative recovery index of "1" indicates full recovery to the pre-fire conifer canopy cover level. Colored lines show individual recovery trajectories, with redder colors indicating higher burn severity (4 = "high burn severity", 3 = medium burn severity). Thick, black lines indicate the mean for each group. "High re-burn fraction" pixels experienced > 50% of their area re-burning at medium severity at some point during the recovery period (2004-2019). "Low re-burn fraction" pixels experienced < 50% of their area re-burning a medium severity during the recovery trajectory. "Not planted" indicates that < 10% of the pixel was planted by the USFS within 6 years of fire. "Planted" indicates that > 50% of the pixel was planted by the USFS within 6 years of fire.

Variation in RRI at the end of recovery period (RRI_{2019}) was best explained by recovery time, burn severity, the fraction of each pixel that re-burned at medium severity, adjusted-folded-northness, post-fire seed availability proxy, mean annual precipitation in the first 3 years after fire and the mean annual daily maximum temperature (adj. R^2 = 0.839, AIC = 1061; Table 2). A Moran's I test confirmed no residual spatial autocorrelation for the top performing model.

Table 2. Model coefficients and significance for the top performing model predicting post-fire conifer canopy recovery. Edf = effective degrees of freedom for the fitted smooth functions; Rdf = reference degrees of freedom for the fitted smooth functions; "Long,lat" = latitude and longitude; Tmax = historical (1980-2010) mean daily maximum temperature; Adj-fold-northness = adjusted, folded, northness; PostFireSAP = seed availability proxy; PPT_Yr1-3 = mean annual precipitation (mm) in the first three years after fire; MediumReburn = the fraction of each pixel that re-burned at medium or high severity during the recovery period (2004-2019).

Predictor	**Estimate**	**Std. Error**	**T-value**	**P-value**
Parametric Coefficients				
Intercept	-0.08255	0.02703	-3.054	0.002361
Recovery time	0.66267	0.19889	3.332	0.000918
MediumReburn	-0.08441	0.03055	-2.763	0.005918
Burn severity	-0.15093	0.02602	-5.802	1.09e-08
Smooth Terms				
Predictor	**Edf**	**Rdf**	**F-value**	**P-value**
Long,lat	72.624	73.248	10.956	< 2e-16
Tmax	9.892	11.738	1.958	0.022314
Adj-fold-northness	4.334	5.499	2.155	0.051146
PostFireSAP	7.142	8.823	3.425	0.000489
PPT_Yr1-3	13.094	13.797	4.278	3.78e-07

As expected, less recovery time, higher burn severity, and a larger medium severity re-burn fraction all predicted lower canopy recovery (RRI_{2019}). Burn severity, mean annual precipitation in the first 3 years after fire, and recovery time were the most significant predictors of recovery (Table 2). Mean annual precipitation in the first 3 years after fire had a non-linear effect on recovery such that as it declines, recovery is unaffected until a critical threshold, ~1,000 mm yr^{-1}, at which point it is severely, negatively affected (Fig. 4a). The mean annual daily maximum temperature had a weak, nonlinear effect on recovery, with some indication that recovery is negatively affected as it exceeds 21.1 ºC (Fig. 3b). Sites at intermediate adjusted-folded-northness (i.e. facing northwest or southeast) had slightly better recovery than sites facing due northeast or southwest (Fig. 3c), but this relationship is weak ($p = 0.051$). Unexpectedly, recovery showed a mostly negative relationship with increasing post-fire seed availability (PostFireSAP; Fig. 3d), except at intermediate levels of seed availability where the relationship was slightly positive.

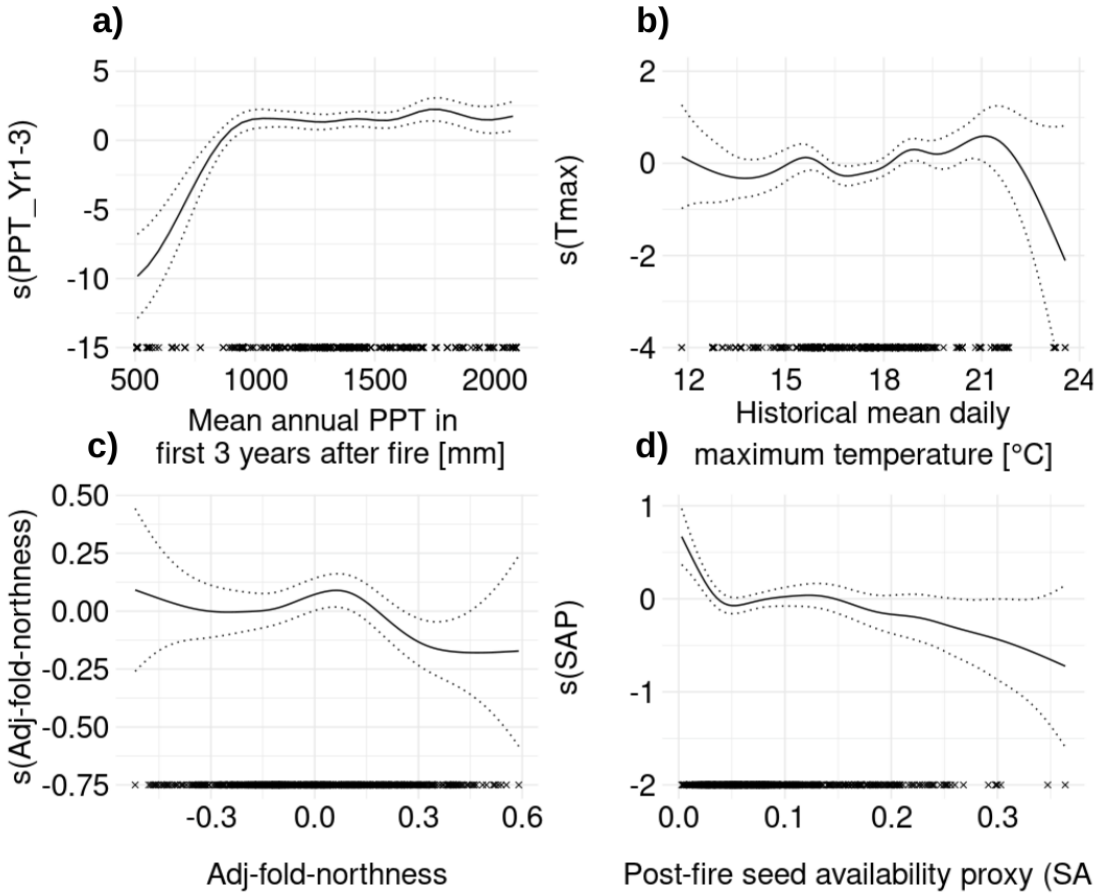


Figure 4. Partial response plots for the smooth terms of the top performing model predicting post-fire conifer canopy recovery. Y-axis values, s(x), show the value of the smooth term at a given value of x while all other terms are held constant. Dotted lines show ± standard error of the fit.

Lower performing recovery models included alternate variables for topography, historical climate, post-fire precipitation, and fire history (Table 3). For example, replacing adjusted-folded-northness with alternative topographic variables such as heat load index or adjusted northness resulted in higher AIC and lower adj. R^2. Surprisingly, topographic position index was not found to be a significant term in any candidate model, indicating that hill slope position (e.g. ridge vs. valley) did not explain additional variation in long-term recovery. Aside from mean annual daily maximum temperature, the inclusion of historical climate variables such as actual evapotranspiration, climate water deficit, and precipitation were rarely significant and did not produce a model that performed as well as the combination of mean annual daily maximum temperature and post-fire precipitation. Alternate post-fire precipitation metrics such as the minimum annual precipitation in the first 3 years after fire, temperature adjusted post-fire precipitation, and post-fire precipitation anomaly, did not perform as well as mean annual post-fire precipitation. Post-fire planting was statistically significant ($p < 0.05$) in lower performing

models (e.g. Model 12), but became statistically insignificant in higher performing models (e.g. Models 2,8). In contrast to the negative effect of medium severity re-burns, low severity re-burns were insignificant for predicting recovery.

Table 3. Candidate models for predicting post-fire conifer canopy recovery. Italicized variables were not significant at p = 0.05. See Table 1 for all variable definitions.

Model Performance			**Predictor Variables**			
Model Rank	**Δ AIC**	**Adj. R^2**	**Topography**	**Historical Climate**	**Post-fire Precipitation**	**Fire History and Seed Availability**
1	0	0.84	Adj-fold-northness	Tmax	PPT_Yr1-3	PostFireSAP, MediumReburn, BurnSev
2	1	0.84	Adj-fold-northness	Tmax	PPT_Yr1-3	PostFireSAP, MediumReburn, BurnSev, *PostFirePlanting*
3	6	0.84	none	Tmax	PPT_Yr1-3	PostFireSAP, MediumReburn, BurnSev
4	6	0.84	*Adj-northness*	Tmax	PPT_Yr1-3	PostFireSAP, MediumReburn, BurnSev
5	8	0.83	none	PPT, Tmax	PPT_Yr1-3	PostFireSAP, MediumReburn, BurnSev
6	19	0.83	none	none	PPT_Yr1-3	PostFireSAP, MediumReburn, BurnSev
7	29	0.83	Adj-fold-northness	*CWD*	PPT_Yr1-3	PostFireSAP, MediumReburn, BurnSev
8	30	0.83	none	*AET*	PPT_Yr1-3	*PostFirePlanting*, PostFireSAP, MediumReburn, BurnSev
9	45	0.84	none	Tmax	PPTmin_Yr1-3	PostFireSAP, MediumReburn, BurnSev
10	51	0.84	none	Tmax	PPT_Yr1-3_T-adjusted	PostFireSAP, MediumReburn, BurnSev
11	53	0.82	Adj-fold-northness	Tmax_Yr1-3	PPT_Yr1-3	PostFireSAP, MediumReburn, BurnSev
12	62	0.81	none	Tmax	PPT_Yr1-3_anomaly	PostFireSAP, MediumReburn, BurnSev
13	232	0.81	HLI	AET	PPT_Yr1-3	PostFirePlanting, PostFireSAP, MediumReburn

Discussion

Performance of the RF algorithm

The RF algorithm was able to predict conifer canopy cover with a RMSE of 0.11, which enabled us to analyze relative differences in recovery outcomes and distinguish between trajectories showing limited recovery versus areas showing a return to pre-fire levels. This level of accuracy does not allow us to reliably distinguish between areas of low conifer canopy cover (~ 0.1 conifer FPC) and areas without any conifer trees so we avoid conclusions that require identifying cases of complete regeneration failure. Our algorithm is an improvement over prior studies that use vegetation indices to quantify post-fire conifer forest recovery (Meng *et al.*, 2015; Frazier *et al.*, 2018) because it provides a more ecologically interpretable metric of forest cover and is less prone to mistaking areas of vigorous shrub cover with recovering forest. The RF algorithm correctly predicted low levels of conifer cover even in validation plots with significant shrub cover (shrub FPC > 0.5; Fig. 1), indicating that it was able to distinguish between these

vegetation types. Future work is needed to assess if spectral mixture analysis or a more advanced machine learning algorithm, such as a deep neural network (Zhang *et al.*, 2016), can improve predictions at the low end of the conifer canopy cover spectrum where conifer FPC < 0.1. We relied on high resolution imagery available from Google Earth for training data and chose images that appeared to be taken as close to a nadir angle as possible. It is likely that some training areas were photographed off-nadir, however this did not appear to systematically bias predictions of canopy cover (Fig. 2). Our training data would likely be improved with the use of higher resolution aerial imagery taken at a nadir angle for the specific purpose of quantifying FPC.

The long-term effect of post-fire precipitation

We found that long-term conifer canopy recovery is negatively affected when mean annual post-fire precipitation drops below 1,000 mm yr^{-1} (Fig. 3a). This result is consistent with prior findings that higher post-fire precipitation increases the probability of conifer seedling presence in California's MCFs (Young *et al.*, 2019; Stewart *et al.*, 2020). It advances prior research by demonstrating that the effect of initial post-fire precipitation goes beyond just influencing the probability of post-fire seedling presence, but also has a multi-decadal influence on the recovery of conifer canopy dominance. This finding is consistent with a recent dendrochronological study spanning the northern Rocky Mountains and the southwest United States which showed that post-fire CWD anomalies affect the juvenile recruitment and demographic structure of ponderosa pine (*Pinus ponderosa*) and Douglas-fir (*Pseudotsuga menziessi*) forests many decades after fire (Littlefield *et al.*, 2020).

Post-fire precipitation has been widely hypothesized to play a significant role in the recovery of California's MCFs, but prior studies have only ever detected a relatively weak effect. For example, Stewart *et al.* (2020) found that models including post-fire precipitation performed marginally better than models that only included historical (1980-2010) mean annual precipitation (Stewart et al., 2020). We found that models performed better when they included post-fire precipitation compared to when they only included historical climate averages such as historical precipitation, CWD, or AET (Table 1). This finding is biologically reasonable because post-fire conditions, not historical averages, directly affect the physical environment of seedlings. Historical climate averages are thought to emerge as statistically significant predictors of regeneration via their correlation with post-fire weather variables or indirectly via "biological inertia" (i.e. propagule pressure; Stewart et al., 2020). Mean historical precipitation was correlated with post-fire precipitation in our dataset (Pearson's $r = 0.87$, $p < 0.0001$). We attribute our detection of a strong effect of post-fire precipitation to our study design, which assessed recovery over a very broad climatic gradient across multiple fire years after competition between seedlings and non-conifer vegetation had already occurred for many decades. Prior findings may not have identified post-fire precipitation as a strong predictor because it does not exert as strong of an influence on the probability of seedling presence as it does on the longer-term competitive balance between shrubs and seedlings.

Shrub competition as a likely mechanism mediating the influence of post-fire precipitation

By focusing on long-term conifer canopy recovery rather than the short-term presence of seedlings, our findings provide new insight into how precipitation may alter the competitive dynamics between conifer seedlings and shrubs. The post-fire precipitation threshold we found (1000 mm yr^{-1}) does not correspond precisely to the biophysical limit of MCFs (although its surprisingly close; Fig. S2a), but it may be important for allowing seedlings to grow fast enough to compete with shrubs and and secure a position in the canopy within 10-15 years after fire. This hypothesis is supported by prior evidence showing that post-fire precipitation affects seedling height dominance relative to competing shrubs, with seedlings more likely to be taller than shrubs during anomalously high post-fire precipitation (Young et al., 2019) and observations that ponderosa pine seedling growth and survivorship decline under shrub canopy cover (Tubbesing *et al.*, 2020). It is also supported by experimental evidence from manipulated field plots showing that soil moisture was the single most important factor releasing white fir (*Abies concolor*) seedlings from shrub competition (Conard & Radosevich, 1982) and more recent manipulative experiments suggesting that shrub competition for soil moisture limits the growth and survival of white fir and sugar pine seedlings (Plamboeck *et al.*, 2008). Our recovery trajectories show a non-linear acceleration of conifer canopy recovery starting between 15 and 20 years after fire (Fig. 2b,d). The time to this inflection point is consistent with results from a data-driven simulation model of seedling-shrub competition which found that conifer seedlingsseedlings required roughly 16-22 years to over-top shrubs (Tubbesing *et al.*, 2022). Sufficient precipitation (> 1,000 mm yr^{-1}) in the first three post-fire years may help conifer seedlings overtop shrubs within this time frame.

Changing fire severity, temperature, and the potential for vegetation transitions

Our findings support prior hypotheses that changing climate and fire regimes may trigger vegetation transitions in California's MCFs. We show that the previously found negative effect of higher burn severity on seedling recovery (Stevens-Rumann & Morgan, 2019; Stewart *et al.*, 2020) is echoed by a negative effect of higher burn severity on conifer canopy cover multiple decades after fire (Table 2). Climate change is contributing to higher severity fires which result in larger patches of bare ground (Steel *et al.*, 2018) devoid of conifer seedlings > 400 m from the edge of surviving forest (Stevens-Rumann & Morgan, 2019). This has led researchers to hypothesize that dispersal limitation in high severity burn scars will result in the long-term persistence of shrub fields (i.e. vegetation transitions) in some areas (Shive *et al.*, 2018). Our findings show that increasing burn severity, which coincides with larger seed-limited areas, results in less conifer canopy recovery 15-32 years later. Future work is required to confirm if the mechanism behind this relationship is due to shrub competition (e.g. Tubbesing et al., 2020, 2022), dispersal limitation, or a combination of these factors.

Our model suggests that conifer canopy recovery is negatively affected when mean annual daily maximum temperature (Tmax) exceeds a threshold of 21.1 ºC (Fig. 3b). However, there were only 35 pixels (255 ha) where Tmax exceeded this threshold and the effect was only weakly significant ($p = 0.02$). A negative effect of hotter mean temperatures is broadly consistent with prior findings that post-fire conifer regeneration declines most substantially in low-to-moderate

elevation (i.e. warmer) forests during drought (Stewart *et al.*, 2020). Although we did not find a strong effect of historical climate water deficit on RRI_{2019}, warmer temperatures result in higher evaporative demand which could tip the competitive balance in favor of more drought tolerant shrubs, such as those it the genus *Arctostaphylos* (Conard *et al.*, 1997), by reducing the amount of post-fire precipitation that remains available in the soil for seedlings. This interpretation is consistent with prior findings that competing shrubs were more abundant when post-fire weather was warmer and drier (Young *et al.*, 2019). Interestingly, historical Tmax was a better predictor of conifer canopy recovery than post-fire Tmax**,** implying that in contrast to post-fire precipitation, hotter temperatures may operate as a longer-term pressure throughout the recovery trajectory. This could be because maintenance respiration increases exponentially with temperature in temperate conifers (Ryan *et al.*, 1995). This is consistent with empirically-based modeling experiments showing that warmer temperatures in California's MCFs are associated with increased mortality and reduced growth (Battles *et al.*, 2008) and findings from a water-balance modeling study showing that increasing temperatures disproportionately reduce ponderosa pine regeneration towards the dry end of its range (Feddema *et al.*, 2013).

Our findings that post-fire precipitation and re-burns limit conifer canopy recovery (Table 2) support prior hypotheses that many regions of California's MCFs, particularly in the warmer, drier parts of the range, may transition to more drought tolerant, pyrogenic shrublands (Tepley *et al.*, 2017; Serra-Diaz *et al.*, 2018). However, there are reasons to suggest that range contractions will not cleanly follow changes in mean climate. Slope, aspect, and soil texture are expected to mediate the influence of increasing climatic aridity on plant moisture stress by providing topographic refugia (McLaughlin *et al.*, 2017; Ackerly *et al.*, 2020). We found some support for this phenomenon via a weak relationship between conifer canopy recovery and adjusted-folded-northness. More southwesterly facing slopes (values closer to -1; Fig. 4c) may be less hospitable to recovery, consistent with prior findings on seedling regeneration (Shive et al., 2018). Sites with the highest values for adjusted-folded-northness (facing northeast) had a slightly more negative influence on recovery than sites with intermediate values (facing northwest or southeast). We speculate that this may indicate the presence of a "recovery sweet spot" between moisture limited (southwest facing) and energy limited (northeast facing) aspects, but these hypotheses require stronger statistical support. Our finding that adjusted-folded-northness performed better than adjusted northness (Table 3) supports its use as the more ecologically relevant topographic metric (McCune & Keon, 2002).

Our findings that burn severity and re-burns both reduce long-term conifer recovery indicates that there will be less conifer canopy cover in the future if fire regimes continue to intensify. Our finding of a post-fire precipitation threshold supports the "storage hypothesis" (Warner & Chesson, 1985) as a potential mechanism for maintaining future patches of MCF even where mean annual precipitation is below 1,000 mm yr^{-1}. The storage hypothesis posits that environmentally sensitive, species-specific recruitment fluctuations occurring on inter-annual timescales result in temporally staggered infusions of diversity into the adult population (Warner & Chesson, 1985; Kelly & Bowler, 2002). The core range of MCF in California has historically experienced a mean annual precipitation > 1000 mm yr^{-1}, but there are some areas near the lower elevation range boundaries where this is not the case (Fig S2a) and where storage effects might

maintain conifer stands. California has experienced remarkable inter-annual and inter-decadal variation in precipitation over the last 1,000 years (Graumlich, 1993) and this pattern is predicted to continue (Allen & Anderson, 2018), making it plausible that future pulses of recruitment occurring when fire preceeds years with > 1,000 mm of precipitation could provide a form of "temporal climate refugia". However, rising temperatures mean that more precipitation will be required to sustain the same level of climate water deficit per unit of precipitation. Additionally, precipitation is anticipated to arrive in fewer, more intense events (Pierce *et al.*, 2013), making it unclear that simply reaching a mean annual threshold will facilitate conifer recovery in the future. The timescale of future precipitation variability will also likely affect future seed sources, with "mega-droughts" having the potential to kill mature seed trees (Park Williams *et al.*, 2013; Stewart *et al.*, 2020). Future work is required to investigate the plausibility of storage effects in maintaining future stands of MCFs near range boundaries by examining the extent to which the current demographic structure of MCFs at range margins indicates more episodic patterns of recruitment than in core range areas.

Post-fire planting and seed availability

The fraction of each pixel planted within the first 6 years of fire was not a significant predictor of regeneration outcomes in our top performing model (Table 3). This may be because the majority of our study area had little or no planting (Fig S3) and that our statistical approach was not specifically designed to detect the extent to which planting can benefit recovery outcomes. A future paired-control study design that controls for other sources of variation in recovery outcomes is required to better understand the efficacy of post-fire planting on long-term conifer canopy recovery. The unexpected negative influence of post-fire seed availability proxy (PostFireSAP) on conifer recovery conflicts with ecological expectations and prior findings that increased seed availability predicts increased post-fire seedling recovery (Puhlick *et al.*, 2012; Tepley *et al.*, 2017; Shive *et al.*, 2018; Rodman *et al.*, 2020; Stewart *et al.*, 2020). A possible reason for this unexpected outcome is the surprisingly strong positive correlation between PostFireSAP and slope in our dataset (Pearson's $r = 0.51$, $p < 0.001$). Sites on steeper slopes typically have higher topographic complexity, which may leave more pixels in the vicinity of surviving trees via "fire refugia" effects (Coop et al., 2019). However, steeper slopes are also less hospitable to seedling establishment (Shive et al., 2018; Stewart et al., 2020), which would likely lead to less conifer canopy recovery. We found that these steeper-sloped regions were also correlated with a higher re-burn fraction (Pearson's $r = 0.43$, $p < 0.001$) which decreases long-term conifer canopy recovery (Table 2).

The unexpected negative effect of PostFireSAP could also be due to limitations associated with quantifying seed availability using remote sensing approaches. Proximity to remnant conifer canopies doesn't necessarily indicate a viable seed supply in the field. Although seed production is shown to increase with parent tree basal area (Wright *et al.*, 2021), large inter-annual variation in seed production (Kelly & Sork, 2002; LaMontagne *et al.*, 2021) can influence post-fire seedling regeneration rates (Stewart *et al.*, 2020). We used a 400 m radius to calculate PostFireSAP following prior studies (Tepley *et al.*, 2017) which were based on findings from the Klamath Mountains that seedling density drops significantly at distances > 400 m from remnant

seed trees (Donato et al., 2009). However, it is possible that a smaller radius would capture seed availability more effectively. Shive *et al.* (2018) investigated the effect of seed availability proxy (SAP) radius on model performance by first fitting an optimal base model without SAP, and then adding SAP with differently sized radii. They determined that a radius of 150 m best predicted seedling regeneration. Future work will test the use of SAPs with differently sized radii and will investigate the effects of PostFireSAP within smaller study domains such as within individual fires and/or stratified by slope classes.

Management implications

This study provides managers with a longer-term view of where conifer canopy cover is likely to return multiple decades after fire. We found that low severity re-burns did not negatively affect long term conifer canopy recovery, indicating that the use of low severity prescribed fire as a fuel reduction tool will not reduce long-term conifer canopy recovery if it is used within 15-32 years after a medium or high severity fire. Unlike low severity re-burns we found that medium severity re-burns negatively affect the return of conifer canopy cover. Where the management goal is to return MCF to its pre-fire conifer canopy cover as quickly as possible, then medium severity re-burns should be avoided. However, the forest management recommendation for large areas of California's MCF is to reduce stand density for fire prevention and resilience (Stephens *et al.*, 2016). Under Euro-American forest management, canopy cover in California's MCFs increased from 25 to 49 percent (Stephens *et al.*, 2015) indicating that a return to the pre-fire canopy cover (i.e. RRI = 1) may not be desirable in many areas. Our analysis did not detect a positive or negative influence of planting on recovery outcomes, but our finding of a critical post-fire precipitation threshold suggests that the timing of planting with respect to variation of inter-annual precipitation will affect the long-term success of planting efforts. Planting in areas that typically receive more than 1,000 mm yr^{-1} (Fig. S2a) will likely have better outcomes than planting in areas receiving less than this amount. Efforts to regenerate areas that receive less precipitation than this on average, such as at low elevation range margins (Fig. 2a), should be timed to coincide with years of above-average precipitation. A recent meteorological analysis found that up to 20% of inter-annual precipitation variance in southern California is predictable using El Nino-Southern Oscillation and Pacific Decadal Oscillation indices (Cheng *et al.*, 2021), indicating that these indices could be used to plan planting efforts to coincide with years of predicted above-average rainfall at dry range margins.

Conclusion

Compared to shorter-term studies of seedling regeneration, our analysis of long-term conifer canopy recovery provided new insight into how topography, fire history, and post-fire precipitation affect forest structure many decades after fire. Conifer canopy recovery is negatively affected when mean annual precipitation in the first 3 years after fire is below 1000 mm yr^{-1}, when fire severity increases, and in areas that experience re-burns within the first 30 years after fire. Our findings indicate that intensifying fire regimes will reduce post-fire conifer canopy recovery in the future, likely causing an increase in shrub-dominated areas within California's MCFs. The timing of fire with respect to inter-annual variation in precipitation will

also drive long-term patterns in conifer canopy recovery. Post-fire planting in the drier parts of the MCF range, where mean annual precipitation is below 1000 mm yr^{-1}, should be planned to coincide with years of above-average precipitation.

Acknowledgments
A. Hanbury-Brown was supported by the National Science Foundation and the National Aeronautics and Space Administration while conducting this research.

Supplementary Information

Figure S1. Example training area used to quantify conifer canopy cover using the point intercept method (Elzinga *et al.*, 2001). The presence or absence of a conifer tree canopy was recorded at each yellow point on the systematic grid.

Figure S2. Areas within our study area where (a) historical (1980-2010) mean annual precipitation is greater than 1,000 mm yr^{-1} and (b) where historical (1980-2010) mean daily maximum temperature is less than 21.1 °C. Green pixels show areas outside these precipitation and temperature thresholds that still have the biophysical potential to support mixed conifer forest according to LANDFIRE (Rollins, 2009). These areas are located at the lower elevation range margins.

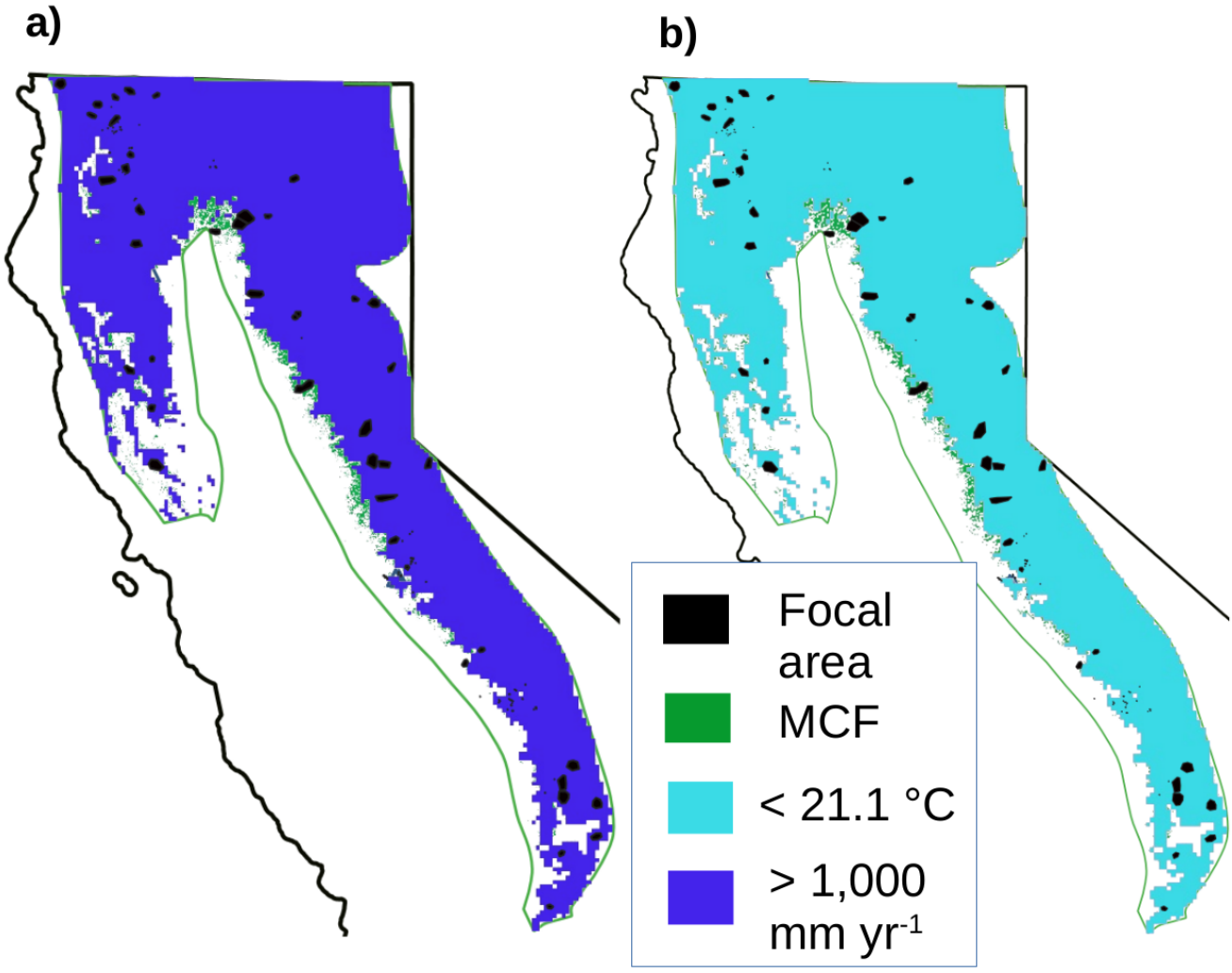


Figure S3. Histogram showing the distribution of the fraction of each 270 m pixel that was planted by the USFS within the first 6 years after fire. Most pixels in our study were not planted at all.

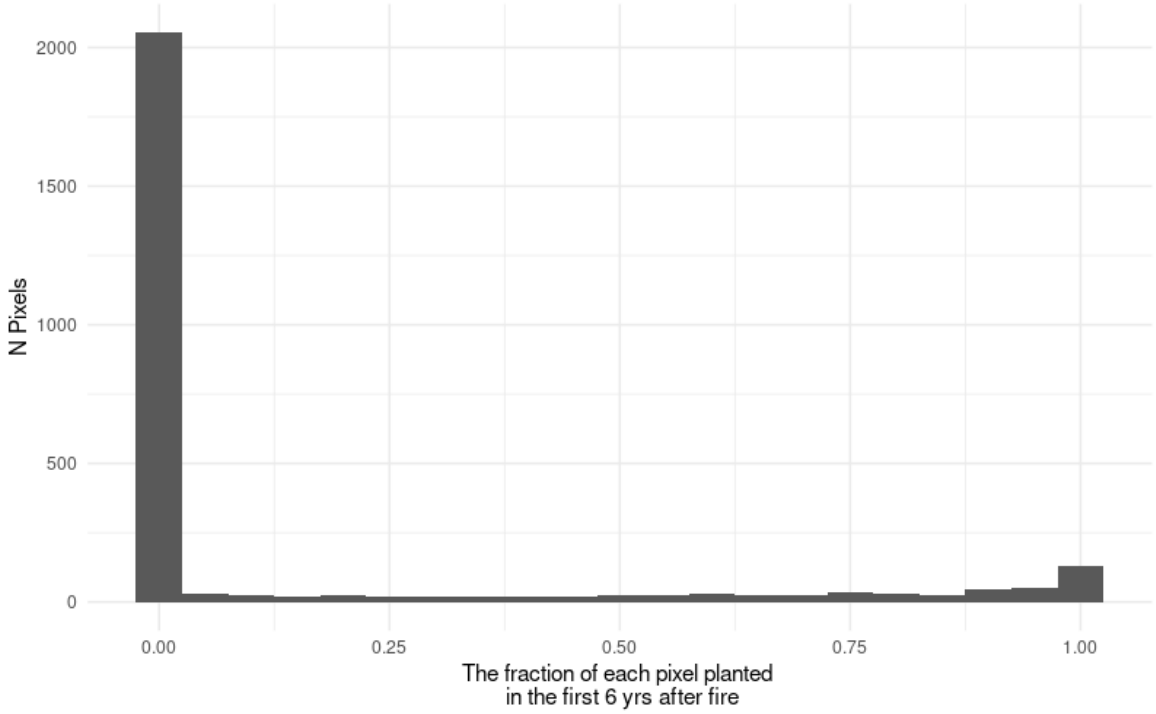
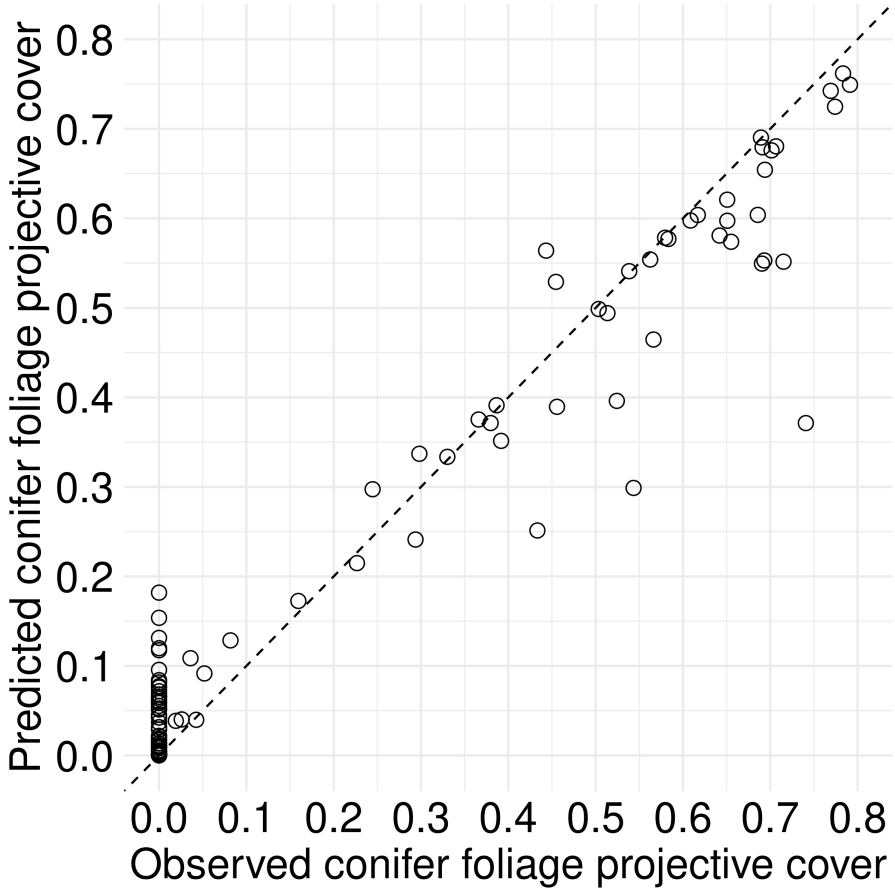


Figure S4. Example of validation imagery taken from the DJI Mavic Mini used to quantify percent conifer canopy cover in 36 validation plots throughout the study area.

Figure S5. Conifer foliage projective cover predicted by the random forest algorithm (Y-axis) plotted against training observations (X-axis) at 106 validation plots distributed throughout the study area. The dashed line (y = x) represents perfect agreement between predictions and field observations. The model predicts the training data with a root mean squared error of 0.06843 and adjusted R^2 of 0.95.



Conclusion

The work presented in this dissertation provides multiple new lines of evidence that forest regeneration processes will mediate changes in forest distribution, structure, and functional composition in response to global change. Sufficiently representing forest regeneration will help reduce uncertainty in projections of global climate. Current vegetation demographic models (VDMs) do not represent regeneration processes with sufficient complexity to capture forest responses to global change. Improved representations of regeneration are required to capture future tree migration, functional turnover, and forest regeneration after landscape-scale disturbance. The development and evaluation of the Tree Recruitment Scheme (TRS), presented in Chapter 2, showed that including an explicit representation of light- and moisture-sensitive germination, seedling mortality, and seedling to sapling transition rates improves predictions of tree recruitment under variable patch-level light, soil moisture and site-level precipitation scenarios. These results indicate that incorporating the TRS into VDMs will improve their ability to capture how variables such as increasing climatic water deficit and intensifying disturbance regimes (Seidl *et al.*, 2017) will mediate changes in forest structure and functional composition.

VDMs must be prepared to capture forest responses to changing precipitation and fire regimes. The work presented in Chapter 3 advanced prior studies on post-fire conifer seedling regeneration in California by showing that long-term conifer canopy recovery is negatively affected by higher intensity fire, when a re-burn occurs within the first 32 years of fire, and when mean annual precipitation in the first 3 years after fire drops below 1,000 mm yr^{-1}. These results indicate that in addition to capturing how fire and precipitation will influence mature tree mortality (Stephens & Moghaddas, 2005; Fettig *et al.*, 2019), it is important for VDMs to capture how these changing variables influence forest resilience. Our results suggest that more frequent, high severity fire will reduce post-fire conifer regeneration, likely causing an increase in shrub dominance in California's mixed conifer forests (MCFs). To understand the current ability of VDMs to capture these forest responses requires compatible data for model evaluation. This study generated a dataset of 2,529 forest recovery trajectories (19,698 ha) which will be used in future work as a benchmark to evaluate how well VDMs currently capture post-fire conifer canopy recovery in California.

The statistical relationships identified in Chapter 3 provide a starting point for improving model capacity to capture long-term recovery outcomes, but a more mechanistic understanding is required to support the development and parameterization of VDM algorithms for California's MCFs. The precipitation threshold identified in Chapter 3 may represent a critical point below which the post-fire competitive balance is tipped in favor of shrubs, but this hypothesis could not be tested with our remote sensing-based approach. A better understanding of how differences in hydraulic traits between conifers and shrubs, such as vulnerability to xylem cavitation, affect relative growth and mortality rates under varying precipitation scenarios may be required to capture and test seedling-shrub competitive dynamics in VDMs. Future empirical work should prioritize experiments that can identify how functional traits mediate physiological or demographic responses to environmental variables such as soil moisture (e.g. Mcdowell et al., 2008; Blackman et al., 2019). This work would support the development of more mechanistic

algorithms based on functional traits as well as support the parameterization of existing TRS algorithms for California's MCFs.

The successful implementation of the TRS in the Functionally Assembled Terrestrial Ecosystem Simulator (FATES-TRS, Appendix A) shows that the TRS is compatible with existing VDM model architecture. Future work will use FATES-TRS to predict California's MCFs under climate change and intensifying fire regimes. Once parameterized for California, FATES-TRS will be well-positioned to conduct modeling experiments that test hypotheses for how changing temperature, precipitation, and CO_2 will affect the future persistence of conifer forests. For example, in Chapter 3 we hypothesized that the post-fire precipitation threshold may result in storage effects (Warner & Chesson, 1985) whereby favorable precipitation years following fires could provide a form of "temporal climate refugia". FATES-TRS could be used to test this hypothesis by simulating conifer recruitment with alternate time series of precipitation and stochastic fire events. The modular implementation of the TRS will allow us to easily compare predictions of California's future vegetation with and without the TRS enabled, providing unique insight into how higher fidelity representations of regeneration affect predictions of forest distribution, structure, and function under global change.

Beyond California, this dissertation provides a starting point for future work needed to improve VDM representations of regeneration processes globally. Simulating global forests will require the representation of additional processes not included in the current version of the TRS. For example, capturing future tree migration and landscape-scale disturbance-recovery dynamics will require the representation of dispersal limitation. Simulating forests in fire- and hurricane-prone ecosystems will require the representation of re-sprouting. Not all processes will be relevant to all forests, highlighting the need for modular complexity as a path forward in VDM development (Fisher & Koven, 2020). Regionally specific modules and parameterizations will allow VDMs to mitigate increasing model complexity while making regional-scale simulations more accurate and relevant to ecosystem management decision-making. For example, a future regeneration module parameterized specifically for California may be able to identify post-fire areas that should be prioritized for planting and/or shrub management versus places where conifer forests will likely never return due to biophysical limitations. VDM-based recommendations would go beyond prior recommendations based on statistical models (Shive *et al.*, 2018) because they can more mechanistically represent future ecophysiology including response to rising CO_2 including response to rising CO_2, demography, and competition. Improving the stewardship of forested ecosystems will benefit from the development of high-performing VDMs. The work presented here demonstrates that regeneration processes should be a critical part of this effort.

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.

Ackerly DD, Kling MM, Clark ML, Papper P, Oldfather MF, Flint AL, Flint LE. **2020**. Topoclimates, refugia, and biotic responses to climate change. *Frontiers in Ecology and the Environment* **18**: 288–297.

Airey C, Skinner CN, Taylor AH. **2016**. High severity fire and mixed conifer forest-chaparral dynamics in the southern Cascade Range , USA. *Forest Ecology and Management* **363**: 74–85.

Alencar AA, Brando PM, Asner GP, Putz FE. **2015**. Landscape fragmentation, severe drought, and the new Amazon forest fire regime. *Ecological Applications* **25**: 1493–1505.

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**.**

Allen RJ, Anderson RG. **2018**. 21st century California drought risk linked to model fidelity of the El Niño teleconnection. *npj Climate and Atmospheric Science* **1**: 1–14.

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.

Andivia E, Madrigal-González J, Villar-Salvador P, Zavala MA. **2018**. Do adult trees increase conspecific juvenile resilience to recurrent droughts? Implications for forest regeneration. *Ecosphere* **9**: e02282–e02282.

Andrus RA, Harvey BJ, Hoffman A, Veblen TT. **2020**. Reproductive maturity and cone abundance vary with tree size and stand basal area for two widely distributed conifers. *Ecosphere* **11**. 10.1002/ecs2.3092

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.

Asner GP. **1998**. Biophysical and Biochemical Sources of Variability in Canopy Reflectance. *Remote Sensing of Environment* **64**: 234–253.

Atondo-Bueno EJ, López-Barrera F, Bonilla-Moheno M, Williams-Linera G, Ramírez-Marcial N. **2016**. Direct seeding of *Oreomunnea mexicana*, a threatened tree species from Southeastern Mexico. *New Forests* **47**: 845–860.

Augspurger CK. **1984**. Light Requirements of Neotropical Tree Seedlings : A Comparative Study of Growth and Survival. *Journal of Ecology* **72**: 777–795.

Balderrama SIV, Chazdon RL. **2005**. Light-dependent seedling survival and growth of four tree species in Costa Rican second-growth rain forests. *Journal of Tropical Ecology* **21**: 383–395.

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.

Batterman SA, Hedin LO, van Breugel M, Ransijn J, Craven DJ, Hall JS. **2013**. Key role of symbiotic dinitrogen fixation in tropical forest secondary succession. *Nature* **502**: 224–227.

Battles JJ, Robards T, Das A, Waring K, Gilless JK, Biging G, Schurr F. **2008**. Climate change impacts on forest growth and tree mortality: a data-driven modeling study in the mixed-conifer forest of the Sierra Nevada, California. *Climatic Change* **87**: 193–213.

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.

Becker P, Rabenold PE, Idol JR, Smith AP. **1988**. Water Potential Gradients for Gaps and Slopes in a Panamanian Tropical Moist Forest's Dry Season. *Journal of Tropical Ecology* **4**: 173–184.

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

Blackman CJ, Creek D, Maier C, Aspinwall MJ, Drake JE, Pfautsch S, O'Grady A, Delzon S, Medlyn BE, Tissue DT, *et al.* 2019. Drought response strategies and hydraulic traits contribute to mechanistic understanding of plant dry-down to hydraulic failure. *Tree Physiology* **39**: 910–924.

Bloor JMG, Grubb PJ. **2003**. Growth and mortality in high and low light: Trends among 15 shade-tolerant tropical rain forest tree species. *Journal of Ecology* **91**: 77–85.

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

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

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

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

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.

Bourg, N.A., W.J. McShea, J.R. Thompson, J.C. McGarvey, and X. Shen. 2013. Initial census, woody seedling, seed rain, and stand structure data for the SCBI SIGEO Large Forest Dynamics Plot. *Ecology* **94** (9): 2111-2112. http://dx.doi.org/10.1890/13-0010.1
Brokaw NVL. **1985**. Gap-phase regeneration in a tropical forest. *Ecology* **66**: 682–687.

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.

Brum M, Vadeboncoeur MA, Ivanov V, Asbjornsen H, Saleska S, Alves LF, Penha D, Dias JD, Aragão LEOC, Barros F, *et al.* 2019. Hydrological niche segregation defines forest structure and drought tolerance strategies in a seasonal Amazon forest. *Journal of Ecology* **107**: 318–333.

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

California Forest Management Task Force. **2021**. California's Wildfire and Forest Resilience Action Plan. https://fmtf.fire.ca.gov/.

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.

Chang-Yang C-H, Needham J, Lu C-L, Hsieh C-F, Sun I-F, McMahon SM. **2021**. Closing the life cycle of forest trees: The difficult dynamics of seedling-to-sapling transitions in a subtropical rainforest. *Journal of Ecology* **109**: 2705–2716.

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

Chapman TB, Schoennagel T, Veblen TT, Rodman KC. **2020**. Still standing: Recent patterns of post-fire conifer refugia in ponderosa pine-dominated forests of the Colorado Front Range. *PLoS ONE* **15**: 1–30.

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

Cheng R, Novak L, Schneider T. **2021**. Predicting the Interannual Variability of California's Total Annual Precipitation. *Geophysical Research Letters* **48**: e2020GL091465.

Chesson PL, Warner RR. **1981**. Environmental Variability Promotes Coexistence in Lottery Competitive Systems Author (s): Peter L . Chesson and Robert R . Warner Published by : The University of Chicago Press for The American Society of Naturalists Stable URL : https://www.jstor.org/s. *The American Naturalist* **117**: 923–943.

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 DA. **2017**. La Selva daily solar radiation (pyranometer) March 1982 through 2016. [WWW document] URL https://bixa.tropicalstudies.org/meteoro/data/lsdata/LS%20daily%20solar%20radiation%20(pyranometer)%20March1992%20through%202016%20(Jan%202017,%20DAC).xlsx [accessed 16 February 2022].

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.

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

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.

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.

Coley PD, Barone JA. **1996**. Herbivory and plant defenses in tropical forests. *Annual Review of Ecology and Systematics* **27**: 305–335.

Comita LS, Hubbell SP. **2009**. Local neighborhood and species shade tolerance influence survival in a diverse seedling bank. *Ecology* **90**: 328–334.

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.

Conard SG, Radosevich SR. **1982**. Growth Responses of White Fir to Decreased Shading and Root Competition by Montane Chaparral Shrubs. *Forest Science* **28**: 309–320.

Conard SG, Sparks SR, Regelbrugge JC. **1997**. Comparative Plant Water Relations and Soil Water Depletion Patterns of Three Seral Shrub Species on Forest Sites in Southwestern Oregon. *Forest Science* **43**: 336–347.

Condit R, Engelbrecht BMJ, Pino D, Pérez R, Turner BL. **2013**. Species distributions in response to individual soil nutrients and seasonal drought across a community of tropical trees. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 5064–8.

Condit R, Perez R, Aguilar S, Lao S, Foster RB, Hubbell S. **2019**. Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years. Dryad, https://doi.org/10.15146/5xcp-0d46.

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

Condit R. **1998**. *Tropical Forest Census Plots*. Springer-Verlag and R. G. Landes Company, Berlin, Germany, and Georgetown, Texas.

Condit, R., D.M. Windsor, and S.P. Hubbell. 2013. NPP Tropical Forest: Barro Colorado, Panama, 1969-2000, R1. Data set. Available on-line [http://daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA doi:10.3334/ORNLDAAC/157

Conlisk E, Castanha C, Germino MJ, Veblen TT, Smith JM, Moyes AB, Kueppers LM. **2017**. Seed origin and warming constrain lodgepole pine recruitment, slowing the pace of population range shifts. *Global Change Biology* **24**: 197–211.

Coop JD, Delory TJ, Downing WM, Haire SL, Krawchuk MA, Miller C, Parisien MA, Walker RB. **2019**. Contributions of fire refugia to resilient ponderosa pine and dry mixed-conifer forest landscapes. *Ecosphere* **10**.

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

d'Oliveira MVN, Ribas LA. **2011**. Forest regeneration in artificial gaps twelve years after canopy opening in Acre State Western Amazon. *Forest Ecology and Management* **261**: 1722–1731.

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.

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.

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

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.

Dewitz J. **2021**. National Land Cover Database (NLCD) 2019 Products.

Dietze MC, Clark JS. **2008**. Changing the gap dynamics paradigm: Vegetative regeneration control on forest response tO disturbance. *Ecological Monographs* **78**: 331–347.

Diffenbaugh NS, Swain DL, Touma D. **2015**. Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences* **112**: 3931–3936.

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.

Donato DC, Fontaine JB, Campbell JL, Robinson WD, Kauffman JB, Law BE. **2009**. Conifer regeneration in stand-replacement portions of a large mixed-severity wildfire in the Klamath–Siskiyou Mountains. *Canadian Journal of Forest Research* **39**: 823–838.

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.

Ediriweera S, Pathirana S, Danaher T, Nichols J, Moffiet T. **2013**. Evaluation of Different Topographic Corrections for Landsat TM Data by Prediction of Foliage Projective Cover (FPC) in Topographically Complex Landscapes. *Remote Sensing* **5**: 6767–6789.

Eidenshink J, Schwind B, Brewer K, Zhu Z-L, Quayle B, Howard S. **2007**. A Project for Monitoring Trends in Burn Severity. *Fire Ecology* **3**: 3–21.

Elzinga CL, Salzer DW, Willoughby JW, Gibbs JP. **2001**. *Monitoring plant and animal populations: a handbook for field biologists*. John Wiley & Sons.

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.

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

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, Wright SJ, De Steven D. **2002**. Survival and ecophysiology of tree seedlings during El Niño drought in a tropical moist forest in Panama. *Journal of Tropical Ecology* **18**: 569–579.

Etchberger RC, Krausman PR. **1997**. Evaluation of Five Methods for Measuring Desert Vegetation. *Wildlife Society Bulletin (1973-2006)* **25**: 604–609.

Evans A, Everett R, Stephens S, Youlz J. **2011**. Comprehensive Fuels Treatment Practices Guide for Mixed Conifer Forests: California, Central and Southern Rockies, and the Southwest. *JFSP Synthesis Reports*.

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].

Feddema JJ, Mast JN, Savage M. **2013**. Modeling high-severity fire, drought and climate change impacts on ponderosa pine regeneration. *Ecological Modelling* **253**: 56–69.

Fetcher N, Oberbauer SF, Strain BR. **1985**. Vegetation effects on microclimate in lowland tropical forest in Costa Rica. *International Journal of Biometeorology* **29**: 145–155.

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, 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**.

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.

Flint LE, Flint AL, Thorne JH, Boynton R. **2013**. Fine-scale hydrologic modeling for regional landscape applications: the California Basin Characterization Model development and performance. *Ecological Processes* **2**: 25.

Flint LE, Underwood EC, Flint AL, Hollander AD. **2019**. Characterizing the Influence of Fire on Hydrology in Southern California. *Natural Areas Journal* **39**: 108–121.

Folke C, Carpenter S, Walker B, Scheffer M, Elmqvist T, Gunderson L, Holling CS. **2004**. Regime Shifts, Resilience, and Biodiversity in Ecosystem Management. *Annual Review of Ecology, Evolution, and Systematics* **35**: 557–581.

Forrest M, Tost H, Lelieveld J, Hickler T. **2020**. Including vegetation dynamics in an atmospheric chemistry-enabled general circulation model: Linking LPJ-GUESS (v4.0) with the EMAC modelling system (v2.53). *Geoscientific Model Development* **13**: 1285–1309.

Foster AC, Martin PH, Redmond MD. **2020**. Soil moisture strongly limits Douglas-fir seedling establishment near its upper elevational limit in the southern Rocky Mountains. *Canadian Journal of Forest Research* **50**: 837–842.

Frazier RJ, Coops NC, Wulder MA, Hermosilla T, White JC. **2018**. Analyzing spatial and temporal variability in short-term rates of post- fi re vegetation return from Landsat time series. *Remote Sensing of Environment* **205**: 32–45.

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

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.

García-Rodríguez A, García-Rodríguez S, Díez-Mediavilla M, Alonso-Tristán C. **2020**. Photosynthetic Active Radiation, Solar Irradiance and the CIE Standard Sky Classification. *Applied Sciences* **10.** 8007.

Garwood NC. **1983**. Seed Germination in a Seasonal Tropical Forest in Panama: A Community Study. *Ecological Monographs* **53**: 159–181.

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

Gilbert GS, Harms KE, Hamill DN, Hubbell SP. **2001**. Effects of seedling size, El Niño drought, seedling density, and distance to nearest conspecific adult on 6-year survival of *Ocotea whitei* seedlings in Panamá. *Oecologia* **127**: 509–516.

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.

Gómez-Aparicio L, Zamora R, Castro J, Hódar JA. **2008**. Facilitation of tree saplings by nurse plants: Microhabitat amelioration or protection against herbivores? *Journal of Vegetation Science* **19**: 161–172.

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. **2017**. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*.

Graumlich LJ. **1993**. A 1000-Year Record of Temperature and Precipitation in the Sierra Nevada. *Quaternary Research* **39**: 249–255.

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.

Hanbury-Brown AR, Ward RE, Kueppers LM. **2022**. Forest regeneration within Earth system models: current process representations and ways forward. *New Phytologist* doi: https://doi.org/10.1111/nph.18131

Hanbury-Brown AR, Powell TL, Muller-Landau HC, Wright SJ, Kueppers LM. **2022**. Simulating environmentally-sensitive tree recruitment in vegetation demographic models. *New Phytologist* doi: https://doi.org/10.1111/nph.18059.

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.

Harms KE, Condit R, Hubbell SP, Foster RB. **2001**. Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *Journal of Ecology* **89**: 947–959.

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.

Hastings A, Abbott KC, Cuddington K, Francis T, Gellner G, Lai YC, Morozov A, Petrovskii S, Scranton K, Zeeman ML. **2018**. Transient phenomena in ecology. *Science* **361**.

Haszpra T, Herein M, Bódai T. **2020**. Investigating ENSO and its teleconnections under climate change in an ensemble view – a new perspective. *Earth System Dynamics* **11**: 267–280.

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.

Ho TK. **1995**. Random decision forests. In: Proceedings of 3rd International Conference on Document Analysis and Recognition. 278–282 vol.1.

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.

Holm JA, Shugart HH, Van Bloem SJ, Larocque GR. 2012. Gap model development, validation, and application to succession of secondary subtropical dry forests of Puerto Rico. *Ecological Modelling* **233**: 70–82.

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.

Hope A, Tague C, Clark R. **2007**. Characterizing post-fire vegetation recovery of California chaparral using TM/ETM+ time-series data. *International Journal of Remote Sensing* **28**: 1339–1354.

Hubbell SP, Foster RB, O'Brien ST, Harms KE, Condit R, Wechsler B, Wright SJ, Loo De Lao S. **1999**. Light-Gap Disturbances , Recruitment Limitation , and Tree Diversity in a Neotropical Forest. *Science* **283**: 554–558.

Hubbell, S.P., Condit, R., and Foster, R.B. 2015. Barro Colorado Forest Census Plot Data. [WWW Document] URL http://ctfs.si.edu/webatlas/datasets/bci [accessed 17 Oct 2021].
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.

Jabot F, Etienne RS, Chave J. **2008**. Reconciling neutral community models and environmental filtering: theory and an empirical test. *Oikos* **117**: 1308–1320.

Jin S, Homer C, Yang L, Danielson P, Dewitz J, Li C, Zhu Z, Xian G, Howard D. **2019**. Overall Methodology Design for the United States National Land Cover Database 2016 Products. *Remote Sensing* **11**: 2971.

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.

Johnson DJ, Condit R, Hubbell SP, Comita LS. **2017**. Abiotic niche partitioning and negative density dependence drive tree seedling survival in a tropical forest. *Proceedings of the Royal Society B: Biological Sciences* **284**: 20172210.

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.

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

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.

Kassambara A. 2020. *ggpubr: 'ggplot2' Based Publication Ready Plots*. [WWW document] URL https://cran.r-project.org/web/packages/ggpubr/index.html [accessed 1 March 2020]
Kassambara A. 2020. *ggpubr: 'ggplot2' Based Publication Ready Plots*. [WWW document] URL https://cran.r-project.org/web/packages/ggpubr/index.html [accessed 1 March 2020]

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.

Kelly C, Bowler M. **2002**. Coexistence and relative abundance in forest trees. *Nature* **417**: 437–440.

Kelly D, Sork VL. **2002**. Mast Seeding in Perennial Plants: Why, How, Where? *Annual Review of Ecology and Systematics* **33**: 427–447.

Kitajima K. **1994**. Relative importance of photosynthetic traits and allocation patterns as correlates of seedling shade tolerance of 13 tropical trees. *Oecologia* **98**: 419–428.

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

Kohyama TS, Kohyama TI, Sheil D. **2018**. Definition and estimation of vital rates from repeated censuses: Choices, comparisons and bias corrections focusing on trees. *Methods in Ecology and Evolution* **9**: 809–821.
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.

Koven CD, Knox RG, Fisher RA, Fisher RA, Chambers JQ, Chambers JQ, Christoffersen BO, Davies SJ, Detto M, Detto M, *et al.* 2020. Benchmarking and parameter sensitivity of physiological and vegetation dynamics using the Functionally Assembled Terrestrial Ecosystem Simulator (FATES) at Barro Colorado Island, Panama. *Biogeosciences* **17**: 3017–3044.
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.

Lackey RT. **1998**. Seven pillars of ecosystem management1Modified from a presentation given at the Symposium Ecosystem Health and Medicine: Integrating Science, Policy, and Management, Ottawa, Ontario, Canada, June 19–23, 1994. This paper has been subjected to scientific peer review but does not necessarily reflect policy positions of the United States Environmental Protection Agency or any other organization.1. *Landscape and Urban Planning* **40**: 21–30.

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.

LaMontagne JM, Redmond MD, Wion AP, Greene DF. **2021**. An assessment of temporal variability in mast seeding of North American Pinaceae. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**: 20200373.

Lancaster N, Baas A. **1998**. Influence of vegetation cover on sand transport by wind: field studies at Owens Lake, California. *Earth Surface Processes and Landforms* **23**: 69–82.

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.

Larocque GR, Archambault L, Delisle C. 2006. Modelling forest succession in two southeastern Canadian mixedwood ecosystem types using the ZELIG model. *Ecological Modelling* **199**: 350–362

Lary DJ, Alavi AH, Gandomi AH, Walker AL. **2016**. Machine learning in geosciences and remote sensing. *Geoscience Frontiers* **7**: 3–10.

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.

Lebrija-Trejos E, Pérez-GarcíA EA, Meave JA, Bongers F, Poorter L. **2010**. Functional traits and environmental filtering drive community assembly in a species-rich tropical system. *Ecology* **91**: 386–398.

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.

Lewis SL, Maslin MA. **2015**. Defining the Anthropocene. *Nature* **519**: 171–180.

Li H, Wan W, Fang Y, Zhu S, Chen X, Liu B, Hong Y. **2019**. A Google Earth Engine-enabled software for efficiently generating high-quality user-ready Landsat mosaic images. *Environmental Modelling & Software* **112**: 16–22.

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, Löffler TJ. 2006. Intra-specific density dependence is required to maintain species diversity in spatio-temporal forest simulations with reproduction. *Ecological Modelling* **198**: 341–361

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.

Littlefield CE, Dobrowski SZ, Abatzoglou JT, Parks SA, Davis KT. **2020**. A climatic dipole drives short- and long-term patterns of postfire forest recovery in the western United States. *Proceedings of the National Academy of Sciences* **117**: 29730–29737.

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 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].

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.

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.

Markl JS, Schleuning M, Forget PM, Jordano P, Lambert JE, Traveset A, Wright SJ, Böhning-Gaese K. **2012**. Meta-Analysis of the Effects of Human Disturbance on Seed Dispersal by Animals. *Conservation Biology* **26**: 1072–1081.

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.

Martínez Cano I, Shevliakova E, Malyshev S, Wright SJ, Detto M, Pacala SW, Muller-Landau HC. **2020**. Allometric constraints and competition enable the simulation of size structure and carbon fluxes in a dynamic vegetation model of tropical forests (LM3PPA-TV). *Global Change Biology* **26**: 4478–4494.

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.

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

McCune B, Keon D. **2002**. Equations for potential annual direct incident radiation and heat load. *Journal of Vegetation Science* **13**: 603–606.

Mcdowell N, Pockman WT, Allen CD, Breshears DD, Cobb N, Kolb T, Plaut J, Sperry J, West A, Williams DG, *et al.* 2008. Mechanisms of Plant Survival and Mortality during Drought : Why Do Some Plants Survive while Others Succumb to Drought ? Published by : Wiley on behalf of the New Phytologist Trust Stable URL : http://www.jstor.org/stable/30149305 REFERENCES Linked refere. **178**: 719–739.

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–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, Ackerly DD, Klos PZ, Natali J, Dawson TE, Thompson SE. **2017**. Hydrologic refugia, plants, and climate change. *Global Change Biology* **23**: 2941–2961.

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.

McMahon S., Parker, G. 2021. Smithsonian Environmental Research Center. [WWW Document] URL https://forestgeo.si.edu/sites/north-america/smithsonian-environmental-research-center [accessed 17 Oct 2021].

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.

Meng R, Dennison PE, Huang C, Moritz MA, Antonio CD. **2015**. Effects of fire severity and post- fire climate on short-term vegetation recovery of mixed-conifer and red fi r forests in the Sierra Nevada Mountains of California. *Remote Sensing of Environment* **171**: 311–325.

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 JD, Thode AE. **2007**. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sensing of Environment* **109**: 66–80.

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

Minor DM, Kobe RK. **2019**. Fruit production is influenced by tree size and size-asymmetric crowding in a wet tropical forest. *Ecology and Evolution* **9**: 1458–1472.

Mitchard ETA. **2018**. The tropical forest carbon cycle and climate change. *Nature* **559**: 527–534.

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

Mok H-F, Arndt SK, Nitschke CR. **2012**. Modelling the potential impact of climate variability and change on species regeneration potential in the temperate forests of South-Eastern Australia. *Global Change Biology* **18**: 1053–1072.

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.

Myers JA, Kitajima K. **2007**. Carbohydrate storage enhances seedling shade and stress tolerance in a neotropical forest. *Journal of Ecology* **95**: 383–395.

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.

Notman E, Gorchov DL. **2001**. Variation in Post-dispersal Seed Predation in Mature Peruvian Lowland Tropical Forest and Fallow Agricultural Sites. *Biotropica* **33**: 621–636.

O'Brien MJ, Peréz-Aviles D, Powers JS. **2018**. Resilience of seed production to a severe El Niño-induced drought across functional groups and dispersal types. *Global Change Biology* **24**: 5270–5280.

Obroucheva N, Sinkevich I, Lityagina S. 2016. Physiological aspects of seed recalcitrance: A case study on the tree *Aesculus hippocastanum*. *Tree Physiology* **36**: 1127–1150.

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.

Park Williams A, Allen CD, Macalady AK, Griffin D, Woodhouse CA, Meko DM, Swetnam TW, Rauscher SA, Seager R, Grissino-Mayer HD, *et al.* 2013. Temperature as a potent driver of regional forest drought stress and tree mortality. *Nature Climate Change* **3**: 292–297.

Paton S. **2019**. Barro Colorado Island, Lutz catchment, Soil moisture, manual. doi:https://doi.org/10.25573/data.10042517.v3. [accessed 16 February 2022].

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.

Pearson RG, Dawson TP. **2003**. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography* **12**: 361–371.

Pearson TRH, Burslem DFRP, Mullins CE, Dalling JW. **2002**. Germination ecology of neotropical pioneers: Interacting effects of environmental conditions and seed size. *Ecology* **83**: 2798–2807.

Pierce DW, Cayan DR, Das T, Maurer EP, Miller NL, Bao Y, Kanamitsu M, Yoshimura K, Snyder MA, Sloan LC, *et al.* 2013. The Key Role of Heavy Precipitation Events in Climate Model Disagreements of Future Annual Precipitation Changes in California. *Journal of Climate* **26**: 5879–5896.

Plamboeck AH, North M, Dawson TE. **2008**. Conifer Seedling Survival Under Closed-Canopy And Manzanita Patches In the Sierra Nevada. *Madroño* **55**: 191–201.

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.

Poorter L. **1999**. Growth responses of 15 rain-forest tree species to a light gradient: The relative importance of morphological and physiological traits. *Functional Ecology* **13**: 396–410.

Poorter L. **2007**. Are Species Adapted to Their Regeneration Niche, Adult Niche, or Both? *The American Naturalist* **169**: 433–442.

Poortinga A, Tenneson K, Shapiro A, Nquyen Q, San Aung K, Chishtie F, Saah D. **2019**. Mapping Plantations in Myanmar by Fusing Landsat-8, Sentinel-2 and Sentinel-1 Data along with Systematic Error Quantification. *Remote Sensing* **11**.

Poulter B, Ciais P, Hodson E, Lischke H, Maignan F, Plummer S, Zimmermann NE. **2011**. Plant functional type mapping for earth system models. *Geoscientific Model Development* **4**: 993–1010.

Powell T, Kueppers L, Paton S. **2017**. Seven years (2008-2014) of meteorological observations plus a synthetic El Nino drought for BCI Panama. doi: 10.15486/ngt/1414275. [accessed 16 February 2022].

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.

Puhlick JJ, Laughlin DC, Moore MM. **2012**. Factors influencing ponderosa pine regeneration in the southwestern USA. *Forest Ecology and Management* **264**: 10–19.

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.

R Core Team. **2020**. R: A Language and Environment for Statistical Computing (version 3.6.3). *R Foundation for Statistical Computing, Vienna, Austria*.

Rodman KC, Veblen TT, Battaglia MA, Chambers ME, Fornwalt PJ, Holden ZA, Kolb TE, Ouzts JR, Rother MT. **2020**. A changing climate is snuffing out post-fire recovery in montane forests. *Global Ecology and Biogeography. doi: 10.1111/geb.13174.*

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.

Rollins MG. **2009**. LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. *International Journal of Wildland Fire* **18**: 235–249.

Rollinson CR, Finley AO, Alexander MR, Banerjee S, Dixon Hamil K-A, Koenig LE, Locke DH, DeMarche ML, Tingley MW, Wheeler K, *et al.* 2021. Working across space and time: nonstationarity in ecological research and application. *Frontiers in Ecology and the Environment* **19**: 66–72.

Roy DP, Zhang HK, Ju J, Gomez-Dans JL, Lewis PE, Schaaf CB, Sun Q, Li J, Huang H, Kovalskyy V. **2016**. A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. *Remote Sensing of Environment* **176**: 255–271.

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.

Rüger N, Huth A, Hubbell SP, Condit R. **2009**. Response of recruitment to light availability across a tropical lowland rain forest community. *Journal of Ecology* **97**: 1–3.

Ruiz Talonia L, Reid N, Gross CL, Whalley RDB. **2017**. Germination ecology of six species of *Eucalyptus* in shrink–swell vertosols: moisture, seed depth and seed size limit seedling emergence. *Australian Journal of Botany* **65**: 22–22.

Running, S., Mu, Q., & Zhao, M. 2015. MOD17A3H MODIS/terra net primary production yearly L4 global 500m SIN grid V006. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD17A3H.006 [accessed 16 February 2022].

Ryan MG, Gower ST, Hubbard RM, Waring RH, Gholz HL, Cropper WP, Running SW. **1995**. Woody tissue maintenance respiration of four conifers in contrasting climates. *Oecologia* **101**: 133–140.

S. A. Soenen, D. R. Peddle, C. A. Coburn. **2005**. SCS+C: a modified Sun-canopy-sensor topographic correction in forested terrain. *IEEE Transactions on Geoscience and Remote Sensing* **43**: 2148–2159.

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.

Sansevero JBB, Garbin ML, Sánchez-Tapia A, Valladares F, Scarano FR. **2020**. Fire drives abandoned pastures to a savanna-like state in the Brazilian Atlantic Forest. *Perspectives in Ecology and Conservation* **18**: 31–36.

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.

Sato H. **2009**. Simulation of the vegetation structure and function in a Malaysian tropical rain forest using the individual-based dynamic vegetation model SEIB-DGVM. *Forest Ecology and Management* **257**: 2277–2286.

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* **201**: 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.

Schonlau M, Zou RY. **2020**. The random forest algorithm for statistical learning. *The Stata Journal* **20**: 3–29.

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.

Serra-Diaz JM, Maxwell C, Lucash MS, Scheller RM, Laflower DM, Miller AD, Tepley AJ, Epstein HE, Anderson-Teixeira KJ, Thompson JR. **2018**. Disequilibrium of fire-prone forests sets the stage for a rapid decline in conifer dominance during the 21stcentury. *Scientific Reports* **8**: 1–12.

Shive KL, Preisler HK, Welch KR, Safford HD, Butz RJ, O'Hara KLL, Stephens SL. **2018**. From the stand scale to the landscape scale: predicting the spatial patterns of forest regeneration after disturbance. *Ecological Applications* **28**: 1626–1639.

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.

Slik JWF. **2004**. El Niño droughts and their effects on tree species composition and diversity in tropical rain forests. *Oecologia* **141**: 114–120.

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, 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.

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

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

Soto-Berelov M, Jones S, Farmer E, Woodgate W. **2015**. Chapter 2. Review of validation standards of biophysical Earth Observation products. In: AusCover Good Practice Guidelines: A technical handbook supporting calibration and validation activities of remotely sensed data product. TERN AusCover, 8–30.

Spear ER, Broders KD. **2021**. Host-generalist fungal pathogens of seedlings may maintain forest diversity via host-specific impacts and differential susceptibility among tree species. *New Phytologist* **231**: 460–474.

Spear ER, Coley PD, Kursar TA. **2015**. Do pathogens limit the distributions of tropical trees across a rainfall gradient? *Journal of Ecology* **103**: 165–174.

Steel ZL, Koontz MJ, Safford HD. **2018**. The changing landscape of wildfire: burn pattern trends and implications for California's yellow pine and mixed conifer forests. *Landscape Ecology* **33**: 1159–1176.

Stephens SL, Lydersen JM, Collins BM, Fry DL, Meyer MD. **2015**. Historical and current landscape-scale ponderosa pine and mixed conifer forest structure in the Southern Sierra Nevada. *Ecosphere* **6**: art79.

Stephens SL, Collins BM, Biber E, Fulé PZ. **2016**. U.S. federal fire and forest policy: emphasizing resilience in dry forests. *Ecosphere* **7**: e01584.

Stevens JT, Collins BM, Miller JD, North MP, Stephens SL. **2017**. Changing spatial patterns of stand-replacing fire in California conifer forests. *Forest Ecology and Management* **406**: 28–36.

Stevens-Rumann CS, Kemp KB, Higuera PE, Harvey BJ, Rother MT, Donato DC, Morgan P, Veblen TT. **2018**. Evidence for declining forest resilience to wildfires under climate change. *Ecology Letters* **21**: 243–252.

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

Stewart JAE, van Mantgem PJ, Young DJN, Shive KL, Preisler HK, Das AJ, Stephenson NL, Keeley JE, Safford HD, Wright MC, *et al.* 2020. Effects of postfire climate and seed availability on postfire conifer regeneration. *Ecological Applications*: 0–1.

Svenning JC, Sandel B. **2013**. Disequilibrium vegetation dynamics under future climate change. *American Journal of Botany* **100**: 1266–1286.

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, Thomann E, Veblen TT, Perry GLW, Holz A, Paritsis J, Kitzberger T, Anderson-Teixeira KJ. **2018**. Influences of fire–vegetation feedbacks and post-fire recovery rates on forest landscape vulnerability to altered fire regimes. *Journal of Ecology* **106**: 1925–1940.

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.

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.

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

Tubbesing CL, York RA, Stephens SL, Battles JJ. **2020**. Rethinking fire-adapted species in an altered fire regime. *Ecosphere* **11**: e03091.

Tubbesing CL, Young DJN, York RA, Stephens SL, Battles JJ. **2022**. Incorporating Shrub Neighborhood Dynamics to Predict Forest Succession Trajectories in an Altered Fire Regime. *Ecosystems* **25**: 136–154.

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.

Turner MG. **2010**. Disturbance and landscape dynamics in a changing world. *Ecology* **91**: 2833–2849.

[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.

Vacchiano G, Ascoli D, Berzaghi F, Lucas-Borja ME, Caignard T, Collalti A, Mairota P, Palaghianu C, Reyer CPO, Sanders TGM, *et al.* 2018. Reproducing reproduction: How to simulate mast seeding in forest models. *Ecological Modelling* **376**: 40–53.

Valdez JW, Hartig F, Fennel S, Poschlod P. **2019**. The recruitment niche predicts plant community assembly across a hydrological gradient along plowed and undisturbed transects in a former agricultural wetland. *Frontiers in Plant Science* **10**: 88–88.

Valentin S, Levin S. 2022. *plotbiomes.* [WWW document] URL https://github.com/valentinitnelav/plotbiomes/blob/develop/DESCRIPTION [accessed 1 March 2020]

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.

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.

Vazquez-Yanes C, Orozco-Segovia A, Medina E, Mooney HA, Vazquez-Yanes C. **1990**. Ecophysiology of Seed Germination in the Tropical Humid Forests of the World: A Review. In: Medina E, Mooney Mooney HA, Vazquez-Yanes C, eds. *Physiological ecology of plants of the wet tropics*. Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-009-7299-5_5.

Verner J, Boss AS. **1980**. California wildlife and their habitats: western Sierra Nevada. *Gen. Tech. Rep. PSW-GTR-37. Berkeley, CA: Pacific Southwest Forest and Range Exp. Stn., Forest Service, U.S. Department of Agriculture: 439 p* **037**.

Visser MD, Bruijning M, Wright SJ, Muller-Landau HC, Jongejans E, Comita LS, de Kroon H. **2016**. Functional traits as predictors of vital rates across the life cycle of tropical trees. *Functional Ecology* **30**: 168–180.

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

Walck JL, Hidayati SN, DIXON KW, THOMPSON KEN, 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.

Warner RR, Chesson PL. **1985**. Coexistence Mediated by Recruitment Fluctuations: A Field Guide to the Storage Effect. *The American Naturalist* **125**: 769–787.

Weissflog A, Markesteijn L, Lewis OT, Comita LS, Engelbrecht BMJ. **2018**. Contrasting patterns of insect herbivory and predation pressure across a tropical rainfall gradient. *Biotropica* **50**: 302–311.

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.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, *et al.* 2019. Welcome to the Tidyverse. *Journal of open source software* **4**: 1686.

Will RE, Wilson SM, Zou CB, Hennessey TC. **2013**. Increased vapor pressure deficit due to higher temperature leads to greater transpiration and faster mortality during drought for tree seedlings common to the forest–grassland ecotone. *New Phytologist* **200**: 366–374.

Wood SN. **2006**. *Generalized Additive Models: An Introduction with R*. New York: Chapman and Hall/CRC.

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.

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.

Wright J, Muller-Landau HC. **2005**. Annual and spatial variation in seed fall and seedling recruitment in a neotropical forest. *Ecology* **86**: 848–860.

Wright MC, van Mantgem P, Stephenson NL, Das AJ, Keeley JE. **2021**. Seed production patterns of surviving Sierra Nevada conifers show minimal change following drought. *Forest Ecology and Management* **480**: 118598.

Wright SJ, Bruijning M, Brassfield D, Cerezo A, Visser MD. **2015**. The Barro Colorado Island Tree Reproduction Dataset. doi:10.5479/si.data.201511251100. [accessed 16 February 2022]

Wright SJ, Bruijning M, Brassfield D, Cerezo A, Visser MD. **2015**. The Barro Colorado Island Tree Reproduction Dataset. doi:10.5479/si.data.201511251100. [accessed 16 February 2022].

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.

Wright SJ, Kitajima K, Kraft NJB, Reich PB, Wright IJ, Bunker DE, Condit R, Dalling JW, Davies SJ, Diaz S, *et al.* 2010. Functional traits and the growth-mortality tradeoff in tropical trees. *Ecology* **91**: 3664–3674.

Wright, SJ and FH Cornejo. 1990. Seasonal drought and the timing of leaf fall in a tropical forest. Ecology 70: 1165-1175.

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.

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.

Yu L, Gong P. **2012**. Google Earth as a virtual globe tool for Earth science applications at the global scale: progress and perspectives. *International Journal of Remote Sensing* **33**: 3966–3986.

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.

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.

Zhang L, Du B. 2016. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geoscience and Remote Sensing Magazine* 4: 22–40.

Zhang T, Niinemets Ü, Sheffield J, Lichstein JW. 2018. Shifts in tree functional composition amplify the response of forest biomass to climate. *Nature* 556: 99–102.

Zimmerman, Jess K., Liza S. Comita, Jill Thompson, María Uriarte, and Nicholas Brokaw. 2010. Patch dynamics and community metastability of a subtropical forest: compound effects of natural disturbance and human land use. *Landscape Ecology* 25: 1099-1111

Appendices

Appendix A: Technical documentation for the implementation of the Tree Recruitment Scheme in the Functionally Assembled Terrestrial Ecosystem Simulator (FATES-TRS)

This technical documentation is adapted from the original published description of the Tree Recruitment Scheme:

Hanbury-Brown AR, Powell TL, Muller-Landau HC, Wright SJ, Kueppers LM. **2022**. Simulating environmentally-sensitive tree recruitment in vegetation demographic models. New Phytologist. https://doi.org/10.1111/nph.18059

The Tree Recruitment Scheme (TRS; Hanbury-Brown et al., 2022) is an optional module for representing environmentally sensitive tree recruitment in Vegetation Demographic Models (VDMs). This documentation describes the implementation of the TRS in the Functionally Assembled Terrestrial Ecosystem Simulator (FATES; Koven et al., 2020). The TRS was developed primarily from studies at Barro Colorado Island (BCI) PanamaBarro Colorado Island (BCI) Panama, with an initial focus on tropical tree PFTs that vary along axes of drought and shade tolerance (Fig. 2), but is designed to be extensible to all tropical forests without the need to alter any algorithms. The implementation of the TRS in FATES (FATES-TRS) must be run in "carbon only mode" (i.e. without nutrients enabled) and without hydrodynamics. The following sections describe how key regeneration processes governing environmentally sensitive tree recruitment in tropical forests are represented in FATES-TRS.

Overview of FATES-TRS

When the DailyPRTAllometricCarbon subroutine is called in each daily timestep the TRS allocates a dynamic fraction of carbon for growth and reproduction (C_{g+r}; positive carbon balance net after tissue turnover and allocation to storage) to reproduction. Regeneration processes, described in detail below, move dynamic fractions of C_{g+r} through a seedbank and seedling pool (Fig. 1) which are tracked in units of carbon. Carbon recruiting out of the seedling pool each day is passed back to FATES's default recruitment subroutine (found in biogeochem/EDPhysiologyMod). The TRS determines how much carbon is available for recruitment and FATES's recruitment subroutine calculates how many new recruits to produce and initializes the new cohort. Carbon in seeds or seedlings that die or that is allocated to non-seed reproductive biomass, moves to the litter pool (Fig. 1). Unlike the offline version of the TRS presented in Hanbury-Brown *et al.* (2022), FATES-TRS uses exponential moving averages (EMAs) of environmental variables in the seedling layer to calculate the rates of regeneration processes. EMAs are tracked on different timescales depending on the process. The EMA of variable X on day i is updated using weighted averaging such that

$$EMA_{x,i} = \left(EMA_{x,i-1}\left(1 - \frac{1}{W_x}\right)\right) + \left(X_i \frac{1}{W_x}\right) \qquad \text{(Equation 1)}$$

where X_i is the new value of X on day i, and W_x is a parameter determining the weights of X_i and $EMA_{x,i-1}$. W_x can be interpreted as the timescale of the EMA, referred to hereafter more intuitively as the "moving window of days" that each EMA is influenced by.

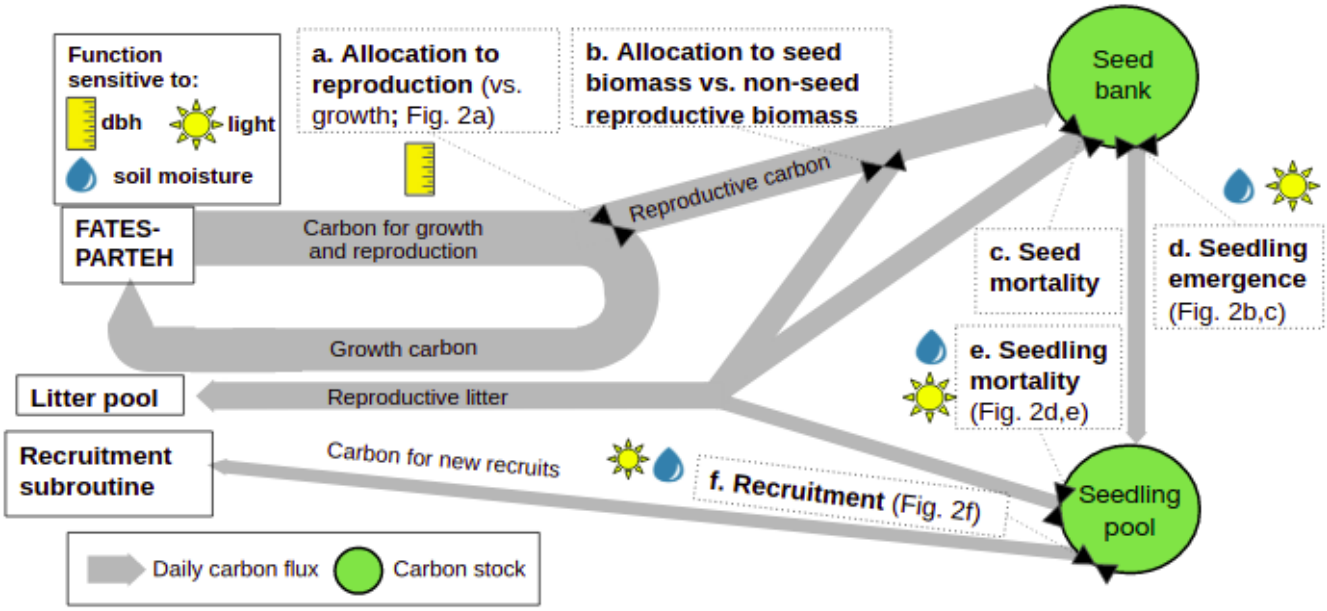


Figure 1. Daily regeneration processes (depicted with hour glasses) represented by FATES-TRS transfer reproductive carbon through seed bank and seedling carbon pools (depicted as circles). Processes are sensitive to DBH or environmental conditions (see inset key). The litter pool receives non-seed reproductive carbon, dead seeds, and dead seedlings. Carbon for new recruits is passed back to FATES's default recruitment subroutine. Parenthetical references to Figure 2 show relationships between regeneration processes and dbh or environmental conditions.

Allocation to reproduction

Allocation to reproduction occurs in FATES's parteh module (parteh/PRTAllometricCarbonMod.F90). FATES without the TRS switched on (i.e. "default FATES") assumes reproductive allocation is either insensitive to size or is a step function of size, depending on the parameterization. In contrast, the TRS allocates a dynamic fraction of cohort-level C_{g+r} (see Table 1 for FATES parameter names) (see Table 1 for FATES parameter names) to reproduction based on the cohort's size and the TRS's reproductive allocation (RA) function (Eqn 2, Fig. 1a, Fig. 2a). This follows observations that the probability a tree is reproductive increases sigmoidally with size within species (Visser *et al.*, 2016b). Each mature cohort contributes to recruitment via the TRS if they are in positive carbon balance. The effective fraction of cohort-level C_{g+r} allocated to reproduction, $F_{E,repro}$ (Eqn 3), is calculated based on a sigmoidal relationship relating the cohort's

current dbh to the probability of being reproductive (P_{repro}; Eqn 2). This formulation assumes that all reproductive individuals in a cohort allocate to reproduction at a constant, PFT-specific rate, F_{repro} (see Table 1 for all TRS parameters), which is modified by P_{repro} to calculate $F_{E,repro}$ (Eqn 3)

$$P_{repro} = \frac{e^{(a_{RA}(dbh)+b_{RA})}}{1 + e^{(a_{RA}(dbh)+b_{RA})}} \quad \text{(Equation 2)}$$

$$F_{E,repro} = (P_{repro})(F_{repro}) \quad \text{(Equation 3)}$$

where a_{RA} and b_{RA} are PFT-specific parameters describing the shape of the sigmoidal curve (Fig. 2a). This functional form is consistent with empirical data (Visser *et al.*, 2016b; Minor & Kobe, 2019). The TRS subsequently multiplies $F_{E,repro}$ by C_{g+r} to get reproductive carbon per cohort.

Allocation to seed vs. non-seed reproductive biomass and seed mortality

In nature, only a subset of the carbon allocated to reproduction becomes seeds, with the rest going to flowers, fruit flesh, capsules, etc. (Wenk *et al.*, 2017b). Default FATES sends all reproductive carbon to an undifferentiated "seed pool" from which carbon is lost and recruits are formed (Fisher *et al.*, 2015b). The TRS partitions each cohort's reproductive carbon into seed carbon and non-seed reproductive carbon (e.g., flowers, fruit flesh, and capsules) based on a prescribed, PFT-specific fraction of reproductive carbon that is seed, F_{seed} (Fig. 1b). This happens in the SeedIn subroutine (biogeochem/EDPhysiologyMod). Seed carbon moves to the seed bank each day and non-seed reproductive carbon moves to litter (Fig. 1b). Seeds in the seed bank die at a PFT-specific, constant rate, S_{mort}, which represents all modes of seed mortality including predation and decay (same as default FATES).

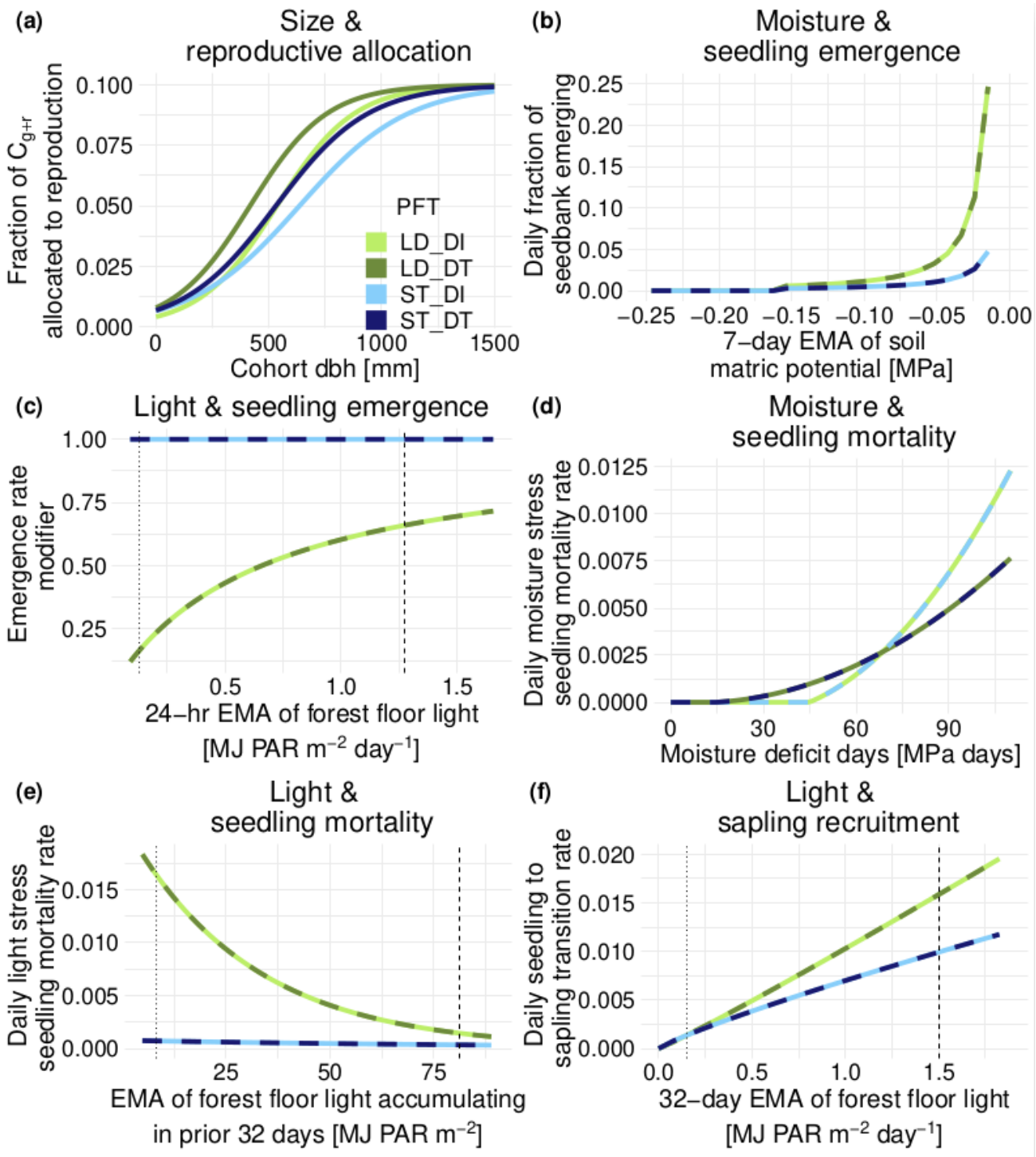


Figure 2. Environmentally sensitive regeneration processes explicitly represented by the Tree Recruitment Scheme, parameterized for four tropical tree PFTs at Barro Colorado Island, Panama: light demanding, drought intolerant (LD-DI); light demanding, drought tolerant (LD-

DT); shade tolerant, drought intolerant (ST-DI); and shade tolerant, drought tolerant (ST-DT). Note that moisture deficit days (panel d) are not equivalent between DT and DI PFTs because they accumulate according to PFT-specific values of Ψ_{crit}. In panels c, e, and f, the dotted vertical lines indicate mean understory light conditions on BCI (2% top of canopy radiation) and the dashed vertical lines indicate conditions in a small-medium sized light gap (20% top of canopy radiation). EMA = exponential moving average.

Seedling emergence

Seedling emergence is sensitive to soil moisture (Garwood, 1983b; Atondo-Bueno *et al.*, 2016; Ruiz Talonia *et al.*, 2017; Foster *et al.*, 2020) and light (Pearson et al., 2002) in nature. Default FATES represents it as an environmentally insensitive constant (Fisher *et al.*, 2015b). In the TRS, emergence depends on both soil moisture and light.

Light-dependence of germination is captured on day i in a Michaelis-Menten rate modifier [0,1]

$$f(PAR_i) = \frac{PAR_i}{PAR_i + PAR_{crit}} \qquad \text{(Equation 4)}$$

based on PAR_i, the 24-hour EMA of photosynthetically active radiation (PAR) at the seedling layer. Unlike in Hanbury-Brown *et al.* (2022) where seedling layer PAR is calculated as a simple fraction of top of canopy irradiance, in FATES-TRS it is calculated as a weighted average of direct and diffuse PARPAR at the lowest leaf layer of the lowest canopy layer on day i[1]. The weights are based on the LAI fraction in the lowest leaf layer that is in full sun versus shade. When there is no vegetation present, PAR at the seedling layer is taken from the boundary conditions (i.e. same as top of canopy). PAR_{crit} is a PFT-specific PAR threshold governing the shape of the germination response to reduced light (Fig. 2c). Most tropical pioneer species exhibit an increase in germination probability with increases in light, whereas germination in shade-tolerant species is insensitive to light (captured by $PAR_{crit} = 0$).

The EMA of soil matric potential on day i ($SMP_{EMA,i}$) at seedling rooting depth, $d_{seeedling}$, is influenced by SMP in a rolling window of days, W_{emerg} (default = 7 days), prior to i according to Eqn 1. If $SMP_{EMA,i}$ is above a critical threshold, Ψ_{emerg}, then seedling emergence occurs. The emergence rate on day i, $F_{emerg,i}$, is dynamically calculated as a function of $SMP_{EMA,i}$ (Fig. 1d, Fig. 2b). The pft-specific moisture response parameter, b_{emerg}, modifies the mean seedling emergence coefficient (a_{emerg}) in response to variation in SMP_{EMA} such that

$$F_{emerg,i} = \begin{cases} 0 & SMP_{EMA,i} < \psi_{emerg} \\ f(PAR_i)(a_{emerg})SMP_{EMA,i}^{b_{emerg}} & SMP_{EMA,i} \geq \psi_{emerg} \end{cases}$$

(Equation 5)

1 Calculating seedling layer PAR as PAR at the lowest leaf layer may underestimate light in patches where only a fraction of the patch is covered with vegetation. Future work is planned to update this calculation to improve estimates of mean patch-level light at the seedling layer in partially vegetated patches.

Eqn 5 produces pulses of seedling emergence in response to seasonal and interannual precipitation events, and stalls seedling emergence under relatively dry conditions.

Moisture and light-sensitive seedling survival

The TRS introduces a seedling pool to FATES that is tracked in the same way as the seed bank in default FATES. Seedling survival decreases differentially at low soil moisture and low light, affecting forest composition across environmental gradients (Kobe, 1999; Engelbrecht et al., 2007). The TRS seeks to capture this with a PFT-specific moisture stress threshold, Ψ_{crit}, below which the seedling pool starts to "accumulate" (mathematically an EMA is tracked according to Eqn 1 with a timescale of W_Ψ days; default = 126) moisture deficit days (MDD) similar to the concept of growing degree days. The new MDD value on day i, MDD_i, is calculated as the difference between the absolute value of site-level SMP on day i, SMP_i, and the absolute value of Ψ_{crit}.

$$MDD_i = \begin{cases} 0 & \psi_j \geq \psi_{crit} \\ |\psi_j| - |\psi_{crit}| & \psi_j < \psi_{crit} \end{cases} \quad \text{(Equation 6)}$$

MDD_i is then used to update an EMA of MDD, MDD_{EMA}, according to Eqn 1. Finally, MDD_{EMA} is multiplied by the timescale of the EMA (in days), W_Ψ, to approximate an "accumulation" of MDD.

This formulation simultaneously captures the magnitude and duration of moisture stress. Observations of seedling wilting points from a manipulative drought experiment at BCI (Engelbrecht & Kursar, 2003b; Engelbrecht *et al.*, 2007a) were used to explore the relationship between MDD accumulation and seedling mortality. Observed drought-induced mortality is 0 up to a critical accumulation of MDD, MDD_{crit}, at which point a convex quadratic relationship best explained drought-induced seedling mortality as a function of MDD (Fig. 1e, Fig. 2d, see SI Methods S1 and Fig. S1 in Hanbury-Brown et al., 2022 for more details). The mortality rate from moisture stress (M_Ψ) on day i is therefore

$$M_{\psi,i} = \begin{cases} 0 & MDD_{EMA,i} < MDD_{crit} \\ a_\psi MDD_{EMA,i}^2 + b_\psi MDD_{EMA,i} + c_\psi & MDD_{EMA,i} \geq MDD_{crit} \end{cases} . \quad \text{(Equation 7)}$$

Seedlings also die from insufficient light, which we refer to as light stress. The light stress mortality rate, M_L, on day i is a function of "cumulative" (mathematically an EMA is tracked according to Eqn 1 with a timescale of W_L days; default = 32) PAR at the seedling layer, $L_{seedlingEMA}$, within a moving window of days, W_L, prior to i (Fig. 1e, Fig. 2e). Similar to the approach used to calculate MDD_{EMA}, $L_{seedlingEMA}$ is calculated by multiplying an EMA of seedling layer PAR, PAR_{EMA}, by W_L to approximate the cumulative light incident at the seedling layer prior to i. Two PFT-specific parameters determine the shape of the negative exponential relationship between mortality and and light

$$M_{L,i} = e^{a_{ML}(L_{seedlingEMA,i})+b_{ML}} \quad \text{(Equation 8)}$$

where a_{ML} is a PFT-specific light response parameter and b_{ML} is the intercept. This function is based on an analysis by Kobe (1999) who tested four functional forms and found that the negative exponential best described light stress mortality for two shade tolerant and one light demanding species that were transplanted into varied light environments. A background seedling mortality rate, $M_{background}$, represents other seedling mortality (e.g. herbivory, pathogens, tree fall, etc.). Total seedling mortality is the sum of moisture-dependent, light-dependent and background mortality.

Recruitment

The rate of transition from seedling to sapling increases with understory light (Brokaw, 1985; Rüger *et al.*, 2009b), but default FATES does not capture this sensitivity (Fisher *et al.*, 2015b). Recruitment in the TRS is represented with a dynamic seedling to sapling transition rate (TR) which is the fraction of total carbon in the seedling pool, $C_{seedling}$, that is available to make new recruits each day (Fig. 1f). The TR on day i is calculated as a power function of $PAR_{EMA,i}$ (Eqn 9; Fig. 2f). If SMP on day i, Ψ_i, is drier than Ψ_{crit} the transition rate goes to zero such that

$$TR_i = \begin{cases} 0 & \psi_i < \psi_{crit} \\ a_{TR}\left(PAR_{EMA,i}\right)^{b_{TR}} & \psi_i \geq \psi_{crit} \end{cases} \quad \text{(Equation 9)}$$

where a_{TR} is a coefficient derived from the mean transition rate at observed mean understory PAR (see SI Methods S1 in Hanbury-Brown et al. (2022) for more information) and b_{TR} is the light response modifier. The light response modifier produces accelerating (i.e. light demanding) or decelerating responses to light (Fig. 2f) depending on if b_{TR} is greater or less than 1. Of a variety of functional forms tested at BCI, a power function with species-specific light response modifiers best explained observed variation in recruitment rates under spatially heterogenous patch-level light (Rüger *et al.*, 2009b). This formulation is more broadly supported by the growth-mortality functional trade-off axis where light demandinglight demanding species can take advantage of higher light conditions through faster relative growth rates (Wright *et al.*, 2010).

Carbon transitioning out of the seedling layer is available to FATES's default recruitment subroutine which converts carbon available for recruitment into a number density of new recruits based on the amount of carbon required to form an individual in the smallest size class tracked by the VDM, Z_0. The number of new recruits predicted on day i, R_i, is

$$R_i = \frac{(TR_i)(C_{seedling,i})}{Z_0}. \quad \text{(Equation 10)}$$

The regeneration processes represented above (Eqns 2-10) introduce 19 new parameters use by FATES-TRS that are not used by default FATES (Table 1).

Table 1. FATES-TRS parameters, grouped by regeneration processes shown in Fig. 1. Parameters that are also used by default FATES are indicated with an asterisk. See SI Methods S1 in Hanbury-Brown et al. (2022), for all parameter value derivations for Barro Colorado Island (BCI). C_{g+r} = carbon for growth and reproduction; RA = reproductive allocation; MDD = moisture deficit days.

Parameter symbol (parameter name)	**BCI Value**	**Units**	**Description**	**Derivation / Source**
Allocation to reproduction				
$*F_{repro}$ (fates_seed_alloc)	0.1 (all PFTs)	-	Fraction of C_{g+r} allocated to reproduction	Harper, 1977; Smith et al., 2001; Fisher et al., 2015
a_{RA} (fates_repro_alloc_a)	LD-DI: 0.0058 LD-DT: 0.0059 ST-DI: 0.0042 ST-DT: 0.0049	Δ RA $[\Delta$ dbh$]^{-1}$	Governs RA as function of dbh (logit function coefficient)	Logistic regression fit to observations Wright et al., 2015; Visser et al., 2016
b_{RA} (fates_repro_alloc_b)	LD-DI: -3.1380 LD-DT: -2.4607 ST-DI: -2.6518 ST-DT: -2.6171	-	Governs RA as function of dbh (intercept in logit function)	Logistic regression fit to observations (Wright *et al.*, 2015; Visser *et al.*, 2016)
Allocation to seed vs. non-seed reproductive biomass and seed mortality				
F_{seed} (fates_repro_frac_seed)	0.24 (all PFTs)	-	Fraction of reproductive C that is seed	Based on previously cited range (Wenk *et al.*, 2017b)
$*s_{mort}$ (fates_seed_decay_rate)	0.51 (all PFTs)	yr^{-1}	Seed mortality rate	Lischke et al., 2006; Fisher et al., 2015
Seedling emergence				
$d_{seedling}$ (fates_seedling_root_depth)	0.06 (all PFTs)	m	Seedling rooting depth	Needs empirical validation
a_{emerg} (fates_a_emerg)	0.0003	day^{-1}	Coefficient for seedling emergence rate	Calculated from Pearson et al., 2002
b_{emerg} (fates_b_emerg)	LD-DI: 1.6 LD-DT: 1.6 ST-DI: 1.2 ST-DT: 1.2	-	Seedling emergence sensitivity to soil moisture	Calibrated to observations of seasonal seedling emergence (Garwood, 1983b)
W_{emerg} (fates_sdlng_emerg_h2o_timescale)	7 (all PFTs)	days	Time window for emergence response to soil moisture	Observations of seasonal seedling emergence (Garwood, 1983)
Ψ_{emerg} (fates_seedling_psi_emerg)	-15744.65	mm H2O suction	Soil moisture required for emergence	This study, see SI Methods S1 in Hanbury-Brown et al. (2022)
PAR_{crit} (fates_par_crit_germ)	0.656	MJ m^{-2} day^{-1}	Critical PAR level for light-sensitive germination	Based on observations of mean irradiance (PAR) in small gaps ($25m^2$) at BCI (Pearson *et al.*, 2002b)
Moisture and light-sensitive seedling survival				
$M_{background}$ (fates_background_seedling_mort)	LD-DI: 0.17 LD-DT: 0.18 ST-DI: 0.19	yr^{-1}	Background seedling mortality rate	Calculated from seedling censuses done at BCI (2003-2012; Johnson et al., 2017)

	ST-DT: 0.11			
Ψ_{crit} (fates_seedling_psi_crit)	DI: -175912.9 DT: -251995.7	mm H2O suction	Seedling moisture stress threshold	Based on Engelbrecht & Kursar 2003; Engelbrecht et al., 2005
MDD_{crit} (fates_seedling_mdd_crit)	DI: 46E5 DT: 14E5	mm H2O suction days	Moisture deficit day threshold for seedling mortality	Based on Engelbrecht & Kursar 2003; Engelbrecht *et al.*, 2005
a_{Ψ} (fates_seedling_h2o_mort_a)	DI: 1.04E-16 DT: 4.07E-17	-	Moisture-based mortality coefficient	Based on Engelbrecht & Kursar 2003; Engelbrecht *et al.*, 2005
b_{Ψ} (fates_seedling_h2o_mort_b)	DI:-5.5E-10 DT:-6.4E-11	-	Moisture-based mortality coefficient	Based on Engelbrecht & Kursar 2003; Engelbrecht *et al.*, 2005
c_{Ψ} (fates_seedling_h2o_mort_c)	DI:3.5E-04 DT:1.3E-05	-	Moisture-based mortality coefficient	Based on Engelbrecht & Kursar 2003; Engelbrecht *et al.*, 2005
W_{Ψ} (fates_sdlng_mdd_timescale)	126 (all PFTs)	days	moving window for MDD	Based on Engelbrecht & Kursar 2003; Engelbrecht *et al.*, 2005
a_{ML} (fates_seedling_light_mort_a)	LD:-0.033 ST:-0.00990	-	Light-based mortality coefficient	Based on Kobe, 1999
b_{ML} (fates_seedling_light_mort_b)	LD:-3.84 ST:-7.15	-	Light-based mortality coefficient	Based on Kobe, 1999
W_{L} (fates_sdlng_mort_par_timescale)	32 (all PFTs)	days	moving window for seedling light response	Based on Augspurger, 1984
Recruitment				
a_{TR} (fates_seedling_light_rec_a)	LD:0.010 ST:0.007	-	Seedling to sapling transition rate coefficient	Derived from parameters used in CLM(ED) and FATES (Fisher *et al.*, 2015b)
b_{TR} (fates_seedling_light_rec_b)	LD: 1.0653 ST: 0.8615	-	Recruitment light response parameter	Based on Rüger et al., 2009
Convert carbon to number density of recruits				
*H_{min} (fates_recruit_hgt_min)	2.345	m	Min height of a new recruit	A min height of 2.345 m produces a new recruit where dbh = 1 cm.
Regeneration mode switch				
*fates_regeneration_model	2	-	Switch to choose regeneration mode (1 = default FATES; 2 = TRS)	NA

FATES-TRS initial test results

Initial testing of FATES-TRS shows that predicted recruitment rates at BCI are broadly consistent with observations and the offline version of the TRS presented in Hanbury-Brown *et al.* (2022; Fig. 3). This indicates that the implementation of the TRS in FATES has not caused unexpected changes in recruitment behavior compared to the offline TRS. Initial testing also

shows that FATES-TRS improves predictions of the magnitude of recruitment at BCI compared to default FATES (Fig. 4). This indicates its potential to improve predictions of mortality and abundance in the smallest size classes. Future FATES-TRS testing will focus on patch-level predictions of recruitment under varying light and soil moisture which will demonstrate is potential to contribute to predictions of functional turnover (along with growth and mortality) in response to environmental variation.

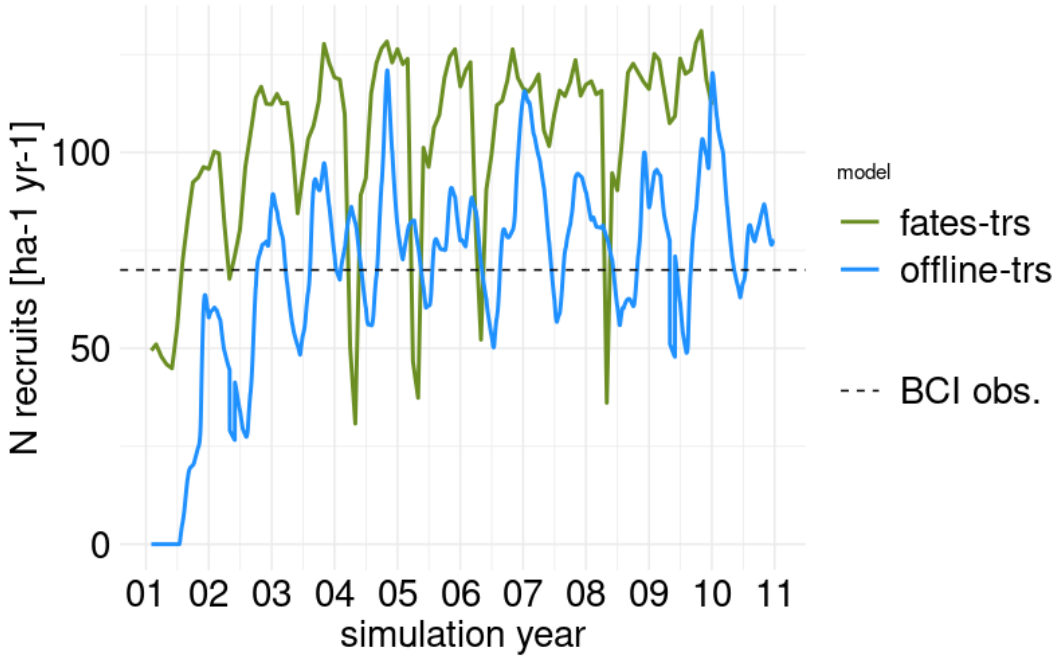


Figure 3. FATES-TRS predictions of recruitment (green line) compared to the offline version of the TRS presented in Hanbury-Brown *et al.* (2022; blue line) and observations at the Barro Colorado Island forest dynamics plot (BCI obs.; dashed line). Simulations were run with observed meteorology (2008–2014 recycled) and were initialized from observed stand structure at BCI during the 2015 census. The "offline-trs" predictions show the sum of recruitment for the four PFTs described in Hanbury-Brown *et al.*, 2022. FATES-TRS predictions are for a single broad-leaved tropical tree PFT with mature tree parameter values that were optimized for BCI in prior unpublished work (pers. commun. R Knox, 2021) and TRS parameter values that are the same as the ST-DT PFT in Hanbury-Brown *et al.* (2022; Table 1).

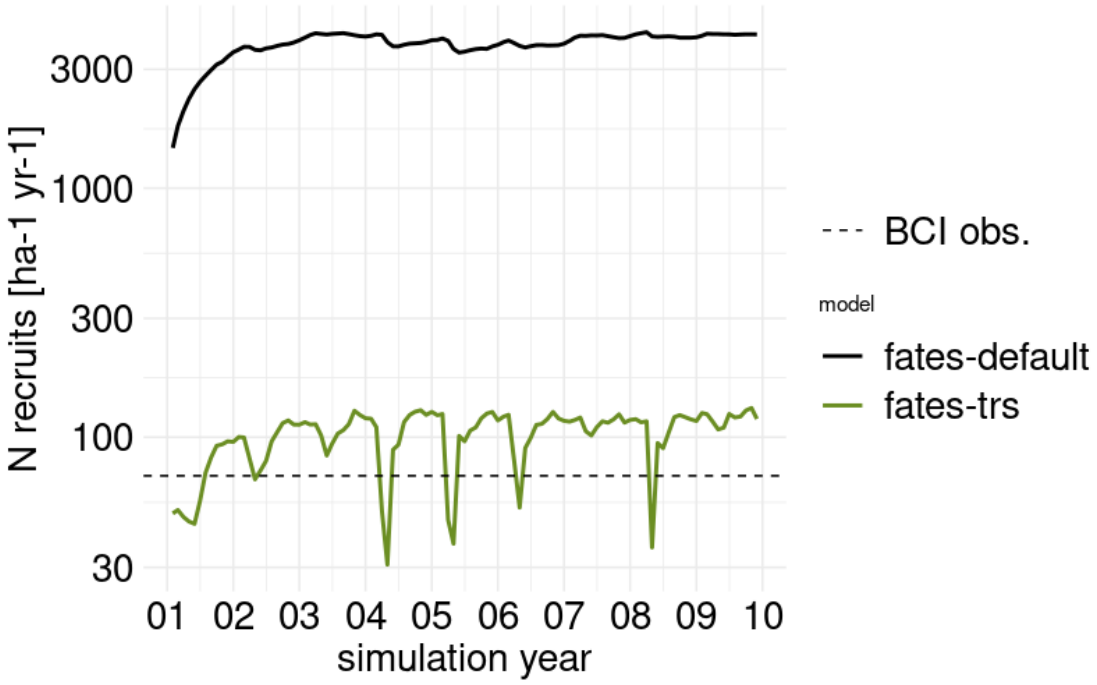


Figure 4. FATES-TRS predictions of recruitment (green line) compared to predictions from default FATES (black line) and observations at the Barro Colorado Island (BCI) forest dynamics plot (dashed line). Simulations were run with observed meteorology (2008–2014 recycled) and were initialized from observed stand structure at BCI during the 2015 census. Predictions are for a single broad-leaved tropical tree PFT with mature tree parameter values that were optimized for BCI in prior unpublished work (pers. commun. R Knox, 2021). The FATES-TRS prediction has TRS parameter values that are the same as the ST-DT PFT in Hanbury-Brown *et al.* (2022; Table 1).