



RESEARCH ARTICLE

10.1029/2023MS004195

Special Collection:

The Energy Exascale Earth
System Model (E3SM):
Advancements in a Decade of
Earth System Modeling

Key Points:

- We calibrate tropical forest coexistence in a state-of-the-art ecosystem demography model using advanced multi-objective (MO) optimization
- The MO algorithm is easy to identify objective functions to achieve forest coexistence and much more effective than random sampling
- We use the calibrated model to analyze the recovery trajectory of forest aboveground biomass under different intensities of deforestation

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. Cheng,
yanyan.cheng@nus.edu.sg

Citation:

Cheng, Y., Wang, W., Detto, M., Fisher, R., & Shoemaker, C. (2024). Calibrating tropical forest coexistence in ecosystem demography models using multi-objective optimization through population-based parallel surrogate search. *Journal of Advances in Modeling Earth Systems*, 16, e2023MS004195. <https://doi.org/10.1029/2023MS004195>

Received 22 DEC 2023

Accepted 11 JUL 2024

Author Contributions:

Conceptualization: Yanyan Cheng

Data curation: Yanyan Cheng

Formal analysis: Yanyan Cheng

© 2024 The Author(s). Journal of Advances in Modeling Earth Systems published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](#), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Calibrating Tropical Forest Coexistence in Ecosystem Demography Models Using Multi-Objective Optimization Through Population-Based Parallel Surrogate Search

Yanyan Cheng¹ , Wenyu Wang¹ , Matteo Detto^{2,3} , Rosie Fisher⁴ , and Christine Shoemaker^{1,5}

¹Department of Industrial Systems Engineering and Management, National University of Singapore, Singapore, Singapore, ²Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey, USA, ³Smithsonian Tropical Research Institute, Balboa, Panama, ⁴CICERO Centre for International Climate Research, Oslo, Norway, ⁵Department of Civil and Environmental Engineering, Cornell University, Ithaca, New York, USA

Abstract Tropical forest diversity governs forest structures, compositions, and influences the ecosystem response to environmental changes. Better representation of forest diversity in ecosystem demography (ED) models within Earth system models is thus necessary to accurately capture and predict how tropical forests affect Earth system dynamics subject to climate changes. However, achieving forest coexistence in ED models is challenging due to their computational expense and limited understanding of the mechanisms governing forest functional diversity. This study applies the advanced Multi-Objective Population-based Parallel Local Surrogate-assisted search (MOPLS) optimization algorithm to simultaneously calibrate ecosystem fluxes and coexistence of two physiologically distinct tropical forest species in a size- and age-structured ED model with realistic representation of wood harvest. MOPLS exhibits satisfactory model performance, capturing hydrological and biogeochemical dynamics observed in Barro Colorado Island, Panama, and robustly achieving coexistence for the two representative forest species. This demonstrates its effectiveness in calibrating tropical forest coexistence. The optimal solution is applied to investigate the recovery trajectories of forest biomass after various intensities of clear-cut deforestation. We find that a 20% selective logging can take approximately 40 years for aboveground biomass to return to the initial level. This is due to the slow recovery rate of late successional trees, which only increases by 4% over the 40-year period. This study lays the foundation to calibrate coexistence in ED models. MOPLS can be an effective tool to help better represent tropical forest diversity in Earth system models and inform forest management practices.

Plain Language Summary Tropical forests have extremely high levels of tree species diversity, which are estimated to harbor about 50% of the world's terrestrial plant species. Representing tropical forest diversity in Earth system models (ESMs) is important to accurately capture and predict the interactions between tropical forests and environmental changes. But simulating coexistence in ESMs is challenging, as only a limited number of models can simulate forest functional coexistence. In addition, only a few algorithms have been developed to calibrate ecosystem fluxes and tree coexistence concurrently. This study applies an advanced multi-objective optimization algorithm to calibrate (a) carbon, water, and energy cycle-related variables and (b) coexistence of two typical tropical forests (i.e., early and late successional forests). Our multi-objective optimization algorithm can satisfactorily capture the dynamics in tropical forest ecosystems and effectively lead many more model runs to successful and stable coexistence than random sampling. The improved parameterization is further applied to investigate the recovery of forest biomass following various intensities of clear-cut deforestation scenarios. Our results have important implications for capturing tropical forest diversity as well as their responses to environmental changes and human interventions such as wood harvest.

1. Introduction

Although they only encompass less than 7% of the world's land area, tropical forests have the highest diversity of tree species (Baccini et al., 2017; Wohl et al., 2012). About 50% of the world's plant species are estimated to exist in tropical forests. For example, there are more than 15,000 plant species in Malaysian Borneo's tropical forests. In some tropical areas, more tree species are found in a single 50 ha plot than that in the entire North America or Europe (Burslam et al., 2001).

Funding acquisition:

Christine Shoemaker

Investigation: Yanyan Cheng

Methodology: Yanyan Cheng

Project administration:

Christine Shoemaker

Resources: Yanyan Cheng, Matteo Detto

Software: Yanyan Cheng, Wenyu Wang

Validation: Yanyan Cheng

Visualization: Yanyan Cheng

Writing – original draft: Yanyan Cheng

Writing – review & editing:

Yanyan Cheng, Wenyu Wang,

Matteo Detto, Rosie Fisher,

Christine Shoemaker

The high tree species richness contributes to the complexity of the forest structure and diversity of functional composition, influencing the responses and feedbacks of tropical forests to climate and anthropogenic changes (Fisher et al., 2010; Lichstein et al., 2007). Given the important role of Earth system models (ESMs) in projecting the impact of tropical forests on Earth system dynamics, inadequate representation of tropical forests in ESMs can lead to systematic biases in these projections. The projections that may be affected include forest carbon and water budgets, as well as the Earth's climate, especially under a changing environment. For example, the misrepresentation of tree species composition can cause a deviation of modeled vegetation biomass from ground measurements by up to two times in an Amazonia tropical forest (Delbart et al., 2010).

The size- and age-structured ecosystem demographic (ED) models (Fisher & Koven, 2020; Fisher et al., 2015, 2018; Ma et al., 2022; Medvigy et al., 2009; Moorcroft et al., 2001), by explicitly simulating many essential vegetation processes (e.g., competition, succession, natural and anthropogenic disturbances) (Chesson, 2000; Freckleton & Watkinson, 2002; Huang et al., 2020), have been recently developed to better represent tropical forest complexity. However, achieving stable forest coexistence in ED models remains an ongoing research challenge due to their computational expense, amplified gap-phase dynamics, potential over-exaggeration of reproduction feedback, rapid shifts in competition outcomes, and insufficient calibration of hydrological parameters (Cheng et al., 2022; Detto et al., 2022; Fisher & Koven, 2020; Fisher et al., 2010; Koven et al., 2019).

In addition, tropical forests have undergone substantial deforestation and degradation (Anderson-Teixeira et al., 2015; Baccini et al., 2012; Longo & Keller, 2019) in the past three decades. Selective logging (i.e., cutting and removal of merchantable timber) of tropical forests contributes to approximately one-eighth of global timber (Huang et al., 2020). It is a major driver of forest loss and degradation in Southeast Asia (Jamhuri et al., 2018) and the Amazon (Tyukavina et al., 2017). The changes due to selective logging can convert tropical forests from a net carbon sink to a net carbon source (Gatti et al., 2021). Therefore, it is important to understand how selective logging operations could affect forest structure and composition, as well as how tropical forests would respond and recover under these anthropogenic disturbances. This has important implications for the conservation and management of forest ecosystems (Asner et al., 2005; Bonan, 2008; Erb et al., 2018; Huang et al., 2020).

In this study, we explore the coexistence of two physiologically distinct tropical forest species—early and late successional tropical forests—in an advanced ED-type model called ELM-FATES, which is the Functionally Assembled Terrestrial Ecosystem Simulator (FATES) implemented in the Energy Exascale Earth System Model (E3SM) Land Model (ELM) (Fisher et al., 2015; Huang et al., 2020; Koven et al., 2019; Leung et al., 2020). The calibration is carried out by using an advanced multi-objective optimization algorithm called Multi-Objective Population-based Parallel Local Surrogate-assisted search (MOPLS). A previous study of calibrating early and late successional plant function types (PFTs) in ELM-FATES used a single-objective parallel global optimization scheme (Cheng et al., 2023). However, the cases where early and late successional PFTs coexist were manually selected from all the evaluation experiments since coexistence was not guaranteed when employing the single-objective global optimization scheme. In contrast, our study utilizes the multi-objective optimization scheme MOPLS to simultaneously calibrate the features and coexistence of early and late successional PFTs. Notably, coexistence calibration is one of the objective functions in our approach, resulting in more robust outcomes with coexisting cases. Furthermore, the flexibility of this multi-objective scheme enables the calibration of coexistence for multiple PFTs, a task that is notably more challenging when using the single-objective global optimization scheme in Cheng et al. (2023). Moreover, this study further addresses the recovery trajectory of aboveground biomass following various intensities of clear-cut deforestation by using an advanced selective logging module within ELM-FATES. These investigations provide valuable insights for forest management practices.

We compare the modeling results against field measurements collected in Barro Colorado Island (BCI), Panama. We find that the MOPLS optimization algorithm exhibits satisfactory performance in calibrating variables, both included and not included in the objective function. The MOPLS scheme is robust in achieving coexistence using different objective functions. Therefore, it is easy to use MOPLS to identify objective functions to calibrate forest coexistence. MOPLS is much more effective than random sampling in searching for coexistence (Cheng et al., 2022). We further analyze the recovery process of tropical forest aboveground biomass under different intensities of clear-cut deforestation. We find that with a 20% clear-cutting, the aboveground biomass can take almost 40 years to return to the initial level. Our improved simulation of tropical forest coexistence contributes to a better understanding and prediction of how future forest carbon stocks will vary under anthropogenic

disturbances. These applications are important for understanding the processes of maintaining tree species diversity (Bonell & Bruijnzeel, 2005).

2. Methods and Data

2.1. Study Area and Field Data for Calibration

Field measurements collected at BCI, Panama, a primary tropical forest site, are used as calibration data in this study. These ecohydrological measurements include tree size distribution (f_{tree}) (Condit et al., 2017), gross primary production (GPP), sensible heat (SH), latent heat (LE) (Detto & Pacala, 2022), soil water content in the top 15 cm (SWC), and runoff (Q) (Cheng et al., 2018). Meteorological data spanning 2003 to 2016, including precipitation, wind speed, air temperature, relative humidity, surface pressure, and solar radiation, is recycled to spin-up each optimization evaluation. Mean annual values for these climate variables are 2350 mm, 1.5 m/s, 299 K, 88.5%, 100 kPa, and 180 W/m², respectively. BCI has distinguished dry (mid-December to mid-April) and wet (late-April to early December) seasons (Cheng et al., 2019; Ogden et al., 2013), with mean seasonal precipitation values of 200 and 2150 mm, respectively.

2.2. Description of ELM-FATES With Selective Logging Module

The Energy Exascale Earth System Model (E3SM) land model (ELM) is the land component of E3SM (Leung et al., 2020). ELM comprises modules that simulate biogeophysical, biogeochemical, and hydrological processes, enabling it to simultaneously simulate the dynamics of terrestrial energy, water, and carbon cycles. The Functionally Assembled Terrestrial Ecosystem Simulator (FATES) is a state-of-the-art ED model that uses size- and age-structures and successional trajectory-based patches to group plants and simulate live vegetation processes (Fisher et al., 2015; Koven et al., 2019). This is based on the ED theory derived from Moorcroft et al. (2001) and incorporates the perfect plasticity approximation proposed by Purves et al. (2008).

Selective logging of tropical forests significantly contributes to forest loss and degradation, yet its realistic representation in ESMs was mostly absent (Huang et al., 2020). In order to address this limitation, an advanced selective logging module has been implemented into ELM-FATES (Huang et al., 2020) to represent realistic wood harvest in field operations at a landscape level. This includes specifying the timing and extent of a logging event and simulating realistic mortality types associated with logging, such as direct-felling, collateral, and mechanical damages. Specifically, the selective logging module in ELM-FATES parameterizes three types of mortality associated with logging in each logging event: direct-felling mortality, collateral mortality, and mechanical mortality. The direct-felling mortality represents the fraction of trees that are greater than or equal to a diameter threshold and are selected for timbers. The collateral mortality represents the fraction of forests affected by the felling of adjacent harvested trees. The mechanical mortality denotes the fraction of forests killed by field infrastructure (e.g., log decks, skids, and trails) used for conducting wood harvest activities and for storing and transporting timbers offsite. Loggers typically cut relatively small trees and avoid relatively large trees for economic reasons. Therefore, another diameter threshold parameter associated with the mechanical mortality is implemented in ELM-FATES, ensuring that only a certain fraction of trees equal to or smaller than that threshold is affected during a logging operation. Please refer to Fisher et al. (2015), Koven et al. (2019), Leung et al. (2020), Huang et al. (2020), and FATES-Development-Team (2019) for a complete overview of ELM-FATES, including the selective logging module.

2.3. Calibrated Parameters for Tropical Forest Coexistence

Following Cheng et al. (2022, 2023), 19 parameters related to plant traits and soil hydrological processes were selected for calibration (Table 1), including the maximum carboxylation rate of RuBisCO at the reference temperature (25°C; $V_{c,max}$), specific leaf area at the top of canopy (sla_{top}), longevity for root and leaf (τ), wood density (ρ_{wood}), background mortality rate (λ), decay parameters in the rooting depth distribution function (α and β), scaling exponent of the Clapp and Hornberger soil retention curve (B_{sw}), saturated soil water content (θ_s), saturated soil matric potential in five soil layers ($\Psi_{s,1-5}$), saturated soil hydraulic conductivity (K_s), maximum fractional saturated area (f_{max}), mean topographic slope (C_s), and decay factors representing the distribution of surface runoff (f_{over}) and subsurface runoff (f_{drain}). In this study, we calibrate the coexistence of early and late successional tropical forests (Figure 1a), based on what has been observed in BCI, Panama. This region typically features a mixture of early and late successional tree species (Condit et al., 1995). More information on these

Table 1

Vegetation (Early and Late Successional Tropical Forests) and Soil Hydrological Parameters Calibrated in This Study

Parameter name		Symbol	Unit	Range	Calibrated value
Maximum carboxylation rate of Rub. at 25C, canopy top	Early PFT	$V_{c,max, early}$	$\mu\text{mol CO}_2/\text{m}^2/\text{s}$	10–106 ^a	106.0
	Late PFT	$V_{c,max, late}$	$\mu\text{mol CO}_2/\text{m}^2/\text{s}$		84.8
Specific Leaf Area (SLA) at top of canopy, projected area basis	Early PFT	$sla_{top, early}$	$\text{m}^2/\text{g C}$	0.007–0.039 ^b	0.008
	Late PFT	$sla_{top, late}$	$\text{m}^2/\text{g C}$		0.007
Longevity for root and leaf (turnover time)	Early PFT	τ_{early}	yr	0.1–1.5 ^c	0.9
	Late PFT	τ_{late}	yr		2.3
Mean density of woody tissue in plant	Early PFT	$\rho_{wood, early}$	g/cm^3	0.2–1.0 ^d	0.42
	Late PFT	$\rho_{wood, late}$	g/cm^3		0.84
Background mortality rate	Early PFT	λ_{early}	1/yr	0.01–0.1 ^e	0.019
	Late PFT	λ_{late}	1/yr		0.01
Root distribution parameter 1	Early PFT	α_{early}	–	0.1–8.0 ^f	2.6
	Late PFT	α_{late}	–		2.1
Root distribution parameter 2	Early PFT	β_{early}	–	0.1–8.0 ^f	3.5
	Late PFT	β_{late}	–		2.8
Scaling exponent of the soil retention curve	–	B_{sw}	–	1–30 ^g	1.0
Saturated hydraulic conductivity	–	K_s	mm/h	0.036–72 ^{g,h}	0.36
Saturated water content (porosity)	–	θ_s	m^3/m^3	0.35–0.6 ^{g,h}	0.6
Saturated soil matric potential at five soil layers	Layer 1	$\Psi_{s,1}$	mm	50–350 ^{g,h}	350.0
	Layer 2	$\Psi_{s,2}$	mm		94.7
	Layer 3	$\Psi_{s,3}$	mm		214.3
	Layer 4	$\Psi_{s,4}$	mm		141.2
	Layer 5	$\Psi_{s,5}$	mm		254.9
Maximum fractional saturated area	–	f_{max}	–	0.01–0.9 ^{g,h}	0.8
Mean topographic slope	–	C_s	–	0.01–0.9 ^{g,h}	0.3
Decay factor that represents the distribution of surface runoff with depth	–	f_{over}	1/m	0.1–5 ^{g,h}	4.8
Decay factor that represents the distribution of subsurface runoff with depth	–	f_{drain}	1/m	0.5–5 ^{g,h}	4.3

^aDomingues et al. (2005). ^bWright et al. (2004). ^cHuang et al. (2020). ^dLongo et al. (2019). ^eLongo and Keller (2019). ^fZeng (2001). ^gHou et al. (2012). ^hHuang et al. (2013).

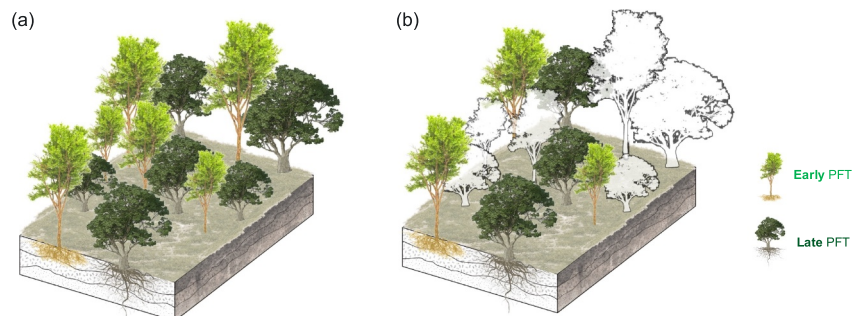


Figure 1. Illustration of (a) coexisted early (light green) and late (dark green) plant functional types (PFTs), and (b) a 50% clear-cut deforestation for early and late PFTs. The white color trees in (b) indicate their removal due to clear-cut deforestation.

Table 2

Parameters for the Selective Logging Module of ELM-FATES, With Their Values Used in Numerical Experiments Under Various Clear-Cut Intensities

Parameter name	Symbol in ELM-FATES	Unit	Value in the clear-cut experiments	Physical meaning
Minimum diameter at breast height (dbh) at which logging is applied	fates_logging_dbhmin	cm	0.0	All trees are selected for harvesting in the clear-cut experiments.
Tree diameter, above which infrastructure from logging does not impact damage or mortality	fates_logging_dbhmax_infra	cm	35	Only trees with a diameter equal to or smaller than 35 cm will be affected by logging infrastructure (e.g., building log decks, skids, and trails).
Fraction of trunk product being shipped offsite; the leftovers will be left onsite as decomposing coarse woody debris	fates_logging_export_frac	%	100	All the logged trunk product will be shipped offsite. No leftovers on the ground.
Fraction of stems killed in the understory when logging generates disturbance	fates_logging_coll_under_frac	%	0.0	No understory stems are killed when logging happens.
Fraction of stems logged directly per event	fates_logging_direct_frac	%	10~100	The clear-cut intensity ranges from 10% to 100%.
Fraction of upperstory large stems that die from logging collateral damage	fates_logging_collateral_frac	%	0.0	No collateral damages in the clear-cut experiments. All damage is caused by direct felling from logging operation.
Fraction of stems killed due to infrastructure and other mechanical damages	fates_logging_mechanical_frac	%	0.0	No infrastructure and other mechanical damages in the clear-cut experiments.

calibrated parameters is given in Table 1. The distinctions in parameterization between early and late PFTs involve the utilization of decay factors following Cheng et al. (2023).

Logged forests often exhibit rapid recovery of carbon and water fluxes within 1–3 years post-logging (Huang et al., 2020; Negrón-Juárez et al., 2020). However, the recovery times for carbon stocks are usually more than 30 years after logging (Huang et al., 2020). In this study, we perform several selective logging simulations to analyze the trajectory of aboveground biomass (AGB) for early and late PFTs after different intensities of clear-cut deforestation. These simulations use the same climate conditions as the calibration data, with intensities ranging from 10% to 100% (Table 2). A 50% clear-cut deforestation indicates that the number of forests is reduced by 50% (Figure 1b).

2.4. Multi-Objective Optimization Scheme

MOPLS (Wang et al., 2023) is a surrogate-assisted algorithm designed for multi-objective optimization problems where the objective functions are evaluated via the black-box and time-consuming simulations. The core idea behind MOPLS is to approximate the objective functions by surrogate models. The surrogate models are analytical functions that are computed from a limited number of simulations and updated every time a new simulation is performed. The surrogate models are used to assist the search of optima in the objective functions, thereby significantly reducing the computational requirements.

Compared to prior multi-objective surrogate-assisted optimization algorithms, MOPLS introduces an innovative tabu mechanism that addresses the common challenge in real-world engineering applications of finding the global optimum in the presence of multiple local optima (Wang et al., 2023). Specifically, to avoid getting stuck in local optima, the tabu mechanism temporarily sets aside some high-quality parameter vectors around which the surrogate-assisted local searches yield poor-quality results. Moreover, MOPLS can take advantage of parallel processing to boost its efficiency. This is because each surrogate-assisted local search and its subsequent objective evaluation are conducted independently, making it possible to assign them to separate worker processors.

MOPLS has demonstrated better performance when compared to other recent state-of-the-art surrogate-assisted multi-objective methods across a suite of widely used benchmark problems (Wang et al., 2023). Additionally, it performs well in the calibration of the Soil & Water Assessment Tool (SWAT) (Wang et al., 2023). In this study, by applying MOPLS to the parameter calibration of ELM-FATES, we aim to identify a set of optimal parameter values with respect to multiple different objectives, which is also known as the “non-dominated set” (Wang et al., 2023). Below, we provide a detailed explanation of the surrogate model adopted by MOPLS and the algorithmic workflow.

2.4.1. Radial Basis Function

In MOPLS, the surrogate model employed is the Radial Basis Function (RBF) with a cubic kernel and a linear tail. Given n distinct d -dimensional sample points $X = [x_1, \dots, x_d] \in \mathbb{R}^d$, and an unknown function $F: \mathbb{R}^d \rightarrow \mathbb{R}$, we generate a training set $V = \{(X_i, F(X_i))\}_{i=1}^n$ in which $F(X_i)$ is the evaluated function value at point X_i . In a noise-free environment, a RBF model approximates the function F in the following augmented form:

$$s(X) = \sum_{i=1}^n \lambda_i \phi(\|X - X_i\|_2) + \sum_{j=1}^{d+1} c_j p_j(X) \quad (1)$$

where the cubic kernel function $\phi(r) = r^3$ is conditionally positive definite, and linear tail $[p_1(X), \dots, p_{d+1}(X)]^T = [1, x_1, \dots, x_d]^T$. $\|\cdot\|_2$ denotes the Euclidean norm. Based on the interpolation principle, Powell (1992) had shown that the coefficients $\lambda = [\lambda_1, \dots, \lambda_n]^T$ and $c = [c_1, \dots, c_{d+1}]^T$ in the augmented RBF model can be uniquely determined by solving a linear system:

$$\begin{bmatrix} \Phi & P \\ P^T & \mathbf{0}_{(d+1) \times (d+1)} \end{bmatrix} \begin{bmatrix} \lambda \\ c \end{bmatrix} = \begin{bmatrix} y \\ \mathbf{0}_{d+1} \end{bmatrix} \quad (2)$$

where the matrix $\Phi \in \mathbb{R}^{n \times n}$ with $[\Phi]_{i,j} = \phi(\|X_i - X_j\|_2)$, matrix $P \in \mathbb{R}^{n \times (d+1)}$ with $[P]_{i,j} = p_j(X_i)$, and column vector $y = [F(X_1), \dots, F(X_n)]^T$.

Furthermore, in the iterative optimization context, the RBF surrogate model s is constantly updated to achieve better prediction accuracy, as additional information (i.e., more evaluated points) about function F becomes available. Therefore, we consider the surrogate value at any unevaluated point X to be a reliable prediction of $F(X)$ and use it to assist the optimization of F , especially when evaluating F demands substantial computational resources due to its reliance on original physical model simulations.

2.4.2. MOPLS Algorithm Workflow

Figure 2 illustrates the algorithmic workflow of MOPLS for the bi-objective calibration of early and late PFTs in ELM-FATES. In the initialization stage (Step 1), MOPLS uses Symmetric Latin Hypercube Design (SLHD) (Ye et al., 2000) to randomly generate a set of distinct parameter vectors and evaluate their objective values through the ELM-FATES simulations. The main loop of MOPLS (Step 2) comprises four sub-steps. First (Step 2.1), it builds a RBF surrogate model for each objective function we are trying to optimize. Second (Step 2.2), it selects high-quality parameter vectors with the assistance of these surrogate models. Third (Step 2.3), it evaluates how well these selected parameter combinations perform in the real ELM-FATES simulations. Lastly (Step 2.4), MOPLS updates a set of parameter vectors that are currently the best found so far, known as the “non-dominated set” (Wang et al., 2023). This process continues until MOPLS reaches a predefined limit on the number of total evaluations, which was set at 192 in our study, following Xia et al. (2020) and Cheng et al. (2023). They utilized an optimization algorithm derived from the same surrogate-assisted optimization toolbox as MOPLS. At that point, MOPLS terminates the loop and returns the set of non-dominated parameter vectors found so far as the final optimization result.

The process of selecting new parameter vectors (i.e., Step 2.2) involves multiple separate local searches aided by surrogate models and is vital for the effectiveness of MOPLS. In this process, we rely on a widely used metric called the Hypervolume improvement indicator (Zitzler et al., 2007) to assess the quality of different parameter

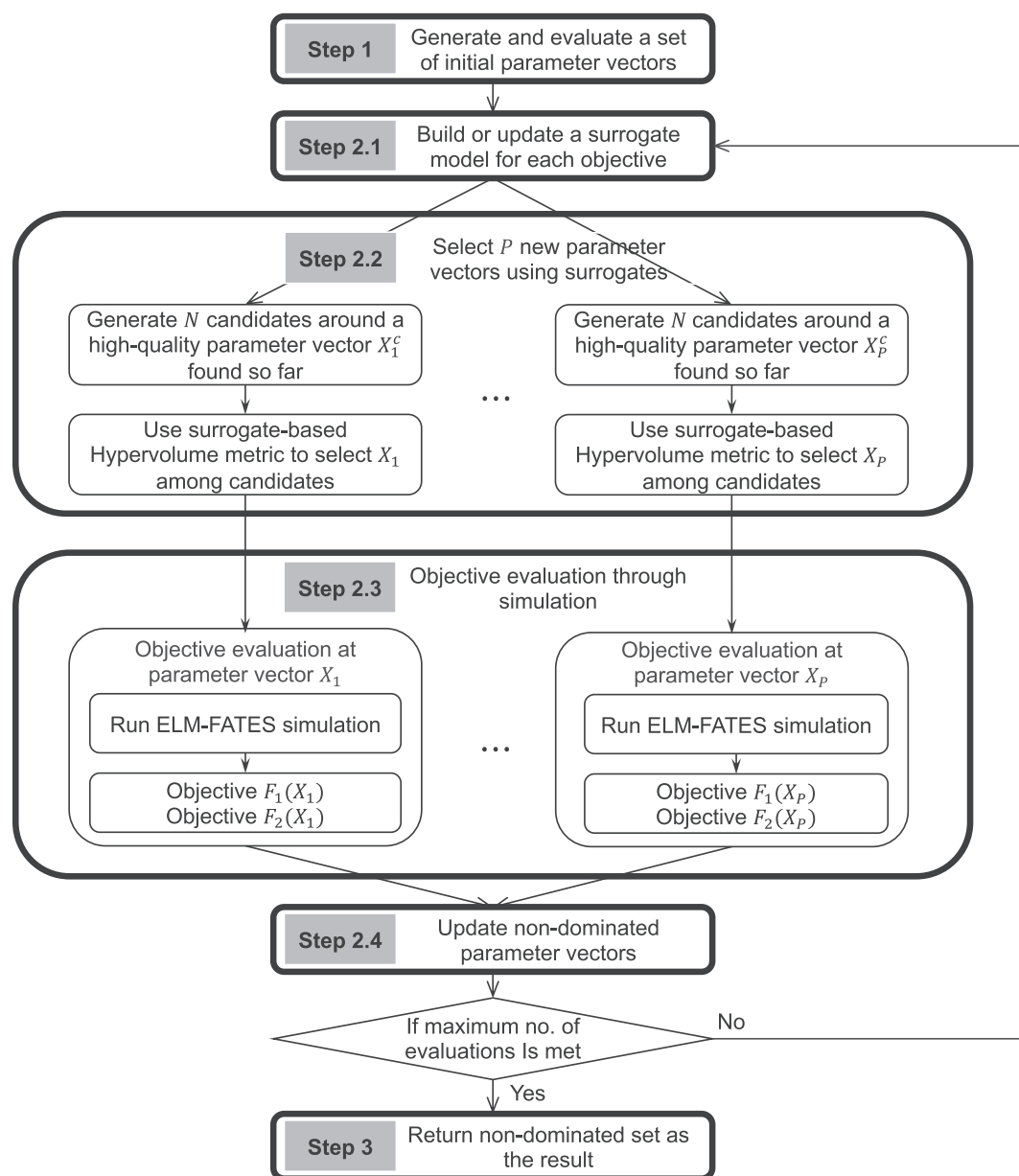


Figure 2. The algorithmic workflow of MOPLS when calibrating early and late plant function types in ELM-FATES. We applied MOPLS with P (number of processors) being 24 in this study. F_1 and F_2 are the two objective functions as defined in Equations 3 and 4, respectively.

combinations across multiple objective functions (more details are provided in Text S1 in Supporting Information S1). The parameter vectors that have a higher Hypervolume improvement indicator are considered better. Each surrogate-assisted local search starts by generating a set of N candidates. These candidates are unevaluated parameter vectors that follow a multivariate Gaussian distribution, with an evaluated parameter vector of high quality serving as the mean. Since the true objective values of each candidate are unknown, the algorithm computes surrogate values of the candidate to approximate its Hypervolume improvement indicator. Finally, we select the candidate with the highest estimated Hypervolume improvement indicator value and send it for objective evaluation through real ELM-FATES simulation in Step 2.3. This step helps us narrow down the best parameter vector choices for our optimization process.

2.5. Multi-Objective Problem Formulation

2.5.1. Objective 1

The first objective of the optimization is to minimize the error between observed and simulated values for f_{tree} , GPP, SH, LE, SWC, and Q in this study. We use Nash Sutcliffe Efficiency (NSE) (Nash & Sutcliffe, 1970) to quantify the error. The first objective function is the sum of negative NSEs for all the six variables, as calculated in Equation 3. An equal weight is assigned to each variable. The original NSE has a range from negative infinity to 1. The closer the NSE to 1, the better the model performance. Therefore, the negative NSE spans from -1 to positive infinity, and the best model performance corresponds to the lowest value for negative NSE, which is -1 . Our optimization algorithm is designed to search for the global minimum, aiming to achieve the best model performance by minimizing the value of the sum of negative NSEs calculated in Equation 3.

$$F_1(X) = \sum_{i=0}^M (-NSE(X)_i) \quad (3)$$

where X is parameter vector, $NSE(X)_i$ is the NSE value for variable i when using parameter vector X , M is the number of calibrated variables (six in this study). Lower $F_1(X)$ values indicate better performance.

2.5.2. Objective 2

The second objective of the optimization is to achieve coexistence for early and late successional PFTs in ELM-FATES. Here we use the biomass percentage of the early successional tropical forest ($perc_{early}$) to represent the degree of coexistence. We define that $perc_{early}$ within 30% and 70% as stable coexistence. The objective function has a minimum at $perc_{early}$ equals to 50%, representing the case where the two PFTs are equal competitors (i.e., early and late successional PFTs occupy equal percentages in the ecosystem) and strongly penalizes cases where $perc_{early}$ for one of the two PFTs exceeds 70% (Figure 1a). Thus, the second objective function is calculated as:

$$F_2(X) = \begin{cases} a * abs(perc_{early}(X) - 50\%)^c & 30\% \leq perc_{early} \leq 70\% \\ b * abs(perc_{early}(X) - 50\%)^c & perc_{early} < 30\% \text{ or } perc_{early} > 70\% \end{cases} \quad (4)$$

where X is parameter vector.

We test three cases for $F_2(X)$ in this study: (1) $a = 500$, $b = 100$, $c = 0.8$ (f1), (2) $a = 1500$, $b = 600$, $c = 0.9$ (f2), and (3) $a = 5500$, $b = 3,000$, $c = 0.5$ (f3) (Figure 3a). Given the absence of a formal formula for identifying coexistence to the best of our knowledge, both the functional form for Objective 2 and the values of parameters a , b , and c are determined through heuristics.

2.6. Optimization Experiment Design

To maximize processor utilization, we employ all processors available within a single computing node ($P = 24$ processors in our case) for the parallel MOPLS algorithm used in this study. To account for the inherent variability in initial conditions within the optimization process, it is imperative to evaluate the performance of an optimization method through multiple trials. We perform five trials for each optimization experiment given the substantial computational demand of calibrating the ELM-FATES model.

3. Results and Discussions

3.1. Impact of Objective Functions on Achieving Coexistence

Figure 3b illustrates the probability densities of the percentage of biomass for early successional PFT ($perc_{early}$) for the three test functions of Objective 2 (Equation 4). The probability of coexistence for early and late PFTs is relatively low when applying the three test functions (Figure 3b). While different objective functions might achieve coexistence more efficiently, the perceived challenge in achieving coexistence does not appear to arise from the choice of objective functions. Instead, it is primarily due to the amplified gap-phase dynamics (Koven et al., 2019) and the potential over-exaggeration of reproduction feedbacks (Fisher

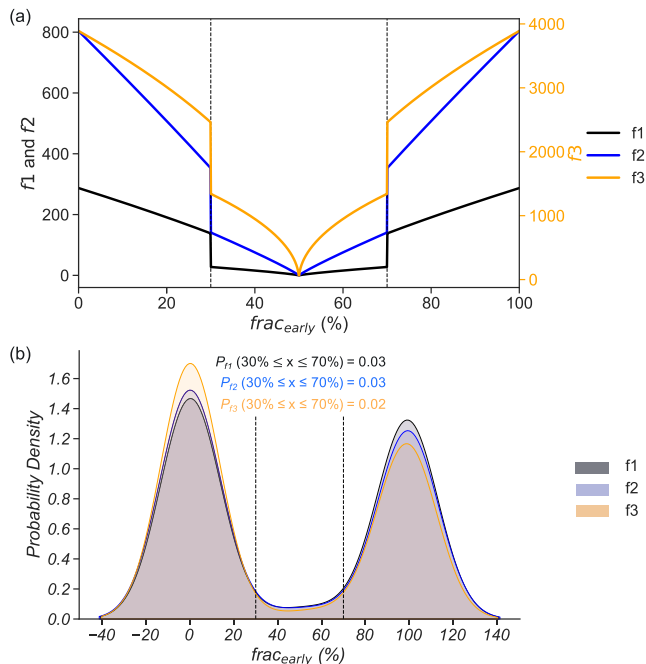


Figure 3. (a) The three objective functions tested for the second objective function ($F_2(X)$ in Equation 4) and (b) the probability density of the biomass percentage of early successional tropical forest ($perc_{early}$).

et al., 2010) in ED models, which may have created conditions that predispose the dominance of early successional tree species. Specifically, in ED models, the forest is represented as a mosaic of patches of varying ages (time since last disturbance). Once a disturbance hits the patch, ED models typically eradicate all trees and reset succession. Early successional species are more efficient to take advantage of the gap formation following disturbances, becoming dominant. Koven et al. (2019) investigated the impact of disturbance representation on the dominance balance between early and late successional PFTs in ELM-FATES. Their findings corroborate that the generation of new patches upon disturbance favored early successional plants. Consequently, the likelihood of coexistence for early and late PFTs is relatively low in our study.

In addition, we find that optimization experiments utilizing various objective functions to characterize coexistence yield similar probability densities for $perc_{early}$ (Figure 3b). This indicates that MOPLS is effective in achieving forest coexistence across a range of objective functions. Therefore, it is easy to identify suitable objective functions for coexistence when using MOPLS.

3.2. Optimization Results

3.2.1. Parameter Distribution

The parameters calibrated by Latin Hypercube Sampling (LHS) and MOPLS exhibit distinct distributions (Figure 4). This disparity arises from the inherent nature of LHS as a sampling method, which enforces a uniform spatial search

during the sampling process, consequently resulting in a uniform distribution of the parameters. In contrast, MOPLS is an optimization algorithm that does not prioritize spatial uniformity; instead, it focuses on localized searches that yield high-quality results, leading to the emergence of localized peaks within the parameter distributions.

In line with a prior study that uses LHS for sensitivity analysis (Cheng et al., 2022), our analysis using MOPLS reveals that f_{tree} , GPP, SH, and LE exhibit greater sensitivity to vegetation parameters, specifically $V_{c,max}$ and τ (Figure S2 in Supporting Information S1). Conversely, SWC exhibits greater sensitivity to hydrological parameters, particularly B_{sw} . Additionally, B_{sw} plays a crucial role in controlling both carbon and water-related variables, consistent with prior studies (Cheng et al., 2022, 2023).

3.2.2. Optimal Parameters

The optimal parameters for the early and late successional tropical forests, as well as the soil hydrological parameters, are similar to the observational values reported for tropical regions (Table 1). For example, the leaf and fine root turnover times are calibrated to be 0.9/2.3 years for early/late PFTs, which is similar to the parameter value calibrated for two Amazonia sites—0.9/2.6 years for early/late PFTs (Detto et al., 2018; Huang et al., 2020). The observed and calibrated ρ_{wood} in Panama are very similar, which are 0.4/0.68 and 0.4/0.8 g/cm³ for early/late successional PFTs, respectively (Condit et al., 2012). The optimal θ_s is 0.6 m³/m³, consistent with field observation in Panama (0.57 m³/m³) (Litt et al., 2020). These consistencies between observed and calibrated parameter values demonstrate the capability of MOPLS in finding the optimal solution for both early and late PFTs in tropical forests.

3.2.3. Model Performance

The seasonal dynamics of observations are adequately captured by the simulations using optimal solutions obtained by MOPLS (Figure 5). The mean errors between the calibrated and observed values for GPP, SH, LH, SWC, and Q are 4.7 g C/m²/month, 3.3 W/m², 7.6 W/m², −0.04 m³/m³, and −4.8 mm/month, respectively. It should be noted that the soil water content in the top 15 cm soil layer is still underestimated during the dry season. This can be potentially attributed to the smaller fraction of deep roots in the optimal solution. Specifically, the

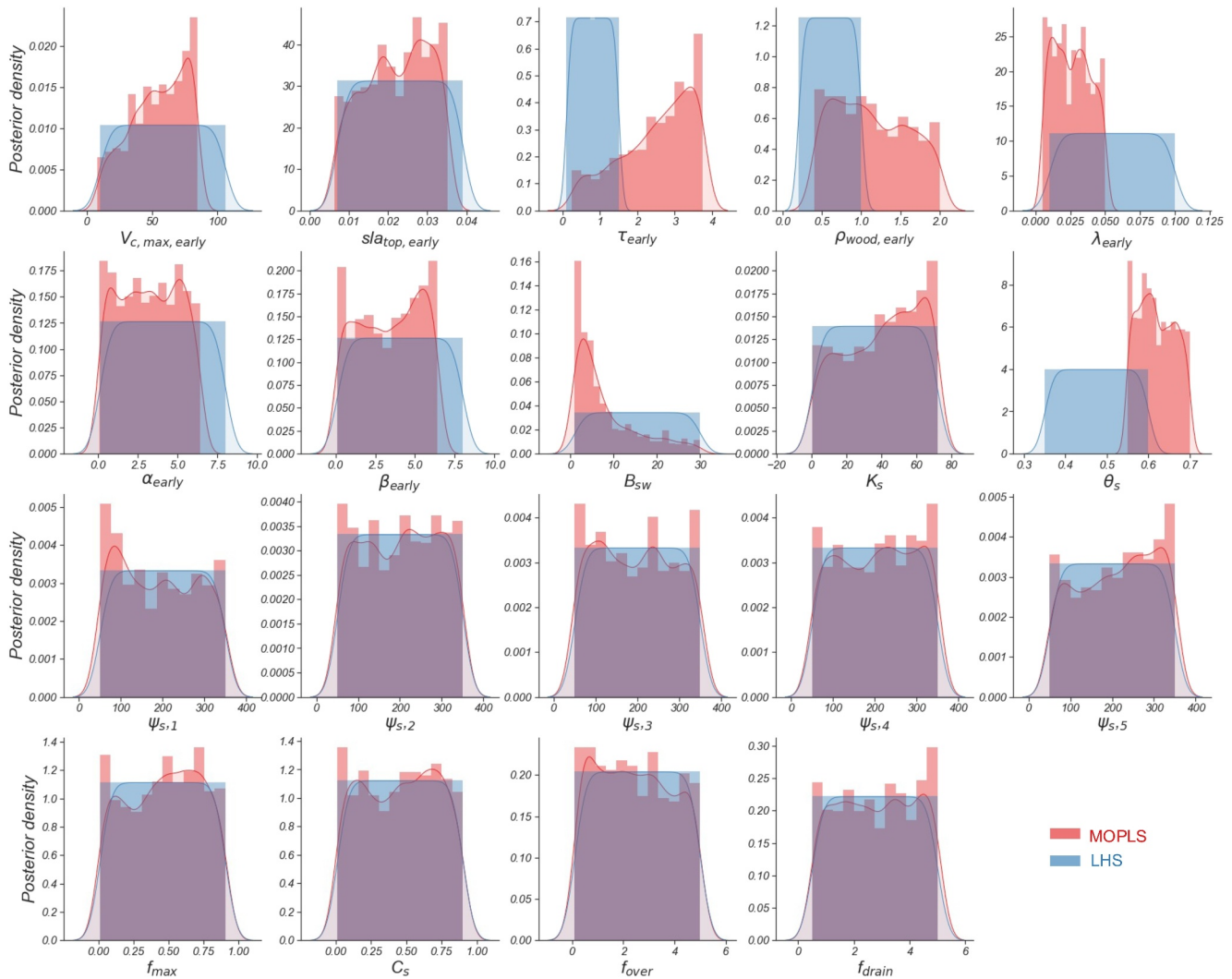


Figure 4. Posterior probability distributions of the 19 vegetation and soil hydrological parameters calibrated using MOPLS (red color) and Latin Hypercube Sampling (LHS, blue color). The vegetation parameters are for early PFTs, while the hydrological parameters are for the top soil layer. Please refer to Table 1 for full names of the parameters.

roots fraction above and below 20 cm of the soil are 0.8/0.6 and 0.2/0.4 for early/late successional PFTs, respectively. Previous studies have also reported the importance of the fraction of deep tree root in simulating dry season shallow layer soil moisture (Bretfeld et al., 2018; Cheng et al., 2022; Chitra-Tarak et al., 2021). In addition, the overestimation of LE during the dry season (Figure 5d) could also contribute to the underestimated shallow layer soil moisture. The underestimated dry-season soil moisture results in underestimated dry-season runoff (Figure 5f), due to the insufficient supply of groundwater recharge. Future attention should be drawn to improve the representation of deep root fraction and seasonal water table movements in the model.

Note that Objective 1 assigns equal weight to each variable (Equation 3), as our goal is to optimize the overall model performance rather than the performance of individual variables. This may not guarantee the optimality for every variable, potentially resulting in suboptimal solutions for certain variables, as observed in the insufficiencies in simulating GPP seasonality (Figure 5b). Future studies can optimize individual variables or apply distinct weights to each variable based on specific priorities or areas of interest. The misrepresentation of GPP seasonality is also possibly stemming from the absence of a plant hydraulics scheme and seasonality in photosynthetic capacity (Wu et al., 2017). Furthermore, a plant hydraulics scheme could potentially redistribute water uptake to shallower soil layers, addressing the underestimated moisture in the shallow soil layer (Figure 5d).

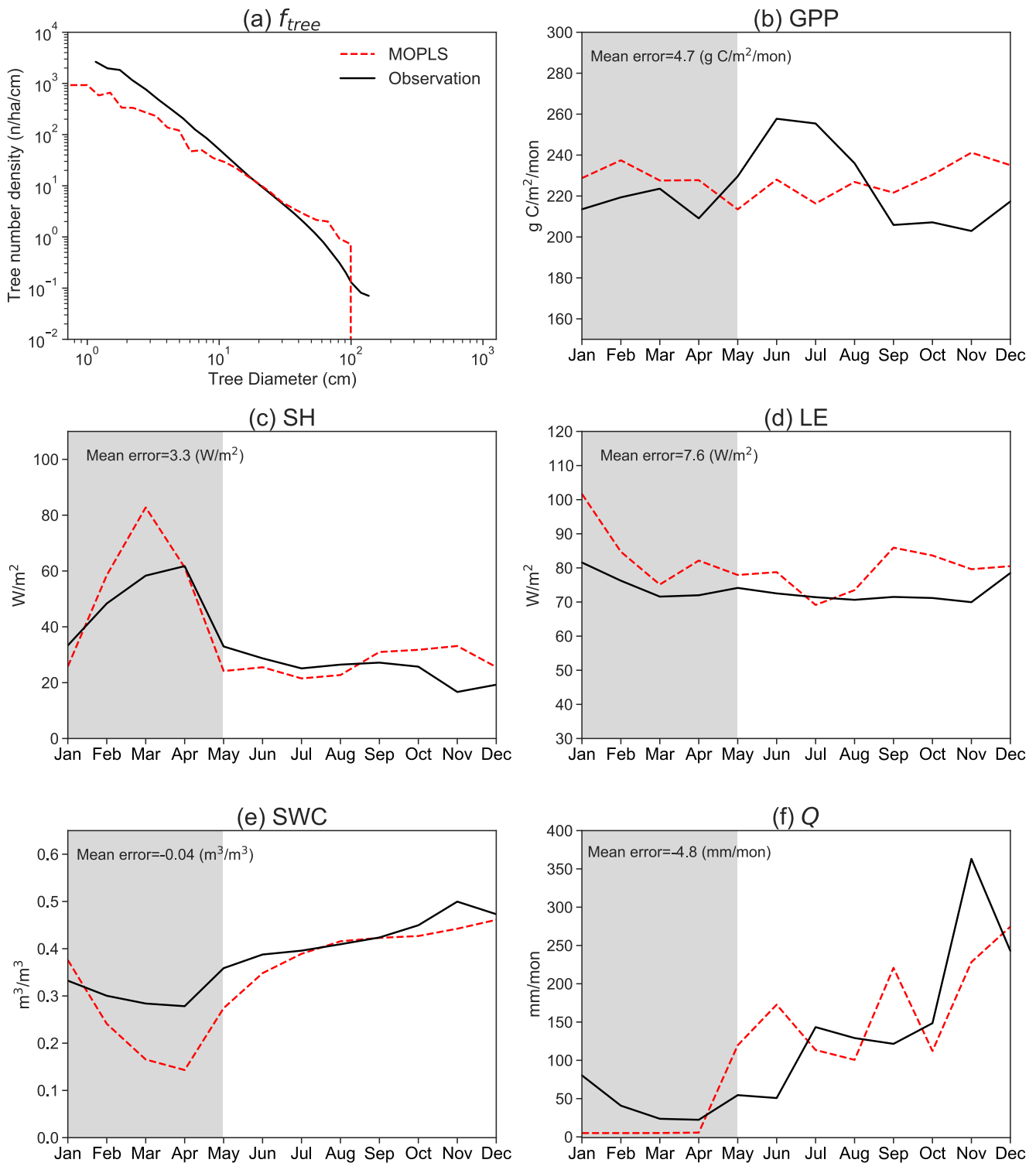


Figure 5. (a) Tree size distribution (f_{tree}) and mean seasonal (b) gross primary productivity (GPP), (c) sensible heat (SH), (d) latent heat (LE), (e) upper-layer (top 15 cm) soil water content (SWC), and (f) runoff (Q) between simulations using optimal solutions obtained by MOPLS (red line) and observations from BCI, Panama (black line). Gray area represents the dry season.

Conducting a more thorough optimization study for ELM-FATES by incorporating a plant hydrodynamic module (Xu et al., 2023) warrants future study.

It is noteworthy that across the Panamanian Isthmus, many canopy species, especially dry-deciduous species, flush new leaves at the transition between dry and wet seasons (Medina et al., 2022). This would confer higher photosynthetic capacity in the early mid wet season. Therefore, the underestimation of early wet-season GPP may also stem from the absence of seasonality in leaf phenology in ELM-FATES, particularly in photosynthetic capacity. In response to this issue, ELM-FATES does incorporate a “leaf age” capability. However, this feature is seldom utilized due to its computational demands. Specifically, it necessitates users to specify variations in $V_{c,max}$ as a function of leaf age and PFT, as well as to track leaf ages in each cohort and calculate the photosynthetic uptake for each of these separately, resulting in substantial computational costs. Alternatively, another approach to addressing this issue could involve utilizing an optimizing model of $V_{c,max}$ that increases in response to high light (Ali et al., 2016). Nevertheless, the version of ELM-FATES used in this study did not implement either of these features, leading to the misrepresentation of GPP seasonality.

In addition to satisfactorily capturing dynamics for calibrated variables, it should be noted that the optimal solution also exhibits satisfactory performance in simulating non-calibrated variables. For instance, the observed and simulated mean annual aboveground biomass (AGB) is 14.0 (Chave et al., 2003) and 13.4 kg C/m², respectively. The biomass for early and late successional PFTs from the optimal solution is 7.1 kg C/m² (52%) and 6.4 kg C/m² (48%), respectively, with each PFT occupying almost an equal percentage. This is consistent with what is observed for coexisted early and late PFTs (Huang et al., 2020).

Furthermore, the ELM-FATES model with MOPLS-calibrated parameters for early and late PFTs can reasonably well capture the size structure and tree density in the BCI tropical forests. For example, the observed and simulated basal areas align well, with the observed/simulated basal area by size classes at 3.8/2.5, 7.3/9.1, and 6.2/7.2 m²/ha for DBH of 1–10 cm, 10–30 cm, and 30–50 cm, respectively. Nevertheless, there is an overestimation in the basal area for large trees (≥50 cm), which is 15.4 and 47 m²/ha for observation and simulation, respectively. Moreover, the simulated stem density values align well with observations for small trees, but there is an overestimation for large trees. Specifically, the observed/simulated stem density by size classes is 4014/1700, 337/310, 54/53, and 30/86 trees per hectare (N/ha) for DBH of 1–10 cm, 10–30 cm, 30–50 cm, and ≥50 cm, respectively. The overestimation of large trees was attributed to the role of a carbon starvation mortality term, which is a common issue in forest demography models (Holm et al., 2020). Specifically, the primary mode of mortality in FATES involves a constant background turnover mortality, along with a carbon starvation mortality term. The latter is mainly responsible for the eradication of trees in the shady understory that are in negative carbon balance, leading to density-dependent mortality as an emergent process. For canopy trees, carbon starvation is rarely triggered, and any overestimate of growth rates tends to result in an overabundance of large trees. To address this issue, Needham et al. (2020) introduced a scheme into FATES to simultaneously track cohort age and size. This modification can mitigate the overestimated biomass of large trees. However, it is important to note that this is not the default configuration of FATES. Future studies can be conducted by applying this new scheme to improve the overestimation of large trees.

3.3. Comparison Between MOPLS and Random Sampling

The probability for achieving coexistence is relatively low, with the probability of $perc_{early}$ within 30%–70% being 0.03, 0.03, and 0.02 for f1, f2, and f3, respectively (Figure 3b). Nevertheless, with the same number of simulations, MOPLS (Figures 6a–6c) performs much better in achieving coexistence than a previous study (Cheng et al., 2022) that used a random sampling strategy (i.e., LHS) (Figures 6d–6f). Specifically, with a total of 1000 ensembles, the number of simulations that can achieve coexistence with $perc_{early}$ within 10%–90%, 20%–80%, and 30%–70% is 52/15, 40/7, and 32/3 for MOPLS/LHS, respectively (Figure 6). With the requirement for coexistence becoming stricter (i.e., the range of $perc_{early}$ increasing from 10%–90% to 30%–70%), MOPLS can still achieve a reasonable number of simulations in which the two PFTs can coexist (i.e., 32 out of 1,000, Figure 6c), while LHS only has three simulations out of 1,000 ensembles that can achieve coexistence for early and late PFTs (Figure 6f). This demonstrates that using MOPLS, an optimization algorithm that employs surrogate approximations for each objective function, is more effective in calibrating coexistence in tropical forests. The surrogated-assist feature of MOPLS enables it to perform a larger number of evaluations within the same computational constraints compared to random sampling. Moreover, MOPLS incorporates several beneficial

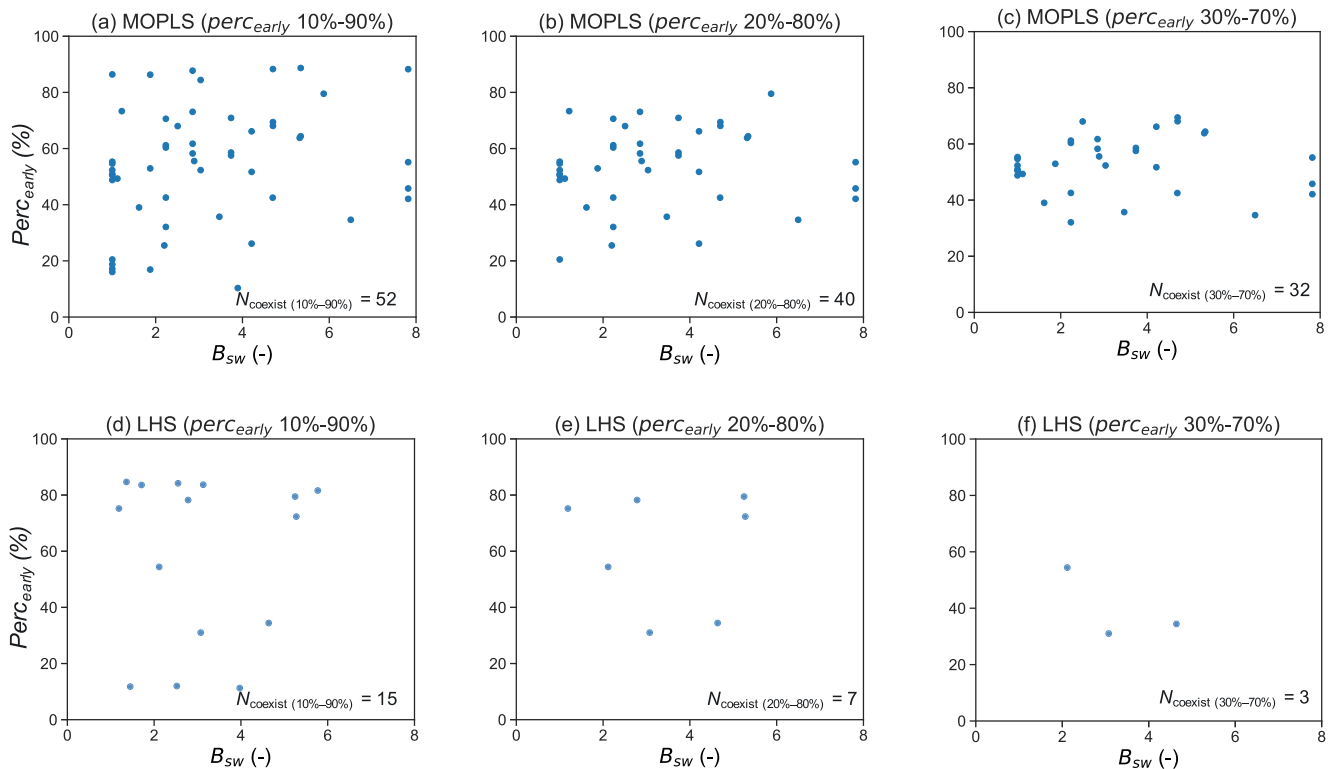


Figure 6. Relationship between the scaling exponent of the soil retention curve (B_{sw}) and the percentage of biomass for early successional PFT ($perc_{early}$) in simulations achieving the coexistence of early and late PFTs. Results are presented for (a–c) Multi Objective Population-based Parallel Local Surrogate-assisted search (MOPLS) and (d–f) Latin Hypercube Sampling (LHS), with distinctions based on $perc_{early}$ within (a) and (c) 10%–90%, (b) and (e) 20%–80%, and (c) and (f) 30%–70%. $N_{coexist}$ (10%–90%), $N_{coexist}$ (20%–80%), and $N_{coexist}$ (30%–70%) represent the number of simulations achieving coexistence with $perc_{early}$ within 10%–90%, 20%–80%, and 30%–70%, respectively.

greedy search strategies that random sampling lacks (Wang et al., 2023). For example, MOPLS employs a surrogate-assisted local search centered around a high-quality evaluated parameter vector. This strategy leverages the assumption that neighboring parameter vectors of high quality are likely to yield high-quality results. In addition, when evaluating how well the parameter combinations perform in the real ELM-FATES simulations, MOPLS consistently select a new parameter combination that is predicted to have the best quality based on knowledge obtained from prior evaluations. These design features set MOPLS significantly apart from pure random sampling and enable it to produce more simulations with stable existence than random sampling (Figure 6).

3.4. Simulated Trajectories of Tropical Forest Aboveground Biomass After Deforestation

We simulated the trajectories of ABG following different intensities of clear-cut deforestation to investigate the recovery of tropical forests under these human-induced disturbances. We find that the damage level in the selective logging practices determines the number of years required for AGB to recover to its original level. The results show that after 20% and 30% clear-cuts, it takes 38 and 65 years for AGB to recover to the initial stage, respectively. Moreover, 50% and 90% clear-cuts require 90 and 153 years, respectively, for AGB to fully recover (Figure 7). The results are consistent with those reported in previous studies (Piponiot et al., 2018), which found that, with initial losses of 10%, 25%, and 50% in forest biomass, it took 12, 43, and 75 years for AGB to recover, respectively. These findings indicate that increasing clear-cut intensities would disproportionately extend the recovery time, warranting careful consideration when intensifying clear-cutting practices.

After fully recovering from logging under various intensities, AGB would continue to increase to a level even higher than the original value (i.e., 14 kg C/m²) (Figure 7a), consistent with previous findings that the AGB accumulation rates in logged forests are higher than those observed in intact forests (Huang & Asner, 2010; Huang et al., 2020; Mazzei et al., 2010). This is mainly attributed to the rapid recovery rates of early successional

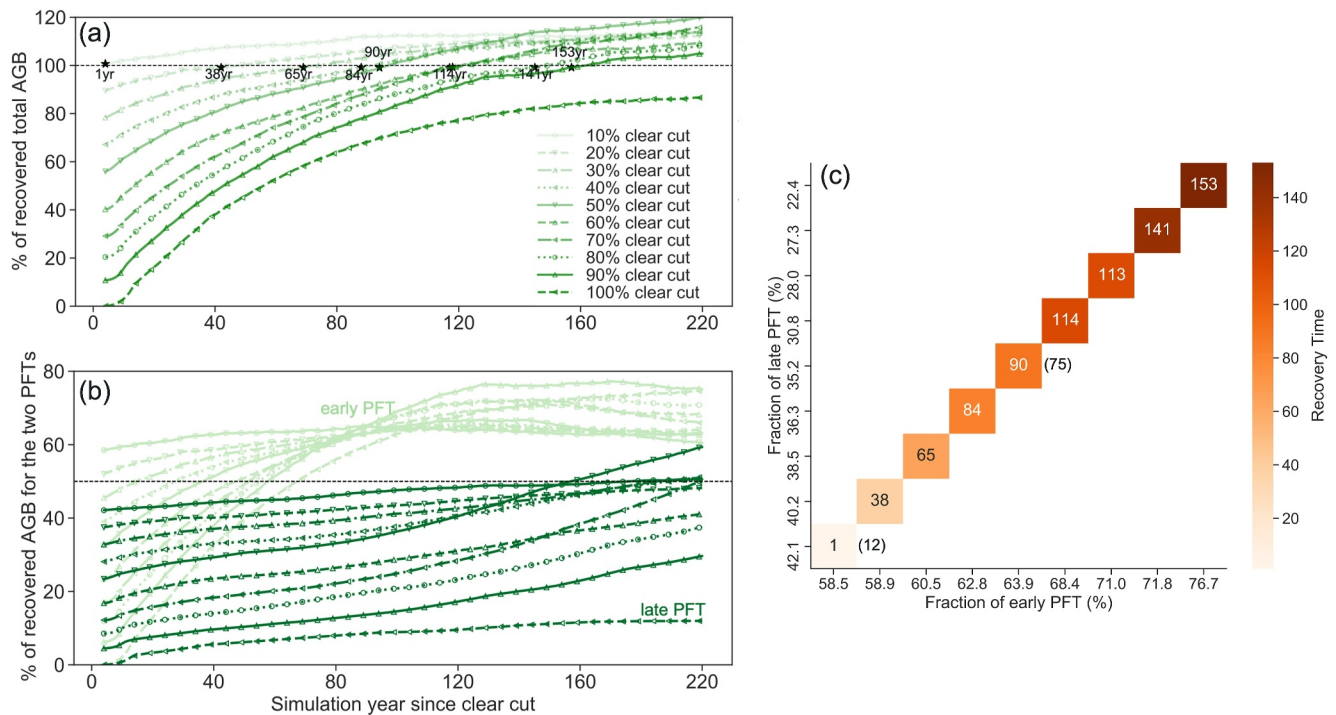


Figure 7. Simulated trajectories of recovered percentages of (a) total aboveground biomass (AGB) and (b) AGB of early and late PFTs following varying intensities of forest clear-cutting (from 10% to 100%). (c) Heatmap illustrates the fraction of early successional (x axis) and late-successional (y axis) PFTs, with color representing the corresponding required recovery time in total. The recovery line in each scenario crossed by the horizontal line of 100% of total AGB in (a) represents the time required for full recovery (e.g., 1 and 38 years as denoted in (a)). The numbers in brackets in (c) represent recovery times from the literature.

trees (Figure 7b), especially during the early stage of the post-logging process when late PFT biomass has not yet fully recovered, and early PFT experiences rapid growth, aided by sufficient sunlight.

The initial recovery rates of AGB for early PFT following clear-cutting are faster after high-intensity logging activities (e.g., >70% clear-cut) (Figures 7b and 7c) because the sufficient light reaching the forest floor favors the light-demanding nature of early PFTs. In contrast, late PFTs take a longer time to recover and do so at a much slower rate, even after low-intensity logging activities (e.g., 10% clear-cut, Figures 7b and 7c).

The findings indicate that in mixed tropical forests, a 20% clear-cut would require nearly 40 years for AGB to recover (Figures 7b and 7c). Specifically, while the early PFT takes only 3 years to recover to its initial fraction, the late PFT requires more than 160 years to recover to the fraction in the original unlogged forests (Figure 7b). These findings have significant implications for forest management practices. This suggests that even a small degree of damage in a tropical forest would demand a long time for carbon stocks to regenerate, thereby exerting a substantial impact on both biogeochemical and biological processes.

Note that the selective logging module in ELM-FATES is specifically designed to simulate real-world wood harvesting in landscape-level field operations. It only accounts for tree mortality linked to human activities and does not consider tree mortality caused by drought or pests/insects. Recovery after drought or pests/insects will probably differ from recovery after selective logging. First, selective logging targets trees of a specific size range, while drought could affect all sizes. Second, drought is known to have delayed effects that might last several years after the events (Trugman et al., 2018).

3.5. Future Directions

As a starting point for calibrating tropical forest coexistence in ED models, this study evaluates the performance of a multi-objective optimization algorithm in calibrating the coexistence for two typical tropical tree species. In the future, MOPLS can be extended to calibrate the coexistence for a greater number of PFTs (i.e., >2 objectives) to better represent tropical forest diversity and better capture the interactions between tropical forests and

environmental changes. Meanwhile, this study only applies MOPLS to ELM-FATES. Future studies can extend the application of the multi-objective optimization algorithm to other ED models, such as ED v3.0 (Ma et al., 2022). While tropical forests worldwide share similarities in climate and soil conditions, each tropical forest is unique. Our study specifically applies MOPLS to a Panama site. Future work can apply the multi-objective optimization algorithm to match other ecosystem observations representing different tropical forest characteristics. Furthermore, improving the representation of deep root fraction and seasonal water table movements in the model deserves attention in future endeavors.

In addition, this study investigates the recovery trajectories of forest biomass after different intensities of deforestation under current climate conditions. Future study can explore forest recovery under different future climate change conditions, such as El Niño-related droughts and changes in wet or dry seasons (T. L. Powell et al., 2018). Moreover, future research can be conducted to understand forest recovery while taking into account the influence of forest regeneration on soil properties (Zhang et al., 2019). Furthermore, future study on forest recovery should encompass not only the restoration of aboveground biomass but also the recovery of forest diversity (Rüger et al., 2020), which is crucial for effective planning of forest restoration and sustainable tropical forest management. Additionally, this study exclusively investigates the recovery time under various clear-cut intensities within the same coexistence scenario. Future research endeavors could explore how recovery time varies under different coexistence scenarios.

4. Conclusions

This study successfully applied a multi-objective optimization called MOPLS (Multi Objective Population-based Parallel Local Surrogate-assisted search) to calibrate the coexistence of tropical forest species in BCI, Panama. Specifically, it focused on calibrating the coexistence of early and late successional forests within an ED model named ELM-FATES (the Functionally Assembled Terrestrial Ecosystem Simulator [FATES] implemented in the Energy Exascale Earth System Model [E3SM] Land Model [ELM]). We performed three optimization experiments using different objective functions to represent the coexistence. MOPLS achieved the coexistence of early and late PFTs with similar probabilities for the three objective functions, indicating that it is easy to identify objective functions for calibrating coexistence when using the algorithm. The optimal solution could not only capture the dynamics of the calibrated variables (i.e., GPP, sensible and latent heat, soil water content, and runoff), but also those of non-calibrated variables (i.e., aboveground biomass, as well as basal area and stem density distributions across tree size classes). This demonstrates the capability of the MOPLS-calibrated parameters for early and late PFTs in simulating the dynamics of tropical forest ecosystems. Furthermore, with an equal number of model evaluations, MOPLS produced a much higher number of simulations achieving coexistence compared to a random sampling strategy. This study establishes the groundwork for calibrating coexistence while attaining satisfactory model performance in ELM-FATES, paving the way for better characterization of tropical forest diversity in advanced ED models.

The optimal parameters were applied to simulate the trajectory of AGB following various intensities of clear-cut deforestation. We found that a 20% clear-cut deforestation would require nearly 40 years for AGB to recover to the initial level, due to the slow recovery rates of late successional forests. Our results demonstrate the promise of MOPLS in identifying model parameters capable of simulating coexistence and diversity in tropical forests, with significant implications for forest management practices. The MOPLS algorithm can be applied to other complex models and employed to calibrate the coexistence of a greater number of PFTs, thereby further improving the representation of tropical forest diversity. The improved representation of tropical forest composition can contribute to better predictions of the trajectory of tropical forest carbon stocks subject to climate and land use changes.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

ELM-FATES has two separate repositories. The ELM model is described in E3SM-Project/E3SM (2019). The ELM code is available at <https://github.com/E3SM-Project/E3SM>. The FATES model is described in FATES-

Development-Team (2019). The FATES code is available at <https://github.com/NGEET/fates>. The MOPLS algorithm is described by Wang et al. (2023). The MOPLS code is based on Python 3.4 (or newer) and available at: <https://github.com/WY-Wang/MOPLS-Opt>. Data used for analysis in this study are available at Cheng (2024).

Acknowledgments

This research was supported by Professor Shoemaker's startup grant from the National University of Singapore (NUS). The computational work for this article was fully performed on resources of the National Supercomputing Centre, Singapore (NSCC). MD was supported by the Carbon Mitigation Initiative at Princeton University and NSF Grant 2017804. The meteorological data were provided by the Physical Monitoring Program of the Smithsonian Tropical Research Institute. We acknowledge the Center for Tropical Forest Science—Forest Global Earth Observatory (CTFS-ForestGEO) that supported the flux tower data collection in BCI, Panama. We thank Dr. Ilija Ilijevski (NUS), Dr. Guozhen Wei (Dalian Maritime University), and Dr. Wei Xia (Politecnico di Milano) for helpful discussion.

References

- Ali, A. A., Xu, C., Rogers, A., Fisher, R. A., Wullschlegel, S. D., Massoud, E. C., et al. (2016). A global scale mechanistic model of photosynthetic capacity (LUNA V1.0). *Geoscientific Model Development*, 9(2), 587–606. <https://doi.org/10.5194/gmd-9-587-2016>
- Anderson-Teixeira, K. J., Davies, S. J., Bennett, A. C., Gonzalez-Akre, E. B., Muller-Landau, H. C., Joseph Wright, S., et al. (2015). CTFS-ForestGEO: A worldwide network monitoring forests in an era of global change. *Global Change Biology*, 21(2), 528–549. <https://doi.org/10.1111/gcb.12712>
- Asner, G. P., Knapp, D. E., Broadbent, E. N., Oliveira, P. J. C., Keller, M., & Silva, J. N. (2005). Ecology: Selective logging in the Brazilian Amazon. *Science*, 310(5747), 480–482. <https://doi.org/10.1126/science.1118051>
- Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., et al. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, 2(3), 182–185. <https://doi.org/10.1038/nclimate1354>
- Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D., & Houghton, R. A. (2017). Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science*, 358(6360), 230–234. <https://doi.org/10.1126/science.aam5962>
- Bonan, G. B. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, 320(5882), 1444–1449. <https://doi.org/10.1126/science.1155121>
- Bonell, M., & Bruijnzeel, L. A. (2005). Forests, water and people in the humid tropics: Past, present and future hydrological research for integrated land and water management. *Forests, Water and People in the Humid Tropics: Past, Present and Future Hydrological Research for Integrated Land and Water Management*, 1–933. <https://doi.org/10.1017/CBO9780511535666>
- Bretfeld, M., Ewers, B. E., & Hall, J. S. (2018). Plant water use responses along secondary forest succession during the 2015–2016 El Niño drought in Panama. *New Phytologist*, 219(3), 885–899. <https://doi.org/10.1111/nph.15071>
- Burslam, D. F. R. P., Garwood, N. C., & Thomas, S. C. (2001). Tropical forest diversity - The plot thickens. *Science*, 291(5504), 606–607. <https://doi.org/10.1126/science.1055873>
- Chave, J., Condit, R., Lao, S., Caspersen, J. P., Foster, R. B., & Hubbell, S. P. (2003). Spatial and temporal variation of biomass in a tropical forest: Results from a large census plot in Panama. *Journal of Ecology*, 91(2), 240–252. <https://doi.org/10.1046/j.1365-2745.2003.00757.x>
- Cheng, Y. (2024). Calibrating tropical forest coexistence in ecosystem demography models using multi-objective optimization through population-based parallel surrogate search [Dataset]. <https://doi.org/10.5281/zenodo.11178560>
- Cheng, Y., Leung, L. R., Huang, M., Koven, C., Detto, M., Knox, R., et al. (2022). Modeling the joint effects of vegetation characteristics and soil properties on ecosystem dynamics in a Panama tropical forest. *Journal of Advances in Modeling Earth Systems*, 14(1). <https://doi.org/10.1029/2021ms002603>
- Cheng, Y., Ogden, F., & Zhu, J. (2019). Characterization of sudden and sustained base flow jump hydrologic behaviour in the humid seasonal tropics of the Panama Canal Watershed. *Hydrological Processes*, 34(3), 569–582. <https://doi.org/10.1002/hyp.13604>
- Cheng, Y., Ogden, F., Zhu, J., & Bretfeld, M. (2018). Land use dependent preferential flow paths affect hydrological response of steep tropical lowland catchments with saprolitic soils. *Water Resources Research*, 54(8), 1–16. <https://doi.org/10.1029/2017WR021875>
- Cheng, Y., Xia, W., Detto, M., & Shoemaker, C. A. (2023). A framework to calibrate ecosystem demography models within Earth system models using parallel surrogate global optimization. *Water Resources Research*, 59(1). <https://doi.org/10.1029/2022WR032945>
- Chesson, P. (2000). General theory of competitive coexistence in spatially-varying environments. *Theoretical Population Biology*, 237(3), 211–237. <https://doi.org/10.1006/tpbi.2000.1486>
- Chitra-Tarak, R., Xu, C., Aguilar, S., Anderson-Teixeira, K. J., Chambers, J., Detto, M., et al. (2021). Hydraulically-vulnerable trees survive on deep-water access during droughts in a tropical forest. *New Phytologist*, 231(5), 1798–1813. <https://doi.org/10.1111/nph.17464>
- Condit, R., Chisholm, R. A., & Hubbell, S. P. (2012). Thirty years of forest census at Barro Colorado and the importance of immigration in maintaining diversity. *PLoS One*, 7(11), 1–6. <https://doi.org/10.1371/journal.pone.0049826>
- Condit, R., Hubbell, S. P., & Foster, R. B. (1995). Mortality rates of 205 neotropical tree and shrub species and the impact of a severe drought. *Ecological Monographs*, 65(4), 419–439. <https://doi.org/10.2307/2963497>
- Condit, R., Pérez, R., Lao, S., Aguilar, S., & Hubbell, S. P. (2017). Demographic trends and climate over 35 years in the Barro Colorado 50 ha plot. *Forest Ecosystems*, 4(1), 1–13. <https://doi.org/10.1186/s40663-017-0103-1>
- Delbart, N., Ciais, P., Chave, J., Viovy, N., Malhi, Y., & Le Toan, T. (2010). Mortality as a key driver of the spatial distribution of aboveground biomass in Amazonian forest: Results from a dynamic vegetation model. *Biogeosciences*, 7(10), 3027–3039. <https://doi.org/10.5194/bg-7-3027-2010>
- Detto, M., Levine, J. M., & Pacala, S. W. (2022). Maintenance of high diversity in mechanistic forest dynamics models of competition for light. *Ecological Monographs*, 92(2). <https://doi.org/10.1002/ecm.1500>
- Detto, M., & Pacala, S. W. (2022). Plant hydraulics, stomatal control, and the response of a tropical forest to water stress over multiple temporal scales. *Global Change Biology*, 28(14), 4359–4376. <https://doi.org/10.1111/gcb.16179>
- Detto, M., Wright, S. J., Calderón, O., & Muller-Landau, H. C. (2018). Resource acquisition and reproductive strategies of tropical forest in response to the El Niño-Southern Oscillation. *Nature Communications*, 9(1), 1–8. <https://doi.org/10.1038/s41467-018-03306-9>
- Domingues, T. F., Berry, J. A., Martinelli, L. A., Ometto, J. P. H. B., & Ehleringer, J. R. (2005). Parameterization of canopy structure and leaf-level gas exchange for an Eastern Amazonian tropical rain forest (Tapajós national forest, Pará, Brazil). *Earth Interactions*, 9(17), 1–23. <https://doi.org/10.1175/EI149.1>
- E3SM-Project/E3SM. (2021). Energy Exascale Earth System Model [Software]. <https://doi.org/10.11578/E3SM/dc.20230110.5>
- Erb, K., Kastner, T., Plutzer, C., Bais, A. L. S., Carvalhais, N., Fetzel, T., et al. (2018). Unexpectedly large impact of forest management and grazing on global vegetation biomass. *Nature*, 553(7686), 73–76. <https://doi.org/10.1038/nature25138>
- FATES-Development-Team. (2019). Technical note for the functionally assembled terrestrial ecosystem simulator (FATES) [Technical Documentation]. <https://doi.org/10.5281/zenodo.3517272>
- Fisher, R., & Koven, C. (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(4). <https://doi.org/10.1029/2018ms001453>
- Fisher, R., Koven, C., Anderegg, W., Christoffersen, B., Dietze, M., Farrior, C., et al. (2018). Vegetation demographics in Earth system models: A review of progress and priorities. *Global Change Biology*, 24(1), 35–54. <https://doi.org/10.1111/gcb.13910>

- Fisher, R., McDowell, N., Purves, D., Moorcroft, P., Sitch, S., Cox, P., et al. (2010). Assessing uncertainties in a second-generation dynamic vegetation model caused by ecological scale limitations. *New Phytologist*, 187(3), 666–681. <https://doi.org/10.1111/j.1469-8137.2010.03340.x>
- Fisher, R., Muszala, S., Versteinst, M., Lawrence, P., Xu, C., McDowell, N., et al. (2015). Taking off the training wheels: The properties of a dynamic vegetation model without climate envelopes. *Geoscientific Model Development*, 8(11), 3593–3619. <https://doi.org/10.5194/gmd-8-3593-2015>
- Freckleton, R. P., & Watkinson, A. R. (2002). Large-scale spatial dynamics of plants: Metapopulations, regional ensembles and patchy populations. *Journal of Ecology*, 90(3), 419–434. <https://doi.org/10.1046/j.1365-2745.2002.00692.x>
- Gatti, L. V., Basso, L. S., Miller, J. B., Gloor, M., Domingues, L. G., Cassol, H. L. G., et al. (2021). Amazonia as a carbon source linked to deforestation and climate change. *Nature*, 595, 388–393. <https://doi.org/10.1038/s41586-021-03629-6>
- Holm, J. A., Knox, R. G., Zhu, Q., Fisher, R. A., Koven, C. D., Nogueira Lima, A. J., 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(3), 1–23. <https://doi.org/10.1029/2019JG005500>
- Hou, Z., Huang, M., Leung, L. R., Lin, G., & Ricciuto, D. M. (2012). Sensitivity of surface flux simulations to hydrologic parameters based on an uncertainty quantification framework applied to the Community Land Model. *Journal of Geophysical Research*, 117(D15), D15108. <https://doi.org/10.1029/2012JD017521>
- Huang, M., & Asner, G. P. (2010). Long-term carbon loss and recovery following selective logging in Amazon forests. *Global Biogeochemical Cycles*, 24(3), 1–15. <https://doi.org/10.1029/2009GB003727>
- Huang, M., Hou, Z., Leung, L. R., Ke, Y., Liu, Y., Fang, Z., & Sun, Y. (2013). Uncertainty analysis of runoff simulations and parameter identifiability in the community land model: Evidence from MOPEX basins. *Journal of Hydrometeorology*, 14(6), 1754–1772. <https://doi.org/10.1175/JHM-D-12-0138.1>
- Huang, M., Xu, Y., Longo, M., Keller, M., Knox, R., Koven, C., & Fisher, R. (2020). Assessing impacts of selective logging on water, energy, and carbon budgets and ecosystem dynamics in Amazon forests using the Functionally Assembled Terrestrial Ecosystem Simulator. *Biogeosciences*, 17(20), 1–46. <https://doi.org/10.5194/bg-2019-129-RC2>
- Jamhuri, J., Samantha, L. D., Tee, S. L., Kamarudin, N., Ashton-Butt, A., Zubaid, A., et al. (2018). Selective logging causes the decline of large-sized mammals including those in unlogged patches surrounded by logged and agricultural areas. *Biological Conservation*, 227(April), 40–47. <https://doi.org/10.1016/j.biocon.2018.09.004>
- Koven, C., Knox, R., Fisher, R., Chambers, J., Christoffersen, B., Davies, S., et al. (2019). Benchmarking and parameter sensitivity of physiological and vegetation dynamics using the functionally assembled terrestrial ecosystem simulator (FATES) at Barro Colorado Island, Panama. *Biogeosciences*, 1–46. <https://doi.org/10.5194/bg-2019-409>
- Leung, L. R., Bader, D. C., Taylor, M. A., & McCoy, R. B. (2020). An introduction to the energy Exascale Earth system model (E3SM) special collection: Goals, science drivers, development, and analysis. *Journal of Advancing Modeling of Earth System*, 12(11). <https://doi.org/10.1029/2019MS001821>
- Lichstein, J. W., Dushoff, J., Levin, S. A., & Pacala, S. W. (2007). Intraspecific variation and species coexistence. *The American Naturalist*, 170(6), 807–818. <https://doi.org/10.1086/522937>
- Litt, G. F., Ogden, F. L., Mojica, A., Hendrickx, J. M. H., Kempema, E. W., Gardner, C. B., et al. (2020). Land cover effects on soil infiltration capacity measured using plot scale rainfall simulation in steep tropical lowlands of Central Panama. *Hydrological Processes*, 34(4), 878–897. <https://doi.org/10.1002/hyp.13605>
- Longo, M., & Keller, M. (2019). Not the same old(-growth) forests. *New Phytologist*, 221(4), 1672–1675. <https://doi.org/10.1111/nph.15636>
- Longo, M., Knox, R. G., Medvigy, D. M., Levine, N. M., Dietze, M. C., Kim, Y., et al. (2019). The biophysics, ecology, and biogeochemistry of functionally diverse, vertically and horizontally heterogeneous ecosystems: The Ecosystem Demography model, version 2.2-Part 1: Model description. *Geoscientific Model Development*, 12(10), 4309–4346. <https://doi.org/10.5194/gmd-12-4309-2019>
- Ma, L., Hurtt, G., Ott, L., Sahajpal, R., Fisk, J., Lamb, R., et al. (2022). Global evaluation of the Ecosystem Demography model (ED v3.0). *Geoscientific Model Development*, 15(5), 1971–1994. <https://doi.org/10.5194/gmd-15-1971-2022>
- Mazzei, L., Sist, P., Ruschel, A., Putz, F. E., Marco, P., Pena, W., & Ferreira, J. E. R. (2010). Above-ground biomass dynamics after reduced-impact logging in the Eastern Amazon. *Forest Ecology and Management*, 259(3), 367–373. <https://doi.org/10.1016/j.foreco.2009.10.031>
- Medina-Vega, J. A., Wright, S. J., Bongers, F., Schnitzer, S. A., & Sterck, F. J. (2022). Vegetative phenologies of lianas and trees in two Neotropical forests with contrasting rainfall regimes. *New Phytologist*, 235(2), 457–471. <https://doi.org/10.1111/nph.18150>
- Medvigy, D., Wofsy, S. C., Munger, J. W., Hollinger, D. Y., & Moorcroft, P. R. (2009). Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. *Journal of Geophysical Research*, 114(1), 1–21. <https://doi.org/10.1029/2008JG000812>
- Moorcroft, P. R., Hurtt, G. C., & Pacala, S. W. (2001). A method for scaling vegetation dynamics: The ecosystem demography model (ED). *Ecological Monographs*, 71(4), 557–586. [https://doi.org/10.1890/0012-9615\(2001\)071\[0557:AMFSVD\]2.0.CO;2](https://doi.org/10.1890/0012-9615(2001)071[0557:AMFSVD]2.0.CO;2)
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282–290. <https://doi.org/10.1080/00750770109555783>
- Needham, J. F., Chambers, J., Fisher, R., Knox, R., & Koven, C. D. (2020). Forest responses to simulated elevated CO₂ under alternate hypotheses of size- and age-dependent mortality. *Global Change Biology*, 26(March), 1–20. <https://doi.org/10.1111/gcb.15254>
- Negrón-Juárez, R. I., Holm, J. A., Faybishenko, B., Magnabosco-Marra, D., Fisher, R. A., Shuman, J. K., et al. (2020). Landsat near-infrared (NIR) band and ELM-FATES sensitivity to forest disturbances and regrowth in the Central Amazon. *Biogeosciences*, 17(23), 6185–6205. <https://doi.org/10.5194/bg-17-6185-2020>
- Ogden, F. L., Crouch, T. D., Stallard, R. F., & Hall, J. S. (2013). Effect of land cover and use on dry season river runoff, runoff efficiency, and peak storm runoff in the seasonal tropics of Central Panama. *Water Resources Research*, 49(12), 8443–8462. <https://doi.org/10.1002/2013WR013956>
- Piponiot, C., Derroire, G., Descroix, L., Mazzei, L., Rutishauser, E., Sist, P., & Hérault, B. (2018). Assessing timber volume recovery after disturbance in tropical forests—A new modelling framework. *Ecological Modelling*, 384(July), 353–369. <https://doi.org/10.1016/j.ecolmodel.2018.05.023>
- Powell, M. J. (1992). The theory of radial basis function approximation in 1990. *Advances in Numerical Analysis*, 2, 105–210. <https://doi.org/10.1093/oso/9780198534396.003.0003>
- Powell, T. L., Koven, C. D., Johnson, D. J., Faybishenko, B., Fisher, R. A., Knox, R. G., et al. (2018). Variation in hydroclimate sustains tropical forest biomass and promotes functional diversity. *New Phytologist*, 219(3), 932–946. <https://doi.org/10.1111/nph.15271>
- Purves, D. W., Lichstein, J. W., Strigul, N., & Pacala, S. W. (2008). Predicting and understanding forest dynamics using a simple tractable model. *Proceedings of the National Academy of Sciences of the United States of America*, 105(44), 17018–17022. <https://doi.org/10.1073/pnas.0807754105>

- Rüger, N., Condit, R., Dent, D. H., DeWalt, S. J., Hubbell, S. P., Lichstein, J. W., et al. (2020). Demographic trade-offs predict tropical forest dynamics. *Science*, 368(6487), 165–168. <https://doi.org/10.1126/science.aaz4797>
- Trugman, A. T., Detto, M., Bartlett, M. K., Medvigy, D., Anderegg, W. R. L., Schwalm, C., et al. (2018). Tree carbon allocation explains forest drought-kill and recovery patterns. *Ecology Letters*, 21(10), 1552–1560. <https://doi.org/10.1111/ele.13136>
- Tyukavina, A., Hansen, M. C., Potapov, P. V., Stehman, S. V., Smith-Rodriguez, K., Okpa, C., & Aguilar, R. (2017). Types and rates of forest disturbance in Brazilian Legal Amazon, 2000–2013. *Science Advances*, 3(4), 1–16. <https://doi.org/10.1126/sciadv.1601047>
- Wang, W., Akhtar, T., & Shoemaker, C. A. (2023). Efficient multi-objective optimization through parallel surrogate-assisted local search with tabu mechanism and asynchronous option. *Engineering Optimization*, 56(7), 1081–1097. <https://doi.org/10.1080/0305215X.2023.2219610>
- Wohl, E., Barros, A., Brunsell, N., Chappell, N. a., Coe, M., Giambelluca, T., et al. (2012). The hydrology of the humid tropics. *Nature Climate Change*, 2(9), 655–662. <https://doi.org/10.1038/nclimate1556>
- Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., et al. (2004). The worldwide leaf economics spectrum. *Nature*, 428(6985), 821–827. <https://doi.org/10.1038/nature02403>
- Wu, J., Serbin, S. P., Xu, X., Albert, L. P., Chen, M., Meng, R., et al. (2017). The phenology of leaf quality and its within-canopy variation is essential for accurate modeling of photosynthesis in tropical evergreen forests. *Global Change Biology*, 23(11), 4814–4827. <https://doi.org/10.1111/gcb.13725>
- Xia, W., Shoemaker, C., Akhtar, T., & Nguyen, M. T. (2020). Efficient parallel surrogate optimization algorithm and framework with application to parameter calibration of computationally expensive three-dimensional hydrodynamic lake PDE models. *Environmental Modelling & Software*, 135, 104910. <https://doi.org/10.1016/j.envsoft.2020.104910>
- Xu, C., Christoffersen, B., Robbins, Z., Knox, R., Fisher, R. A., Chitra-tarak, R., et al. (2023). Quantification of hydraulic trait control on plant hydrodynamics and risk of hydraulic failure within a demographic structured vegetation model in a tropical forest (FATES—HYDRO V1. 0). *Geoscientific Model Development*, 16(21), 6267–6283. <https://doi.org/10.5194/gmd-16-6267-2023>
- Ye, K. Q., Li, W., & Sudjianto, A. (2000). Algorithmic construction of optimal symmetric Latin hypercube designs. *Journal of Statistical Planning and Inference*, 90(1), 145–159. [https://doi.org/10.1016/s0378-3758\(00\)00105-1](https://doi.org/10.1016/s0378-3758(00)00105-1)
- Zeng, X. (2001). Global vegetation root distribution for land modeling. *Journal of Hydrometeorology*, 2(5), 525–530. [https://doi.org/10.1175/1525-7541\(2001\)002<0525:gvrdf>2.0.co;2](https://doi.org/10.1175/1525-7541(2001)002<0525:gvrdf>2.0.co;2)
- Zhang, J., Bruijnzeel, L. A., Quiñones, C. M., Tripoli, R., Asio, V. B., & van Meerveld, H. J. (2019). Soil physical characteristics of a degraded tropical grassland and a “reforest”: Implications for runoff generation. *Geoderma*, 333, 163–177. <https://doi.org/10.1016/j.geoderma.2018.07.022>
- Zitzler, E., Brockhoff, D., & Thiele, L. (2007). The hypervolume indicator revisited: On the design of pareto-compliant indicators via weighted integration. *Lecture Notes in Computer Science*, 4403. https://doi.org/10.1007/978-3-540-70928-2_68