ABSTRACT

Title of Dissertation:	ADVANCED MODELING USING LAND-USE HISTORY AND REMOTE SENSING TO IMPROVE PROJECTIONS OF TERRESTRIAL CARBON DYNAMICS
	Lei Ma, Doctor of Philosophy, 2021
Dissertation directed by:	Professor George Hurtt, Department of Geographical Sciences

Terrestrial carbon dynamics are an important component of the global carbon cycle. Quantifying, attributing, and projecting terrestrial carbon dynamics can provide valuable information in support of climate mitigation policy to limit global warming to 1.5 °C. Current modeling efforts still involve considerable uncertainties, due in part to knowledge gaps regarding efficient and accurate scaling of individual-scale ecological processes to large-scale dynamics and contemporary ecosystem conditions (e.g., successional states and carbon storage), which present strong spatial heterogeneity. To address these gaps, this research aims to leverage decadal advances in land-use modeling, remote sensing, and ecosystem modeling to improve the projection of terrestrial carbon dynamics at various temporal and spatial scales. Specifically, this research examines the role of land-use modeling and lidar observations in determining contemporary ecosystem conditions, especially in forest, using the latest land-use change dataset, developed as the standard forcing for CMIP6, and observations from both airborne lidar and two state-of-the-art NASA spaceborne lidar

missions, GEDI and ICESat-2. Both land-use change dataset and lidar observations are used to initialize a newly developed global version of the ecosystem demography (ED) model, an individual-based forest model with unique capabilities to characterize fine-scale processes and efficiently scale them to larger dynamics. Evaluations against multiple benchmarking datasets suggest that the incorporation of land-use modeling into the ED model can reproduce the observed spatial pattern of vegetation distribution, carbon dynamics, and forest structure as well as the temporal dynamics in carbon fluxes in response to climate change, increased CO_2, and land-use change. Further, the incorporation of lidar observations into ED, largely enhances the model's ability to characterize carbon dynamics at fine spatial resolutions (e.g., 90 m and 1 km). Combining global ED model, land-use modeling and lidar observation together can has great potential to improve projections of future terrestrial carbon dynamics in response to climate change and land-use change.

ADVANCED MODELING USING LAND-USE HISTORY AND REMOTE SENSING TO IMPROVE PROJECTIONS OF TERRESTRIAL CARBON DYNAMICS

by

Lei Ma

Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park, in partial fulfillment of the requirements for the degree of Doctor of Philosophy
2021

Advisory Committee:
Dr. George Hurtt, Chair
Dr. Shunlin Liang
Dr. Joseph Sullivan
Dr. Ritvik Sahajpal
Dr. Lesley Ott
Dr. Benjamin Poulter

Dedication

To my parents

Fengsheng Ma (马凤生) *and Dongmei Zhang* (张冬梅)

For their unconditional love and support

Acknowledgements

Pursuing PhD abroad is not an easy journey for me. Looking back at these years at University of Maryland, I feel so lucky to being accompanied by tons of encouragement, support and help which make these years a precious, joyful and memorable journey in my life.

First, I'd like to thank my advisor Dr. George Hurtt. Your tireless passion for science, critical thinking and profound and extensive knowledge has inspired me to devote to be a scientist. Thanks for your guidance and mentorship over years, and also for believing in me and offering me opportunities to collaborate with many brilliant researchers from various fields.

I'd like to thank my committee members, Dr. Lesley Ott, Dr. Benjamin Poulter, Dr. Ritvik Sahajpal, Dr. Shunlin Liang and Dr. Joseph Sullivan. Your insights and expertise have improved this dissertation from various perspectives. Thanks for funding support from NASA Carbon Monitoring System projects led by Dr. George Hurtt and Dr. Abhishek Chatterjee, NASA Interdisciplinary Science project led by Dr. Lesley Ott, and the Ann G. Wylie Dissertation Fellowship from UMD Graduate School.

Thanks to all current and former GEL lab mates, Maosheng Zhao, Justin Fisk, Steve Flanagan, Wenli Huang, Katelyn Dolan, Jennifer Kennedy, Zhen Zhang and many others, for your previous work that make this research possible. Special thanks to Louise Chini, for your assistance and advice with publishing Chapter 2. To Ritvik Sahajpal for all your guidance and suggestions on modeling and programing, and for your patience and encouragement during every discussion at GEL lab. To Hao Tang, for our numerous brainstorming discussions on interesting science questions. I am glad to have you as a close friend at the same lab who is always very positive to my ideas and progress. To Donal Seán O'Leary, for your encouragement as my 'academic sibling' and affability to international students like me. To Rachel Lamb, for your endless help and empathy, for encouragement to apply for fellowships, and for collaborations and efforts on this dissertation and. It is my honor to share the journey with you.

Thank my friends, Yanjia Cao, Yuhan Rao, Yuan Zhou, Jiaying He, Liangping Sun, Zheng Liu, Diyang Cui, Aolin Jia, Junchuan Fan, Xiaoxue Zhou and many others, who helped me and added much joy to these years. Special thanks to Moying Li and Yao Li, for being encouraging and supportive to me. Special thanks to Jiaming Lu, for your unconditional support, care and companionship on every aspect. Your enthusiasm in life makes me stay positive and mentally healthy when I was drowning in work. Also, thanks for bringing 'Carbon' (our super cute cat) to our home, who makes our work-from-home days during Covid a lot more joyful.

Finally, I'd like to thank my parents, Fengsheng Ma and Dongmei Zhang, who always stand behind me and support me to stay in school for as long as I want. Your love has always been and will always be my endless source of passion to pursue my dreams.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAD	Average absolute difference
AGB	Aboveground biomass
ATLAS	Advanced Topographic Laser Altimeter System
CAMS	Copernicus Atmosphere Monitoring Service
CHM	Canopy Height Model
CLM	Community Land Model
CMIP	Coupled Model Intercomparison Project
CSP	Carbon sequestration potential
CSPG	Carbon sequestration potential gap
CSPTG	Carbon sequestration potential time gap
DGVM	Dynamic Global Vegetation Model
ED	Ecosystem Demography
ESA CCI	European Space Agency Climate Change Initiative
ESM	Earth System Model
FAO	Food and Agriculture Organization
FATES	Functionally Assembled Terrestrial Ecosystem Simulated
FIA	Forest Inventory and Analysis
HYDE	History Database of the Global Environment
HWSD	Harmonized World Soil Database
GCB	Global Carbon Budget
GEDI	Global Ecosystem Dynamics Investigation
GFC	Global Forest Change
GFED	Global Fire Emissions Database
GHG	Greenhouse gas
GLAS	Geoscience Laser Altimeter System
GLC	Global Land Cover
GLCC	Global Land Cover Characteristics
GLM2	Global Land-use Model 2
GPP	Gross primary productivity
IBM	Individual-based model
ICESat-2	Ice, Cloud, and land Elevation Satellite-2
IPCC	Intergovernmental Panel on Climate Change
ISS	International Space Station
LAI	Leaf area index
LEAF	Land Ecosystem Atmosphere Feedback
Lidar	Light detection and ranging
LUH	Land Use Harmonization
LULCC	Land-use and land-cover change
MERRA2	Modern-Era Retrospective analysis for Research and Applications, version 2
MODIS	Moderate Resolution Imaging Spectroradiometer
MRV	Monitoring, reporting and verification
MSTMIP	Multi-scale Synthesis and Terrestrial Model Intercomparison Project
MvG	Mualem-van Gehuchten
NAIP	National Agricultural Imagery Program
NASA	National Aeronautics and Space Administration

NLCD	National Land Cover Dataset
NOAA CT	National Oceanic and Atmospheric Administration CarbonTracker
NBP	Net biome production
NPP	Net primary productivity
PDE	Partial differential equations
PFT	Plant functional type
POLARIS	Probabilistic Remapping of SSURGO
RGGI	Regional Greenhouse Gas Initiative
RH	Relative height
SAS	Size- and Age-Structure
SIF	Sun-induced fluorescence
TCCF	Tree Cover Continuous Fields
USCA	United State Climate Alliance
US EPA	United States Environmental Protection Agency

Chapter 1 Introduction

1.1 Background and Motivation

Terrestrial ecosystems play a fundamental role in global carbon dynamics, storing about 2000 ~ 3000 Pg C in vegetation and soil and uptaking one-third of anthropogenic fossil fuel and cement emissions over the last decade (Ciais et al., 2014; Friedlingstein et al., 2020). A key challenge in terrestrial carbon cycle science is quantifying and attributing of terrestrial sink. There are several factors that are believed to jointly contribute to terrestrial sink. Specifically, elevated atmospheric CO2 concentrations could enhance the photosynthetic rate of terrestrial ecosystems (Hickler et al., 2008; Keenan et al., 2016; Schimel et al., 2015). Nitrogen deposition could further support this enhanced productivity by meeting the nitrogen demand of plants, particularly in regions where nutrients are limited (Finzi et al., 2007, Thomas et al., 2010). Climate change could lengthen the growing season of temperate forests, with the potential for increased carbon uptake (Friedl et al., 2014). Regrowth over secondary forests could sequester carbon until an equilibrium state is reached (Hurtt et al., 2002; Pugh et al 2019; Williams et al., 2012). However, attributing the terrestrial sink to each of these factors involves considerable uncertainties, which in turn affect projections of future carbon dynamics under climate change and resulting climate mitigation policy.

These uncertainties are partly related to a lack of accurate information on initial forest conditions, particularly regarding forest age and contemporary carbon storage. Historical human activities have largely altered natural ecosystems in terms of plant structure, age, and species composition and these alterations vary in location, time, and magnitude (Hurtt et al., 2011). For example, nearly 60% of global land has been impacted by human land use, and

92.9% of forest land in the United States is secondary (Hurtt et al., 2020). As a result, contemporary forests are dynamic mosaics of stand patches with different ages resulting from prior human disturbance events. In addition, carbon stocks and rates of carbon sequestration vary strongly with forest age (Law et al., 2004), emphasizing the spatial heterogeneity of carbon stocks and fluxes.

Another source of uncertainty relates to the inaccurate representation of demographic processes within ecosystem models. Ecosystem establishment involves many processes at the individual plant scale, including photosynthesis, reproduction, resource competition, gap formation and post-disturbance recovery and population dynamics. However, these processes have not been fully considered in most of the current and widely used Earth System Models (ESMs) and Dynamic Global Vegetation Models (DGVMs) due to computational challenges. In each grid-cell, current ESMs and DGVMs abstract each plant function type as a big-leaf canopy plant with associated area and depict disturbance impact as area changes in big-leaf canopies. This simplification fails to characterize structure and age changes caused by disturbances and complicates post-disturbance competition between tall and short individuals (Fisher et al., 2018).

Recent advances in land-use modeling, remote sensing technology, and ecosystem modeling provide opportunities to improve understanding of both contemporary conditions and scaling of individual-scale process across broad spatial domains. To obtain current information on initial forest conditions, two approaches are presently available. The first approach is to spin up a ecososytem model with land-cover change information. Considerable efforts have been made to model global historical land-use history (Hurtt et al., 2006, 2011, 2020). The newly developed global land-use transition dataset (LUH2) includes comprehensive information about deforestation, reforestation, cropland expansion, and shifting cultivation, and incorporates constraints from satellite-based land cover and forest change. LUH2 has great potential to generate information on historical land-cover changes and improve characterization of land-use impacts on forests (Hurtt et al., 2020). In addition, advances in lidar remote sensing provide by far the most accurate measurements of forest structure over large spatial domains (Drake et al., 2002; Dubayah and Drake, 2000; Huang et al., 2019; Tang et al., 2012, 2021). For example, the USGS 3D Elevation Program (3DEP) provides freely available airborne lidar data across the United States (USGS 2019). Two recent NASA lidar missions, GEDI and ICESat-2, also provide spaceborne observations of forest structure at the global scale (Dubayah et al., 2020a; Neuenschwander and Pitts 2019). These lidar observations can be used to approximate contemporary forest age and carbon storage for ecosystem models. One such model, the Ecosystem Demography (ED) model, has been continuously developed over the last two decades to improve understanding of forest carbon dynamics (Albani et al., 2006; Fisk et al., 2013, Flanagan et al., 2016; Hurtt et al., 1998, 2002; Moorcroft et al., 2001). ED model is an individual-based prognostic ecosystem model which integrates submodules of growth, mortality, hydrology, carbon cycle and soil biogeochemistry. Regional studies have demonstrated ED's advantages in mechanistically simulating plant competition for light, water and nutrients, and efficiently scaling the physiological processes of individual plants to ecosystem scales (Hurtt et al 2002, 2004, 2010, 2016, Fisk et al 2013, Flanagan et al 2019). ED has also been used with lidar data to establish current forest conditions and project fine scale sequestration potentials at the state and regional scales (Hurtt et al., 2019b). All of these advances together provide an opportunity to improve carbon modeling at the global scale.

1.2 Research objectives

The overarching goal of this research is to improve projections of terrestrial carbon dynamics by integrating advances and opportunities from land-use modeling, remote sensing and ecosystem modeling. This research examines the role of these advances in improving our understanding of current forest initial conditions and assesses the global performance of an individual-based ecosystem model. To do so, I have set four specific objectives described in Chapters 2, 3, 4, and 5 respectively. The first is to generate a land-cover change history between 850-2015 by identifying a translation rule to translate land-use change history to land cover. The second is to develop, calibrate, and evaluate a global version of the ED model with land-cover history as input to spin-up the model to contemporary conditions. The third is to integrate airborne lidar observations and global ED into projections of future forest carbon dynamics at the regional scale. The fourth is to explore the potential of spaceborne lidar observations in ED initialization at the global scale.

To achieve these four objectives, the corresponding research questions are as follows:

1) What is historical land-cover change, and how could it be determined from a land-use change dataset and constraints of contemporary forest cover and biomass?

2) How can global ecosystem modeling be improved to incorporate advances of land-use history and remote sensing?

3) How can advanced modeling improve projections of future carbon sequestration with remote sensing?

4) What is the potential of spaceborne lidar observations to improve baseline estimates of forest carbon in models?

1.3 Dissertation Outline

This research is presented in five chapters. Specifically, Chapter 2 identifies a translation rule to generate land-cover change history from the latest land-use change dataset (LUH2). Chapter 3 develops and calibrates a global version of the ED model and evaluates simulations of carbon dynamics, vegetation distribution, and structure by spinning up with land-cover history. Chapter 4 develops a regional forest carbon modeling system by integrating the global ED model and airborne lidar observations and provides spatially-explicit estimates of baseline forest carbon and future carbon sequestration potential. Chapter 5 develops a global ED initialization approach utilizing spaceborne lidar observations from GEDI and ICESat-2 and evaluates resulting AGB estimates by forest inventory. Chapter 6 concludes with a summary of the major findings across all chapters and potential future research. Additional supporting figures, tables and analysis can be found in the Appendices.

Chapter 2 Global rules for translating land-use change (LUH2) to land-cover change for CMIP6 using GLM2

Abstract

Anthropogenic land-use and land-cover change activities play a critical role in Earth system dynamics through significant alterations to biogeophysical and biogeochemical properties at local to global scales. To quantify the magnitude of these impacts, climate models need consistent land-cover change time series at a global scale, based on land-use information from observations or dedicated land-use change models. However, a specific land-use change cannot be unambiguously mapped to a specific land-cover change. Here, nine translation rules are evaluated based on assumptions about the way land-use change could potentially impact land cover. Utilizing the Global Land-use Model 2 (GLM2), the model underlying the latest Land-Use Harmonization dataset (LUH2), the land-cover dynamics resulting from land-use change were simulated based on multiple alternative translation rules from 850 to 2015 globally. For each rule, the resulting forest cover, carbon density and carbon emissions were compared with independent estimates from remote sensing observations, U.N. Food and Agricultural Organization reports, and other studies. The translation rule previously suggested by the authors of the HYDE 3.2 dataset, that underlies LUH2, is consistent with the results of our examinations at global, country and grid scales. This rule recommends that for CMIP6 simulations, models should (1) completely clear vegetation in land-use changes from primary and secondary land (including both forested and non-forested) to cropland, urban land and managed pasture; (2) completely clear vegetation in land-use changes from primary forest and/or secondary forest to rangeland; (3) keep vegetation in land-use changes from primary non-forest and/or secondary non-forest to rangeland. Our analysis shows that this rule is one of three (out of nine) rules that produce comparable estimates of forest cover,

vegetation carbon and emissions to independent estimates and also mitigate the anomalously high carbon emissions from land-use change observed in previous studies in the 1950s. According to the three translation rules, contemporary global forest area is estimated to be 37.42×10^6 km^2, within the range derived from remote sensing products. Likewise, the estimated carbon stock is in close agreement with reference biomass datasets, particularly over regions with more than 50 % forest cover.

2.1 Introduction

Historical land-use activities have been significantly affecting the global carbon budget in both direct and indirect ways and changing Earth's climate through altering land surface properties (e.g., surface albedo, surface aerodynamic roughness and forest cover) (Betts, 2001; Bonan, 2008; Brovkin et al., 2006; Claussen et al., 2001; Feddema et al., 2005; Guo and Gifford, 2002; Pongratz et al., 2010; Post and Kwon, 2000). It has been estimated that, during the past 300 years, > 50 % of the land surface has been affected by human land-use activities, > 25 % of forest has been permanently cleared and $10-44\times10^6$ km^2 of land are recovering from previous human land-use disturbances (Hurtt et al., 2006). Impacts on the carbon cycle result from several of various other processes: deforestation removes natural forest, and its corresponding carbon biomass is used for wood products, burning or decay by microbial decomposition (DeFries et al., 2002). Afforestation/reforestation, in contrast, recovers forest which accumulates carbon, but sequestration potentials are constrained by water and nutrient availability (Smith and Torn, 2013). Wood harvesting is one of the largest sources, contributing gross carbon emissions by modifying the litter input into various soil pools, stand age and biomass of secondary forest (Dewar, 1991; Hurtt et al., 2011; Nave et al., 2010). Cumulatively, models estimate that land use and land-use change have contributed to a net flux of 205±60 Pg C to the atmosphere during 1850–2018 (Friedlingstein et al., 2019). While emissions from land use and land-use change only account for 10 % of current anthropogenic carbon emissions, they were a dominant contributor to increasing the atmospheric CO_2 above preindustrial levels before 1920 (Ciais et al., 2014).

Quantification of historical land-use and land-cover change (LULCC) is important because it serves as the basis for examining the role of human activities in the global carbon budget and the resulting impacts to Earth's climate system. For this purpose, LULCC reconstructions enter Earth system models (ESMs) (Lawrence et al., 2016), dynamic global vegetation models (DGVMs) (Friedlingstein et al., 2019) and bookkeeping models (Hansis et al., 2015) to quantify biogeochemical and biophysical impacts of historical land-use change as part of historical simulations (DECK and CMIP6 historical simulations), future projections (scenarioMIP), impacts studies (ISIMIP), paleoclimate studies (PMIP), land-use-specific simulations (LUMIP) and biodiversity studies (IPBES). Considerable efforts have been devoted to modeling historical land-use states (Klein Goldewijk et al., 2017; Kaplan et al., 2009; Pongratz et al., 2008; Ramankutty and Foley, 1999) and land-use transitions (Houghton, 1999; Hurtt et al., 2006, 2011). In particular, the recent Land-Use Harmonization 2 (LUH2) dataset (Hurtt et al., 2020) has been developed to provide global gridded land-use states and transitions in a consistent format for use in ESMs as part of CMIP6 experiments. However, large uncertainties still exist in the carbon/climate studies based on many of the above LULCC products (Chini et al., 2012; Houghton et al., 2012; Pongratz et al., 2014). For example, the global carbon budget reports that the spread of cumulative LULCC carbon emissions during 1850–2018 estimated by DGVMs are as large as 60 Pg C though all models are forced by the LUH2 (Friedlingstein et al., 2019). LULCC carbon emissions in CMIP5 have an anomalous spike during the years 1950–1960. These anomalous emission estimates by ESMs (hereinafter referred to as the “pasture anomaly”) are caused by an implausible high conversion rate of primary and secondary vegetation to pasture, with the 1950s having double the conversion rate of the 40s or 60s. Because of this, the simulated terrestrial land flux has a 2-decade delay in the switch from a land carbon source to a land carbon sink compared to observations (Shevliakova et al., 2013).

Standardization of LULCC data is critical for CMIP6 to simplify intercomparison of the ESMs and facilitate model analysis. The CMIP6 requires the LUH2 as standard land-use input for all ESMs; however, the data standardization could be undermined if models implement the LUH2 differently, such as applying different rules to translate the LUH2 into land-cover change, which is essential for models. Identifying the consistent rules between models for the LUH2 use is critical for two reasons. First, although land-use changes are generally associated with a change in land cover and carbon stocks (see Figure 1 in Pongratz et al., 2018), these two changes are not always equivalent, and the degree of land-cover alteration varies with the types of land-use changes and the location where land-use changes happen. An inconsistent land-cover translation from the same land-use products will potentially produce variance in land-cover dynamics across models and in turn impact the land surface biophysical and biochemical processes. Second, the HYDE 3.2 data underlying LUH2 has redefined the former pasture category used in CMIP5 into the two subcategories of "managed pasture" and "rangeland" (with the total being termed "grazing land"). This redefinition intends to mitigate the pasture anomaly by suggesting different treatments of vegetation and carbon removal in models for these two types of land-use changes (Klein Goldewijk et al., 2017). However, explicit suggestions are not yet provided for land cover resulting from these newly defined land-use types. Therefore, a consistent rule across models for the LUH2 translation is needed, with potential to reduce impacts of inconsistent LUH2 usage on studying land-use effects through CMIP6.

To recommend a rule for translating historical land-use changes from the LUH2 for CMIP6 models, this study investigates the impacts of land-use change on land cover by proposing several alternative sets of translation rules, which are then integrated into the Global Land-use Model 2 (GLM2) (Hurtt et al., 2019a, 2020) to simulate the forest cover and carbon dynamics. These simulations are then evaluated against estimates of contemporary forest cover and carbon density from remote sensing observations, and the resulting cumulative LULCC carbon emissions are compared with a range of independent estimates.

2.2 Methodology

In this study, two key land-cover properties (i.e., forest cover and vegetation carbon) are simulated by combining historical land-use change with translation rules. The historical land-use change information is specified by the LUH2 dataset (v2h, available at https://doi.org/10.22033/ESGF/input4MIPs.10454),which serves as the forcing data for a new generation of advanced ESMs as part of CMIP6. Section 2.2.1 describes the details of land-use change characterization, and Section 2.2.2 defines each translation rule. The resulting forest cover and vegetation carbon is tracked in each grid cell (0.25°×0.25°) for the years 850 to 2015 using methods described in Section 2.2.3 and 2.2.4. The simulated forest cover and vegetation carbon are then compared with multiple published datasets of land-cover, carbon stock and estimates of land-use change emissions (see details in Section 2.2.5).

2.2.1 Land-use change characterization

The LUH2 dataset was generated with the GLM2 (Hurtt et al., 2019a, 2020), which, like its predecessors (Hurtt et al., 2006, 2011), estimates annual sub-grid-cell land-use states and transitions by including multiple constraints such as gridded patterns of historical land use from the HYDE database (Klein Goldewijk et al., 2017), historical national wood harvest reconstructions, potential biomass and recovery rates, and others. Building upon previous work from CMIP5, for which the original LUH1 dataset was used, the LUH2 has extended the time span to 850–2100 and increased spatial resolution to 0.25°×0.25°. In addition, the LUH2 includes 12 different land-use types (i.e., forested and non-forested primary and secondary land, cropland of C3 annual, C3 perennial, C4 annual, C4 perennial and C3 nitrogen-fixing, urban, managed pasture, and rangeland) and includes transitions between all combinations of these categories.

In the LUH2, "primary" refers to land previously undisturbed by any human activities since 850, while "secondary" refers to land undergoing a transition or recovering from previous human activities. Global secondary land area was specified as zero in 850. Note that primary and secondary lands are further subdivided into forested and non-forested grids using a definition based on the potential aboveground biomass density (forested land requiring an aboveground biomass density ≥ 2 kg C m^{-2}).

2.2.2 Translation rules

Nine translation rules are proposed (Table 2.1) to analyze the effects of land-use change on land-cover dynamics, whereby each rule differs in treatment of vegetation cover and vegetation carbon stock during land-use changes. Rules 1–4 all assume complete clearance of vegetation for cropland and vary on vegetation clearance for managed pasture and rangeland. Rules 5–9 are added for analytical purposes, rather than as realistic possibilities. For example, Rule 3 presumes all land-use changes alter land cover and reduce carbon stock, and this rule would produce the least global forest cover and carbon stock. Rules 1 and 3 differ in treatment of vegetation in non-forested land when converted to rangeland, and the resulting difference between their carbon stocks indicate the impact of rangeland expansion on non-forests and also tests whether the disaggregation of grazing land into managed pasture and rangeland will address the pasture anomaly issue in 1950–1960. Rule 1 (clearance of all vegetation for cropland and managed pasture; and only forest clearance for rangeland) is in fact the rule suggested in the underlying HYDE dataset and its distinction between pasture and rangeland (Klein Goldewijk et al., 2017). For simplicity, we do not consider partial removal of vegetation in this study; vegetation is either fully removed or fully remains as these land-cover transitions represent the maximum and minimum bounds for land-cover alteration. In this study, the translation rules are applied to all regions and are constant across the whole simulation period. Although the impacts of land-use change on land cover may vary in different regions, the discussion of region-varied and time-varied translation rules is beyond the scope of this study.

It is important to note that these nine rules are not equally realistic, and the purpose of including Rules 5–9 is to investigate individual or joint contributions of cropland, managed pasture and rangeland expansion on forest and carbon. For example, forest and carbon dynamics resulting from Rule 6 could suggest the individual impact of cropland expansion.

2.2.3 Simulation of land-cover change

In this study, land-cover change is simulated by performing a modified GLM2 simulation in which the computed land-use transition rates (using the same methodology as LUH2) are supplemented with a set of translation rules (Table 2.1) to track forest cover change and carbon dynamics at 0.25º spatial resolution. Note that the modified GLM2 still generate and track the exact same land-use transitions of the LUH2 and has additional function to track associated land-cover change in terms of forest cover and vegetation carbon. GLM2 uses a statistical model to estimate ecosystem stocks and fluxes with temperature and precipitation as inputs (see (Hurtt et al., 2002) for details). The annual temperature and precipitation maps from MSTMIP (Wei et al., 2014) were averaged over 1901 and 2000 to generate the spatially varied and temporally static climatological temperature and precipitation, which was then used to spin up the GLM2 globally at 0.25x0. 25˚ resolution for 500 years. The climatology stays as constant over the spin up period, and other environmental factors were not taken into consideration such as CO_2 fertilization, nitrogen limitation and climate variability.

When land is converted to cropland, managed pasture, and/or rangeland, each translation rule indicates that vegetation in primary and secondary may be cleared or remain intact as the result of land-use changes. For example, for a given land-use transition rate from forest to pasture, if the applied translation rule indicates to clear the vegetation completely, then the resulting grid cell vegetation fraction in forest land-use type is reduced equal to the amount of pasture gained. If the rule indicates not to clear vegetation, then only the land-use type will be changed to pasture and the vegetation area will be unchanged, but the vegetation will be influenced by the management in terms of stand age/biomass, which are assumed to cease growing due to pressure from subsequent human management. If this pasture land is further converted to other non-primary and non-secondary land (e.g. cropland, rangeland or urban), the vegetation remaining from previous forest-pasture conversion then will be totally cleared. Therefore, the vegetation fraction existing within the cropland, managed pasture, rangeland and urban of each grid-cell can be tracked via the following equation:

$$f(i,t+1) = f(i,t) + f^{\,gained}(i,t) - f^{lost}(i,t), (i = 5,6,7,8)\ , \qquad (2.1)$$

Where $f(i,t)$ is the fraction of grid-cell that is vegetated in land-use type *i* (i.e., classes 5-8: cropland, managed pasture, rangeland, urban) at time *t*, $f^{gained}(i,t)$ and $f^{lost}(i,t)$ are gained and lost vegetation fractions respectively. The vegetation fraction could only be gained in land-use change from primary and secondary land (both forested and non-forested), and be lost in land-use change to any other land-use types except forested and non-forested primary land.

Transition Rule	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Rule 8	Rule 9
->Crop	X	X	X	X	X	X	O	O	O
->Managed pasture	X	F	X	X	O	O	X	X	O
->Rangeland	F	F	X	O	X	O	X	O	X

$$f^{gained}(i,t) = \sum_{j=1}^{4} a_{ij}(1-\gamma_{ij})\,, (i = 5,6,7,8; j = 1,2,3,4)\,, \quad (2.2)$$

$$f^{lost}(i,t) = \frac{f(i,t)}{l(i,t)} \sum_{k=1,k\neq i}^{8} a_{ki}\,, (i = 5,6,7,8; k = 3,4,\cdots,8)\,, \quad (2.3)$$

The possible values of i, j and k are 1, 2, … , 8 representing primary forested land, primary non-forested land, secondary forested land, secondary non-forested land, cropland, managed pasture, rangeland and urban respectively. a_{ij} is the land-use transition fraction estimate by LUH2 from land-use type j (i.e., primary forested land, primary non-forested land, secondary forested land, secondary non-forested land) to land-use type i, γ_{ij} represents the translator factor to convert land-use change to land-cover change, it equals to 1 if the translation rule in Table 2.1 indicates an 'X' or 'F' for this land-use change. For example, γ_{ij} is 1 for land-use change from primary land (forested, non-forested grids) to cropland in Rules 1 and 2, but 0 for the same type of change in Rules 8 and 9. This translator factor is 1 for all types of land-use change in Rule 3 since all vegetation is cleared during all land-use changes. $l(i,t)$ is the land-use fraction estimate by LUH2 for type i at time t, and this fraction is larger than or equal to its vegetation fraction $f(i,t)$.

Table 2.1 Rules for vegetation clearance during cropland, pasture and rangeland expansion. 'X' indicates complete removal of vegetation if the primary and secondary land state is altered. 'O' indicates no vegetation removal when land-use change occurs. 'F' indicates that vegetation is only removed if the preceding land cover is forested primary or forested secondary land.

Vegetation in primary and secondary land can remain or be lost in land-use changes to cropland, pasture or rangeland depending on translation rules. According to the definition of primary land in the LUH2, its transition to other land-use types is unidirectional, thus primary land could not gain vegetation from any land-use changes. Wood harvest on primary land will result in vegetation loss and a change of land-use type to secondary land, but harvest on secondary land will not change the land-use type. Furthermore, vegetation in secondary land could be gained from harvest on primary land and may be gained through the process of abandonment of cropland, pasture or rangeland depending on translation rules. Note that reforestation but not afforestation is also considered in this study. The former is to re-establish forest on the land which has been forested before, while the latter is an anthropogenic activity to establish forests on land which has never been forested. Thus, the vegetation of primary and secondary land is tracked by the following equations:

$$f(i,t+1) = f(i,t) - f^{lost}(i,t) + f^{gained}(i,t), (i = 1,2,3,4) \ , \tag{2.4}$$

$$f^{lost}(i,t) = \begin{cases} \sum_{j=5}^{8} a_{ji} + b_i, (i = 1,2; j = 5,6,7,8) \\ \sum_{j=5}^{8} a_{ji}\gamma_{ji} \qquad , (i = 3,4; j = 5,6,7,8) \end{cases} \tag{2.5}$$

$$f^{gained}(i,t) = \sum_{k=5}^{8} \frac{f(k,t)}{l(k,t)} a_{ik} + b_j, (i = 3,4; j = 1,2;\ k = 5,6,7,8) \tag{2.6}$$

Where $f(i,t)$ is fraction of vegetation at land-use category i (primary forested land, primary non-forested land, secondary forested land, secondary non-forested land) at time t. a_{ji} is land-use transition fraction from primary and secondary land to cropland, managed pasture, rangeland and urban in LUH2. b_i or b_j is wood harvest fraction from primary or secondary (forested or non-forested) land. $f(k,t)$ and $l(k,t)$ are vegetation fraction and land-use fraction in land-use type k (i.e., cropland, managed pasture, rangeland, urban), and a_{ik} is land-use transition due to land-use abandonment.

2.2.4 Simulation of vegetation carbon dynamics

Vegetation carbon stocks fluctuate through releasing and accumulating carbon in response to natural growing conditions, disturbances, and anthropogenic land-use changes, which can vary widely in terms of their carbon impacts. For land-use changes associated with clearing or harvesting vegetation, the forest biomass is either released immediately (e.g. burning) or stored in soil pools or as timber products (both of which eventually decay over decades). However, when managed land is abandoned and allowed to recover, the vegetation takes up CO_2 from the atmosphere through photosynthesis, resulting in increasing carbon stocks in vegetation and possibly soils. The magnitude of each of these bi-directional carbon flows ultimately determine if the land is a net carbon sink or carbon source. In this study, the temporal dynamics of carbon fluxes after land-use change are simplified, with all biomass (above- and below-ground) being released instantaneously to the atmosphere. Note that the biomass stock change is a rough proxy of actual net land-use change fluxes, for which delayed emissions from litter and soil carbon and product pools needed to be accounted for as well as instantaneous emissions from burning biomass. Changes in soil carbon associated with loss of vegetation biomass are usually associated with carbon losses, but are likely less important than biomass changes, as are net fluxes from product pool changes (Erb et al., 2018).

Similar to land-cover change simulation in section 2.2.3, if translation rules indicate vegetation clearing at expansion of cropland, managed pasture, rangeland or urban land, vegetation biomass is totally released as a carbon emission, and its age is set as zero. If vegetation is not cleared based on translation rules, the biomass remains but ceases to increase, and the age of this vegetation also remains unaffected, because the age is used in this model only for the calculation of biomass density. Keeping age fixed corresponds to keeping biomass from further growing, which represents the influences of management. If the land is abandoned and converted back to secondary land, a mean age is calculated over all vegetation with different ages, then the mean age increases year by year and biomass regrows towards equilibrium. Thus, the biomass density in secondary vegetation at time *t* is calculated for each grid cell using its mean age, potential biomass, and potential NPP:

$$B(t) = B_0\left(1 - e^{-NPP_0 \times G(t)/B_0}\right), \tag{2.7}$$

Where $B(t)$ is the aboveground biomass density of vegetation at secondary land at time *t*, and B_0 is the potential aboveground biomass density from the GLM2 model and varied by grid location, and NPP_0 is the potential NPP of the wood fraction that is allocated to cumulate stem and branch biomass annually, and $G(t)$ is the mean age of secondary vegetation. Note that B_0 and NPP_0 are estimated by a statistical model in GLM2 using climatological temperature and precipitation and are spatially varied but temporally constant over simulation period of 850 to 2015. Above- to below-ground biomass ratio is assumed as 3:1 when converting aboveground biomass to total biomass (above- and belowground), and biomass density is converted to carbon by a ratio of 0.5.

Plants cultivated by human management (e.g. crops and orchards) are not tracked in this study; zero biomass is assigned to cropland, managed pasture, rangeland and urban use types.

However, carbon is tracked for vegetation remaining from primary or secondary due to the translation rules, as well as lands that convert from human management back to natural lands. Thus, the total carbon stocks in this study are expected to be lower than other estimates (Houghton, 2003; Saatchi et al., 2011), especially in the grids with a higher fraction of non-primary and non-secondary land-use.

2.2.5 Diagnostics for evaluating translation rules

To evaluate which translation rules best translate land-use changes to land-cover changes, the simulation results were compared with contemporary forest cover and carbon density maps from remote sensing observations and other estimates, as well as LULCC carbon emissions from other studies using different models. Contemporary values of forest cover and carbon density are used for two reasons. First is the lack of multiple diagnostics of forest cover and carbon density across the whole simulation period (i.e., 850 to 2015). Second is that contemporary values could potentially reflect cumulative error in converting land-use change to land-cover change since 850. We assume that if a translation rule produces a best match with the diagnostic maps of forest cover and carbon density, then it would also produce the best estimate for the historical period.

Diagnostics of contemporary forest cover consist of six widely used satellite-based land-cover and tree coverage datasets (Bartholomé and Belward, 2005; Bicheron et al., 2008; DeFries et al., 2000; Friedl et al., 2010; Hansen et al., 2013; Loveland et al., 2000) (see Table 2.2) and the Global Forest Resources Assessment (FRA) 2015 (FAO, 2015). In Table 2.2, GLC, GLC2000, GlobCover and MODIS LC are land-cover datasets rather than tree cover and were produced based on different classification schemes resulting in different land-cover legends. Prior to being used as diagnostics in this study, they needed further reclassification of their land-cover legends into a common representation of forest canopy cover at the same spatial resolution (0.25˚) by the following procedures: First, the GLCC, GLC2000, GlobCover and MODIS LC were converted to tree cover fraction based on Table A.1 at their native resolutions (Song et al., 2014). Then, all six datasets were resampled to 1 km resolution and translated to a binary (forest versus non-forest) map by applying a 30% tree-cover threshold (Sexton et al., 2016). Through counting the percentage of pixels marked as forest within each 0.25x0.25˚ grid cell, six global gridded forest cover maps at 0.25º spatial resolution were generated, and resulting global forest area of each dataset are shown in Table 2.2. As these satellite-based datasets were developed from different sensors (e.g., AVHRR, SPOT-4, MERIS, MODIS, Landsat) and models (regression trees, decision tree, clustering labels and random forests), an averaged map (hereinafter referred to as 'Averaged satellite-based forest cover') was generated in accompany with the six forest cover maps to examine spatial pattern of contemporary forest cover simulated by each translation rule. In addition, since FAO only reports national forest cover (not spatially explicit), these data were only used for comparison at the country level.

Table 2.2 Summary of land cover products used in this study including six satellite-based datasets and FAO FRA report.

Product	Global Forest Area (10^6 km^2)	Time	Publication	Data Type/Classification Scheme
GLCC	40.89	1992-1993	Loveland et al., 2000	Land Cover (IGBP)
GLC2000	38.22	1999-2000	Bartholome et al., 2005	Land Cover (GLC 2000)
GlobCover	35.66	2004-2006	Bicheron et al., 2008	Land Cover (GlobCover)
MODIS LC	41.05	2001	Friedl et al., 2010	Land Cover (IGBP)
1 Kilometer Tree Cover Continuous Fields (TCCF)	42.74	1992-1993	DeFries et al., 2000	Tree Percentage
Global Forest Change (GFC)	41.71	2000	Hansen et al., 2010	Tree Percentage
FAO	40.55	2000	FRA 2015	National Censuses

Carbon density maps are employed as the second metric to evaluate the translation rules. Two datasets were employed: the IPCC Tier-1 biomass carbon map for the year 2000 (Ruesch and Gibbs, 2008) and a pantropical biomass map (hereinafter referred to as the Baccini's product (Baccini et al., 2012). The former, a global above- and below-ground carbon density map, is created by dividing the globe into 124 carbon zones by land-cover, continental regions, eco-floristic zones, and forest age and assigning each zone a unique carbon stock value. The latter is estimated by combining ground plots, GLAS LiDAR observations and optical reflectance of MODIS. This dataset employs the empirical relationship between aboveground biomass and tree diameter at breast height and estimates aboveground biomass density for pantropical regions (40˚S-30˚N). Both carbon density maps were resampled to 0.25˚ before evaluation.

In addition, the ability of the translation rules to reproduce LULCC carbon emissions is also assessed. The estimates of LULCC carbon emissions were compiled from published papers

(Table 2.3) (Houghton, 2010; Houghton and Nassikas, 2017; Le Quéré et al., 2018; Pongratz et al., 2009; Reick et al., 2010; Shevliakova et al., 2009; Stocker et al., 2011). These studies have significant discrepancy in emissions estimates as they employed various methods (e.g., book-keeping methods and different process-based models), LULCC datasets, and considered different types of land-use change activities. They also differ in treatment of environmental change, for example, (Pongratz et al., 2009; Reick et al., 2010; Shevliakova et al., 2009; Stocker et al., 2011) include effects of evolving climate or atmospheric CO_2 concentration on LULCC emissions, which is not accounted for in bookkeeping model based studies (Houghton, 2010; Houghton and Nassikas, 2017). In this study, only the range of these estimates during the pre-industrial and industrial periods are chosen to evaluate the translation rules. We posit that the recommended translation rule should not produce anomalous carbon emissions that are outside the compiled range.

Table 2.3 Summary of carbon emissions due to LULCC from available studies at pre-industrial and industrial period.

Reference	Time span	Carbon Emissions (Pg C)	LULCC types
		Pre-industrial Period	
Reick et al., 2010 (bookkeeping model)	1100-1850	80	Cropland/Pasture Change
Reick et al., 2010 (DGVM)	1100-1850	47	
Pongratz et al., 2009	850-1850	53	Cropland/Pasture Change
Stocker et al., 2011	until 1850	69	Cropland/Pasture Change, Urban
		Industrial Period	
Houghton 2010	1850-2005	156	Cropland/Pasture Change, shifting cultivation in tropics, and wood harvest
Houghton and Nassikas, 2017	1850-2015	145	Cropland/Pasture Change, shifting cultivation in tropics, and wood harvest
Shevliakova et al.,2009	1850-2000	164 - 188	Cropland/Pasture Change, shifting cultivation in tropics, and wood harvest
Pongratz et al.,2009	1850-2000	108	Cropland/Pasture Change
Reick et al.,2010 (bookkeeping model)	1850-1990	153	Cropland/Pasture Change
Reick et al.,2010 (DGVM)	1850-1990	110	Cropland/Pasture Change
Stocker et al., 2011	1850-2004	164	Cropland/Pasture Change, Urban
Le Quéré et., 2018	1850-2014	195	Cropland/Pasture Change, shifting cultivation in tropics, and wood harvest

In summary, the GLM2-based estimates of forest cover and carbon density in the year 2000 and LULCC carbon emissions during the periods 850-1850 and 1850-2000, based on nine different translation rules are compared with the above three types of diagnostics (i.e., contemporary forest cover/area and carbon density maps, LULCC emissions). The final recommended translation rules should produce: 1) the forest cover with the smallest difference with diagnostic maps at global, country and grid scale, the total forest cover at global and country level should be comparable to the range of diagnostics, and spatial pattern should also be close to diagnostics; 2) the closest carbon density map compared to diagnostics with the smallest difference, comparable spatial pattern and total carbon stock as well; and 3) reasonable LULCC carbon emissions within the range from other diagnostic estimates and minimizing the anomalous emissions during 1950-1960.

2.3 Results

2.3.1 Potential forest cover and biomass carbon

The GLM2 estimates global vegetation carbon stock (including above- and belowground) in 850 as 718 Pg C, and the resulting potential biomass map is shown in Figure 2.1a. For comparison, global potential vegetation carbon stock was estimated as 557 Pg C in (Kucharik et al., 2000), 772 Pg C in (Pan et al., 2013) and 923 Pg C in (Sitch et al., 2003). Forested land in GLM2 is defined as land which has aboveground potential biomass of at least 2 kg C/m^2 (Hurtt et al., 2006, 2011). With this definition, global potential forest area was estimated as 47.82 million km^2, and the resulting potential forest cover map is shown in Figure 2.1b. For comparison, global potential forest area was estimated as 48.68 million km^2 in (Pongratz et al., 2008), and potential forests and woodlands area was 55.3 million km^2 in (Ramankutty and Foley, 1999).

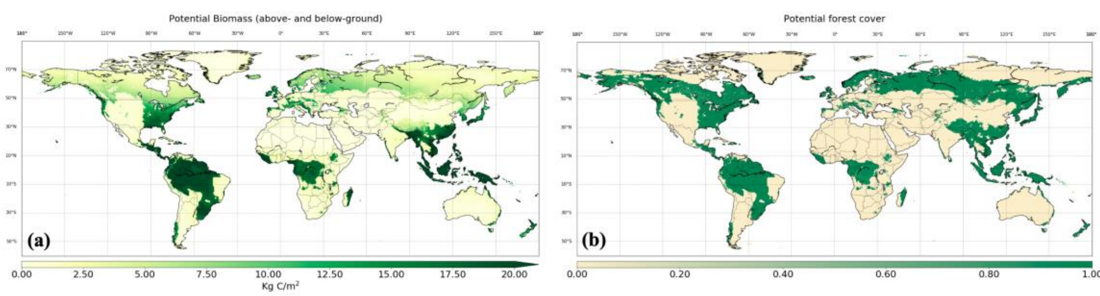


Figure 2.1 Potential biomass density (a) and potential forest cover (b) in 850 estimated by GLM2 model.

2.3.2 Forest cover evaluation

The global gridded forest cover maps resulting from Rules 1-9 in 2000 are generally consistent in forest extent with satellite-based observations (shown in Figure 2.2 and Figure A.6). For example, they all estimate high forest cover in tropical rainforests and northern boreal forests but low cover in Western USA, Eastern Europe and Central Asia. As Rules 1, 2, and 3 only differ in whether to clear vegetation and carbon in the conversion from non-forest to pasture or rangeland, the forest cover resulting from Rules 1, 2, and 3 are the same. All rules of 1-9 consistently estimate higher forest cover than the averaged satellite-based forest cover in West Siberia and South China, and lower forest cover in African savannas and East Siberia, Western Mexico and Argentina. Separately, Rules 4, 6, 7, 8 and 9 shows larger forest cover than Rules 1, 2, 3 and 5 in South and Southeast of Brazil and Tiber in China.

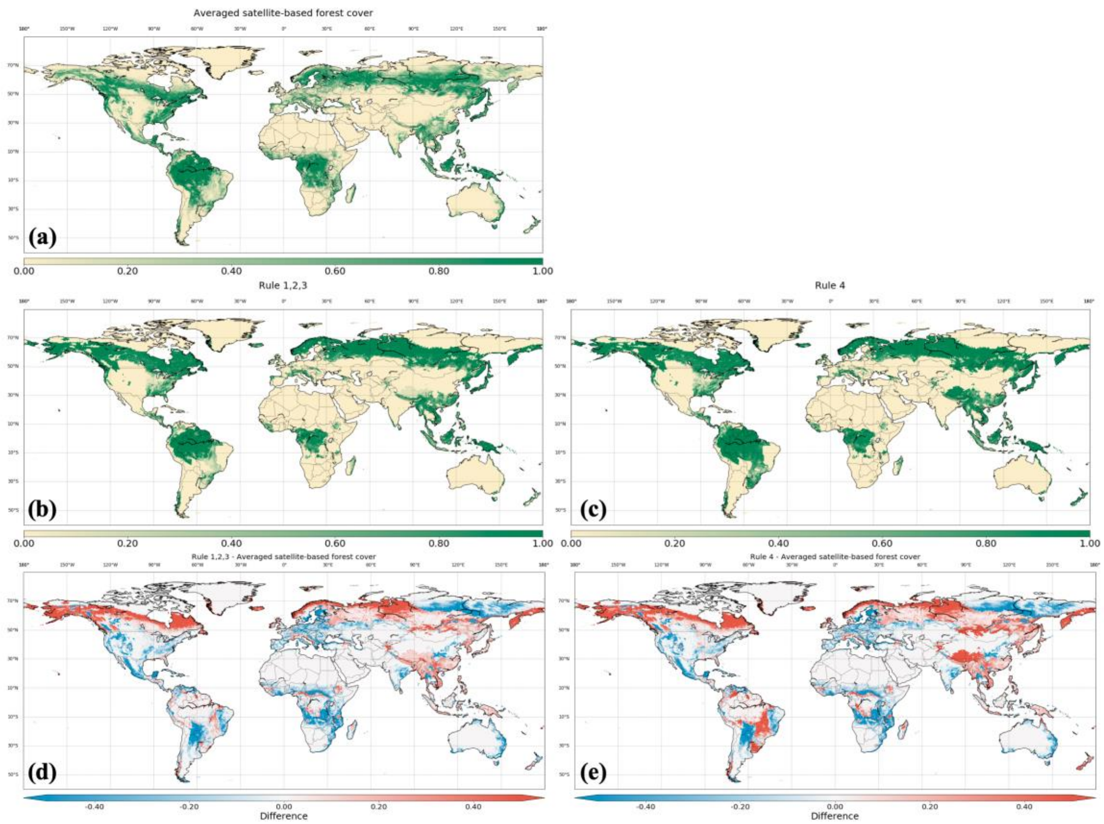


Figure 2.2 Forest cover in 2000 from the Averaged satellite-based forest cover in (a), Rule 1, 2, 3 in (b) and Rule 4 in (c). (d) and (e) are maps of forest cover difference between (b) and (a), and (c) and (a) respectively.

The total area of global forest in 850 amounts to 47.82 million km^2 according to the GLM2 model (Figure 2.1b and Figure 2.3a) when all forested lands were in a primary state by definition and decreased thereafter (Figure 2.3a). Forest loss has accelerated since the beginning of the Industrial Revolution and shows relatively high annual change rates (shown in Figure 2.3c). The translation rules produce a wide range of global forest cover in 2000 from 37.42 to 45.89 million km^2. In Rules 1, 2, and 3, the global forest is lost at the highest rate due to all land-use change activities on forested land resulting in the clearing of forest, and only 37.42 million km^2 of global forest is left in 2000 under these three rules. In contrast, under Rule 4 forest remains during rangeland expansion, and this would result in greater forest cover (e.g., 41.80 million km^2 in 2000, Table 2.4). The forest losses in Rules 6, 8, and 9 indicate the individual contribution of cropland, managed pasture and rangeland expansion. For example, rangeland and cropland expansion results in the most and second most of forest loss with an area of 4.34 million km^2 and 4.06 million km^2 respectively during 850-2000.

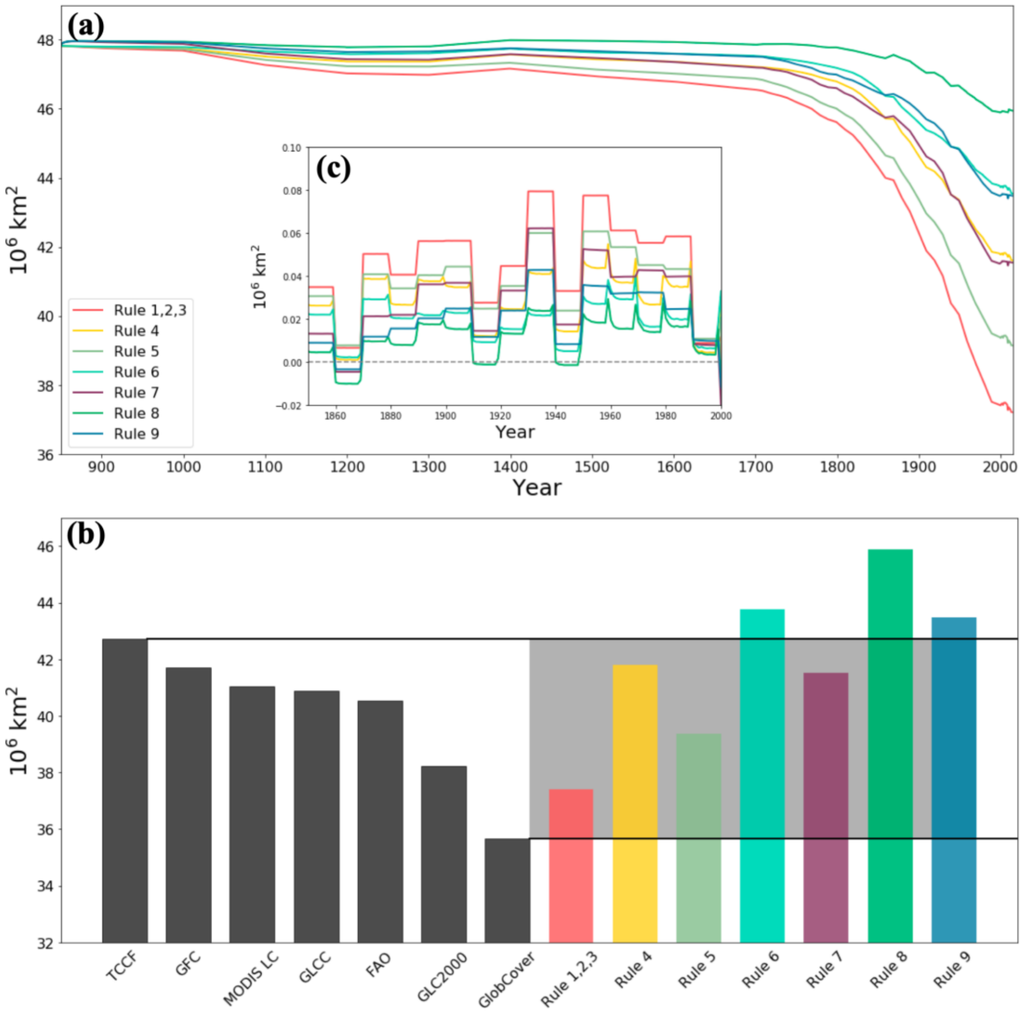


Figure 2.3 (a) Global forest area resulting from translation rules from 850 to 2015; (b) Comparison of global forest area in 2000 between remote sensing and FAO (shown as black bars) and results of translation rules (colored bars); (c) Annual change rate from 1850 to 2000. Positive values indicate forest loss.

Six satellite-based forest cover datasets and FAO data report the global forest area around the year 2000 ranging from 35.66 to 42.74 million km^2. One of major reasons underlying the discrepancy in global forest area is the difference in defining 'forest', particularly in the regions with intermediate tree cover (Sexton et al., 2016). The global forest area in the year 2000 resulting from the translation rules are compared to the range of seven diagnostic estimates (Figure 2.3b). The forest cover based on Rules 6, 8 and 9 is beyond the range of the diagnostics, indicating that these rules underestimate the impacts of land-use change on land-cover and overestimate the global forest existing in the present day. The excessive remaining forest cover in these three rules also rejects these rules' assumptions that only a particular type of land-use change would alter the land-cover. In contrast, Rules 1-4, 5 and 7 produced estimates of global forest area within the range of diagnostics.

The forest cover estimation from translation rules is further compared with diagnostic datasets at the country level (Table 2.4). In the diagnostic forest cover datasets, three-fourths of global forest cover lies within eight countries: the Russian Federation, Brazil, Canada, United of States of America, China, Democratic Republic of the Congo, Indonesia and Peru. The forest cover estimates from Rules 1-4 are generally well within the range of diagnostics. For example, 6 of 8 countries have estimates within the range for Rules 1, 2, and 3, and 5 of 8 countries for Rule 4. China and Brazil are the two countries where Rules 1-3 and Rule 4 have relatively larger difference between their estimates, the difference between Rules 1, 2, 3 and Rule 4 are 1.17 million and 1.08 million for China and Brazil respectively. Rule 5 and 7 overestimated forest area of China, Russian Federation and Canada though their global forest areas are within the range of diagnostic and are within range for Brazil, Democratic Republic of the Congo, Indonesia, and Peru.

Table 2.4 Forest area (10^6 km^2) in 2000 of eight countries with the largest forest area, and all other countries combined ('Others'), estimated by the 9 translation rules, range compiled from satellite-based datasets and FAO report.

Country	**Forest Area (10^6 km^2)**							**Range from satellite-based products and FAO**
	Rule 1, 2, 3	**Rule 4**	**Rule 5**	**Rule 6**	**Rule 7**	**Rule 8**	**Rule 9**	
Russian Federation	8.72	9.15	8.80	9.23	9.01	9.44	9.10	6.65-8.62
Brazil	4.61	5.69	4.89	5.96	5.05	6.12	5.33	4.19-5.92
Canada	5.59	5.63	5.59	5.64	5.76	5.81	5.77	3.27-4.36
United States of America	2.81	2.94	3.06	3.19	3.62	3.76	3.87	2.65-3.36
China	2.04	3.22	2.44	3.61	2.45	3.63	2.85	1.34-2.14
Democratic Republic of the Congo	1.57	1.61	1.60	1.64	1.63	1.67	1.66	1.57-2.11
Indonesia	1.30	1.33	1.36	1.38	1.58	1.60	1.64	0.99-1.64
Peru	0.76	0.78	0.78	0.80	0.77	0.79	0.79	0.69-0.79
Others	10.02	11.47	10.86	12.31	11.63	13.08	12.48	12.21-17.08
World	**37.42**	**41.80**	**39.38**	**43.76**	**41.52**	**45.89**	**43.48**	**35.66-42.74**

These comparisons evaluate the resulting gross forest cover of the translation rules at global and country level. Further examination at the grid level is also needed. Since the FAO report only provides national forest cover, the averaged satellite-based forest cover map and each of the six satellite-based forest cover maps were used to calculate the average of absolute difference across global grids (Figure 2.4) respectively. Rules 1, 2, and 3 consistently produce the smallest overall difference than Rule 4 and other rules regardless of which satellite-based forest cover is chosen as the reference. The average absolute difference (AAD) of Rule 1, 2, 3 is under 90 km^2 comparing to the averaged satellite-based forest cover map, and even smaller comparing to the GFC. The smallest difference of all rules across six reference forest maps indicates the GLC2 may have more similar spatial distribution to the GLM2 estimate. Regional comparison of average of absolute difference (Figure A.1) suggests Rules 1, 2, 3 give better estimate of forest cover at the north and south temperate zones (i.e., 60°N ~ 23°N and 23°S ~ 60°S) than tropical zone (23°N ~ 23°S). All rules have similar AAD at 60°N ~ 90°N zone.

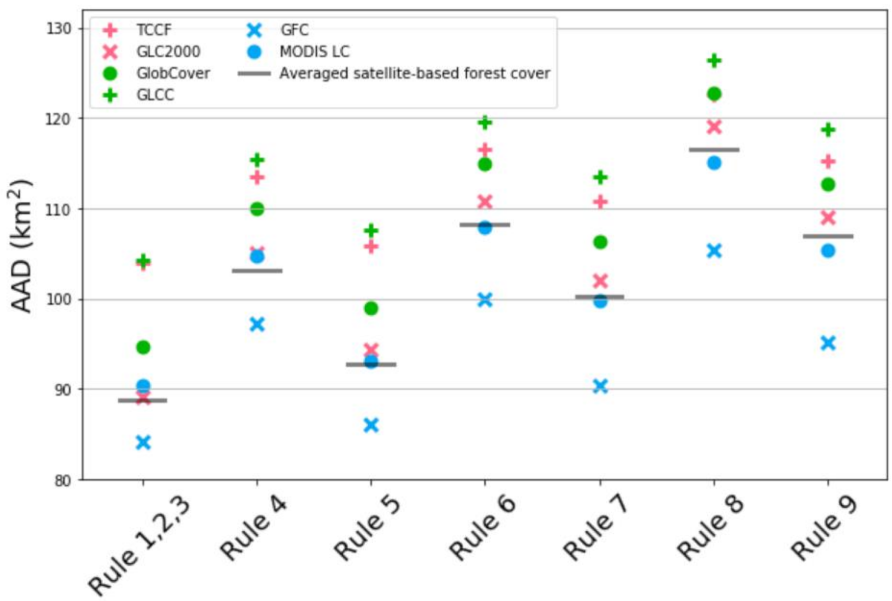


Figure 2.4. Global average of absolute difference in forest area between maps estimated by translation rules, and each of the six satellite-based forest cover maps as well as the averaged satellite-based forest cover map.

2.3.3 Evaluation of carbon dynamics

The net carbon emissions of the nine translation rules were calculated over two periods (850 to 1850 and 1850 to 2000) and compared to other studies (Table 2.5). Rules 1-4 produced similar patterns to other studies, specifically that global carbon emissions of 1850-2000 are twice as large as that of 850-1850. However, the emissions estimates of each period varied among Rules 1-4, from 55 to 77 Pg C during 850-1850 and from 142 to 185 Pg C during 1850-2000, due to the assumptions for clearing vegetation during land-use change. For example, Rule 3 produced the largest emissions as the carbon in both forested and non-forested land is released for all land-use changes, and Rule 1 produces fewer emissions since the vegetation is not cleared and carbon is not released when non-forested land is converted to rangeland. In general, Rule 1, 2, 3 and 4 estimated comparable emissions with other studies, while the emissions of the Rules 6-9 are out of range (Table 2.5).

Table 2.5 Summary of LULCC carbon emissions estimated by the 9 translation rules and those from other studies in Table 2.3.

Translation Rule	Carbon Emissions Estimation (Pg C)			Emission Range from Table 2.3		Estimation using LUH1
	850-1850	1850-2000	1950-1960	850-1850	1850-2015	1950-1960
Rule 1	72	175	20			
Rule 2	70	170	19			
Rule 3	77	185	22			
Rule 4	55	142	16			
Rule 5	63	146	17	47-80	108-195	26
Rue 6	41	104	11			
Rule 7	28	107	13			
Rule 8	5	65	7			
Rule 9	13	67	7			

Carbon emissions from pasture expansion were calculated for LUH1 (Hurtt et al., 2011) and this is used as a baseline to assess the improvement of translation rules on the pasture anomaly. Rules 1-4 estimate fewer emissions during this decade and decrease the anomaly between 4 to 10 Pg C. Rule 1 reduces anomalous emissions by 6 Pg C, indicating the sole contribution of the LUH2 to mitigate pasture anomaly. In LUH1, the anomalous emissions spike during 1950-1960 mainly arises from overestimating the emissions from pasture expansion, especially in three regions (i.e., Africa, East, South and Central Asia, and North America). The carbon flux from expansion of managed pasture and rangeland in LUH2 was reduced at global (Figure 2.5) and regional (Figure 2.6) scales in simulations based on Rules 1, 2, and 3. Note that the pasture land in LUH1 corresponds to rangeland and managed pasture together in LUH2. Rule 2 reduces more anomalous emissions than Rule 1 (reduced 6 Pg C in Rule 1 and 7 Pg C in Rule 2), because Rule 1 completely clears vegetation when transitioning to managed pasture, whereas Rule 2 only removes vegetation if the preceding land cover is primary or secondary forest.

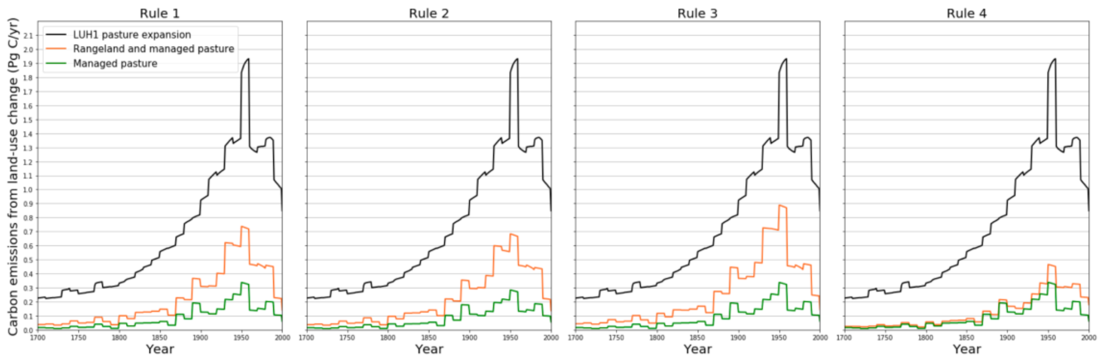


Figure 2.5. Carbon emission due to vegetation (forests and non-forests) removal in expansion of managed pasture and rangeland. Black line represents emissions from pasture expansion in LUH1. Orange and green lines represent emissions from expansion of managed pasture and rangeland and from expansion of just managed pasture respectively in LUH2. Note that the pasture category in LUH1 corresponds to managed pasture and rangeland together in LUH2.

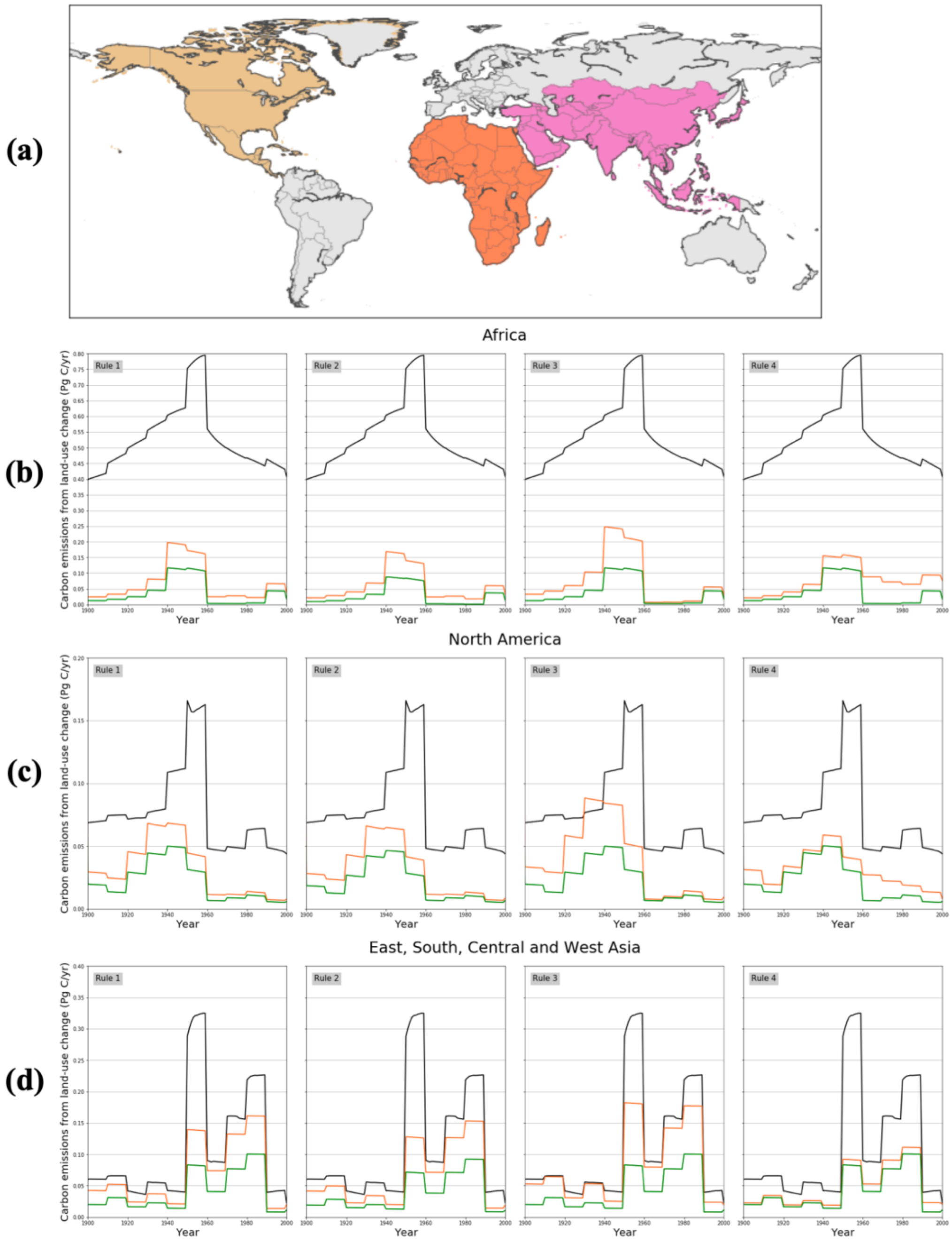


Figure 2.6. As in Figure 2.5 but three regions: (b) Africa; (c) East, South, Central and West Asia; (d) North America. (a) illustrates the defined boundaries of (b) - (d).

Rules 1-4 generally capture the spatial pattern that carbon density in tropical rainforest regions is much higher than northern boreal forests (Figure 2.7). These four rules overestimate carbon density at high latitudes of the Northern Hemisphere, in South China and in the Amazon rainforests but underestimate density across much of Sub-Saharan Africa, Mexico and the Southwestern part of the United States (Figure A.2 and Figure A.3). To further examine the spatial pattern of estimated carbon density, the estimates from all rules were compared to the carbon density maps of IPCC Tier-1 (above- and belowground) globally and the Bacchini's dataset (only aboveground) at the pantropical scale by calculating averaged absolute difference (Figure 2.8). According to this comparison, Rules 1-3 best capture the carbon density globally (Figure 2.8). Regional comparison of the IPCC Tier-1 biomass map and rule estimates indicate Rules 1-4 have comparable AAD of carbon density at the zone of 90°N ~ 60° N, the AAD difference between four rules is largest at 23°S ~ 60°S, followed by 23°N ~ 23°S and 23°N ~ 60°N (Figure A.4). Carbon density estimates of Rules 1-3 were further examined at regions where their estimates have difference (shown in Figure A.5a). The spatial pattern (Figure A.5c-A.5f) and histogram (Figure A.5b) of carbon density difference between rules and IPCC Tier-1 biomass estimates shows that all of these three rules underestimate carbon density and more grids are less underestimated in Rules 1-2 than Rule 3. The underestimation is expected because biomass of human cultivated vegetation is not tracked, and nor is growth of natural vegetation on cropland and pasture and rangeland. However, uncertainty level of the IPCC Tier-1 biomass should be taken into account when determining rule performance. Three bias levels of IPCC Tier-1 biomass map (i.e., ±10%, ±20% and ±30%) were considered (Figure A.5b). At these levels of uncertainty in the reference, Rules 1-3 could not be distinguished in performance. Finally, the carbon stock comparison between Rules 1-3 (Figure 2.9) shows these three rules underestimate carbon stock at low forest fraction, but give better agreement with diagnostics as forest fraction increases.

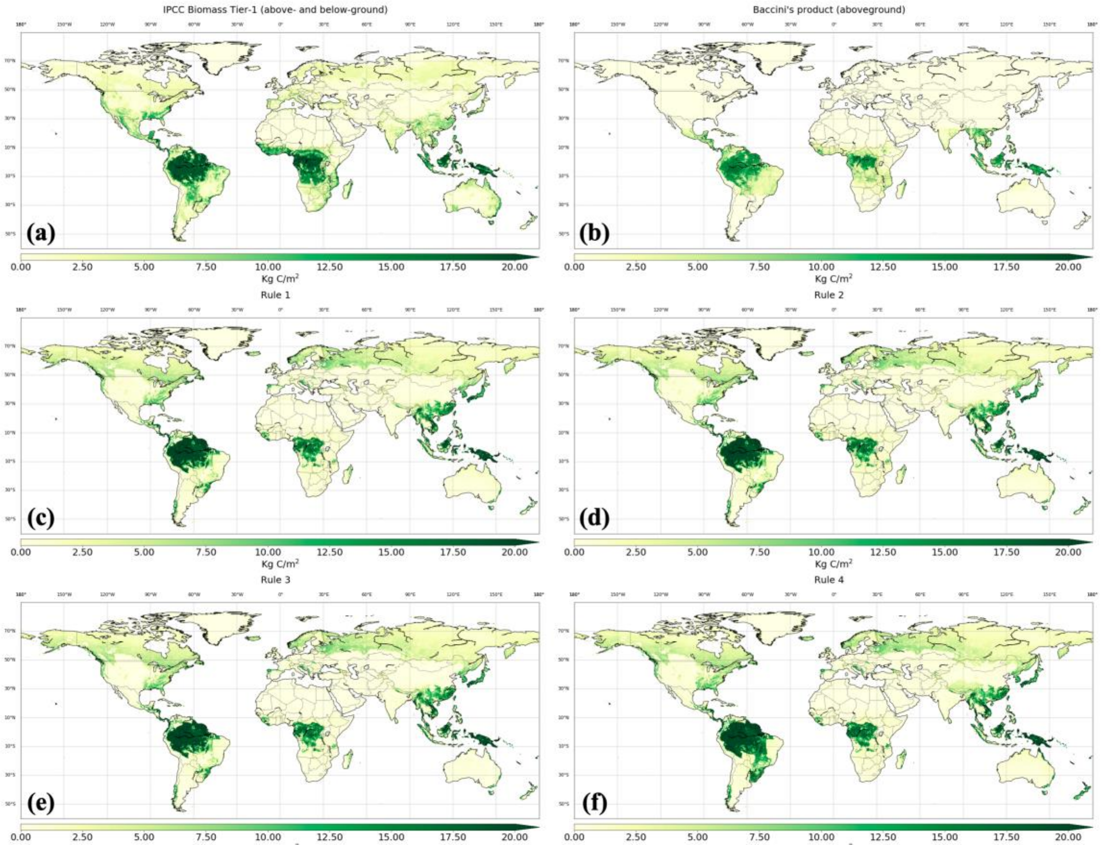


Figure 2.7. (a) IPCC Biomass Tier-1 density; (b) Baccini's product (only aboveground) at pantropical; global carbon density (above- and below-ground) maps estimated by Rules 1-4 from (c) to (f).

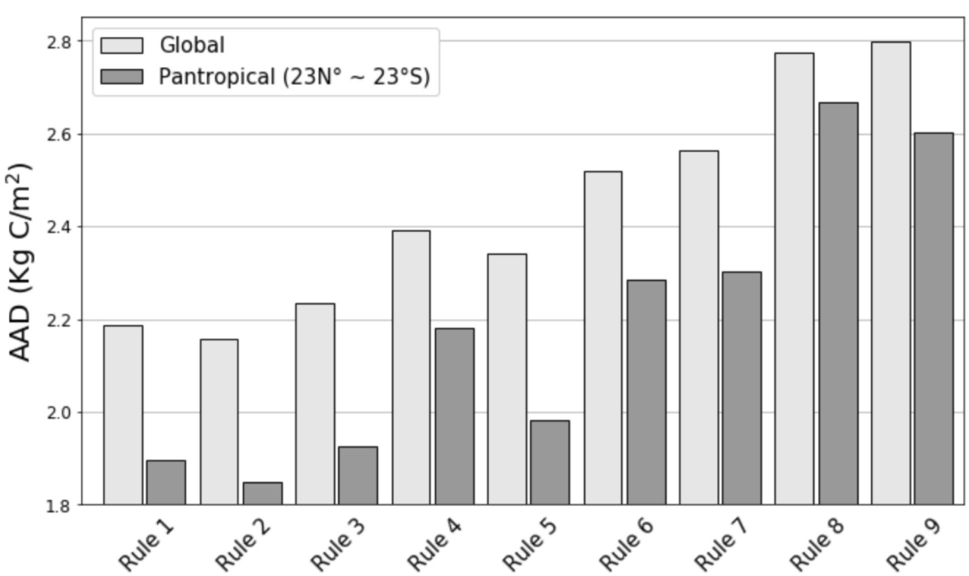


Figure 2.8. Average of absolute difference in carbon density between estimations of the 9 translation rules and two diagnostic maps: global comparison with IPCC Tier-1 biomass

density map (incl. above- and below-ground); tropical comparison with Baccini's carbon density map (only aboveground).

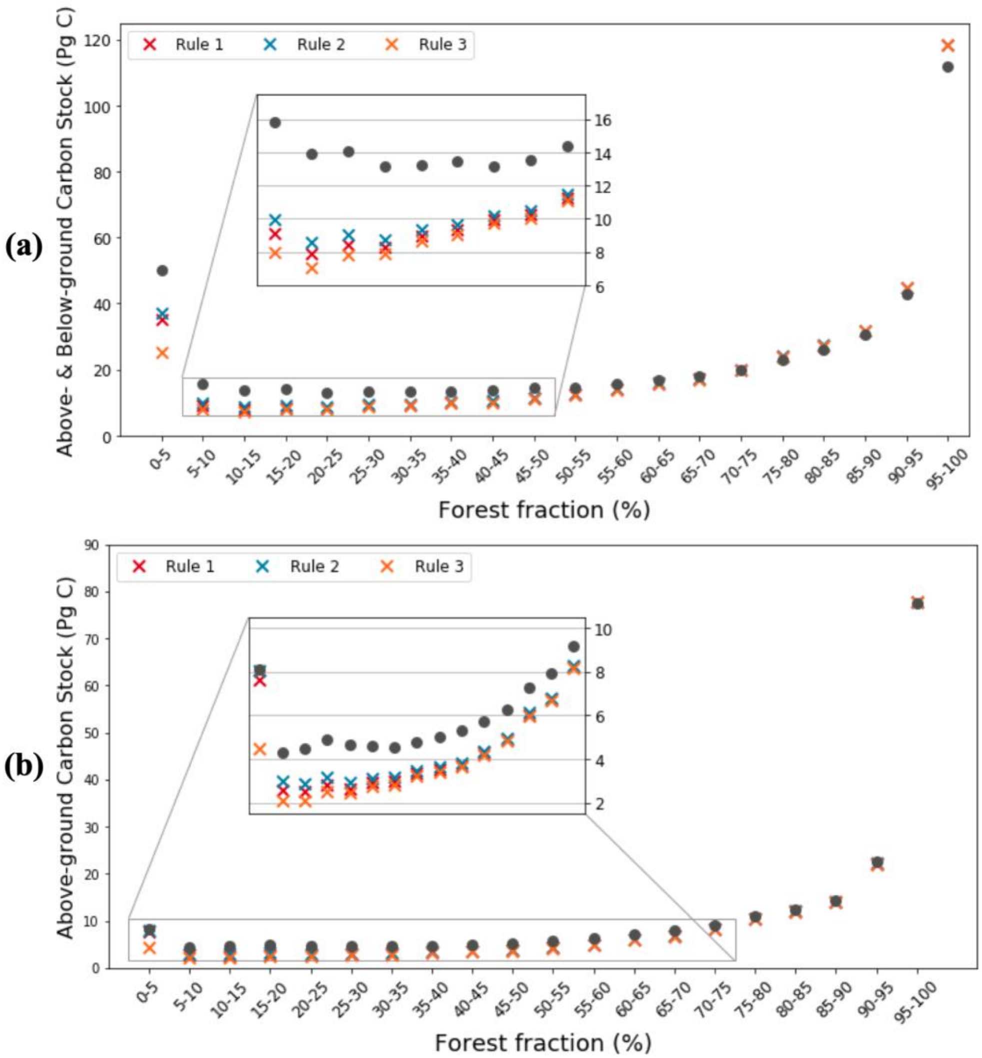


Figure 2.9. Total carbon stock grouped by forest fraction from the averaged satellite-based forest cover map. (a) global (above- and below-ground); (b) pantropical (aboveground).

2.4 Discussion and Conclusion

This study quantified the results of multiple alternative translation rules for estimating the potential effects of land-use change on land-cover utilizing the LUH2 dataset, and the underlying land model embedded in it (GLM2). The evaluations of forest cover and carbon indicate that Rules 1-3 on average and globally outperform other rules and are able produce the closest estimates of contemporary forest cover and carbon to diagnostics. The evaluations also confirm that prior recommendation of translation rule from HYDE 3.2 (Goldewijk et al., 2017) corresponding to the Rule 1 could produce comparable estimates of forest cover and vegetation carbon relative to diagnostics. Differentiation between Rules 1-3 depends largely on estimates of vegetation carbon because these rules produce equivalent estimates of forest cover. Comparisons of carbon stock and gridded difference in carbon density have shown that Rule 2 produces closer estimates of carbon density than Rules 1 and 3 relative to diagnostics. However, given underlying uncertainty of the carbon density reference map, the difference between Rules 1, 2 and 3 is small implying the differentiation of these rules is not possible in this study based on the difference alone.

A key feature of this study is to explicitly link land-use change and land-cover change and to provide insights into the consequences of choosing different land-use translation rules in ESMs. This study quantitively characterizes historical land-cover change using the same underlying model of the LUH2, namely the GLM2. Estimates of forest cover and vegetation carbon between translation rules could provide information about sensitivities of ESMs to the LUH2 implementation. For example, despite of same land-use transitions from the LUH2, Rules 1-4 still have a difference of 43 Pg C in LULCC emissions during 1850-2000. Such difference solely from land-use translation accounts for about 24% of the range of estimated vegetation carbon changes during 1850-2005 between CMIP5 models (Jones et al., 2013). Another feature is the relatively extensive evaluation of the LUH2 translation with multiple diagnostic datasets. The diagnostic datasets used in this study could serve to evaluate ESMs such as forest cover range at global and country level. Besides, this study also emphasizes the necessarity of improving vegetation carbon estimates, especially in regions with low forest cover or vegetation carbon in order to further differentiate translation rules.

In additional to the nine rules designed in this study, many other designs of translation rules are possible for LUH2 implementation in CMIP6 models such as spatially or temporally varied rules. It is important to note that the designed translation rules of this study are spatially and temporally constant meaning land-use changes at different regions or years will result in the same land-cover change for a given translation rule and given land-use transitions. This simplification may result in errors in land-use change translation because impacts of land-use change on land-cover could vary by regions and time. Combination of spatially/temporally varied rules and LUH2 may produce better estimates of forest cover and carbon density than these nine rules of this study. However, spatially/temporally varied translation rules will potentially add complexity to the LUH2 implementation in ESMs. Meanwhile, identification of such rules is sophisticated and also requires diagnostics with historical coverage. Uncertainties in these diagnostics should be small enough in order to differentiate various translation rules.

The estimated forest cover and carbon dynamics are subject to the several assumptions being made, the land-use change dataset being used, the land-cover properties being evaluated, reference datasets, and the models. This study used the LUH2 dataset because of its required used in CMIP6 and widespread used in other studies. The land cover properties addressed here include two critical variables (i.e., forest cover and carbon stock) due to their biophysical and biogeochemical significance. Multiple datasets based on remote sensing and other sources were selected for evaluation with the intention to provide a robust reference. The use of GLM2 model was selected to provide the most internally consistent treatment of these issues given its role in producing the LUH2 dataset. Given these considerations, it is possible that different results could be obtained for different systems. Although multiple of satellite-based land-cover datasets were included, they disagree the presence or absence of forest over low forest cover regions such as shrublands and semi-arid savannahs, and the discrepancies due to technical challenges and disagreement of forest definition. In addition, global vegetation carbon mapping is still challenging and uncertain mainly because of indirect proxies of biomass and paucity of in situ measurements and observations from space. Uncertainties in vegetation carbon diagnostics limit the evaluation of translation rules such as differentiation of Rules 1-3. Furthermore, dynamics of forest cover and vegetation carbon from past to present interact with climate change and increasing atmospheric CO_2, which are not considered in this study. Finally, the carbon emission estimates using the same translation rules and land-use change dataset may be different using other ESMs/DGVMs.

Future research is needed to investigate both the robustness of these findings, and potentially identify even better implementations. The CMIP6 LUMIP study is designed to quantify some of these effects (Lawrence et al., 2016) through model inter-comparison. Additional work on translation rules should include possible spatial/temporal varying rules, partial land clearing, and more land cover variables (e.g., forest age, height, soil carbon, energy balance) and focus on Rules 1-3 differentiation with better diagnostics such as annual land-cover maps from ESA Climate Change Initiative (CCI) (Lamarche et al., 2017) and lidar-based aboveground biomass from NASA's Global Ecosystem Dynamics Investigation (GEDI) mission (Dubayah et al., 2020a).

Acknowledgement

We gratefully acknowledge the support from DOE-SCIDAC DESC0012972, NASA-TE NNX13AK84A and NASA-IDS 80NSSC17K0348.

Chapter 3 Global Development and Evaluation of Ecosystem Demography model

Abstract

Terrestrial ecosystems play a critical role in the global carbon cycle and climate mitigation. Understanding and quantifying underlying ecological processes is essential for projecting future responses and feedbacks between the climate and terrestrial ecosystems. A new generation of vegetation models with a focus on demographic processes is important for meeting this need at the global scale. Here, we present the global development and evaluation of the Ecosystem Demography (ED) model, which features mechanistic competition between individual vegetation and formal scaling of the physiological processes of individual-based vegetation dynamics to ecosystem scales. Building upon series of regional versions of ED, we introduce several modifications such as refining plant functional types and leaf physiology and including product pools for wood harvesting. We collect a set of benchmarking datasets from remote sensing and field measurement to evaluate global ED performance with respect to simulating vegetation distribution, structure, and carbon stocks and fluxes across different temporal and spatial scales. Model evaluation suggests that global ED predicts well: (i) general patterns of broadleaf and needleleaf trees (dominance and coexistence); (ii) global total GPP, spatial variation, seasonal cycle and interannual trends; (iii) interannual variability of NBP in El Niño and La Niña events; and (iv) vertical structure of leaf area and global spatial pattern of canopy height.

3.1 Introduction

Terrestrial ecosystems and associated carbon cycle are of critical importance in providing ecosystem services and regulating global climate. Plants store approximately 450-650 Pg C as biomass globally and remove approximately 120 Pg C from the atmosphere each year through photosynthesis (Beer et al., 2010; Ciais et al., 2014). Human activities over past centuries have significantly impacted terrestrial ecosystems through biophysical and biogeochemical mechanisms including rising atmospheric CO_2, a warming climate, and alteration of structure, demography and functioning of ecosystems (Cramer et al., 2001; Walther et al., 2002; Brovkin et al., 2004; Pielke Sr. et al., 2011). Currently, terrestrial ecosystems are estimated to be a net carbon uptake of 1.9 ± 1.1 Pg C yr^{-1} for the past decade (2009-2018) (Friedlingstein et al., 2020). Quantification and attribution of the terrestrial carbon sink requires in-depth understanding of underlying ecological processes and their sophisticated responses and feedbacks to climate change, elevated CO_2, and land use and land cover change (LULCC) across multiple biomes and spatial and temporal scales (Canadell et al., 2007; Erb et al., 2013; Keenan and Williams, 2018). This demand has driven the emergence and development of dynamic global ecosystem models (DGVMs), which simplify the structure and functioning of global vegetation into several plant functional types and simulate vegetation distribution and associated biogeochemical and hydrological cycles with ecophysiological principles (Prentice et al., 2007; Prentice and Cowling, 2013). The first generation of DGVMs have been used successfully to address a variety of carbon cycle related questions and also integrated into Earth System Models (ESMs) (Cramer et al., 2001; Sitch et al., 2008). However, despite the broad applications, the first generation of DGVMs has been critiqued for their poor vegetation representation in a grid cell by several homogeneous patches of plant functional types (PFTs). These models lack representation of competition between plant individuals and explicit characterization of demographical processes (e.g., recruitment and mortality) (Quillet et al., 2010). These limitations may increase uncertainties arising from ecosystem demographics in projection of future ecosystem dynamics and responses and feedbacks to climate (Huntingford et al., 2008; Purves and

Pacala, 2008; Friend et al., 2014; Friedlingstein et al., 2014; McDowell et al., 2020). New generation of vegetation models is in demand with a focus on demographics and fine-scale heterogeneity at individual-level (Scheiter et al., 2013; Fisher et al., 2018).

Individual-based models (IBMs) or a special category called forest gap models have a long history as tools to understand and predict long-term dynamics of ecosystem succession, structure and composition by scaling up from individual-level processes (Botkin et al., 1972; Shugart and West, 1977; Urban et al., 1991; Pacala et al., 1996; Köhler and Huth, 1998; Bugmann and Solomon, 2000; Fischer et al., 2016; Shugart et al., 2018). In IBMs, each individual plant is explicitly simulated for its fate throughout the life cycle as functions of local resources (light, water and nutrients) and competition between neighbour plants (Shugart et al., 2018). Plant starts from seedling, grows in size (e.g., height, diameter and biomass), share resources with other plant and also modifies local resource availability. Larger plants dominate canopy and shade surrounding smaller plants, death of larger plants forms canopy gaps and a new light regime. Plant establishment and mortality processes are stochastically determined. Species-specific parameterizations and functions determine plant characteristics in terms of birth, dispersal, growth, shade tolerance and death as well survivor strategy. Because of representing plant at individual scale, IBMs can usually use forest inventory measurements to initialize and parameterize the models. Forest gap models have been intensively developed and applied in a variety of forests over globe, the model family has now grown to hundreds of models from early twelve pioneer models (Shugart et al., 2018). However, applications of forest gap models are limited to local and regional domains because of intensive computation demand. The inherent stochasticity nature of IBMs requires multiple realizations to obtain mean behaviour of the models, and each realization has to handle hundreds of thousands of plant individuals in a landscape-level simulation. Such computational complexity is quite challenging for continental and global applications, particularly when coupling with other sophisticated models of plant physiology, biophysics and hydrology.

In response to this challenge, the Ecosystem Demography (ED) model has been developed with an economic approach (Size- and Age-Structured, SAS) to approximate forest gap models (Hurtt et al., 1998; Moorcroft et al., 2001). In ED, plants are discretized into patches according to succession age (years since the last disturbance), and plants within the same patch are further discretized into cohorts according to size (e.g., height or biomass). Hence, all individuals in a cohort are treated as identical, modeling entity becomes cohort instead of each individual as forest gaps models do. Moreover, in ED, establishment and mortality increase or reduce individual density of a cohort rather than stochastically add or remove individuals. This SAS approximation largely reduce computational complexity, allowing ecological processes at individuals scale to be efficiently scaled up ecosystem scale dynamics. In addition to increasing computation efficiency, ED retain the ability to capture fine-scale processes: it tracks individual growth, recruitment, mortality, competition, and recovery from disturbance, and capture spatial heterogeneity resources (e.g., light, water and nutrient); it explicitly tracks forest structure (e.g., vertical leaf area, canopy height, etc) along with ecosystem succession.

Since its emergence, the original ED has been continuously developed and applied at various regions and spatial scales by different research groups. The original model currently has branched into three general derivatives as summarized in Fisher et al., 2018. One branch is the ED2 which was started from Medvigy (2006) and Medvigy et al., (2009) and subsequently developed to recent version 2.2 (Longo et al., 2019a, 2019b). The ED2 branch incorporates canopy and soil biophysical scheme from the Land Ecosystem Atmosphere

Feedback (LEAF-2) to solve carbon, water and energy cycle at patch- and cohort-level. The implementation of biophysics gives rise to sub-grid abiotic heterogeneity which is important to track short-term fluxes of CO_2, water and energy. This branch also includes regional development such as empirically constraining phenology scheme for temperate and tropical regions (Kim et al., 2012; Jeong and Medvigy, 2014); introducing trait-driven plant hydraulic scheme for tropical regions (Xu et al., 2016); data-constrained parametrizations for north America ecosystems (Medvigy and Moorcroft, 2012). ED2 also has been used to predict resilience of tropical forests to climate change (Zhang et al., 2015; Levine et al., 2016; Longo et al., 2018) and response of temperate forest succession to elevated CO_2 (Miller et al., 2016); More details of current ED2 and the summary of sequent developments could be found in Longo et al., (2019a). Another branch is the FATES (Functionally Assembled Terrestrial Ecosystem Simulated) model, which was started from (Fisher et al., 2015). This branch introduces ED concept (i.e., SAS approximation) into the Community Land Model version 4.5 (CLM4.5) and also merged other modifications (Fisher et al., 2015). FATES model alters the original land surface representation in CLM4.5 by replacing PFT-based tiling structure with cohort and patch based tiling structure. FATES has been used to predict biome boundaries in eastern North America and to explore its sensitivity plant trait variations.

The final branch is ED, which remain the original naming in subsequent developments. This branch ED mainly focus on characterization of land-use driven demographic dynamics and direct connection of ecosystem successional stages to forest structure observations. Land-use characterization was introduced by Hurtt et al., (2002), which allows ED to track additional sub-grid heterogeneity associated with different types of land use and the transitions between types. With this development, ED can simulate long-term ecosystem dynamics with effects of historical land use activities (e.g., deforestation, harvesting, shifting cultivation and reforestation). In addition, Hurtt et al., (2004) proposed to directly link ED simulations with forest structure observations from remote sensing. This development uses canopy height from lidar measurements as proxy to infer contemporary successional stages of ecosystems. Several subsequent studies have carried out to demonstrate the efficacy of lidar remote sensing in improving estimates of carbon fluxes and vegetation carbon stocks (Hurtt et al., 2010, 2016, 2019b; Ma et al., 2021). Both these two developments advances ED in the way to obtain contemporary demographic conditions, which is critically important to quantify contribution of demographic processes to terrestrial carbon sink and project associated future carbon sequestration in future (Hurtt et al., 2019b; Ma et al., 2021). Other applications and developments include investigation of net impacts of tropical cyclones on the carbon balance of eastern US between 1851 and 2000 (Fisk et al., 2013), and prediction of spatially-explicit plant migration in response to climate change (Flanagan et al., 2019). Currently, this branch ED has integrated submodules from leaf physiology at rapid temporal scales, phenology, growth, reproduction and mortality at intermediate temporal scales, and vegetation composition, soil biogeochemical cycles and LULCC, ED can be used to predict large-scale ecosystem dynamics and associated carbon and nitrogen cycles to the ecosystem and community levels, capturing responses to environmental changes including natural disturbance (e.g., storms, fire and etc), climate change, rising CO_2 and LULCC. This branch ED has been used by NASA Carbon Monitoring System as the tool for high spatial resolution (e.g., 90 m) regional forest carbon modeling and monitoring (Hurtt et al., 2019b; Ma et al., 2021) and also by NASA Global Ecosystem Dynamics Investigation mission for quantification of land carbon sequestration potential (Dubayah et al., 2020a; Ma et al., 2020b).

While numerous studies have utilized the branch of ED model at various spatial scales and regions, none have done so globally. The global development and evaluation has become increasingly important in the context of: studying the roles of demographic processes in the global terrestrial carbon cycle; developing global carbon modeling and monitoring system by leveraging satellite observations from ongoing lidar missions (GEDI and ICESat-2) (Dubayah et al., 2020a; Markus et al., 2017); identifying model uncertainties and limitations of global scale applications and prioritizing future developments in need. Therefore, in this study, we develop the global version of ED model by building upon previous regional versions of the last branch ED and introducing several modifications. Specifically, we have designed an experimental protocol of model spin-up to contemporary ecosystem conditions by taking account climate change, rising CO_2 and land use history. We then evaluate global ED's simulation of vegetation distribution, carbon fluxes and stocks, and vertical vegetation structure against benchmarking datasets from remote sensing observations and field measurements. Our purpose here is not only to provide an overview of model performance, but also insights into the next stage of model development. We will summarize the core principles of the global ED model and newly introduced modifications in this study in section 3.2. We will then describe our experiment design, model simulation, and benchmarking datasets for evaluations in section 3.3. Finally, we will discuss current strengths and limitations of global ED and future development needed in sections 3.4 and 3.5, respectively.

3.2 Model Development

Global ED is built upon series of previous developments (Moorcroft et al., 2001; Hurtt et al., 2002; Albani et al., 2006; Fisk, 2015; Flanagan et al., 2019). To extend ED's capabilities globally, several modifications are introduced to capture global vegetation distribution across biomes and related carbon stocks and fluxes. This section starts with introduction of ED principle and dynamic characterization and then summarizes the modifications. The full descriptions of each submodules in ED can be found in the Appendix B.

3.2.1 ED Principle

ED characterizes ecosystem heterogeneity using an SAS approach which consists of two Partial Differential Equations (PDE) to capture vertical and horizontal heterogeneity in resource availability and vegetation structure. The SAS approach represents vegetation by a hierarchical structure of patches and cohorts as illustrated in Figure 3.1, a given grid/site is separated into patches according to age since the last disturbance or LULCC, and a patch is further separated into cohorts according to plant size (e.g., height and biomass). In the SAS approach, horizontal heterogeneity is captured by tracking patch demography, patch dynamics in terms of ageing and disturbance are depicted by a PDE:

$$\frac{\partial}{\partial t}p_i(a,t) = -\frac{\partial}{\partial a}p_i(a,t) - \lambda_i(a,t)p_i(a,t) - \sum_j \lambda_{i,j}(a,t)p_i(a,t) \tag{3.1}$$

Where $p_i(a,t)$ is proportion (or area) of a patch in grid cell with age a in land use type i at time t. $\lambda_i(a,t)$ is the natural disturbance rate of the patch, $\lambda_{i,j}(a,t)$ is land use transition rate from land type i to j at time t. Terms on the right-hand side of the equation represent proportion of change due to ageing, natural disturbance, and LULCC, respectively. In addition to age, area and land use type, patches have respective resource availability of light, water and nutrients and cohort compositions.

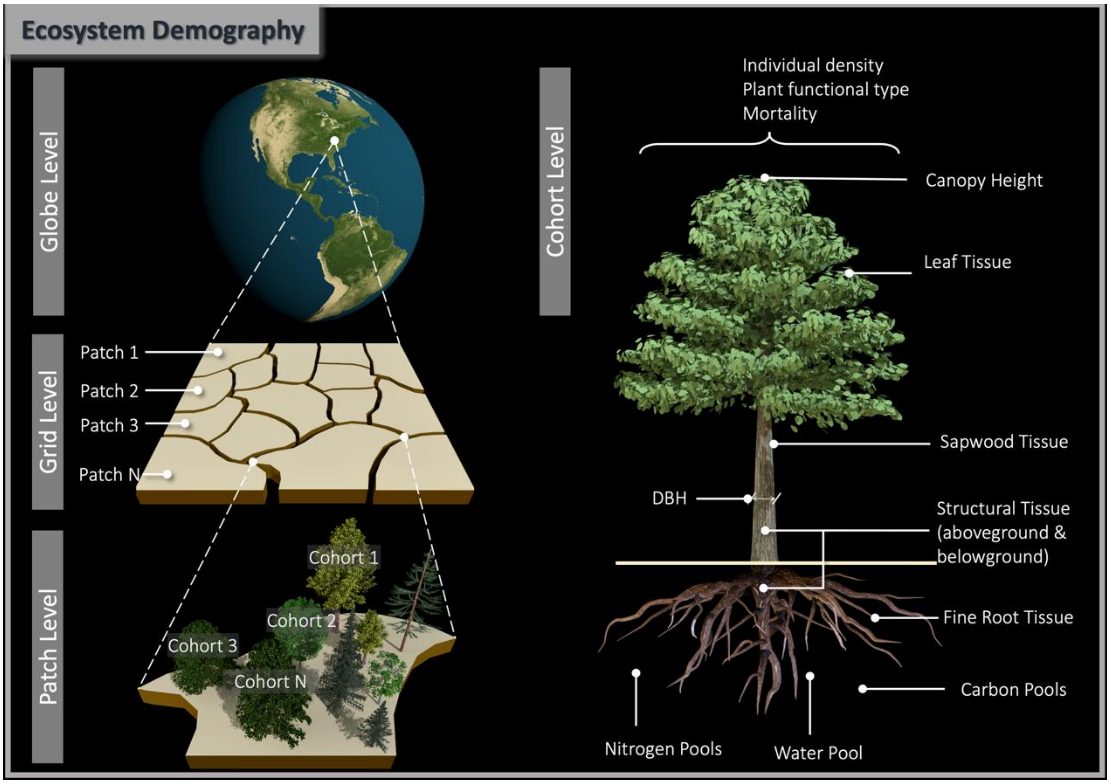


Figure 3.1. Diagram of vegetation representation scheme in ED model. Globe consists of land grids with fixed spatial resolution. A grid consists of patches with different ages from last disturbance and land use types, and patch areas dynamically change over time as a result of disturbance and land use changes. A patch consists of consists with different plant functional types and sizes. Plants in a cohort are depicted by properties including individual density, canopy height, diameter at breast (DBH), and biomass in leaf, sapwood, structural tissue and fine roots, and all these properties are simulated as a result of interaction with environment and other cohorts. Note that not all properties are shown here.

Vertical heterogeneity is captured by tracking plant size (e.g., height and biomass) in cohorts. Cohort contains plants with the same function types and size, and cohorts compete with each other for light, water and nutrients. Plant size and individual density is tracked for each cohort in each patch. Cohort dynamics in terms of growth, mortality and ageing are depicted by:

$$\frac{\partial}{\partial t} n_i(\mathbf{z}, \mathbf{x}, a, t) = -\frac{\partial}{\partial x}[G(\mathbf{z}, \mathbf{x}, a, t) n_i(\mathbf{z}, \mathbf{x}, a, t)] - \frac{\partial}{\partial a} n_i(\mathbf{z}, \mathbf{x}, a, t) - \mu(\mathbf{z}, \mathbf{x}, \bar{r}, t) n_i(\mathbf{z}, \mathbf{x}, a, t) \tag{3.2}$$

Where $n_i(\mathbf{z}, \mathbf{x}, a, t)$ is plant individual density of a cohort with size **z**, functional type **x**, and age a at time t in land use type i. $G(\mathbf{z}, \mathbf{x}, a, t)$ and $\mu(\mathbf{z}, \mathbf{x}, \bar{r}, t)$ are growth in size and mortality rates, respectively, and are nonlinear functions depending on resource and competition outcomes. Terms on the right-hand side of the equation represent individual density changes due to growth, ageing, and mortality, respectively.

The above two PDE equations depict general dynamics of patch demography and plant individual density in cohort, with detailed dynamic processes implemented by submodules. Processes represented in ED are illustrated in Figure 3.2. Specifically, for cohort dynamic, plants in a cohort assimilate carbon from atmosphere through photosynthesis, which is described in leaf physiology submodule. Modeling carbon assimilation takes into account the light, water and nutrient availability of a patch and environmental conditions (e.g., air temperature, humidity, and CO_2 level). Carbon uptake is then allocated to growth of height and biomass, autotrophic respiration, and decay of tissues (leaves, stem and, roots), which are described in the growth and allocation submodule, and also to the germination and dispersal of seedlings between patches, which is described in reproduction submodule. Plants change leaf biomass as seasons shift or when environmental conditions are unfavourable, which is described in leaf phenology submodule. Plant individual density changes over time due to mortality from either natural death or carbon starvation, which is described in the mortality submodule. Carbon from dead plants and decayed tissue of living plants are transferred to the soil pools and subsequent decomposition processes are tracked by the soil biogeochemical submodule. In addition to carbon, plants also uptake water and nitrogen from the soil and lose water through evapotranspiration and return nitrogen back to soil. The water and nitrogen cycles are respectively described in the hydrological and soil biogeochemical submodules. For patch dynamics, disturbance diversifies patch demography, resulting in a grid as mosaic of patches with different ages, areas and land use types. Disturbance also alters plant composition, competition, succession and other processes in cohorts and sequent carbon, water and nitrogen in soil. In addition to natural disturbance (e.g., fire and wind), a variety of land use activities (e.g., wood harvesting, deforestation and reforestation) is characterized in the submodule of disturbance and land use.

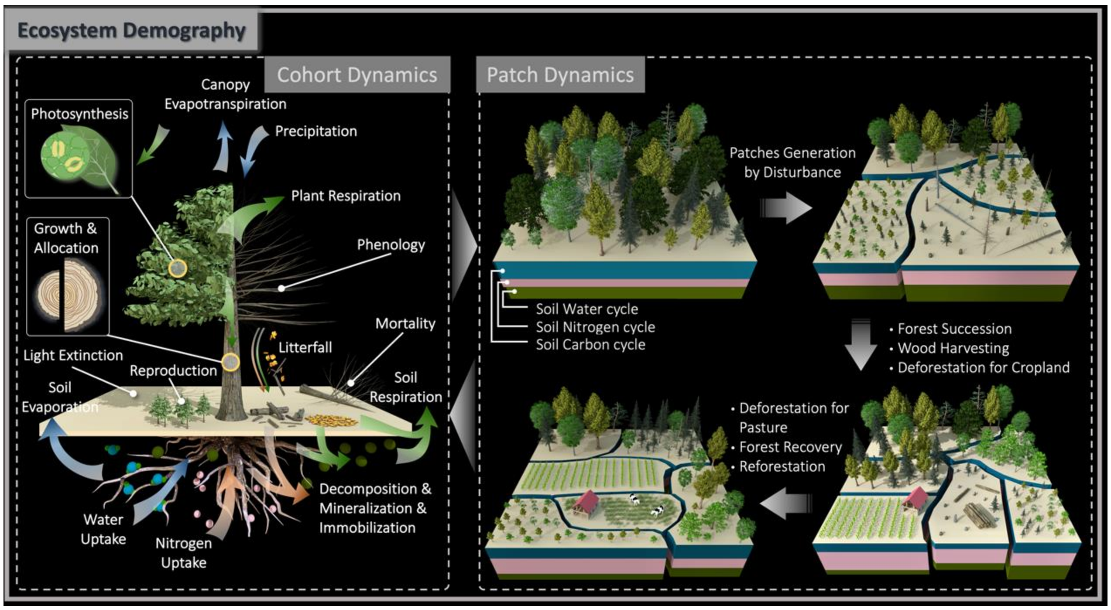


Figure 3.2. Schematic diagram of processes represented in ED model. Dynamics at cohort level consists of carbon-related flow (green arrow), water-related flow (blue arrow) and nitrogen-related (orange arrow). Carbon dynamics include carbon assimilation by photosynthesis, carbon allocation for plant growth in height/DBH, reproduction and respiration, carbon translocation between plants and soil through tissue turnover as litterfall and dead plants due to mortality, and carbon decomposition and respiration in soil carbon pools. Water dynamics include water inputs from precipitation and infiltration into soil, uptake by vegetation and evaporation and transpiration of soil and canopy. Nitrogen dynamics includes nitrogen uptake from soil pools, translocation from vegetation to soil through litterfall and dead plants, and mineralization and immobilization in soil. Note that not all processes that ED characterize are depicted here. Dynamics at patch level consist of consequences from a variety of disturbance events both natural and anthropogenic. Patch dynamics include disturbance-driven patch heterogenization in age and areas, forest succession, wood harvesting, deforestation for cropland and pasture expansion, and forest recovery and reforestation from abandoned cropland, harvested forest and pasture.

3.2.2 Refinement of plant function type

In Global ED, we refine PFTs previously developed in Moorcroft et al., (2001), Hurtt et al., (2002) and Albani et al., (2006). Here we include seven major types, namely early-successional broadleaf trees (EaSBT), middle-successional broadleaf trees (MiSBT), late-successional broadleaf trees (LaSBT), northern and southern pines (NSP), late-successional conifers (LaSC), C3 shrubs and grasses (C3ShG), and C4 shrubs and grasses (C4ShG). The broadleaf PFTs (i.e., EaSBT, MiSBT, and LaSBT) are distinguished between tropical and non-tropical subtypes in terms of leaf traits (e.g., leaf lifespan, specific leaf area, and leaf photosynthesis rate) and mortality rate. The boundary of tropical and non-tropical subtypes is delineated by whether the multidecade average air temperature during the coldest month of the year is above or below 18 °C.

PFTs differ in phenology, leaf physiological traits, allometry, mortality rate and dispersal distance. For example, for phenology, needleleaf PFTs (i.e., NSP and LaSC) are evergreen, and broadleaf PFTs (both tropical and non-tropical subtypes) and grass PFTs are cold- and drought-deciduous; For leaf traits, broadleaf tropical subtypes have longer lifespan but lower specific lead area and carboxylation rate than non-tropical subtypes. needleleaf PFTs have longer lifespan than broadleaf PFTs; For allometry, the seven major PFTs all use different allometric equations, but broadleaf PFT subtypes share the same allometry; For mortality, grass PFTs have the largest mortality, followed by broadleaf PFTs and needleleaf PFTs. broadleaf tropical PFTs are larger than non-tropical subtypes.; For dispersal distance, the EaSBT disperse more seedlings to non-local patch than the MiSBT and LaSBT, and the NSP is more than the LaSC. Detailed parameterizations of which can be found within each submodule section of the Appendix B.1.

Spatial distribution of PFTs is mechanistically determined by individual competition for light, water and nutrients. No quasi-equilibrium climate–vegetation relationships, such as satellite-based PFT maps or climatic envelope thresholds, are used to constrain presence or absence of PFTs. All PFTs could potentially coexist in any location over globe and are initialized with the same density; the subsequent competition determines when and where specific PFTs dominate the ecosystems. The competitive advantage of each PFT results from plant traits such as photosynthesis efficiency, height growth rate, and reproduction strategies. These advantages vary with climate conditions and ecosystem successions as well. For example, leaf physiological traits exhibit trade-off across PFTs (Reich et al., 1997). Comparing to needleleaf PFTs, broadleaf PFTs have a relatively larger leaf area per leaf weight and higher carbon assimilation rate per leaf area, but higher carbon demand for leaf turnover. Moreover, the early-successional PFT rapidly accumulates carbon, quickly grows in height, and disperses seeds over long-distances. These characteristics lead to its dominance at early successional state of recently disturbed ecosystems. However, its intolerance of shade makes it less competitive as the canopy close, eventually being replaced by mid- and late-successional PFTs which have lower morality in shade but grow more slowly in height.

3.2.3 Freezing injury

Exposure to low temperatures can cause tissue damage to twigs and buds, affecting sequent carbon balance and survival (DeHayes, 1992; Gu et al., 2008; Sakai and Larcher, 2012; Sakai and Weiser, 1973; Vitasse et al., 2014). In global ED, we characterized injury effects by introducing leaf loss at low temperatures. When monthly average air temperature drops below the defined freezing resistance threshold varying with PFTs, the carbon in leaf biomass is reduced as well as in active biomass pool. The resulting leaf loss could affect ongoing carbon assimilation and height growth, and also may result in competitive disadvantage over others PFTs with more resistance to freezing.

3.2.4 Leaf physiology

We refine leaf physiology submodule such as reformulating functional forms of photosynthesis calculation in C3 and C4 pathways (Farquhar et al., 1980; Von Caemmerer and Furbank, 1999), adding boundary layer conductance for diffusing water vapor and CO_2 between ambient air and leaf surface, and parameterizing temperature dependence in Arrhenius-based functions (Bernacchi et al., 2001; von Caemmerer et al., 2009; Kattge and Knorr, 2007; Massad et al., 2007; Von Caemmerer, 2000). In this physiology submodule, photosynthesis, stomatal conductance and leaf energy balance are coupled to mechanistically quantify leaf-level carbon and water exchange in response to environmental conditions (air temperature, shortwave radiation, air humidity, wind speed, CO_2 level). Specifically, photosynthesis sub-model depicts any biochemical limitations to carbon assimilation by considering temperature-dependent enzyme activities, CO_2 supply and light availability; stomatal conductance sub-model estimates stomatal openness in response to air humidity and resulting water vapor and CO_2 exchange between ambient air and leaf intercellular space; leaf energy balance sub-model solves the energy budget equation of absorbed and emitted radiation, and estimates heat loss due to convection and transpiration to estimate leaf temperature at equilibria with environmental conditions. Three sub-models are simultaneously solved by numerical iterations to determine photosynthesis, leaf respiration and transpiration rates per unit of leaf area. Detailed equations and associated parameterizations can be found in Appendix B.3.

3.2.5 Evapotranspiration

We introduce evaporation from soil and wet canopy to estimation of total evapotranspiration in a patch, which is absent in previous models. Thus, in global ED, total evapotranspiration consists of soil evaporation E_{soil}, wet canopy evaporation E_{canopy} , and canopy transpiration T_{canopy}:

$$ET = E_{soil} + E_{canopy} + T_{canopy} \tag{3.3}$$

Canopy transpiration is estimated by scaling the transpiration rate of all leaves up to the canopy-level. Evaporation from soil and wet canopy is estimated by an approach from Mu et al., 2011. Specifically, in this approach, soil evaporation comes from both saturated soil surfaces and moist surfaces. Evaporation from both surface types is modelled using the Penman-Monteith (P-M) equation (Monteith, 1965), but moist surface is further constrained by an empirical soil moisture function (Fisher et al., 2008). Canopy evaporation is also estimated by the PM equation with associated parameters of absorbed radiation and aerodynamic and surface resistance. Soil and canopy evaporation are mediated by partition of net incoming radiation between the canopy and soil, and the partitioning fraction is a function of leaf area in the canopy; Dense canopy allows less radiation to reach the ground and in turn, results in less evaporation from soil. Soil and canopy evaporation are computed separately for daytime and nighttime to reflect different meteorological conditions, including shortwave radiation, air temperature and humidity. Details of evapotranspiration calculation can be found in Appendix B.9.

3.2.6 Soil hydrology

In the hydrology submodule, we still use a single-layer bucket model to track soil water availability within a patch as developed in Moorcroft et al., (2001) but additionally consider processes of snowpack formation and snow melt and add an option to use Mualem-van Gehuchten (MvG) based soil hydraulic parameters. In this submodule, soil gains water from precipitation and snowmelt and loses water via evapotranspiration, percolation and runoff. Snowpack formation and snow melt currently is depicted in a simple and empirical way that snowpack forms by precipitation when monthly average air temperature is below the freezing point and starts to melt with a rate linearly related to the air temperature until depletion. Therefore, taking into account precipitation, snow melt, percolation, runoff, and evapotranspiration from soil and canopy, the change rate of soil water content can be expressed as:

$$dW/_{dt} = P + SM - perc - ET \tag{3.4}$$

Where P is precipitation and set to zero when monthly minimum air temperatures are lower than 0 °C, SM is snowmelt, ET is total evapotranspiration, and $perc$ is percolation and runoff, as estimated by the Mualem-van Gehuchten (MvG) equation. Soil hydraulic properties (e.g., soil depth and saturated hydraulic conductivity) and other parameters used in MvG equation are specified by Montzka et al., (2017), which employed Miller-Miller theory to scale the state-of-the-art soil dataset SoilGrids1km for Earth System Model to provide spatially consistent hydraulic parameters. See Appendix B.9 for more details regarding the soil hydrology module.

3.2.7 Wood product pools & crop calendar

Modifications on land use submodule include adding wood products pools and crop calendar. Previously developed land use submodule only tracks changes in vegetation and soil carbon during various land use activities, but not track decay process after removal. However, removed carbon from wood harvestings and deforestations are used for various purposes, resulting in different lifetimes and temporal emissions to atmosphere. Therefore, we add three wood product pools to track the lifecycles of harvested wood and associated decay process. Product pools gain wood from land use activities such as wood harvesting or deforestation. The biomass of vegetation involved are then allocated to litter pool and wood product pools. Allocation fractions vary with PFTs types, land use activities, and spatial location (Appendix B.11). Wood stored in product pools decays over time and ends up as emissions to atmosphere. These three pools differ in their decay rates (1, 10 and 100 years, respectively). For example, the 10-yr product pool generally depletes within about 10 years if no additional carbon is loaded. Decay rates are estimated using the exponential function from Harris et al., (2015).

A climate-driven crop calendar (Sacks et al., 2010) is used to determine the planting and harvesting date of crop-type PFTs (i.e., C3ShG and C4ShG). Crop growing areas are delineated to temperature- or precipitation-limited regions depending on whether the average air temperature of the coldest month is below or above 10 °C. In temperature-limited regions, crop PFTs are planted in or harvested from crop patches relative to the defined temperature threshold. For precipitation-limited regions, planting and harvesting occurs relative to the monthly total precipitation is threshold (i.e., 100 mm/month).

3.3 Model Experiment and Evaluation

Global ED consists of two separate runs at 0.5° spatial resolution with consideration of climate variability, elevated CO_2, and land use change. The first run (called the equilibrium simulation) spun up global ED to obtain model initial conditions representing ecosystem states at A.D. 850. This run was performed for 1000 years by which PFT composition and carbon pools of vegetation and soil reached a dynamic equilibrium. The second run (called transient simulation) restarted from the initial conditions at A.D. 850 and continued running for 1166 years corresponding to A.D. 851 – A.D. 2016 with varying CO_2 levels, land-use change, and climate variability. Both runs were driven with meteorological forcing from NASA Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA2) (Gelaro et al., 2017) and surface CO_2 concentration from NOAA CarbonTracker Database version 2016 (NOAA CT2016) (Peters et al., 2007, with with updates documented at http://carbontracker.noaa.gov). Additionally, the transient simulation run prescribed burned area using the Global Fire Emissions Database, version 4 (GFED4) (Randerson et al., 2015) and forced land-use change with Land Use Harmonization version 2 (LUH2) (Hurtt et al., 2019a, 2020). Sections 3.1 and 3.2 detail the equilibrium and transient simulations, respectively, and section 3.3 details the meteorology, CO_2, and burned area and land-use forcings.

3.3.1 Equilibrium simulation

The equilibrium simulation was started from bare ground where the soil and vegetation carbon pools were set at zero, and all PFTs were initialized with equal seedling density for all patches and all grid cells over globe. This run was driven for 1000 years with MERRA2 climatology of 1981-1990 and NOAA CT2016 surface CO_2 average of 2001-2014 (spatial varying and global average rescaled to 280 ppm). No climatic envelope or potential biome maps were used to constrain PFT spatial distribution. The land-use change module was disabled in this run of simulation.

3.3.2 Transient simulation

The transient simulation was restarted from equilibrium and driven with the LUH2, MERRA2, NOAA CT2016 and GFED4 datasets. The land-use change submodule was activated, and all land-use transition types from LUH2 were incorporated into the simulation at annual time steps including shift cultivation, deforestation, and wood harvesting. MERRA2, NOAA CT2016 and GFED were used throughout the simulation with varying temporal settings depending on data availability. Specifically, for MERRA2, a climatology between 1981-1990 was used until 1981, and annual meteorology was used subsequently. For NOAA CT2016, a surface CO_2 concentration average between 2001-2014, which varies spatially and grows over time, was used until 2000, while annual NOAA CT2016 surface CO_2 concentration was used subsequently. For GFED4 burned area, an averaged between 1996-2016 was used until 1996, after which annual burned area were used.

3.3.3 Forcing data

Meteorological variables utilized from MERRA2 include surface air temperature (TLML), surface specific humidity (QLML), precipitation (PRECTOTCORR), incident shortwave radiation (SWGDN), surface wind speed (SPEED), and multi-layer soil temperature (TSOIL1-TSOIL3). Original estimates of surface air temperature, surface specific humidity, incident shortwave radiation and surface wind speed were averaged from daily hourly to monthly hourly for each year between 1981 to 2016. The resulting annual monthly average of diurnal meteorological variables were used to drive the leaf physiology submodule in ED.

Hourly surface air temperature, precipitation and soil temperature were also aggregated to monthly averages for each year between 1981 to 2016, and then used to drive the soil hydrology, phenology, evapotranspiration, and biogeochemical modules in ED.

Surface CO_2 concentration was extracted from the lowest vertical level of NOAA CT2016 CO_2 mole fraction which is temporally and spatially varying. The original datasets were first linearly interpolated from 3°x2° (longitude x latitude) to 0.5°x0.5° and from 3-hour to hourly, and then averaged to monthly hourly estimates for each grid cell and each year between 2001 and 2014, resulting in surface CO_2 concentration maps with 4032 fields (14 years, 24 hours, 12 months) for each 0.5°x0.5° grid. The surface CO_2 concentration maps were used to drive the transient simulation from 850 to 2000, retaining average spatial variation between 2001 and 2014 and applying a scaling factor to force the global annual average CO_2 concentration to remain at 280 ppm before 1850, then grow linearly to 310 ppm in 1950 and to 375 ppm in 2000. This increasing trend in global average matches observed CO_2 growth rates from (Keeling, 2008).

LUC forcing was derived from the LUH2 (version v2h) for years 850-2015. The original land use state and land use transitions were aggregated from a spatial resolution of 0.25°x0.25° to 0.5°x0.5° for each year between 850 and 2015. Subtypes of land use states and associated transitions were grouped to major land use types of its predecessor version (LUH1). Specifically, sub crop types of C3 annual crops (c3ann), C3 perennial crops (c3per), C4 annual crops (c4ann), C4 perennial crops (c4per) and C3 nitrogen-fixing crops (c3nfx) were all merged as cropland. Forested primary land (primf) and non-forested primary land (primn) were merged as primary land, forested secondary land (secdf) and non-forested secondary land were merge as secondary land, managed pasture (pastr) and rangeland were merged as pasture. Note that all types of land use transitions and gross rate were used in the model's land use module.

3.3.4 Benchmarking evaluation

A comprehensive benchmarking package (Table 3.1) was collected to evaluate ED performance in simulating ecosystem dynamics and the terrestrial carbon cycle's response to climate change, land-use change, and rising CO_2. Multiple critical variables that govern ecosystem functioning, including PFT distribution, carbon stocks in vegetation and soil, carbon and water fluxes, and vegetation structures in terms of canopy height and vertical LAI, were compared against the benchmarking package. ED evaluation was carried out at different spatial (grid, latitudinal, and biome) and temporal scales (climatological, seasonal and interannual).

Table 3.1. Summary of benchmarking datasets used for evaluation of global ED model.

Variable	Source	Description	Reference
		Vegetation distribution	
PFT	ESA CCI	Global gridded, 300-m, 2015	ESA (2017)
Carbon stocks			
AGB	GEOCARBON	Global gridded, 0.01-degree, 2010	Avitabile et al., (2016); Santoro et al., (2015)
	Santoro et al., (2018)	Global gridded, 100-m, 2010	Santoro et al., (2018)
	Spawn et al., (2020)	Global gridded, 300-m, 2010	Spawn et al., (2020)
Soil carbon	HWSD	Gridded, 0.05 degree, 2000	Wieder et al., (2014)
		Carbon and water fluxes	
GPP	FLUXCOM (RS+METEO, CRUJIA and ERA5)	Global gridded, 0.0833-degree, 1979-2017 monthly	Jung et al., (2020)
	FluxSat	Global gridded, 0.05-degree, 2001-2018 monthly	Joiner et al., (2018)
NBP	CAMS (v17r1)	Global gridded, 1.875x3.75-degree, 1979-2017 monthly	Chevallier et al., (2005)
	Jena CarbonScope (s81oc_v2020)	Global gridded, 2.5x2.0 degree, 1981-2016 daily	Rödenbeck et al., (2008)
	CarbonTracker Europe (CTE)	Global gridded, 1x1 degree, 2000-2016 monthly	van der Laan-Luijkx et al., (2017)
	GCB2020 DGVMs	Global total, 1959-2019 yearly	Friedlingstein et al., (2020)
	GCB2020 Residual sink	Global total, 1959-2019 yearly	Friedlingstein et al., (2020)
ET	FLUXCOM (RS+METEO, CRUNCEP and GSWP3)	Global gridded, 0.0833-degree, 1981-2014 monthly	Jung et al., (2020)
		Vegetation structure	
Tree height	GEDI L2A (v001)	51°N ~ 51°S, 20-m footprint, 2019-2020	Dubayah et al., (2020b)
	ICESat-2 ATL08 (v003)	51°N ~ 51°S, 100-m footprint, 2018-2020	Neuenschwander et al., (2020)
LAI	MODIS MCD15A3H (v006)	Global gridded, 500-m, 2003-2016 4-day	Myneni et al., (2015)
	GEOV2	Global gridded, 1/3-km, 1999-2016 10-day	Verger et al., (2014)
Vertical LAI	GEDI L2B (v001)	51°N ~ 51°S, 20-m footprint, 2019-2020	Dubayah et al., (2020c)

The satellite-based land cover product, ESA CCI, was used to examine the distribution of three modelled PFTs, grass, broadleaf trees, and needleleaf trees. Many satellite-based land cover datasets differ largely in PFT definition from the global ED. For example, no successional types of PFTs exist in ESA CCI land cover types. Thus, the native PFTs in global ED and ESA CCI both have to be aggregated to broader categories such as broadleaf PFTs, needleleaf PFTs and grass PFTs. To do this, the 22 native land cover classes of ESA CCI were first reclassified to 'broadleaf evergreen tree', 'broadleaf deciduous tree', 'needleleaf evergreen tree', 'needleleaf deciduous tree', 'natural grass' and 'manned grass' by using a cross-walk table (Poulter et al., 2015). They were then further merged by phenology type and aggregated to 0.5 degree, resulting in PFT fraction maps of broadleaf PFTs, needleleaf PFTs, and grass and shrub PFTs. ED PFTs of EaSBT, MiSBT and LaSBT were merged as broadleaf PFTs, NSP and LaSC were merged as needleleaf PFTs, and C3ShG and C4ShG were merged as grass and shrub PFTs.

Modelled gross primary productivity (GPP) was evaluated with respect to spatial pattern, seasonality, and interannual variability using satellite data-driven GPP datasets FLUXCOM (Jung et al., 2020) and FluxSat (Joiner et al., 2018), and the satellite-retrieved sun-induced chlorophyll fluorescence dataset (CSIF) (Zhang et al., 2018). The FLUXCOM and FluxSat datasets are derived from a data-driven approach which combines carbon fluxes measurements from FLUXNET and satellite observations from MODIS. One major difference between FLUXCOM and FluxSat is meteorological forcing and which specific approach was used. FLUXCOM used meteorological forcing and a machine learning approach, while FluxSat used a simplified light-use efficiency model that does not rely upon meteorological forcing. The FluxSat also used satellite-based SIF to delineate highly productive regions. Satellite measurements of sun-induced chlorophyll fluorescence (SIF) have recently been suggested as a promising proxy of terrestrial GPP, exhibiting high sensitivity to plant photosynthetic activities (Lee et al., 2013; Guanter et al., 2014; Yang et al., 2015). In this study, we chose CSIF dataset for its improved spatiotemporal continuity. CSIF is generated by fusing Orbiting Carbon Observvotory-2 (OCO-2)-retrieved SIF and MODIS reflectance data using a machine learning approach. FLUXCOM, FluxSat and CSIF were all resampled to monthly estimates at 0.5x0.5 spatial resolution before the evaluation.

ED modelled net biome productivity (NBP) was compared against multiple sources including estimates from process-based models, atmospheric inversions, and the 2020 global carbon budget (GCB2020) (Friedlingstein et al., 2020). For process-based models, 17 DGVMs reported in the GCB2020 were used to calculate the respective net land sink by differencing land uptake and land use emissions estimates (i.e., S_{LAND} - E_{LUC}). For atmospheric inversions, three systems are used, namely CarbonTracker Europe (CTE) (van der Laan-Luijkx et al., 2017), Jena CarboScope (version s81oc) (Rödenbeck et al., 2008) and the Copernicus Atmosphere Monitoring Service (CAMS) (Chevallier et al., 2005). The three inversions both derive surface carbon fluxes with atmospheric CO_2 measurements, prior constraints on fluxes, uncertainty and atmospheric transport model, but vary with choices for specific data, prior constraints and transport models (Peylin et al., 2013). In the GCB2020, the residual terrestrial sink was used, which was calculated as total emissions from fossil fuel and land use change minus atmospheric CO_2 growth rate and the ocean sink (i.e., $E_{FF} + E_{LUC} – G_{ATM} – S_{OCEAN}$). Multiple source estimates are used here for various purposes: comparing to DGVMs can examine agreement between global ED and other bottom-up process-based approaches in temporal variability of NBP; comparing to the residual sink of global carbon budget report can examine if global ED estimates could reduce current budget imbalance; and comparing to atmospheric inversions can examine consistency in spatial attributions and temporal variability of NBP between bottom-up and top-down approaches.

ED-modelled carbon pools are evaluated with regards to vegetation aboveground biomass (AGB) and soil carbon. The reference AGB data includes estimates from GEOCARBON (Santoro et al., 2015; Avitabile et al., 2016), Santoro et al., (2018), and Spawn et al., (2020). These three AGB datasets provide high spatial resolution (e.g., 100 m to 1000 m) wall-to-wall global estimates of the year 2010 but differ in their methodologies. Specifically, GEOCARBON AGB is a forest biomass map produced by harmonizing the pan-tropical biomass map from Avitabile et al., (2016) with the boreal forest biomass map from Santoro et al., (2015). AGB from Santoro et al., (2018) is produced by combining spaceborne SAR (ALOS PLASAR, Envisat ASAR), Landsat-7, and Lidar observations from ICESat. AGB from Spawn et al., (2020) is biomass not only for forest but also other woody non-forest plants. Reference soil carbon is from the Harmonized World Soil Database (HWSD) (Wieder et al., 2014), including soil carbon for topsoil (0 to 30 cm) and subsoil (30 to 100 cm).

Evaluation of ED-modelled forest structure focuses on total and vertical distribution of leaf area index (LAI) and tree canopy height. Two reference LAI products, namely MODIS MCD15A3H (Myneni et al., 2015) and GEOV2 LAI (Verger et al., 2014), are used for evaluating total LAI in terms of spatial distribution, seasonality, and interannual variability. The MODIS and GEOV2 LAI datasets are both derived from passive optical observations with empirical-based inversion methods which relates leaf area with optical canopy reflectance or vegetation indices, but these two products vary with source of optical observations and choices for inversion methods. Reference vertical LAI is from the Global Ecosystem Dynamics Investigation (GEDI) L2B products, which retrieves leaf vertical distribution from lidar waveform return (Dubayah et al., 2020c). Reference canopy height are direct forest structure observation from GEDI L2A (Dubayah et al., 2020b) and the Ice, Cloud and land Elevation Satellite (ICESat-2) ATL08 products (Neuenschwander et al., 2020). Mean canopy height is generated at 0.5 degree from the relative height 98th percentile (RH98) of all GEDI L2A footprints and canopy top height (h_canopy) of all ICESat-2 ATL08 segments of good quality.

3.4 Results

Here, we present simulation results of global ED across four primary categories, PFT distribution, vegetation and soil carbon pools, carbon and water fluxes, and vegetation structure, including associated spatial and temporal evaluations with benchmarking datasets.

3.4.1 PFT distribution

The spatial fraction and corresponding latitudinal areas of broadleaf PFTs, needleleaf PFTs and grass and shrub PFTs are shown in Figure 3.3. Native PFTs for both global ED and ESA CCI have been re-grouped via a crosswalk table (details in section 3.5). Global ED generally represents the observed distributions of both PFTs, where needleleaf PFTs dominate at high latitudes, broadleaf PFTs dominate in the tropics, and grass and shrub PFTs widespread all over the globe. For example, global ED predicts the observed coexistence of broadleaf and needleleaf PFTs in southern China and eastern US. However, global ED predicts the existence of needleleaf PFTs along the Andes Mountains in South America and in southern Australia, a pattern is not seen in the ESA CCI but in ground observations (Farjon and Filer, 2013). ED also predicts more broadleaf PFTs in eastern Europe and southern China, less broadleaf PFTs in Africa savanna, less needleleaf PFTs in east Siberia, and less grass and shrub PFTs both in Africa savanna and northern China. In terms of latitudinal area per PFT, the smallest discrepancies global ED and ESA CCI appear in broadleaf PFTs, followed by needleleaf PFTs, and grass and shrub PFTs.

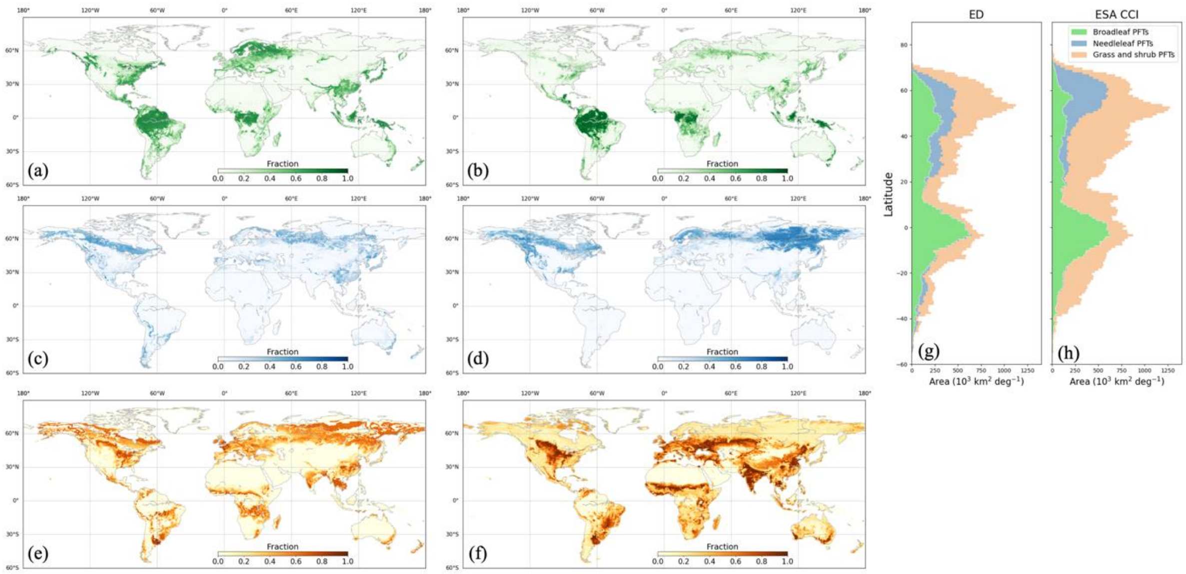


Figure 3.3. Spatial fraction of broadleaf PFTs, needleleaf and PFTs and grass and shrub PFTs in year of 2015 from ED (a), (c) and (e), and from ESA CCI (b), (d) and (f). Corresponding latitudinal area is compared in (g) and (h).

3.4.2 Vegetation and soil carbon

ED estimates global total aboveground vegetation carbon (including forest and non-forest) at 298 Pg C in 2010. This compares to 283 Pg C and 297 Pg C estimated by Spawn et al., (2020) and Santoro et al., (2018), and 220 Pg C estimated by the GEOCARBON which only includes forest, respectively. ED AGB well captures the spatial pattern of the three reference AGBs, showing the highest biomass densities areas across the tropics (i.e., the Amazon rainforest, the Congo river basin, and southeast Asia) with declining biomass densities northward towards the temperate and boreal regions. For example, average AGB density is about 15 kg C/m^2 in the tropics and under 2.5 kg C/m^2 across temperate and boreal regions (Figure 3.4e). The established biomass transition along the African forests-savanna zone is well reproduced by ED, albeit with relatively lower values than the reference data across the savanna. Discrepancies between ED and reference AGB data appear in southern China, Southeast Asia and southeast Brazil, where AGB overestimation in southern China and Brazil may result from land-use forcing. The LUH2 v2h underestimates harvesting area on primary forest in high-latitudes, southern China, and Southeast Asia, and also underestimates cropland area in southern Brazil (Chini et al., 2021).

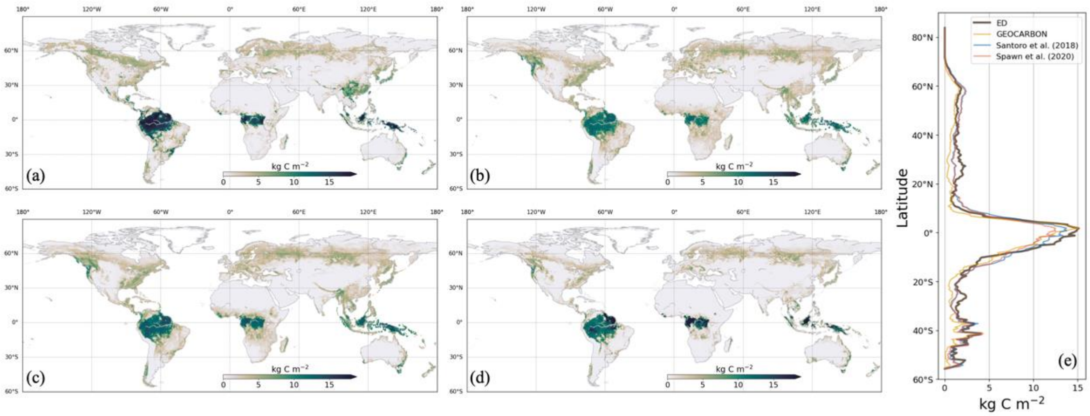


Figure 3.4. AGB in 2010 from ED (a), Spawn et al., (2020) (b), Santoro et al., (2018) (c), and GEOCARBON (d), with latitudinal average AGB compared in (e). Note (a)-(c) include AGB of both forest and non-forest area, and (d) only includes forest biomass.

ED estimates global total soil carbon at 671 Pg C in 2000, which is within the range of CMIP5 ESMs (510 - 3040 Pg C) (Todd-Brown et al., 2013) but lower than the HWSD estimate of 1201 Pg C. Comparing total stocks at the biome level (Figure 3.5d) shows that ED generally reproduces soil carbon variation across biomes, but notably underestimates carbon in boreal forest/taiga, deserts and xeric shrublands, tropical and subtropical grasslands, savannas and shrubland. The soil carbon map from ED reveals different spatial patterns compared to HWSD, with relatively less spatial heterogeneity and fewer regions with densities above 30 kg C/m^2. This bias may arise from poor representations of biophysical conditions such as water-saturated soils, where soil carbon is well preserved and likely missing critical mechanisms. Similar to most DGVMs/ESMs, soil carbon distribution is primarily driven by NPP and soil temperature. However, these two drivers can only explain a small amount of spatial variation in the HWSD map, and other more important drivers are not well characterized in ED and other ESMs (Todd-Brown et al., 2013). It should be noted that there are also substantial uncertainties with current empirical soil carbon maps in terms of both global totals and spatial distribution (Todd-Brown et al., 2013).

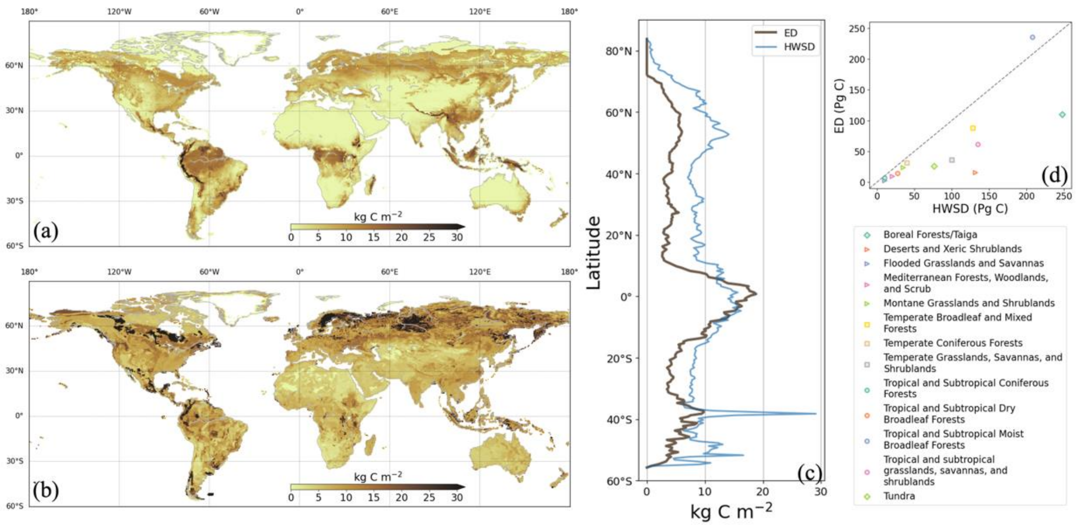


Figure 3.5. Soil carbon density in 2000 from ED (a) and HWSD (b). Latitudinal average density and total stocks per biome are compared in (c) and (d), respectively.

3.4.3 Carbon and water fluxes

Globally, ED estimates average annual GPP at 134 Pg C yr^{-1} between 2001-2016, which compares to 120 Pg C yr^{-1} from FLUXCOM and 136 Pg C yr^{-1} from FluxSat over the same period (Figure 3.6). Global ED does well in capturing the spatial pattern of mean annual GPP at the grid and latitudinal scales. Areas of highest productivity occur in the tropics, followed by temperate and boreal regions. Average annual GPP is about 4 kg C/m^2/yr at the tropics and 1 kg C/m^2/yr in temperate regions. For the tropics, global ED is 0.5 kg C/m^2/yr higher than FLUXCOME but closer to the FluxSat (within 0.2 kg C/m^2/yr), and lower at the Africa Savanna. Additionally, ED has relatively higher GPP in southern China and Brazil than either reference dataset. A notably increasing annual trend in total global GPP can be seen in both global ED and FluxSat estimates between 2001-2016 as well as from globally averaged CSIF (Figure 3.7). ED also reproduces GPP interannual variability from FluxSat, FLUXCOM and CSIF, dipping in the years 2005, 2012 and 2015, and peaking in 2006, 2011 and 2014 (Figure 3.8). Regarding latitudinal seasonality at the biome scale, ED captures GPP timing for most latitudinal zones including 60° ~ 90°N, 45° ~ 60°N, 15° ~ 30°N and 60° ~ 30°S. Major differences appear in 30° ~ 45°N, where ED shows decrease from July-September, and in 15°S ~ 0°, where ED shows delayed monthly timing of lowest annual GPP values.

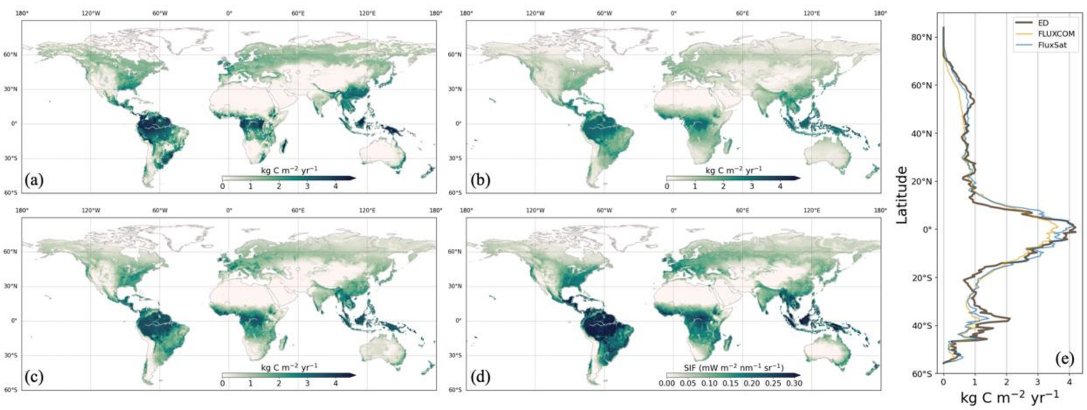


Figure 3.6. Average annual GPP between 2001 and 2016 from ED (a), FLUXCOM (b), FluxSat (c) and CSIF (d). Comparison of latitudinal average GPP is shown in (e).

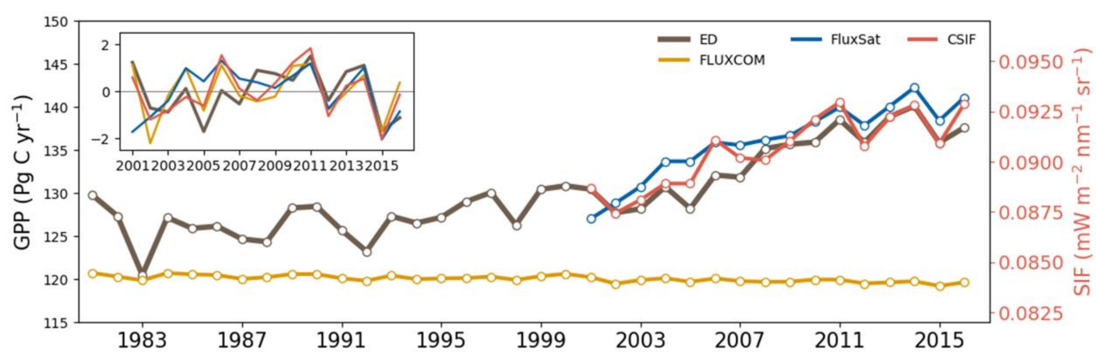


Figure 3.7. Time-series of global annual total GPP from ED, FLUXCOM and FluxSat, and global annual average CSIF. Their interannual anomaly is shown in the inset.

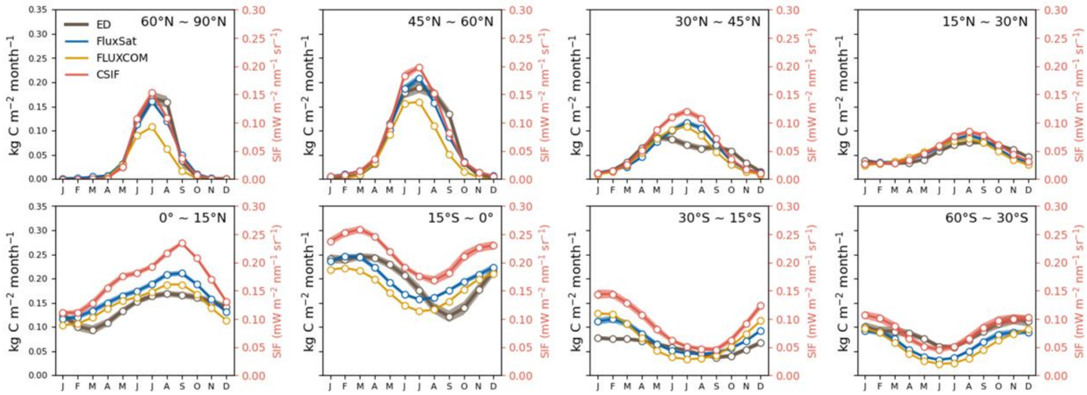


Figure 3.8. Average seasonal cycle (2001-2016) of GPP from ED, FLUXCOM, FluxSat, and CSIF by latitudinal band.

All NBP estimates shows an increasing trend but also substantial interannual variation during the 1981-2015 period (Figure 3.9). Global ED generally shows consistent variability with other estimates at the global scale; all results show strong NBP reductions in El Niño years (such as 1983, 1998 and 2015), and strong NBP increases in La Niña years (such as 1989, 2001-2002 and 2011). An exception is 1991-1992, where global ED and DGVMs are both lower than atmospheric inversions. This is probably due to the Mt. Pinatubo eruption, which is not included in the shortwave radiation forcing of GCB2020 DGVMs or global ED (Mercado et al., 2009; Friedlingstein et al., 2020). In addition, global ED shows a continued increasing trend over the 2007-2016 period, as reflected in the mean of atmospheric inversions but not the mean of DGVMs. For example, ED NBP is 2.34 Pg C/yr from 2007-2016, which is higher than either the mean of DGVMs (1.40 Pg C/yr) or the GCB2020 residual terrestrial sink (1.81 Pg C/yr), but within the range of the atmospheric inversions estimate (1.77 ~ 2.64 Pg C/yr) and closer to the upper bound of the DGVMs range (i.e., 0.58 ~ 2.82 Pg C/yr). Latitudinal value comparison between global ED and atmospheric inversions indicates contrasting attribution of the global sink (Figure 3.10). In comparison to the atmospheric inversions, global ED predicts a stronger sink in tropics and relatively weaker sink in the Northern Hemisphere. Such a pattern is highlighted in the global carbon budget, where process-based models and the atmospheric inversions generally show less agreement on the magnitude of the carbon sink in these two regions.

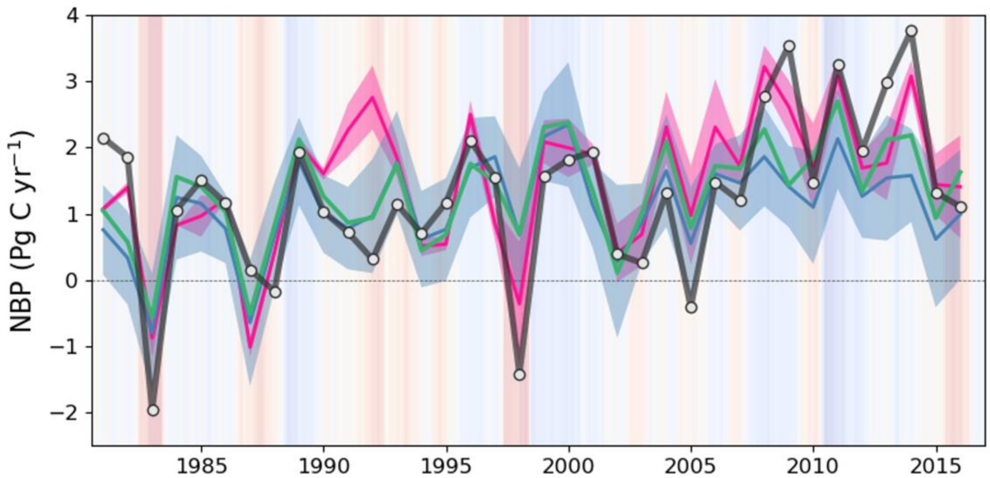


Figure 3.9. Global annual NBP between 1981 and 2016 from ED (black line), DGVMs from the GCB2020 (ensemble average shown in blue line with ±1σ spread shown in blue shading), the ensemble of atmospheric inversions (ensemble average shown in pink line with ±1σ spread shown in pink shading), and the terrestrial residual sink of the GCB2020 (green line). Positive values indicate net carbon uptake from land. Background shading represents the bi-monthly Multivariate El Niño/Southern Oscillation (ENSO) index, where red indicates El Niño and blue indicates La Niña.

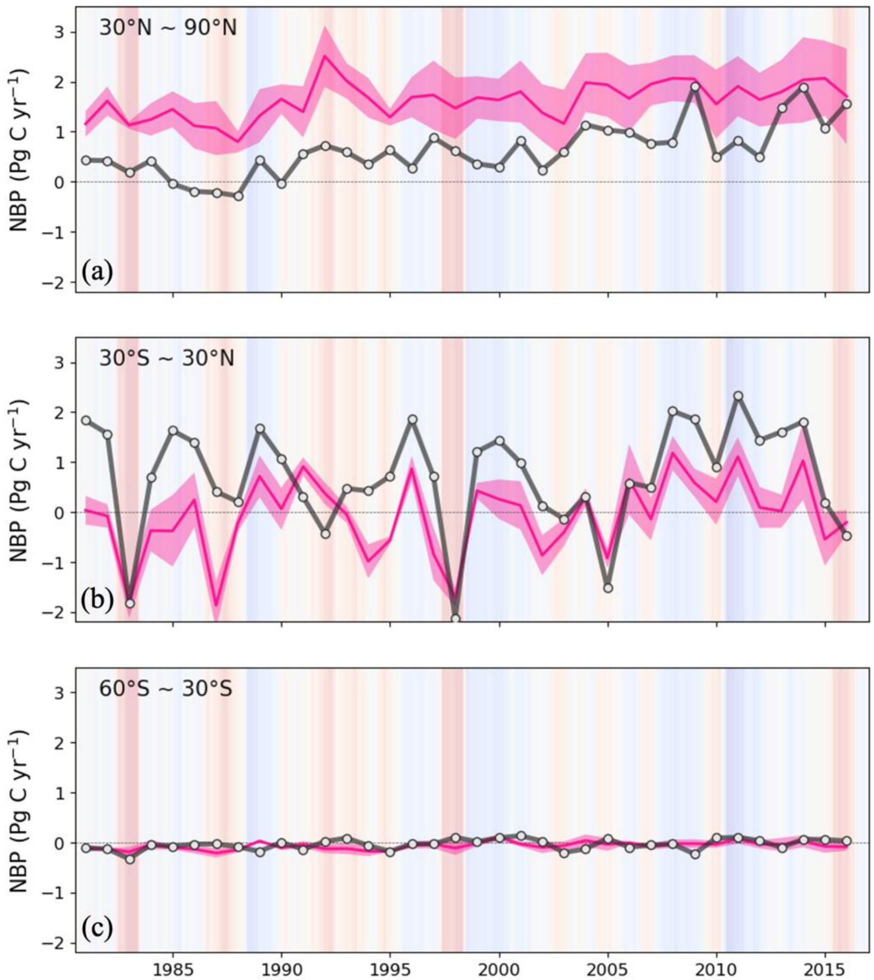


Figure 3.10. Annual NBP between 1981 and 2016 from ED and ensemble of atmospheric inversions for the Northern Hemisphere (>30°N) (a), tropics (30°N ~ 30°S) (b) and the Southern Hemisphere (<30°S) (c). Black line is ED, and the pink line and pink shading are the inversion ensemble average and ±1σ spread of atmospheric inversions, respectively.

ET estimates from ED and FLUXCOM are compared with respect to grid and latitudinal distribution (Figure 3.11). ED reproduces the general spatial pattern of ET, with the highest rates located across the tropics and slow decreases towards high latitudes. This pattern also follows the spatial distribution of precipitation. ED shows closer alignment with FLUXCOM at the tropics (i.e., 1500 mm/yr) as well as latitudes above 60°N and below 35°S (i.e., below 500 mm/yr), but notably underestimates average annual ET in other latitudes. ED shows smaller ET than FLUXCOME in dry regions such as southern Africa and interior Australia.

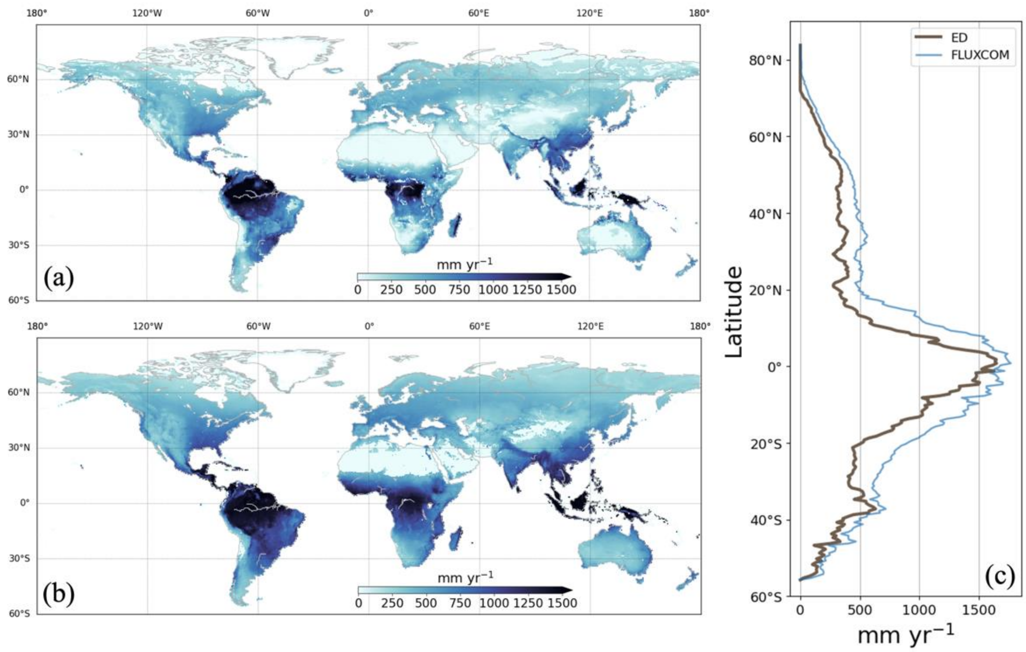


Figure 3.11. Average annual ET between 1981 and 2016 from ED (a) and FLUXCOM (b) with corresponding latitudinal average comparison (c).

3.4.4 Vegetation structure

Evaluation of vegetation structure estimates focus on leaf area and canopy height. Figure 3.12 presents the spatial distribution of growing season LAI from global ED, GEOV2, and MODIS. Growing season LAI is chosen for comparison because winter snow in the northern region (e.g., boreal forests) might affect LAI retrieval and cause uncertainties in remote sensing estimates (Murray-Tortarolo et al., 2013). There is good agreement in spatial pattern between ED and reference LAIs (Figure 3.12d), showing peaks in the tropics and boreal region (near 50°N), and relatively low estimates across temperate regions. In the tropics, ED has an average LAI of 6.0 m^2/m^2, which is similar to GEOV2 but higher than MODIS. However, ED shows higher LAI in temperate and boreal regions than both reference LAIs, specifically in southern China and Brazil. Despite the magnitude discrepancy of growing season LAI between ED and reference LAIs, there is a general agreement in the greening trend between 1999 and 2016 (as shown in Figure 3.13). The linear fitted LAI trend is 0.058 m^2/m^2 per decade for ED, 0.090 m^2/m^2 for GEOV2 and 0.046 m^2/m^2 for MODIS. However, ED shows larger interannual variation than the references. LAI seasonality is also compared across latitudinal bands in Figure 3.14, suggesting ED captures peak season in latitudinal bands 60° ~ 90°N, 45° ~ 60°N, and 60° ~ 30°S, but shows less agreement with the references in the tropics (0° ~ 15°N and 15S° ~ 0°). In addition, ED LAI in winter is larger than either reference LAI; at latitudes above 45°N, and between 30°N and 45°N, ED LAI is higher for all seasons. Similarly, higher LAI also appears in 60°S ~ 30°S, across southern China and Brazil.

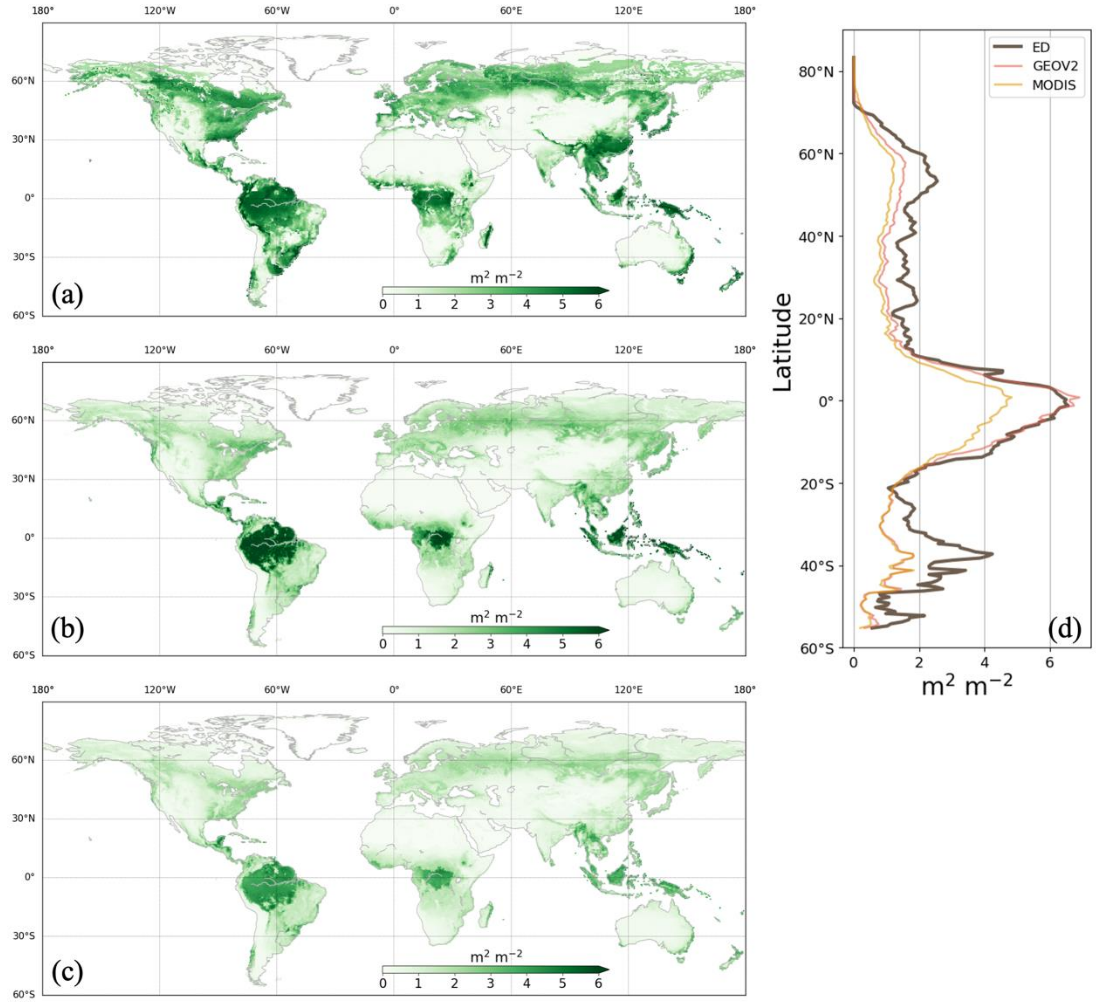


Figure 3.12. Average LAI during the growing season between 2003 and 2016 from ED (a), GEOV2 (b), and MODIS (c). Corresponding latitudinal averages are compared in (e). Growing season is defined as the months during which the average air temperature of MERRA2 is above 0°C.

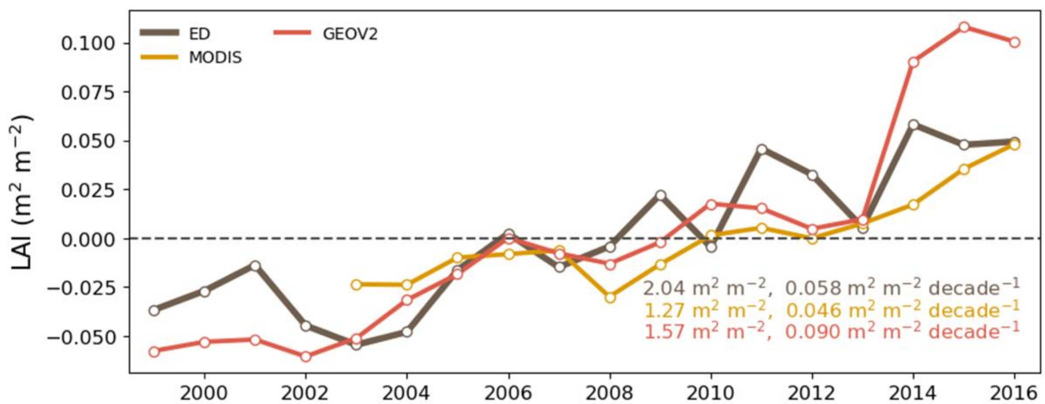


Figure 3.13. Interannual global average growing season LAI from ED, MODIS and GEOV2. The anomaly is calculated by subtracting annual LAI by multi-year average.

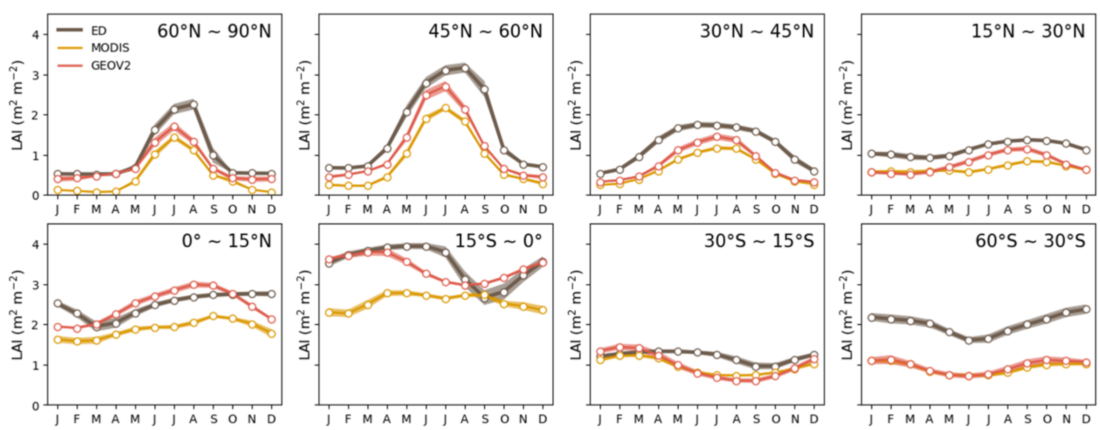


Figure 3.14. Seasonal LAI by latitudinal band from ED, MODIS and GEOV2.

ED's vertical LAI profile is compared to GEDI L2B at different height intervals between 51°N ~ 51°S where GEDI has observations. Spatially, ED and GEDI L2B both show regions where LAI becomes smaller as height increases (Figure 3.15). Consequently, most vegetated regions have LAI estimates at heights between 0 and 10 m. Only tropical rainforest and parts of southern China and the US have LAI above 30 m. The relative fraction of vertical LAI by latitudinal band indicates ED broadly agrees with GEDI L2B, where the relative fraction follows an exponential decay curve as height increase (Figure 3.16). The majority of LAI is in low canopy area (e.g., <= 15 m), where LAI under 10 m is at least 40% of total LAI in the entire canopy. Discrepancies can be seen at the 0-5 m and 10-15 m LAI interval along most of latitudinal bands.

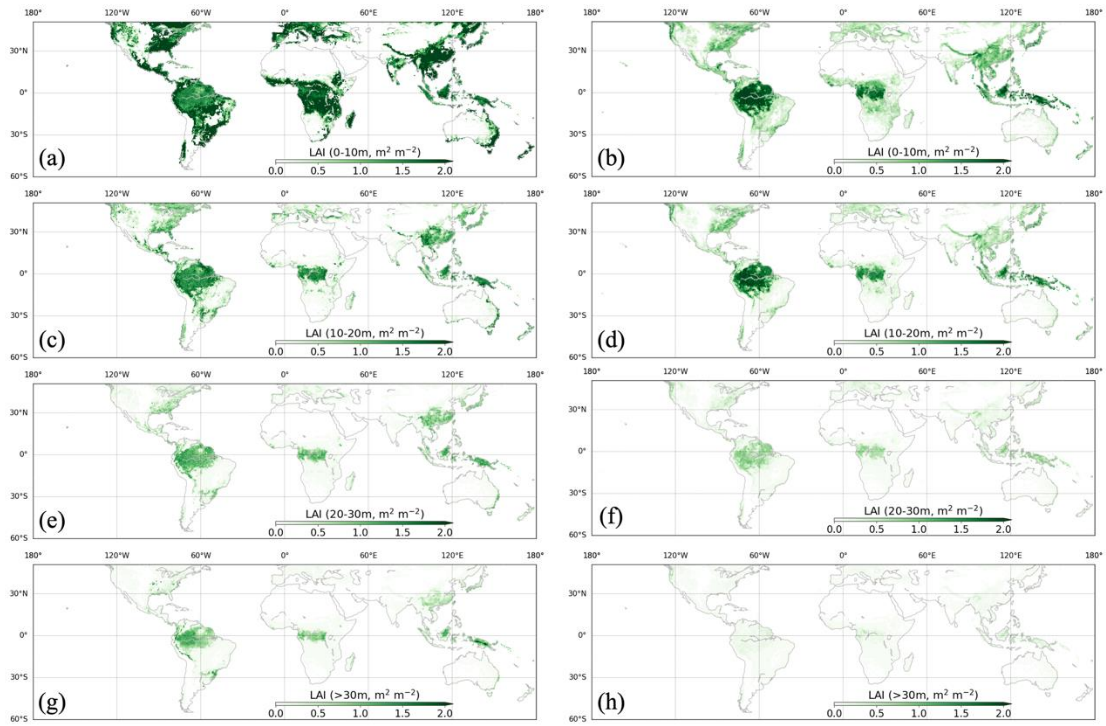


Figure 3.15. Vertical LAI from ED and GEDI L2B at height (0-10 m) in (a) and (b), 10-20 m in (c) and (d), 20-30 m in (e) and (f), and above 30 m in (g) and (h), respectively.

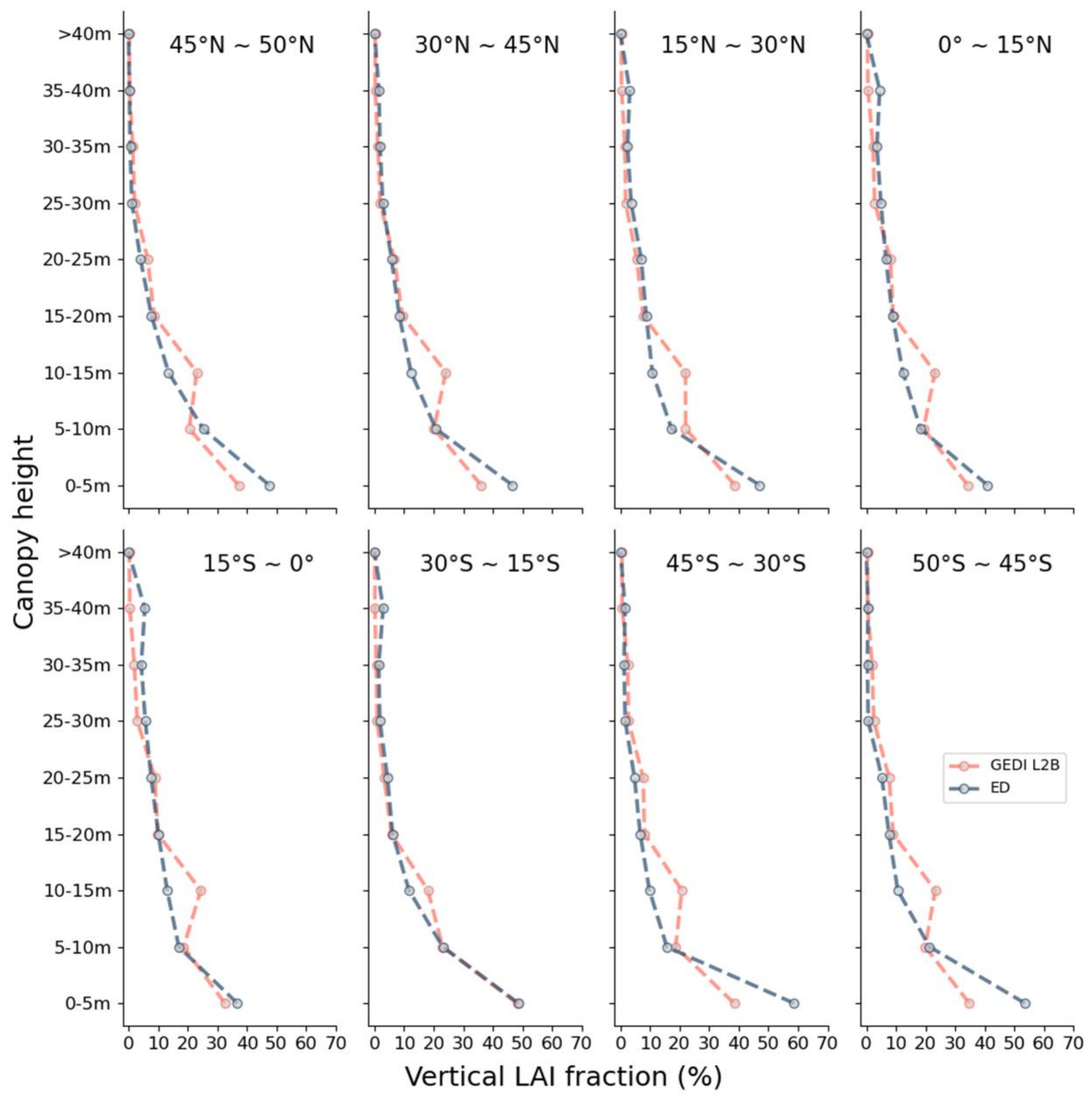


Figure 3.16. Relative fraction of vertical LAI by latitudinal band between ED and GEDI L2B.

Tree canopy height estimates from ED are compared with satellite lidar observations of GEDI and ICESat-2 in Figure 3.17. ED agrees with lidar canopy height estimates with taller trees in tropical rainforests, southern China and the eastern US. The canopy height gradient from forests to savannas in South America (northwest to southeast) and in Africa (central to north and south) are also captured by ED generally. However, ED tree height in southern China, and Brazil is higher than the references. This is a similar pattern to AGB, which may be related to land use forcing where cropland area and harvesting area of primary forest is underestimated in these regions. ED tree height is smaller than references across African savanna.

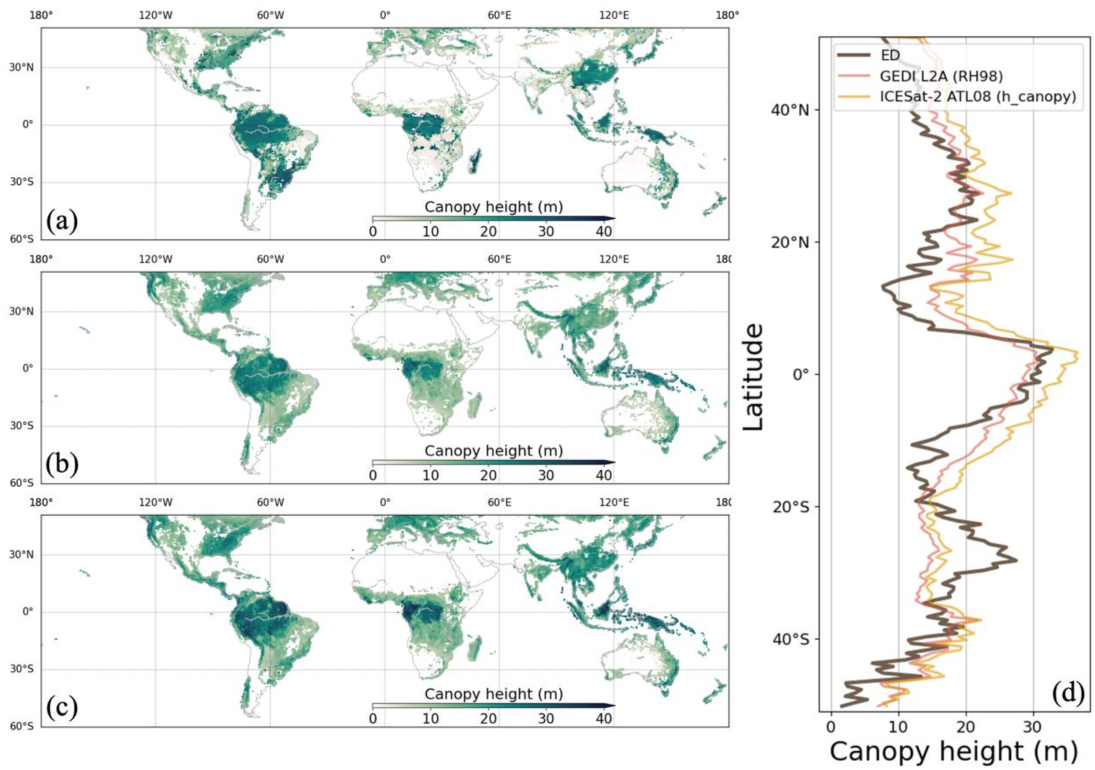


Figure 3.17. Canopy height from ED (a), GEDI L2A (b), and ICESat-2 ATL08 (c). Latitudinal averages are compared in (d). ESA CCI data grids with tree fractions below 5% are masked.

3.5 Discussion and Conclusions

In this study, we developed global ED and evaluated its capabilities to simulate vegetation dynamics and the terrestrial carbon and water cycles. Model development is built on series of regional versions and introduces modifications on submodules of PFTs, leaf physiology, soil hydrology and evapotranspiration and land use change.

The developed model is spun up from bare ground to present, simulating forest succession and associated carbon and water cycle dynamics with considerations of historical LULCC, rising CO_2 levels and climate change. Evaluations against a comprehensive benchmarking dataset suggest that the global ED well captures spatial variation of PFT, AGB, GPP, ET, LAI (both vertical and total) and canopy height across latitudes and biomes; estimates global total AGB and GPP within range of references but underestimate soil carbon; agree with references in temporal trend and interannual variability of global GPP and NBP in response to climate variability and rising CO_2.

Evaluation of AGB, GPP and NBP suggests that ED can characterize the spatial and temporal variability of the terrestrial carbon cycle. First, ED did well in its spatial representation of AGB, including transitions from the tropics to the boreal regions and from rainforests to the savanna across South America and Africa. ED estimates of global AGB (298 Pg C) in 2010 is close to reference AGB values in the same year (i.e., 283 and 297 Pg C). Second, ED reproduced the spatial distribution of reference GPPs, where the global total is close to FluxSat GPP and falls within the range of TRENDY DGVMs (Anav et al., 2015). More importantly, ED estimated annual global GPP reveals a positive trend, reflecting both FluxSat GPP and the new GPP proxy SIF. This positive trend coincides with estimates of other process-based models and satellite-driven datasets (Anav et al., 2015; O'Sullivan et al., 2020). A corresponding positive trend is also found in global average LAI values between 1999 and 2016, concurring canopy greening trend evident in other studies (Zhu et al., 2016; Piao et al., 2020). The predicted enhancement of productivity is likely a response to combined effects of CO_2 fertilization, extended growing seasons due to increasing temperatures, and forest recovery from prior disturbances, since the ED transient simulation takes into consideration of rising CO_2, transient meteorology, and LULCC. Third, ED's NBP estimate closely agrees with land sink estimates from multiple references in terms of global total magnitude and interannual variability. For example, ED reproduces the impacts of El

Niño and La Niña on the land sink, revealing the land as carbon source at strong El Niño events in 1982-1983 and 1997-1998. ED estimated NBP is generally within the range of GCB2020 DGVMs before 2007, and closer to top-down atmospheric inversions from 2008 to 2016. Yet, causes of the estimated increase in 2008-2016 NBP is unclear and additional experiments are needed to attribute it to climate- or demographic-driven sink enhancement.

Simulating physiological processes at the individual level enables ED to characterize competition between individuals and represent these processes as large-scale forest dynamics. For example, the prediction of broadleaf PFT dominance in the tropics and the large fraction of needleleaf PFTs in the boreal regions is a result of PFT competition at the individual level. Such competition affects the gain-cost carbon balance of PFT individuals and in turn, affects reproduction and mortality within ecosystems. In ED, broadleaf PFTs are defined with relatively larger leaf area, higher photosynthetic efficiency, and faster growth in height, but also with high carbon demand and turnover of active carbon components such as leaves, roots and sapwoods. The growing season length for broadleaf PFTs is determined by environmental conditions. For example, water stress and low temperatures would cause leaf senescence. In contrast, needleleaf PFTs have smaller but long-lived leaves which can photosynthesize all year round even under unfavourable environments. Therefore, warm and moist conditions in the tropics support broadleaf PFTs with high photosynthetic productivity and non-interrupted growing seasons. Such conditions enable them to grow faster in height and shade other trees, giving them a competition edge and placing competitive pressure on needleleaf PFTs growth. In contrast, cold and dry conditions in high latitudes and altitudes favour needleleaf PFTs. These cold and dry conditions limit the growing seasons and may cause freezing-induced leaf loss to broadleaf PFTs, which prevents them from maintaining a gain-cost carbon balance and reproduction. As a result, needleleaf PFTs outcompete broadleaf PFTs, as indicated by their larger spatial domain and global fraction.

Characterizing individual vegetation structure provides useful connections between ED and forest structure observations from lidar remote sensing. This connection provides specific opportunities for assessing the model's ability to capture forest structure and initialize with land use history. In ED, the vertical structure (i.e., canopy height and leaf area) of each cohort is explicitly modelled and tracked over natural succession, from seedlings to mature forests, and also over disturbance (both natural and anthropogenic) and recovery processes. In this study, following an approach commonly employed by DGVMs/ESMs to initialize the model with land use history, we spun up ED using equilibrium and transient simulations. The equilibrium simulation obtained primary ecosystems conditions corresponding to the year 850. The transient simulation obtained contemporary ecosystem conditions resulting from historical changes in land use, climate and atmospheric CO_2. The resulting forest structure, in terms of canopy height and vertical LAI, was evaluated against novel and direct forest structure observations from GEDI and ICESat-2. The evaluation suggests that ED's spin-up process generally captures observed spatial patterns of canopy height and vertical LAI. For example, ED predicted higher level LAI (>20 m) in tropical forests than forests in other regions, and LAI under 10 m accounts for 40% of total LAI regardless of biome and region. ED also reproduced the overall height gradient across transition zones from forest to savanna in South America and Africa. However, this evaluation also shows that ED produces an uneven spatial distribution at the transition zone and may not predict the height of sparse tree such as southern savanna in Africa. ED may also overestimate canopy height in southern Brazil; similar overestimates appear for GPP, AGB and total LAI.

Benchmark evaluation also suggests future developments could focus on improving ED through aspects including soil hydrology, soil carbon and nitrogen cycle. First, evaluation of GPP and LAI seasonality in the tropics indicates disagreement with reference data with regards to the timing of peak productivity and total leaf area. ED estimated GPP and LAI reach their minimum in March at 0° ~ 15°N and in September at 15°S ~ 0°, and reach maximum in July ~ September at 0° ~ 15°N and in February-April at 15°S ~ 0°. The timing of GPP maximum and minimum appears to be delayed by about a month from FLUXCOM, FluxSat and SIF estimates. A similar delay also occurs in LAI. Despite ED NBP responsiveness to El Niño, which reduces the sink or indicates a carbon source in some years, NBP response in ED appears stronger than the reference data, and this primarily comes from tropics (Figure 3.10b). The mismatch in GPP and LAI timing between ED and references and ED' strong response to El Niño may be related to over-intensified soil water stress on leaf productivity during dry seasons owing to current soil hydrology submodule. Current submodule is a single layer bucket model with soil depth up to 1-m, this submodule reduces soil water holding capacity and could not characterize mechanisms like deep roots assessing water at depth and hydraulic redistribution between deep and near-surface soil layers (Baker et al., 2008; Poulter et al., 2009).

Second, ED estimate of global soil carbon is lower than the HWSD and shows large disagreement in spatial distribution at the grid scale. Underestimation primarily comes from cold regions and biomes. Spatial correlation at the grid-scale between global ED and HWSD is very low, which is a well-known issue in CMIP5 ESMs (Todd-Brown et al., 2013). Most ESMs and DGVMs presume soil carbon as a product of NPP-related litter accumulation and decomposition primarily controlled by temperature and moisture, however, NPP and soil temperature could only explain 10% of spatial variation in HWSD. Future model development could explore the additional mechanisms of the soil carbon cycle as well as uncertainties related to empirical soil carbon maps including HWSD.

Third, future model could improve nitrogen cycle by introducing processes including biological fixation and atmospheric decomposition, and also by further parameterizing and evaluating at global scale. The nitrogen cycle submodule has been developed in prior versions of ED and coded in this global ED model, however, this submodule was deactivated and not calibrated in our simulations. Current submodule contains several belowground pools (i.e., structural metabolic, slow and mineralized), and decomposition rates vary over time and space as function of soil moisture, temperature, texture, and lignin-to-nitrogen ratio. The submodule includes mineralization processes of organic nitrogen from vegetation and nitrogen loss via leaching, however, other processes such as biological fixation and atmospheric decomposition are not included, which may result in nitrogen pools unable to each equilibrium.

Our development and evaluation of global ED demonstrates the model's ability to characterize essential aspects of terrestrial vegetation dynamics and the carbon cycle. This model has also been integrated with NASA's Goddard Earth Observing System, Version 5 (GEOS-5) to forecast seasonal biosphere-atmosphere CO_2 fluxes in 2015-16 El Niño (Ott et al., 2018), and used in NASA's Carbon Monitoring System to develop regional forest carbon modeling estimates for climate mitigation planning (Ma et al., 2021). Future work will focus on addressing limitations as discussed above and making direct connections with lidar forest structure observations from GEDI and ICESat-2 to improve demographic processes, and the quantification and attribution of the terrestrial carbon cycle.

Chapter 4 High-resolution forest carbon modeling for climate mitigation planning over the RGGI region, USA

Abstract

The inclusion of forest carbon in climate change mitigation planning requires the development of models able to project potential future carbon stocks - a step beyond traditional monitoring, reporting and verification (MRV) frameworks. Here, we updated and expanded a high-resolution forest carbon modeling approach previously developed for the state of Maryland to 11-states in the Regional Greenhouse Gas Initiative (RGGI) domain, which includes Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. In this study, we employ an updated version of the Ecosystem Demography (ED) model, an improved lidar initialization strategy, and an expanded calibration/validation approach. High resolution (90 m) wall-to-wall maps of present aboveground carbon, aboveground carbon sequestration potential, aboveground carbon sequestration potential gap, and time to reach sequestration potential were produced over the RGGI domain where airborne lidar data were available, including 100% of eight states, 62% of Maine, 12% of New Jersey, and 0.65% of New York. For the eight states with complete data, an area of 228,552 km^2, the contemporary forest aboveground carbon stock is estimated to be 1,134 Tg C, and the forest aboveground carbon sequestration potential gap is estimated to be larger at >1,770 Tg C. Importantly, these estimates of the potential for added aboveground carbon sequestration in forests are spatially resolved, are further partitioned between continued growth of existing trees and new afforested/reforested areas, and include time estimates for realization. They are also assessed for sensitivity to potential changes in vegetation productivity and disturbance rate in response to climate change. The results from this study are intended as input into regional, state, and

local planning efforts that consider future climate mitigation in forests along with other land-use considerations.

4.1 Introduction

Over the past decade, terrestrial ecosystems have sequestered one-third of global fossil fuel emissions, with forests especially critical for mitigating climate change (Friedlingstein et al., 2019, Pan et al., 2011). Initiatives to increase the capacity of the terrestrial carbon sink through afforestation and reforestation are actively being implemented from local to global scales (Griscom et al., 2017, Fargione et al., 2018). Such efforts utilize a range of scientific approaches to quantify forests' carbon capture benefit, often relying on national or continental level estimates (Depro et al., 2008, Roxburgh et al., 2006, Rhemtulla et al., 2009). Higher-resolution information on present and potential future forest carbon stocks could further support policy development and management activities and provide important insights for climate mitigation planning (Cook-Patton et al., 2020, Goldstein et al., 2020). In particular, there is a growing need for advanced forest carbon models to quantify potential future options for additional carbon storage in forests by integrating scientific advances in remote sensing and modeling (Gibbs et al., 2007, Achard et al., 2007, Hurtt et al., 2014, Reinmann et al., 2020).

To be most useful, forest models should contain multiple features. First, they should be highly accurate. While the accuracy of future projections is difficult to ascertain by itself, abundant data exists for the present across a wide range of heterogeneous conditions that can be used in calibration and validation. Moreover, mapping present and modeling future carbon within the same framework can help ensure data consistency and maintain logical consistency between present and future ecosystem condition estimates. This capability may be particularly relevant for carbon budgets, which seek to track emission reductions relative to established mitigation goals and baselines.

Second, models should be able to quantify forest carbon at high spatial resolution. Land-use decisions must ultimately be implemented at local scales, as climate mitigation activities for terrestrial carbon relate directly to land ownership. Furthermore, historical and ongoing deforestation and degradation has intensively fragmented forests and resulted in high spatial heterogeneity in forest carbon distribution (Haddad et al., 2015, Turner et al., 2003, Ordway and Asner 2020).

Third, models should operate across a range of policy relevant spatial domains including local, state, regional, and ultimately national and global scales. Modeling across large domains supports consistent carbon estimates over space and minimizes discrepancies between jurisdictional estimates, offering greater flexibility for data users. A system crossing large spatial domains could particularly benefit multi-actor collaboratives, which may require a common scientific approach in support of forest carbon trading and reporting schemes.

Advances in forest ecosystem modeling and remote sensing over the past two decades offer the potential to begin to meet these needs. A new generation of forest ecosystem models have been developed to efficiently track more detailed ecological processes over large spatial scales (Hurtt et al., 1998, Moorcroft et al., 2001, Fisher et al., 2018). In remote sensing, optical data have increased in resolution, and in combination with Light Detection and Ranging (lidar) has enabled accurate and high-resolution measurements of forest vertical structure (Dubayah and Drake 2000, O'Neil-Dunne et al., 2014a, Drake et al., 2002). Together, these advances have been combined to provide accurate initial conditions for forest models (Hurtt et al., 2004, 2010, 2016, Thomas et al., 2008, Antonarakis et al., 2011, 2014).

In parallel, these same advances have enabled high-resolution empirical biomass mapping which can provide important baseline maps and be used to validate forest model estimates (Huang et al., 2015, Huang et al., 2019, Tang et al., 2021). Operationally, a prototype of this integration has been developed for and implemented over the state of Maryland (USA), where the Ecosystem Demography model (ED) was initialized with 1 m airborne lidar forest height and optical imagery to generate 90 m 'wall-to-wall' maps of contemporary forest aboveground carbon stocks, future aboveground carbon sequestration potential, and the time to achieve it in years (Hurtt et al., 2019b).

This study aims to improve and expand the approach to integrated forest modeling and lidar remote sensing another order of magnitude, to an important multi-state region. Here, we focus forest aboveground carbon stocks and future forest aboveground carbon sequestration potential over an 11-state region consisting of 10 members of the Regional Greenhouse Gas Initiative (RGGI), the nation's first mandated cap-and-trade program for CO_2 emissions, plus Pennsylvania, which is expected to join RGGI by 2022. We refer to these states collectively as the 'RGGI' region. The RGGI region encompasses a land area nearly 10 times that of Maryland and includes a large gradient of topography, temperature and precipitation.

4.2 Data and Methods

4.2.1 Study area

The 11-state RGGI domain of this study encompasses about 281,695 km^2 of total land area, including 12,542 km^2 in Connecticut, 5,047 km^2 in Delaware, 23,187 km^2 in New Hampshire, 2,362 km^2 in New Jersey, 804 km^2 in New York, 49,977 km^2 in Maine, 25,142 km^2 in Maryland, 20,202 km^2 in Massachusetts, 115,883 km^2 in Pennsylvania, 2,678 km^2 in Rhode

Island, and 23,871 km^2 in Vermont, respectively. Maine, New Jersey and New York land area is only partially included because of limited lidar coverage over these states.

There are five US EPA level II ecoregions including Mixed Wood Plains, Southeastern USA Plains, Ozark, Ouachita-Appalachian Forests, Mississippi Alluvial and Southeast USA Coastal Plain, and Atlantic Highlands. Across the domain, the annual average temperature ranges from -1.4 °C to 14.4 °C, total annual precipitation ranges from 795 mm/yr to 2,178 mm/yr. Most of this area belongs to the Appalachian Highlands with dense forest, and the Atlantic Coastal Plain, which is generally flat and low in elevation. Elevation ranges from sea level to 1,917 m at Mount Washington in the state of New Hampshire. Dominant forest types are deciduous forests (i.e., maple-beech-birch, oak-hickory and aspen-birch).

Similar to most of the eastern US, forests in the RGGI region were intensively cleared for agricultural expansion by the mid-19th century, with partial reforestation and restoration after agricultural abandonment and westward relocation. In the mid to late 20th century, significant loss and fragmentation occurred due to urbanization. As a result, land-cover in the RGGI region is comprised of 59% forests and 16% cropland and pasture according to the 2011 National Land Cover Dataset (NLCD).

All states in the region have greenhouse gas (GHG) reduction goals based on state legislation or gubernatorial directives. While the states are actively pursuing efforts in support of climate mitigation, the degree to which forest carbon is currently included in climate planning is highly variable (Lamb et al., 2020a). Ten of the eleven states in the study domain are members of the US Climate Alliance (USCA) (excepting New Hampshire) and currently participate in the USCA's National and Working Lands Challenge, which commits states to integrating priority actions and pathways for land-based carbon into state GHG mitigation plans. Furthermore, as members of the NASA Carbon Monitoring System's Multi-State Working Group, all eleven states have signaled an interest in better utilizing advanced tools and technologies to better estimate forest carbon stocks and fluxes.

4.2.2 Model

This study uses the ED model which is an individual-based prognostic ecosystem model. By integrating submodules of growth, mortality, hydrology, carbon cycle and soil biogeochemistry, ED can track plant dynamics including growth, mortality and reproduction. Along with plant dynamics, ED can track the carbon cycle, including carbon uptake by leaf photosynthesis, carbon allocation to biomass growth in leaves, roots and stems, carbon redistribution from plants to soil based on plant tissue turnover from dead plants due to mortality and disturbance, carbon decomposition in various pools (metabolic litter pool, structural litter pool, soil slow pool, soil passive pool, wood product pool, harvested crop pool, etc.) as well as carbon combustion from fire. Over the last two decades, ED has been continuously developed and combined with lidar and land-use change data to predict ecosystem dynamics and associated water and carbon fluxes across spatial scales (e.g., site, regional and global) and temporal scales (short-term seasonal to long-term decadal and century) (Hurtt et al., 2002, 2004, 2010, 2016, Fisk et al., 2013, Flanagan et al., 2019). ED distinguishes itself from most other ecosystem models by explicitly tracking vegetation structure and scaling fine-scale physiological processes to large-scale ecosystem dynamics (Hurtt et al., 1998, Moorcroft et al., 2001, Fisher et al., 2018). In ED, vegetation structure (e.g., height and diameter at breast height) and physiological processes (e.g., leaf photosynthesis and phenology) are modelled at the individual scale, where individual plants compete mechanically for light, water and nutrients. Explicitly modeling vegetation height facilitates a potential connection to lidar data. The most advanced version of ED was used in this study and it has been recently calibrated and evaluated globally by various benchmarking datasets such gross primary productivity (GPP), leaf area index (LAI), aboveground biomass (AGB), and net biome productivity (NBP) (Ma et al., 2020b).

4.2.3 Data

4.2.3.1 Model drivers

Meteorological variables used to drive ED come from NASA Daymet, available from 1980 to 2017 at daily temporal resolution and 1km spatial resolution (Thornton et al., 2016), and MERRA2, available from 1980 to 2017 at hourly temporal resolution and 0.5° spatial resolution (Gelaro et al., 2017). As ED requires hourly meteorological variables to compute leaf carbon assimilation and transpiration, the climatology data from Daymet was interpolated to hourly metrics using the MERRA2 climatology. Specifically, hourly air temperature, humidity and shortwave radiation were calculated using following equations:

$$T_{hr} = \frac{T_{max,D} - T_{min,D}}{T_{max,M} - T_{min,M}}\left(T_{hr,M} - T_{min,M}\right) + T_{min,D} \tag{4.1}$$

$$e_{hr} = \frac{\bar{e}_{hr,D}}{\bar{e}_{hr,M}} e_{hr,M} \tag{4.2}$$

$$SW_{hr} = \frac{\overline{SW}_{hr,D}}{\overline{SW}_{hr,M}} SW_{hr,M} \tag{4.3}$$

Where $T_{hr,M}$, $e_{hr,M}$ and $SW_{hr,M}$ are MERRA2 air temperature, air humidity and shortwave radiation at hour (hr), respectively. $\bar{e}_{hr,M}$ and $\overline{SW}_{hr,M}$ are MERRA2 daily average air humidity and shortwave radiation, respectively. $T_{max,M}$ and $T_{min,M}$ are MERRA2 daily maximum and minimum air temperature, respectively. $\bar{e}_{hr,D}$ and $\overline{SW}_{hr,D}$ are Daymet daily average air humidity and shortwave radiation, respectively, and $T_{max,D}$ and $T_{min,D}$ are Daymet daily maximum and minimum air temperature, respectively.

CO_2 concentration was held constant at 360 ppm both spatially and temporally, a value near the middle of the CO_2 range between 1981 and 2017 (Keeling 2008). Physical soil and hydraulic property inputs are from Probabilistic Remapping of SSURGO (POLARIS) (Chaney et al., 2016) and CONUS-SOIL (Miller and White 1998). The POLARIS dataset remaps the SSURGO database and fills gaps using a machine learning model and high-resolution geospatial environmental data. The soil water module of ED calculates water content and percolation rate based on saturated hydraulic conductivity, saturated water content and Van Genuchten water retention curve parameters from POLARIS and depth to bedrock data from CONUS-SOIL.

The annual average air temperature, annual total precipitation from NASA Daymet and soil depth from CONUS-SOIL are shown in Figure C.2.

4.2.3.2 Canopy cover and height

ED was initialized with canopy height and tree canopy cover maps to generate aboveground biomass (AGB). The canopy height map was derived from a 1 m lidar Canopy Height Model (CHM) (O'Neil-Dunne et al., 2014a, 2014b). Utilizing suggested height metrics from Hurtt et al., 2019b and ED's native 10 diameter canopy scale, 1 m CHM was first aggregated to 10 m by extracting the max canopy height and then averaging to 90 m (hereafter referred to as lidar canopy height). The tree canopy cover map was derived using object-based approach that integrated the lidar data and multi-spectral optical images from the National Agricultural Imagery Program (NAIP). NAIP optical images with 1 m spatial resolution was first classified as tree cover in conjunction with lidar canopy height, and then aggregated to 90 m (O'Neil-Dunne et al., 2014a, 2014b). The lidar canopy height and NAIP tree canopy cover at 90 m resolution are shown in Figure 4.1, with lidar acquisition year shown in Figure C.24. Lidar and NAIP data were available for 0.65% of New York, 12% of New Jersey, 62% of

Maine and 100% of Connecticut, Delaware, Maryland, Massachusetts, New Hampshire, Pennsylvania, Rhode Island and Vermont. Specifically, lidar and NAIP data cover the New York counties of Bronx, Kings, New York, Queens and Richmond as well as the New Jersey counties of Morris, Sussex, and Warren.

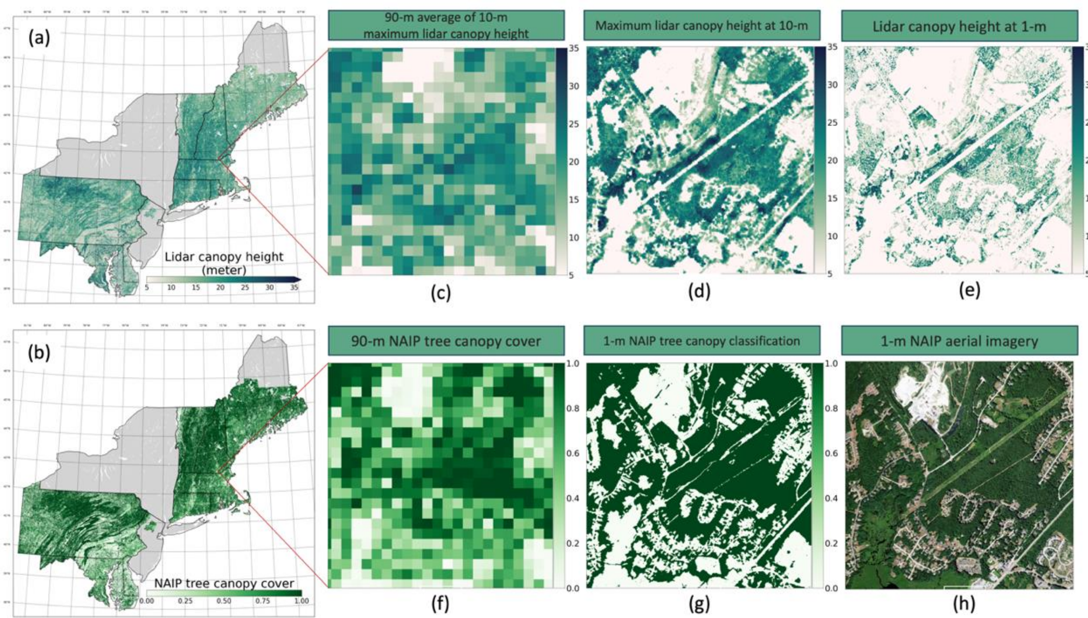


Figure 4.1 90 m lidar canopy height (a) and NAIP tree canopy cover (b) over the RGGI region. Using a sample region in New Hampshire, (c)-(h) illustrate the process of 90 m lidar canopy height and NAIP tree canopy cover generation, where (e) and (h) are lidar canopy height and NAIP aerial imagery at 1 m resolution; (d) utilizes (e) to identify the maximum lidar canopy height over 10 m land cells; (g) is NAIP tree canopy classification of (h) at 1 m. (c) and (f) are derived by averaging (d) and (g) respectively to 90 m resolution.

4.2.3.3 Land cover

Land cover of non-forested wetland, inland water, and impervious surface was excluded from the analysis. Specifically, land-cover types of open water and herbaceous wetland in NLCD 2011 were aggregated from 30 m to 90 m by counting the percentage of each land-cover type, and Percent Developed Imperviousness in NLCD 2011 was aggregated to 90 m by averaging the related 30 m fraction (Jin et al., 2013, Xian et al., 2011). Then the aggregated 90 m data was used to proportionally exclude water, wetland and impervious surface from each land cell.

4.2.4 Model Initialization, Projection and Evaluation

This study generally follows the initialization and projection approach used in Hurtt et al., 2019b, but proposes a modification to the initialization method (here defined as weighting-based initialization method) to improve AGB estimates where ED-modelled canopy height saturates. Full details of this initialization and projection process can be found in Appendix C.1. We simulated ecosystem succession from bare ground to mature forest by running the ED model for 500 years with meteorology, CO_2 and soil properties as inputs (described in section 4.2.3.1). This model run generated a gridded lookup table that stores a time-series trajectory of AGB and canopy height over the course of succession for each 90 m land cell. In post-processing, ED's stored AGB-height lookup table was then initialized and projected with maps of lidar canopy height, NAIP canopy tree cover and NLCD impervious fraction.

ED initialization combined the stored lookup table with lidar height and tree canopy cover to estimate contemporary aboveground biomass (hereafter referred as ED initialized AGB). The ED initialized AGB represents the present aboveground carbon stock of existing trees at the time of lidar acquisition, and it sets the successional baseline for future carbon sequestration. ED then projections AGB growth from ED initialized AGB to the maximum potential AGB based on current meteorology, CO_2 and soil properties for every 90 m land cell. This projection does not reduce current impervious surface area, nor does it consider land-use change projections or local laws restricting areas for afforestation or reforestation. Future carbon sequestration was calculated from both the continued growth of existing trees, as identified by the NAIP canopy cover map (hereafter referred as continued growth), and newly reforested or afforested trees, as simulated on any proportion of the 90 m grid cell not otherwise covered by impervious surface, open water, or herbaceous wetland (hereafter referred as regrowth). Following Hurtt et al., 2019b, several projection-related metrics were defined: 95% of the future maximum AGB the land can reach was defined as the carbon sequestration potential (CSP), the difference between CSP and ED initialized AGB was defined as carbon sequestration potential gap (CSPG), the time in years required to reach CSP from present AGB was defined as carbon sequestration potential time gap (CSPTG). ED initialization and projection processes were only completed for areas where lidar canopy height and NAIP tree canopy cover data were available (Section 4.2.3.2). Note that this study's projections of CSP, CSPG and CSPTG only involve aboveground carbon, because it is observable component of forest carbon from lidar remote sensing.

ED initialized AGB was directly validated using aboveground live biomass estimates from US Forest Service Forest Inventory and Analysis (FIA) plots (hereafter referred to as FIA plot AGB). For this study, 4,540 FIA plots were extracted from the FIA database based on three criteria: (1) geographically located within RGGI regional domain; (2) estimates were reported within period of 2004 to 2015; and (3) sites are in forest condition and with AGB larger than zero. For each FIA plot, all overlapping 90 m land cells of ED initialized AGB were averaged to compare against FIA plot AGB. Spatial mismatches between FIA plots and ED initialized AGB may affect evaluation, especially in highly fragmented forest, because the FIA plot footprint (about 0.4 ha) is smaller than the land cell size of ED initialized AGB maps (about 0.8 ha).

ED initialized AGB was also compared to wall-to-wall AGB maps including AGB estimated from lidar-informed empirical models (Huang et al., 2019, Tang et al., 2021), hereafter as Lidar empirical AGB, and to other existing AGB products (Blackard et al., 2008, Saatchi et al., 2012, Kellndorfer et al., 2013, Wilson et al., 2013, Santoro et al., 2018). These wall-to-wall AGB maps vary widely in terms of input data, modeling method and spatial resolution. However, comparison to these maps allowed for further evaluation of ED initialized AGB over regions with a lack of FIA plots with trees measured, such as non-forested areas, and allowed us to examine the efficacy of using high-resolution (1 m) lidar and NAIP imagery data to capture fine-scale AGB heterogeneity.

4.3 Results

4.3.1 ED Initialization and evaluation

ED initialized AGB is shown in Figure 4.2a. The AGB spatial pattern corresponds well to landcover, lidar canopy height and topography. Generally, high AGB occurs in mountainous areas along the Appalachians Mountains in western Maryland, northcentral Pennsylvania, and southern Vermont and New Hampshire. Low AGB occurs in populated and cultivated areas such as eastern Maryland, south-eastern Pennsylvania, most of the Atlantic Coastal Plain, eastern Massachusetts, and the Connecticut River Valley. Table 4.1 summarizes average AGB density and aboveground carbon stocks from ED by state. Southern states such as Maryland, Delaware and Pennsylvania have relatively lower AGB densities (< 100 Mg/ha) than the New England states; this pattern was also identified by inspection of the NAIP imagery. Aboveground carbon stock in the RGGI region is estimated as 1,134 Tg C for 228,552 km^2, excluding the states of Maine, New York and New Jersey because of their partial coverage of lidar and NAIP data.

Table 4.1 Statewide average NAIP tree cover, average AGB density and carbon stocks of ED initialized AGB, as well as CSP, CSPG and average CSPTG for the states of Connecticut, Delaware, Maryland, Massachusetts, New Hampshire, Pennsylvania, Rhode Island and

Vermont.

	NAIP Tree Canopy Cover (%)	ED Initialized AGB		CSP		CSPG		CSPTG (year)
		Density (Mg/ha)	Stocks (Tg C)	Density (Mg/ha)	Stocks (Tg C)	Density (Mg/ha)	Stocks (Tg C)	
Connecticut	68.8	119.2	75.4	247.2	156.3	139.8	81.3	231
Delaware	35.8	43.7	11.1	251.8	64.0	212.6	53.0	292
Maryland	48.8	80.1	101.7	270.0	343.1	203.5	242.1	262
Massachusetts	69.7	111.2	114.1	221.3	227.0	121.6	113.6	229
New Hampshire	85.5	133.8	156.5	224.6	262.7	100.9	107.2	207
Pennsylvania	63.7	90.7	528.1	265.8	1547.5	183.8	1022.0	279
Rhode Island	69.5	111.8	15.3	237.0	32.4	134.0	17.2	242
Vermont	79.6	109.7	131.8	225.6	270.9	122.4	139.7	246

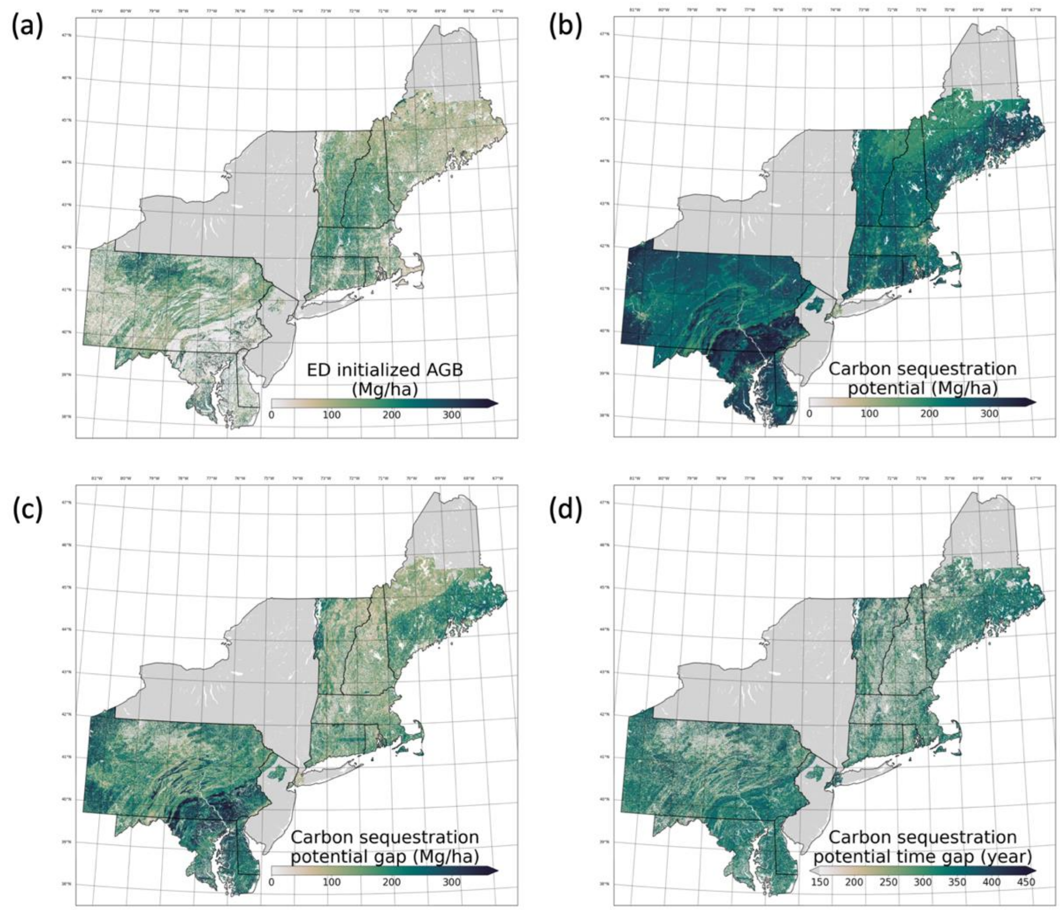


Figure 4.2 RGGI region maps of (a) ED initialized AGB; (b) CSP; (c) CSPG, defined as the difference between CSP and initialized AGB; (d) CSPTG, defined as time in years to reach to carbon sequestration potential from initialized AGB.

ED initialized AGB correlates with AGB from FIA plots and lidar wall-to-wall maps at grid-level comparison (Figure 4.3). Using over 4,000 plots, Figures 4.3a and b suggest that ED initialized AGB explains moderate variations in FIA plot AGB (R^2 of 0.35), and also shows close alignment with the 1:1 line with a bias of 7.22 Mg/ha and RMSE of 61.87 Mg/ha. Relatively stronger agreement is shown in Figures 4.3c and d, with comparison of ED to lidar empirical AGB; there was a higher R^2 of 0.85 and smaller RMSE of 29.54 Mg/ha. The comparison of ED to wall-to-wall lidar empirical AGB includes an expansive spatial domain (about 34.9 million 90 m land cells) as well as non-forest area not otherwise sampled by FIA plots with measured trees. Despite the overall high correlation between these two maps, ED initialized AGB differs from lidar empirical AGB for larger ED-based AGB (> 200 Mg/ha), where there are fewer lidar-based AGB estimates above 200 Mg/ha. Further comparison of lidar empirical AGB to FIA plot AGB in Figure C.3 indicates that lidar empirical AGB is likely to underestimate AGB where estimates exceed 250 Mg/ha. ED initialized AGB using the mid-point method is also compared to FIA plot AGB in Figure C.4, which shows larger RMSE and bias than the weighting-based method in Figures 4.3a and b. Comparison between Figures 4.3a-b and Figures C.4a-b highlights the improvements gained by using the weighting-based method, such as increased correlation between ED initialized AGB and FIA plot AGB and correction for the overestimation of AGB after height saturation.

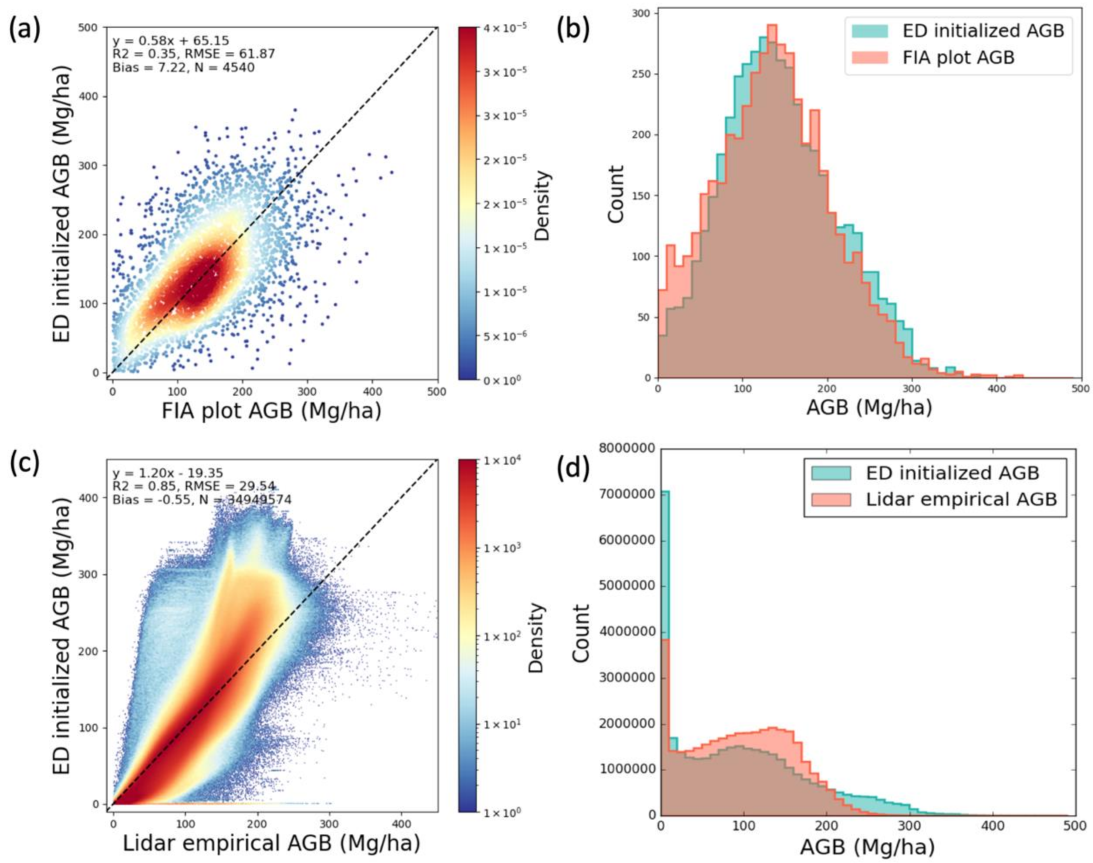


Figure 4.3 Density scatter plots and histograms comparing ED initialized AGB to FIA plot AGB in (a) and (b), and to lidar empirical AGB in (c) and (d) for all 90 m land cells. For (a) and (b), the corresponding ED initialized AGB is obtained by averaging original 90 m ED initialized AGB over overlapping land cells within the bounded circle area of four FIA subplots (about 40 m in radius).

ED initialized AGB was also compared to the lidar empirical AGB and existing wall-to-wall AGB maps at both the state and county levels. Figure 4.4 illustrates that, at county level, ED initialized AGB has relatively more agreement (i.e., RMSE and bias) with lidar empirical AGB, NBCD (Kellndorfer et al., 2013), Saatchi et al., 2012 and GlobBiomass (Santoro et al., 2018) than with the Blackard et al., 2008 and Wilson et al., 2013 maps. Stronger correlation among datasets can be found when comparing carbon stocks rather than average densities at the county level. Aboveground carbon stocks at the state level (Table 4.2) also show closer agreement among ED initialized AGB, lidar empirical AGB, NBCD, Saatchi et al., 2012, and

GlobBiomass. Aboveground carbon stock from ED initialized AGB is estimated as 1,134 Tg C in the RGGI region (excl. Maine, New Jersey and New York because of incomplete lidar and NAIP data coverage), which is within range of other AGB maps (939 ~ 1,152 Tg C).

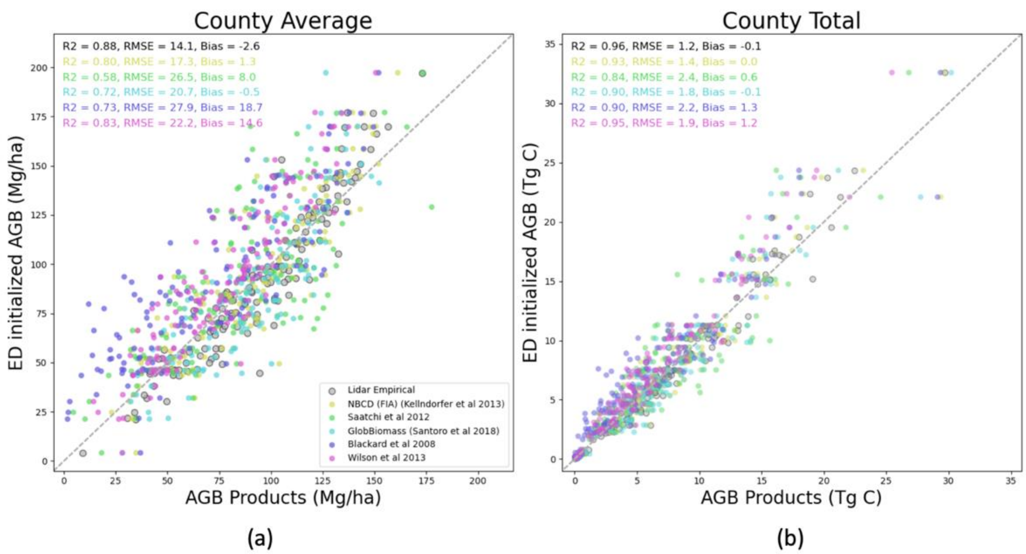


Figure 4.4 Comparison of ED initialized AGB to lidar empirical AGB and existing AGB products, including NBCD (Kellndorfer et al., 2013), Saatchi et al., 2012, GlobBiomass (Santoro et al., 2018), Blackard et al., 2008, and Wilson et al., 2013, for county-wide (a) average AGB density and (b) carbon stocks.

Table 4.2 Statewide aboveground carbon stocks (Tg C) estimated by ED initialized AGB, lidar empirical AGB, and existing AGB products, including NBCD, Saatchi et al., 2012, GlobBiomass, Blackard et al., 2008 and Wilson et al., 2013. Superscripts represent deviation degree between ED and other AGB products. The + indicates that the estimate is greater than ED initialized AGB, the * indicates that estimate is lower than ED initialized AGB. The number of +/* symbols next to each estimate represents the relative difference at the intervals of 0-10%, 10-20%, 20-30%, 30-40%, 40-50%. For example, ***/+++ represents either a -20 to -30% or 20 to 30% difference from ED initialized AGB.

State	ED Initialized AGB (Tg C)	Lidar Empirical AGB (Tg C)	NBCD (FIA) (Kellndorfer *et al.*, 2013) (Tg C)	Saatchi et al., 2012 (Tg C)	GlobBiomass (Santoro *et al.*, 2018) (Tg C)	Blackard et al., 2008 (Tg C)	Wilson et al., 2013 (Tg C)
Connecticut	75.4	71.9*	71.9*	53.5***	66.4**	50.6****	54.3***
Delaware	11.1	12.8^{++}	10.8*	11.2^{+}	13.0^{++}	8.1***	11.9^{+}
Maryland	101.7	101.7^{+}	81.3***	87.3**	98.7*	64.9****	79.7***
Massachusetts	114.1	110.3*	111.7*	82.0***	111.0*	80.9***	87.8***
New Hampshire	156.5	149.3*	150.0*	120.8***	151.8*	138.2**	126.6**
Pennsylvania	528.1	560.8^{+}	548.8^{+}	561.4^{+}	550.4^{+}	453.2**	474.0**
Rhode Island	15.3	15.1*	12.0***	10.6****	13.2**	9.0*****	9.5****
Vermont	131.8	130.4*	139.6^{+}	116.6**	138.0^{+}	134.0^{+}	119.3*

4.3.2 ED Projection of carbon sequestration potential

The ED-projected CSP for the RGGI region is shown in Figure 4.2b. The spatial pattern of CSP more generally reflects the heterogeneity of environmental conditions rather than current landcover. Areas of higher sequestration potential tend to appear in regions with deep soil or warmer air temperatures (Figure C.2). For example, warmer air temperatures together with deeper soil in central Maryland and south-eastern Pennsylvania result in larger sequestration potential estimates than in other areas of the region. Relatively cooler air temperatures in the White Mountains lead to lower sequestration potentials in northeastern Vermont, northern New Hampshire and western Maine. The statewide total aboveground CSP is estimated at 156.3 Tg C in Connecticut, 64.0 Tg C in Delaware, 343.1 Tg C in Maryland, 227.0 Tg C in Massachusetts, 262.7 Tg C in New Hampshire, 1547.5 Tg C in Pennsylvania, 32.4 Tg C in Rhode Island and 270.9 Tg C in Vermont. Statewide CSP of Maine, New Jersey and New York is not reported here because of their incomplete lidar and NAIP data coverage.

The projected maps of CSPG and CSPTG are shown in Figure 4.2c and 4.2d, indicating strong spatial variation across the RGGI region. As expected, large sequestration gaps and longer sequestration time gaps generally appear where present AGB is low, such as in eastern and central Maryland, southeastern Pennsylvania, and southeastern Maine. Relatively smaller gaps are located in northcentral Pennsylvania and western Maryland, most of Rhode Island, Connecticut, Massachusetts, Vermont and New Hampshire. Statewide average CSPTG ranges from 207 to 292 years, and correlates to the relative fraction of CSPG to CSP. The longest and shortest CSPTG appear in Delaware and New Hampshire, respectively. Statewide, Delaware is currently at 1/5 of its aboveground carbon sequestration potential, Maryland and Pennsylvania are at about 1/3, Rhode Island, Connecticut, Massachusetts, Vermont and New Hampshire are at about 1/2. The relatively larger CSPG in Delaware is likely due to its low tree cover and young forest. As Table 4.1 indicates, Delaware has the lowest tree cover (i.e., 35.8%) in the region, less than half that of highly forested states such as New Hampshire and Vermont. Despite substantial crop abandonment in 2008 and 2016 (Lark et al., 2020), tree cover in DE is only 30% lower than the adjacent state of Maryland, but its AGB density is only half that of Maryland. This difference implies that Delaware forests are younger. The CSPG is also stratified by continued growth areas and regrowth area for each state (shown in Figure 4.5) and each county (shown in Figure C.8-C.15). The stratification in Figure 4.5 indicates that large gaps are primarily located in regrowth area for Maryland, Delaware and Pennsylvania, but primarily in continued growth areas for the other six states.

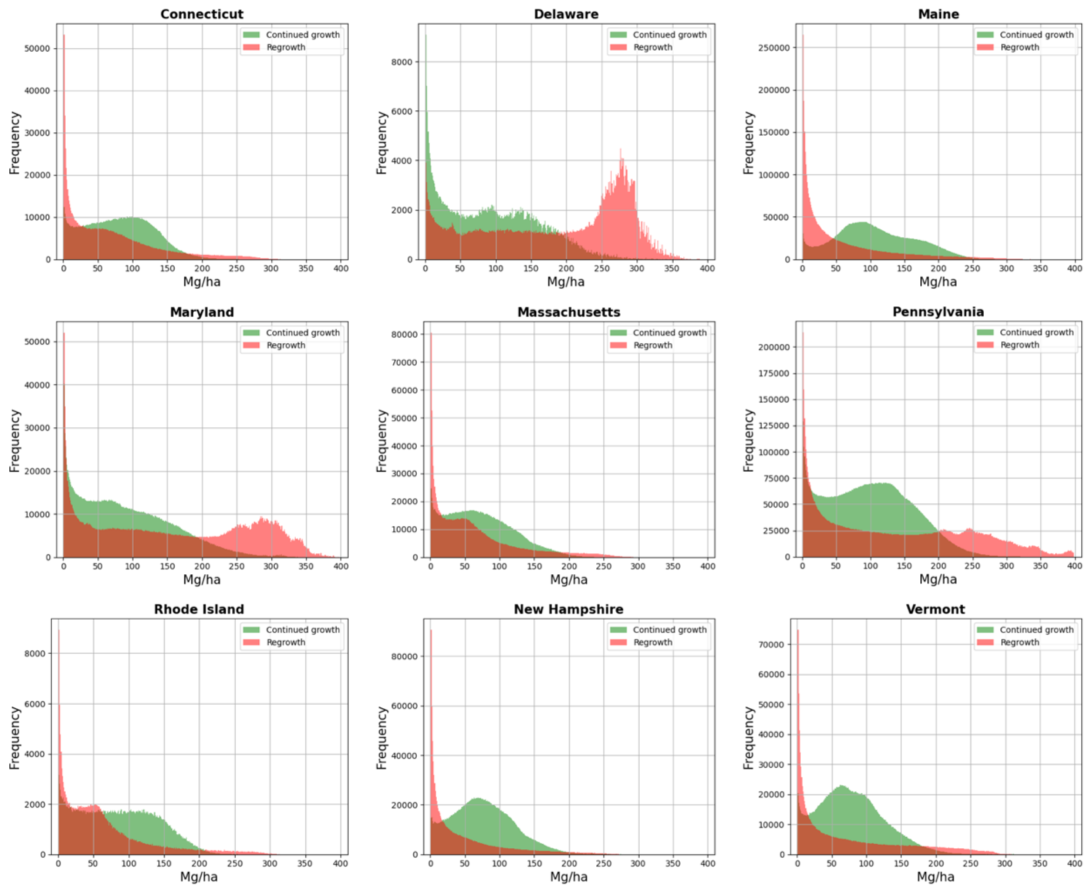


Figure 4.5 CSPG over areas with continued growth (green) vs that over regrowth (red) in Maryland, Delaware, Pennsylvania and Rhode Island, Connecticut, Massachusetts, Vermont, New Hampshire and part of Maine.

In addition to aboveground carbon sequestration potential, we also project the potential future growth trajectory of aboveground carbon. Figure 4.6a shows the potential annual aboveground carbon stock from present to 300 years in the future for each state, and 4.6b are corresponding maps of potential AGB in years 2050, 2100, 2200 and 2300, respectively. The relative contribution of AGB from continued growth and regrowth varies by state. For example, the contribution of regrowth to newly gained AGB by 2300 is 46.3% in Connecticut, 70.7% in Delaware, 65.9% in Maryland, 43.0% in Massachusetts, 26.1% in New Hampshire, 51.1% in Pennsylvania, 35.9% in Rhode Island and 34.9% in Vermont. Delaware has the highest contribution from regrowth among all states; the second and third highest contributions are in Maryland and Pennsylvania, respectively. In contrast, continued growth is the largest contributor to the aboveground carbon stocks in all other states. Annual estimates of aboveground carbon stocks from present to 300 years for each county in the RGGI region (except Maine, New Jersey and New York) can be found in Figures C.16-C.23, and county-level aboveground carbon stocks in 2020, 2030, 2040 and 2050 are summarized in tables C.1-C.8.

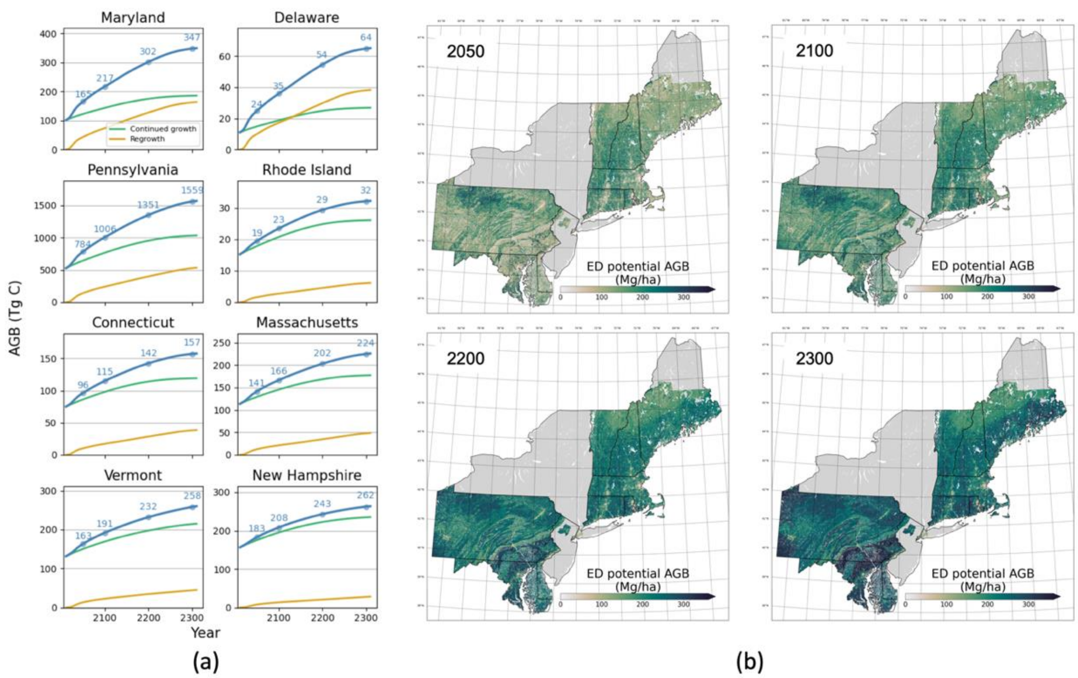


Figure 4.6 ED potential AGB from present to 300 years in the future. Blue line in (a) represents statewide annual aboveground carbon stocks for Maryland, Delaware, Pennsylvania and Rhode Island, Connecticut, Massachusetts, Vermont and New Hampshire. The four numbers along each curve correspond to the stock value at years 2050, 2100, 2200 and 2300. Corresponding maps of AGB density are also mapped in (b). Green and yellow lines in (a) represent the relative contributions of continued growth and regrowth to the carbon stocks.

Projections of CSP and CSPTG are further assessed for sensitivity to changes in Net Primary Productivity (NPP) and disturbance rate. As an example, Figure 4.7 examines how the average CSP and CSPTG across Maryland, Delaware and Pennsylvania, which together account for 50% the land area in the RGGI region, respond to different NPP and disturbance rates. As expected, averaged CSP and CSPTG of Maryland, Delaware and Pennsylvania are generally proportional to NPP and inversely proportional to disturbance. High CSP and CSPTG appear at high NPP but low disturbance, representing conditions where a forest may gain carbon quickly and lose less of it over time to disturbance. In contrast, low NPP and high disturbance result in low CSP and CSPTG because of a slowing carbon sequestration rate and high losses due to disturbance.

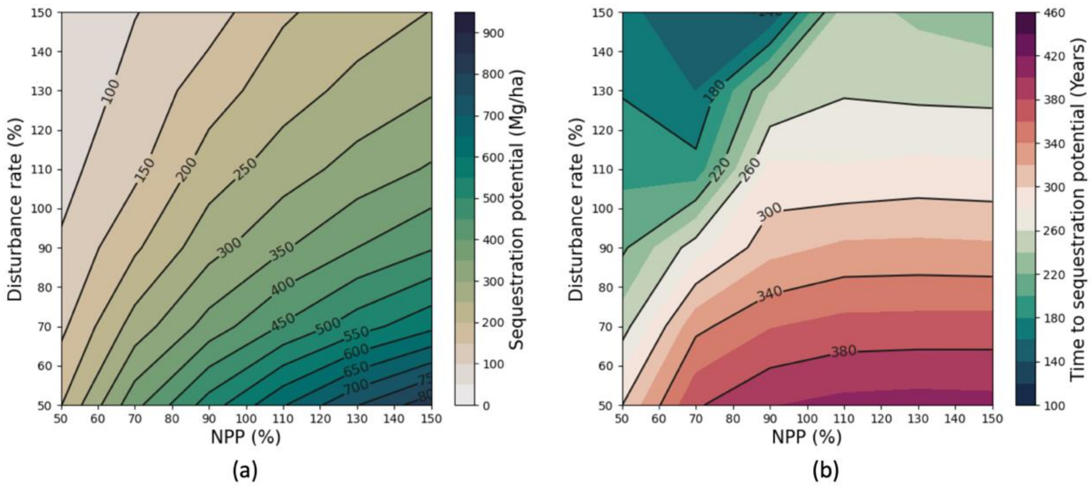


Figure 4.7 Sensitivity of average CSP and CSPTG over the states of Maryland, Delaware and Pennsylvania in response to percent changes in NPP and disturbance rate. NPP and disturbance rates are changed from 50%–150% at an increment of 20%.

4.4 Discussion

Forests play a crucial role in climate mitigation. Avoided deforestation, improved forest management and reforestation could provide two-thirds of the cost-effective nature-based climate mitigation needed to hold warming to below 2 °C (Griscom et al., 2017), with the regrowth of natural forest the single largest natural climate solution both globally and within the United States (Cook-Patton et al., 2020, Fargione et al., 2018). In this context, accurate high-spatial resolution estimates of the potential for additional aboveground carbon sequestration in forests is needed. This work combined advances in ecosystem modeling and remote sensing to estimate present forest aboveground carbon stocks and project future forest aboveground carbon sequestration potential at 90 m resolution across the eleven states in RGGI region, including 34.9 million 90 m land cells over an area of 283,000 km^2.

The RGGI region has large aboveground carbon sequestration potential gap compared to its present aboveground carbon stocks. We found that present AGB stocks in Delaware are at one-fifth of the potential, Maryland and Pennsylvania are at one-third, Connecticut, Massachusetts, New Hampshire, Rhode Island and Vermont are at about half. The significant gap between present AGB and aboveground carbon sequestration potential provides opportunities for regional climate mitigation. Maximum potential gains in AGB by 2050 would be 21.2 Tg C in Connecticut, 13.8 Tg C in Delaware, 28.0 Tg C in Massachusetts, 64.2 Tg C in Maryland, 26.6 Tg C in New Hampshire, 257.7 Tg C in Pennsylvania, 4.2 Tg C in Rhode Island, and 32.0 Tg C in Vermont. Together, these eight states have the potential to gain an additional 209.1 Tg C through continued growth of existing trees and 238.6 Tg C through regrowth of new trees. The average annual sequestration rate (i.e., 11.5 Tg C/yr) is equivalent to 9.6% of these eight states' average annual fossil fuel emissions between 2011-2017 (i.e., 119 Tg C/yr) according to US EPA fossil fuel combustion inventories (U.S. Environmental Protection Agency 2019). The high spatial resolution underlying these region-wide estimates may help decision-makers prioritize areas for reforestation. For example, this work indicates that counties of Lancaster, Crawford, Bradford, Washington and York in Pennsylvania, Sussex in Delaware, and Coos in New Hampshire, where the CSPG is over 20 Tg C have exceptionally high potential to gain additional forest aboveground carbon. These counties with the high CSPG vary in terms of dominate land cover type. The six

Pennsylvania and Delaware counties are predominately pasture and cropland, whereas the New Hampshire county is mostly forest.

Historical and present land-use activities and natural disturbance have resulted in a diverse landscape of heavily fragmented forest, mixed with non-forest patches. While this land cover matrix poses a great challenge for biomass mapping, this study incorporates forest structure data as detailed as 1 m spatial resolution to address this challenge. To illustrate, three areas representing typical forest, residential and agricultural area were chosen as examples. NAIP aerial imagery in Figures C.5, C.6 and C.7, respectively show a forested area as a mixture of trees and gaps (black shadows), a residential area as a mixture of houses and trees along roadsides and backyards, and an agricultural area as a mixture of cropland and scattered trees. Across all three heterogeneous landscapes, treed versus non-tree areas are easily identified using high-resolution (1 m) lidar canopy height and tree cover classification. The resulting variations in AGB are consequently well captured in ED initialized AGB and lidar empirical AGB estimates. These fine-scale detailed maps also identify a substantial amount of land carbon being stored in trees that are scattered across non-forest areas, which emphasizes the importance of trees outside of forests to regional and global carbon budgets (Huang et al., 2015, Zomer et al., 2016, Chapman et al., 2020, Spawn et al., 2020). These high-resolution capabilities are incredibly valuable for carbon modeling and climate mitigation planning across regions with strong AGB spatial heterogeneity.

This study expands the previous carbon modeling system prototyped for the state of Maryland (Hurtt et al., 2019b) to a larger domain using an updated version of the ED model and an improved initialization method. We compared new results to prior results for MD, where both studies used the same lidar canopy height and NAIP tree cover data inputs for initialization. This study yields comparable estimates to Hurtt et al., 2019b. In particular, our results concur with Hurtt et al., 2019b's spatial pattern of sequestration potential and sequestration gap. For example, the Maryland counties of Frederick and Baltimore have the largest CSP (>25 Tg C) in both studies. Further, this study shows agreement with the location of Maryland's largest aboveground carbon stocks, first in Garrett County, followed by Charles, Baltimore, Allegany and Frederick counties, respectively. Although our study includes woody wetland, we have excluded it here to provide a direct comparison to Hurtt et al., 2019b. This study projects total CSP in Maryland as 304.0 Tg C, which is within 3.5% of CSP estimates in Hurtt et al., 2019b (314.8 Tg C). There are, however, differences in contemporary aboveground carbon stock estimates between this study and Hurtt et al., 2019b; this study estimates Maryland's present carbon stocks at 84.5 Tg C, which is ~26 Tg C lower than Hurtt et al., 2019b. The possible causes for this difference in present aboveground carbon stocks include different model driver inputs (e.g., soil depth, air temperature etc.) as well as the change in initialization method. The weighting-based initialization method in this study corrects overestimation at high AGB area and yields improved correlation with FIA plots. For example, switching the initialization method used in this study back to the prior method results in an estimate of aboveground carbon stocks of 94.7 Tg C, implying about 40% of the difference may be due to the initialization method.

We also compared our estimates of natural forest regrowth rate with a recent global estimate from Cook-Patton et al., 2020, which combines field measurements across multi-biomes and environmental covariates to produce a global 1-km map of potential AGB growth rates for the first 30 years of natural forest regrowth. We averaged the potential aboveground carbon growth rate between 5 and 30 years to align with Cook-Patton et al. Average growth rates across the RGGI region are similar between this study and Cook-Patton et al, at 0.96 Mg C/ha/yr and 1.05 Mg C/ha/yr, respectively. Spatial comparisons are shown in Figure C.25.

We find that absolute differences for more than half of all land area (59.2%) are below 0.30 Mg C/ha/year, with 40.5% within 0.20 Mg C/ha/yr and 20% of land area within 0.10 Mg/C/ha/year. Figure C.26a shows the average potential growth rates stratified by soil depth, suggesting the strong dependence of AGB growth on soil depth in ED. Most area has a soil depth between 100cm and 160cm, and the two carbon estimates are closest at this depth range. The Figure C.26b shows ED potential growth rates as a function of stand age, suggesting the growth rate varies with age during the first 30 years of natural forest regrowth. This nonlinear age dependence is included in ED estimates but not in Cook-Patton et al., 2020.

Understanding uncertainty sources is essential to applications of our estimates of climate mitigation potential and also to provide insights into subsequent modeling development. Uncertainties regarding the estimation of present aboveground carbon stocks arise from ED drivers and initialization inputs. First, due to lack of available meteorology datasets with both high spatial and temporal resolution, the meteorological driver is interpolated by both fine spatial resolution (Daymet) and coarse spatial resolution (MERRA2) data products. Imperfect interpolation, inherent mismatches between two meteorology products or underlying uncertainties within them may propagate errors to ED estimates. Second, in this study, ED is run with average meteorology and CO_2 level between 1981 and 2017 despite temporal variability and positive increasing trends. Some states may also have lower disturbance rates than the spatially constant value used in ED (1.2% yr^{-1}). These factors may jointly lower AGB growth rate and aboveground carbon sequestration potential, especially if these average rates change over time. Third, the uncertainty associated with initialization inputs is inconsistent acquisition times of the lidar canopy height data. Our lidar data comes from acquisition between 2004 to 2018, and some archived lidar data may have less sampling density than others (Huang et al., 2019b, Tang et al., 2021).

Uncertainties regarding the estimation of future forest carbon stocks arise from future climate change, CO_2 levels, and the disturbance regime. Future climate change and continued increases in CO_2 levels will likely have complicated impacts on vegetation growth. For example, field experiments found that increasing CO_2 could enhance tree growth indicating more carbon will be sequestered (Norby and Zak 2011, Walker et al., 2019). However, faster growth may also accelerate harvesting events, reducing tree longevity, and resulting in less carbon residence time in trees (Körner 2017). Warming could also lengthen growing seasons and in turn enhance annual vegetation growth (Menzel and Fabian 1999); however, it may also inversely increase evapotranspiration and autotrophic and heterotrophic respiration (White et al., 1999, Piao et al., 2008, Crowther et al., 2016), decrease soil moisture and increase fire risk (Westerling et al., 2006). Additionally, future disturbance is likely to increase as warming continues. Warmer and drier conditions facilitate disturbance related to fire, drought and insect outbreaks and decrease disturbance related to snow and ice, while warmer and wetter conditions increase disturbances from wind and pathogens (Seidl et al., 2017). Given the unclear net impact of climate change and lack of climate change scenarios harmonized at high spatial resolutions, this study does not explicitly consider changes in future climate, CO_2 and disturbance, but alternatively investigates how potential alterations in NPP and disturbance rate could propagate to estimates of future carbon sequestration.

Finally, this study has focused on forest aboveground carbon stocks from present to future. The work has capitalized on the qualitative advances in data and modeling of aboveground forest structure, and focused on producing new high-resolution estimates of above ground carbon that are most directly related to these measurements and also important for policy.

This work does not yet include other forest carbon pools such as soil carbon or wood products, nor the dynamics related to forest management (Birdsey et al., 2006, Lippke et al., 2011). Our projections of CSP and CSPG are thus conservative underestimates of ecosystem's total carbon sequestration potential (Domke et al., 2020a, 2020b), and future work could build off these advances above ground to address these other components. Looking ahead, the forest carbon modeling approach used in this study is currently being expanded to US CONUS and global domains and to include soil carbon and associated carbon fluxes. Key to expanding this work's spatial coverage is to switch from airborne remote sensing to orbital observations. Despite the demonstrated capability of 1 m airborne lidar data to capture fine-scale heterogeneity, leveraging this existing data at national and global scales encounters several challenges, including inconsistency of instrument quality and acquisition time as well as limited spatial coverage. New airborne lidar collection can also be expensive. Future work should utilize satellite-based forest structure measurements from the ongoing NASA missions of Global Ecosystem Dynamics Investigation (GEDI) (Dubayah et al., 2020a), Ice, Cloud, and land Elevation Satellite-2 (ICESat-2) (Markus et al., 2017) and Landsat-based tree cover from Global Forest Change dataset (Hansen et al., 2013). Key for the inclusion of soil carbon will be sufficient data to constrain model estimates and account for the effects of disturbance and land use history. Relevant soil carbon inputs include the Harmonized World Soil Database (Wieder et al., 2014) and SoilGrids250m dataset (Hengl et al., 2017). Relevant products for disturbance and land use history include the Global Forest Change dataset (Hansen et al., 2013), the North American Forest Dynamics dataset (Goward et al., 2016), and the Land Use Harmonization 2 dataset (Hurtt et al., 2020).

Acknowledgements

This work was funded by NASA Carbon Monitoring System (NASA-CMS) project (80NSSC17K0710), NASA New Investigator Program (NASA-NIP) (80NSSC18K0708) and also supported by the University of Maryland Ann G. Wylie Dissertation Fellowship. We also acknowledge in-kind support from the US Forest Service, Forest Inventory and Analysis Program.

Chapter 5 Demonstrative use of spaceborne lidar (GEDI/ICESat-2) in the Ecosystem Demography model

Abstract

Characterizing terrestrial carbon dynamics with processed-based models is important for climate mitigation, but can involve high levels of uncertainty regarding current forest conditions. Accurate representation of contemporary forest structure and carbon storage within processed-based models is critical for estimating realistic short-term and long-term carbon dynamics. Two recent NASA spaceborne lidar missions, Global Ecosystem Dynamics Investigation (GEDI) and ICE, Cloud, and Land Elevation Satellite 2 (ICESat-2), are offering an unprecedented volume of forest structure observations on a global scale. In this study, we explore and assess the potential for these emergent lidar observations to improve estimates of initial forest conditions within ecosystem models. Here, we specifically initialize a process-based ecosystem model (Ecosystem Demography (ED) model) at 0.01° resolution using global estimates of forest canopy height derived from GEDI and ICESat-2 observations. Compared to previous global ED results from Chapter 3, AGB estimates from the initialized ED show enhanced spatial heterogeneity, capturing effects from fine scale disturbance, logging, and deforestation. Evaluations against FIA plot data from US Forest Service show favorable results, with an average bias lower than at 10 Mg/ha. The results of this study demonstrate the promising value of combining space-borne lidar observations in ecosystem modeling.

5.1 Introduction

Observing, understanding, and predicting terrestrial carbon dynamics is essential for supporting climate change mitigation policy and planning (Canadell and Raupach, 2008;

Keenan and Williams, 2018; Schimel et al., 2015; Tharammal et al., 2019; Xiao et al., 2019). Ecosystem models, especially process-based models, have demonstrated their value for studying large-scale interactions and consequences of underlying ecological processes, and making projections of future carbon dynamics under climate change and land use change (Ahlström et al., 2012; Quesada et al., 2018; Sitch et al., 2008). However, results from current ecosystem models can be highly uncertain, due in part from a lack of accurate information on contemporary forest conditions, particularly regarding forest age and carbon status (Besnard et al., 2018; Houghton et al., 2009; Mitchard et al., 2013; Pregitzer and Euskirchen, 2004). Concurrently, remote sensing technologies have rapidly advanced offering observation data at increased temporal and spatial resolutions, over expanded domains, and with more types of retrieved surface properties (Drake et al., 2002; Hansen et al., 2013; Simard et al., 2011; Tang et al., 2012). Such advances in remote sensing observations provide opportunities to improve characterization of initial conditions in ecosystem models and in turn reduce uncertainties of model estimates.

The world's forests present strong spatial heterogeneity in successional age and vertical structure, posing challenges to ecosystem models regarding representation of initial conditions. The spatial heterogeneity resulting from cumulative impacts from environmental conditions, historical disturbance, land use change, regrowth, etc (Townsend et al., 2008; Vieira et al., 2004). Contemporary heterogeneity would in turn result in contrasting subsequent ecosystem functioning and carbon dynamics under future climate change (Fadrique et al., 2018; Levine et al., 2016). To obtain initial conditions in ecosystem models, several initialization approaches have been explored, and their strengths and limitations depend on the scale of application. One approach is to assume forests in their 'potential' state, with final conditions determined by environmental conditions (e.g., climate and soil) (Cramer et al., 2001; Sitch et al., 2008). This approach is highly uncertain and has been challenged as it contradicts current conditions of many forests that have been out of equilibrium (Carvalhais et al., 2010; Pappas et al., 2015; Williams et al., 2009). A second approach is to spin-up models from past to present with land-use history, which can inform the location, type, and magnitude of various land-use changes (Hurtt et al., 2002, 2011, 2020; Shevliakova et al., 2009). While employed by multiple Earth System Model (ESMs) in CMIP6, this approach is usually limited by the spatial resolution of land-use history (>25 km) (Hurtt et al., 2020; Lawrence et al., 2016). A third approach is to initialize with field-based measurements (Medvigy et al., 2009; Pacala et al., 1996; Pappas et al., 2015). One example is the US Forest Service's Forest Inventory and Analysis (FIA) program, which takes regular measurements of a variety of forest metrics over thousands of permanent plots nationwide (Bechtold and Patterson, 2005; Smith, 2002). Similar data, however, does not exist worldwide and potential inconsistencies may exist between different forest inventories. More importantly, these FIA plots only cover limited areas and are small in size (<1 ha).

Lidar remote sensing offer novel opportunities for model initialization at large spatial scales and at high spatial resolutions. Through decades of development, lidar has demonstrated its unique ability to accurately, consistently, and efficiently measure forest structure over large spatial domains (Drake et al., 2002; Dubayah and Drake, 2000; Huang et al., 2019; Tang et al., 2012, 2021). Initializing models with lidar data was pioneered in local study sites in Costa Rica, where lidar-based forest canopy height was shown to effectively constrain ecosystem model estimates of carbon stocks and fluxes (Hurtt et al., 2004). This lidar-based initialization approach was then further tested in various studies, demonstrating its potential as a robust approach to advance ecosystem modeling at scale (Antonarakis et al., 2011; Hurtt et al., 2010, 2016, 2019b; Ma et al., 2021). One recent example is the development of a high-resolution regional forest carbon modeling system for projecting future carbon sequestration potential (Hurtt et al., 2019b; Ma et al., 2021). In this system, a forest canopy height map at 90 m spatial resolution is derived from airborne lidar observations, and then linked to an ecosystem model, called the Ecosystem Demography (ED) model, to determine baseline estimates of current aboveground carbon and remaining carbon sequestration potential from this baseline. This modeling system further demonstrates the power of lidar in observing fine-scale heterogeneity in forest conditions and carbon.

Two ongoing NASA spaceborne lidar missions, namely GEDI (Global Ecosystem Dynamics Investigation) and ICESat-2 (ICE, Cloud, and Land Elevation Satellite 2), provide huge potential for advancing global ecosystem modeling. GEDI is was launched on December 5th, 2018 and is located on board the International Space Station (ISS) through 2021. GEDI is the first spaceborne lidar instrument specifically optimized to measure forest vertical structure (Dubayah et al., 2020a). Over two-year of mission lifetime, GEDI aims to provide over 10 billion cloud-free waveforms over temperate and tropical forests between 51.6°N and 51.6°S. The collected waveforms are processed to generate a suite of data products including canopy height, canopy cover, canopy leaf area index and profile, topography, and footprint-level and gridded aboveground biomass density (AGB). In parallel, ICESat2 was launched on September 15th, 2018 as a free-flying satellite. On board is a lidar instrument ATLAS (Advanced Topographic Laser Altimeter System), which uses a novel laser ranging technology called photon-counting. Despite its primary mission focus on measuring ice sheet elevation and sea ice thickness, ICESat-2 also collects global measurements of forest canopy height over its three-year lifetime (Neuenschwander and Pitts, 2019). While both missions have their own respective instrument designs and mission concepts, they share the goal of mapping global forest structure and forest carbon. Data products from these missions will benefit ecosystem modeling by providing direct forest structure observations for model initialization or benchmarking model estimates of carbon stock and forest structure.

To leverage emergent spaceborne lidar observations, we developed an initialization approach for the ED model at 0.01° spatial resolution (approximately 1 km at the equator) over a near global domain (51°N ~ 51°S). Here, we examine the benefits of using both GEDI and ICESat-2 datasets within ED to better capture fine-scale forest heterogeneity and generate a carbon baseline for future projections. Additionally, we evaluate the wall-to-wall data coverage currently provided by both NASA missions and consider their consistency in canopy height measurements. To do this, we generate a gridded canopy height histogram from raw observations at the footprint-level and use it to initialize the ED model. We also compare the resulting AGB estimates against field inventory data and lidar empirical AGB estimates.

5.2 Methods

ED initialization employs a height-based approach, which was first proposed in Hurtt et al., (2004) and subsequently applied in regional applications (Hurtt et al., 2010, 2016, 2019b; Ma et al., 2021; Antonarakis et al., 2011). The height-based initialization approach requires an AGB and canopy height growth trajectory pre-simulated by the ED model and tree canopy height and canopy fraction data from remote sensing observations. The initialization determines for each grid cell, the forest's current time at its growth trajectory and corresponding AGB. The overall workflow of ED initialization with GEDI and ICESat-2 is illustrated in Figure 5.1. Section 5.2.1 describes the core principles of ED model and details the initialization process; section 5.2.2 describes generation of canopy height and tree cover from GEDI and ICESat-2 observations and drivers required to run ED model.

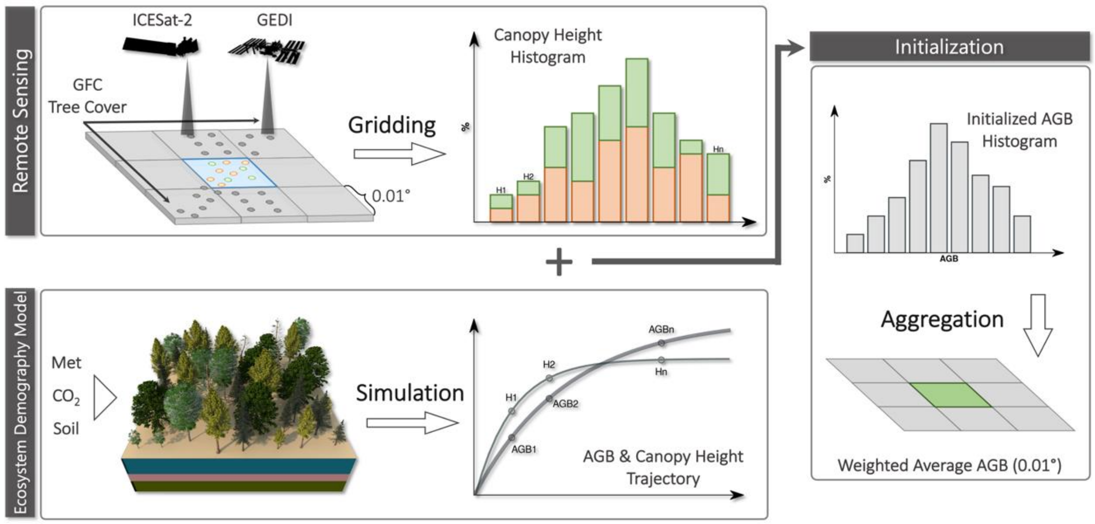


Figure 5.1. Illustration of ED initialization at a grid. Top box depicts the process of generating gridded canopy height histogram (ranging from 5 m to 50 m and bin size of 0.5 m) and average tree cover for the blue grid with size of 0.01°. Color circles present GEDI and ICESat-2 footprint/segment-level observations. Note that not every grid has observations from either or both missions. The bottom box depicts the process of generating AGB-height trajectory for the blue grid by running ED with drivers of meteorology, CO_2, soil properties. The right box depicts process of initializing with simulated AGB-height growth trajectory and canopy height histogram.

5.2.1 ED and Initialization

5.2.1.1 ED Principle

ED is individual-based prognostic ecosystem model that integrates submodules of growth, mortality, hydrology, carbon cycle and soil biogeochemistry (Hurtt et al., 1998; Moorcroft et al., 2001). ED can characterize plant dynamics at individual-levels including growth, mortality, reproduction and competition for light, water and nutrients. ED can also simulate the carbon cycle, including carbon uptake by leaf photosynthesis to carbon allocation for growth in leaves, roots, stem and seedlings, as well as carbon decomposition in various soil carbon pools. The model can further characterize changes in individual plant density and composition under natural disturbance and land-use and land cover change. ED has been used to characterize regional carbon dynamics in response to climate change, elevated CO_2, land use and land cover change, and natural disturbance (Hurtt et al., 2002; Fisk et al., 2013; Flanagan et al., 2019).

Explicit characterization of vertical structure during ecosystem succession is a feature that distinguishes ED from most other ecosystem models. Specifically, each plant has own structure attributes such as associated canopy height, diameter at breast height. Canopy height dynamics are then tracked as a result of competition between plants and cumulative carbon balance between photosynthesis and respiration with given environmental conditions (e.g., temperature, precipitation, radiation, and soil moisture). Explicit tracking of canopy height facilitates the potential connection between the ED model with external forest structure data from remote sensing observations and field measurements. Linking ED with forest structure data can help determine the contemporary successional state of each location in the gridded domain.

5.2.1.2 Lidar initialization

Height-based initialization determines each location's successional state from the pre-simulated AGB-height trajectory using lidar canopy height and canopy cover as inputs (Ma et al., 2021). This initialization process is comprised of two major steps: 1) generate gridded AGB-height trajectory, also called the "lookup table," by running ED for a certain period with drivers of meteorology, CO_2, disturbance and soil properties. The AGB-height trajectory is a subset of outcomes from ED simulation of ecosystem dynamics from bare ground to equilibrium state; and 2) index the AGB-height lookup table with lidar canopy height and canopy cover. This step matches lidar canopy height with ED-simulated canopy height to identify successional state at time of lidar acquisition. The corresponding AGB is subsequently defined as initialized AGB.

In this study, lidar canopy height was taken from a 0.01° gridded canopy height histogram derived from GEDI and ICESat-2 footprint/segment-level observations. For canopy cover, we used 0.01° average tree canopy cover in 2010 derived from the Global Forest Change dataset (Hansen et al., 2013). The use of a canopy height histogram instead of average canopy height is based on the sampling nature of two lidar missions, spatial variation of canopy height within 0.01°, and nonlinear relationship between canopy height and AGB. As both GEDI and ICESat-2 missions only provide ground samples of canopy height instead of wall-to-wall coverage, these samples could vary largely within 0.01°. A canopy height histogram can take height variation into account and avoid nonlinearities resulting from scaling footprint/segment-level observations to a 0.01° grid.

The initialization indexed the AGB-height lookup table for each bin of the canopy height histogram and then weighted by associated frequency, resulting in a AGB corresponding to a 0.01° grid cell with 100% tree canopy cover. This AGB is then linearly adjusted 0.01° average tree canopy cover by assuming zero AGB at non-tree fraction.

5.2.2 Data

5.2.2.1 GEDI/ICESat-2 gridded canopy height histogram

GEDI is specifically designed for forest vertical structure measurements (Dubayah et al., 2020a). The laser instrument is a geodetic-class waveform lidar consisting of 3 lasers which produce 8 ground tracks of 25 m footprint samples. The 8 ground tracks are spaced approximately 600 m apart in the cross-track direction and 60 m in the along track direction on the Earth’s surface. By processing returned full waveforms, L2A product provide canopy height metrics and topographic surface elevation at footprint level (Dubayah et al., 2020b).

ICESat-2's ATLAS instrument utilizes photon-counting laser technology Markus et al., (2017). It is comprised of 3 pairs of strong and weak beams with each pair being separated by 3.3 km and a pair width of 90 m. The 3 pairs of beams produce 6 ground tracks with footprint size about 17 m ~20 m, and each footprint is spaced by 70 cm. By processing all retuned photons, ATL08 product provides estimates of terrain height, canopy height and canopy cover at fixed segment size of 100 m (Neuenschwander and Pitts, 2019).

We generated a spatially gridded canopy height histogram using canopy height metric *RH100* (i.e., relative height at 100th energy percentile) from GEDI L2A footprint-level measurements (algorithm 2) and canopy height metric *h_canopy* from ICESat-2 ATL08 segment-level measurements. Specifically, 6751 GEDI L2A granules (acquired between 2019 April and 2020 September) and 95825 ICESat-2 ATL08 granules (acquired between and 2018 October and 2020 July) were downloaded from the NASA LP DAAC. For both GEDI and ICESast-2, only footprints/segments that meet each filtering criteria are used for gridding. For GEDI, the filtering criteria are: 1) *quality_flag* value of 1; 2) *sensitivity* > 0.95; 3) *RH100* within range of 5 m and 50 m. For ICESat-2, the filtering criteria: 1) *dem_removal_flag* value of 0; 2) *canopy_flag* value of 1; 3) *cloud_flag_atm* < 4; 4) *h_canopy* within range of 5 m and 50 m. The filtering criteria are set to exclude bad observations or those returned by non-forest areas. The filtered footprints/segments were then mapped into 0.01° grids using associated coordinate variables (i.e., using *lat_lowestmode* and *lon_lowestmode* for GEDI, *latitude* and *longitude* for ICESat-2). This gridding step results in maps with 90 layers and each layer records the footprint count within the range of the height bin.

5.2.2.2 Tree cover

Tree canopy cover is from the Global Forest Change (GFC) dataset which provides at 30 m the latest fractional tree cover in 2010 (Hansen et al., 2013). The GFC dataset also provide pixel level estimates of annual forest losses and decadal gains. Combining tree cover 2010 and gain/loss could potentially produce tree cover estimate in 2019. However, as the gain layer has not been updated since 2012, this approach will likely result in biased low tree cover at areas with plantation after 2012 (e.g., southern US). Therefore, this study uses tree cover in 2010 and averages it to 0.01° from the original 0.00025° resolution (approximately 30 m at the equator).

5.2.2.2 ED Drivers

ED model runs are driven by meteorological forcings, soil characteristics, and atmospheric CO_2 concentration. Meteorological forcing data comes from the NASA Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) (Gelaro et al., 2017). All variables, including surface air temperature, surface specific humidity, incident shortwave radiation, wind speed, precipitation and multi-layer soil temperature, were spatially interpolated to 0.5° and averaged to monthly diurnal estimates. Soil hydraulic properties come from Montzka et al., (2017), which provides spatial parameter maps of soil depth and saturated hydraulic conductivity at 0.25°. CO_2 is held spatially and temporally constant at 400 ppm.

5.3 Results

5.3.1 Gridded canopy height

Gridded canopy height histograms and average canopy height derived from GEDI and ICESat-2 observations are shown in Figure 5.2 and 5.3. Both missions show height variability within 0.01°. Since both missions likely take measurement samples from different areas in a grid, it is expected that the resulting canopy height histograms may shows some discrepancies. The average canopy height maps show similar spatial distributions between the two missions, with high values appearing in the tropical rainforest, Pacific Northwest and eastern coast of North America, eastern Europe, eastern Himalayas, southern Asia, and the far east of Russia.

Due to the sampling nature of each lidar mission and unfinished acquisitions, both gridded maps have coverage gaps (as shown in inset axis of Figure 5.3). Generally, high latitudes have better coverage than lower latitudes due to orbital convergence. Moreover, GEDI and ICESat-2 reveal different track coverages and distributions. Because of the orbital resonance of ISS, GEDI's tracks are not evenly spaced, resulting in dense coverage (nearly full coverage) in areas where tracks are clustered and large gaps between track clusters. In contrast, ICESat-2's tracks are relatively evenly spaced but coverage gaps exist between each individual track due to its repeating orbit and cyclical off-nadir pointing system. ICESat-2 also excludes observations with tree cover less than 5%.

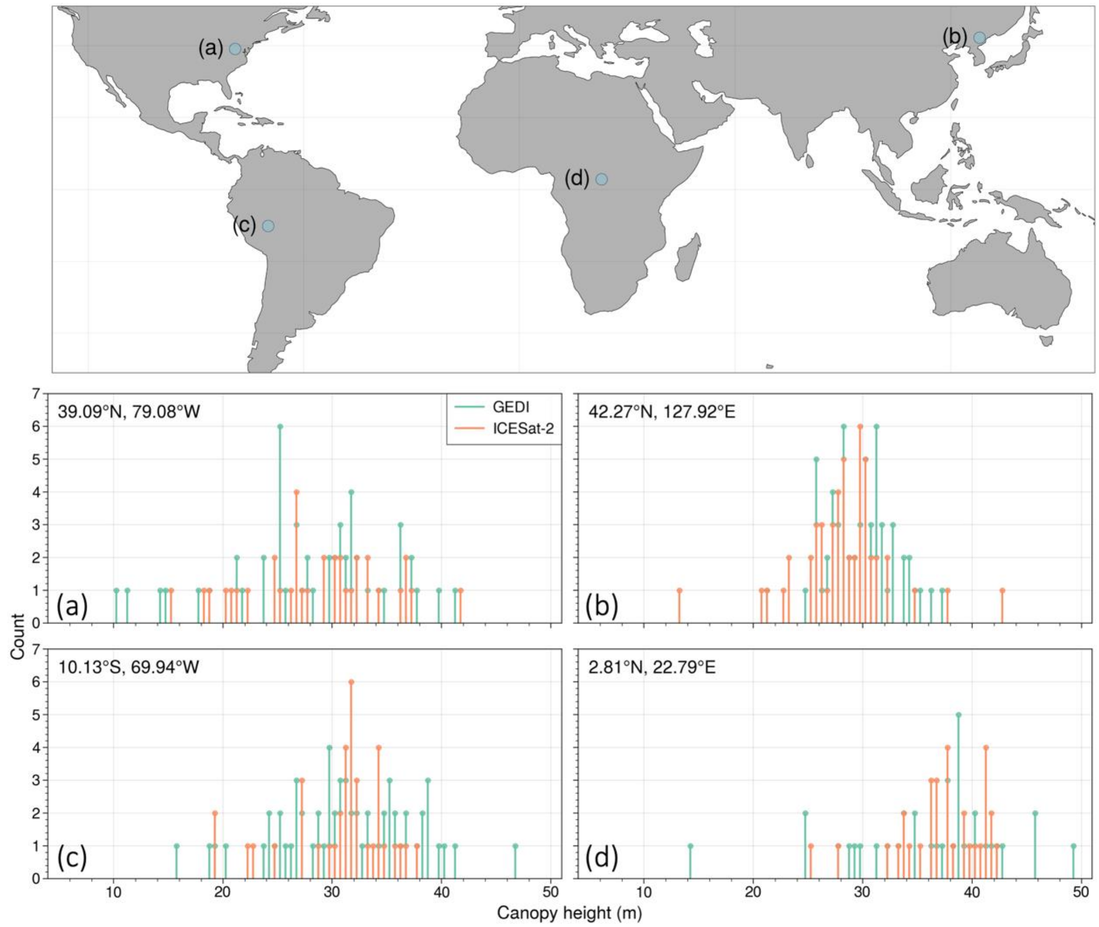


Figure 5.2. Canopy height histograms at 0.01° at four example grid locations, produced by gridding footprint/segment-level observations from GEDI L2A and ICESat-2 ATL08 datasets.

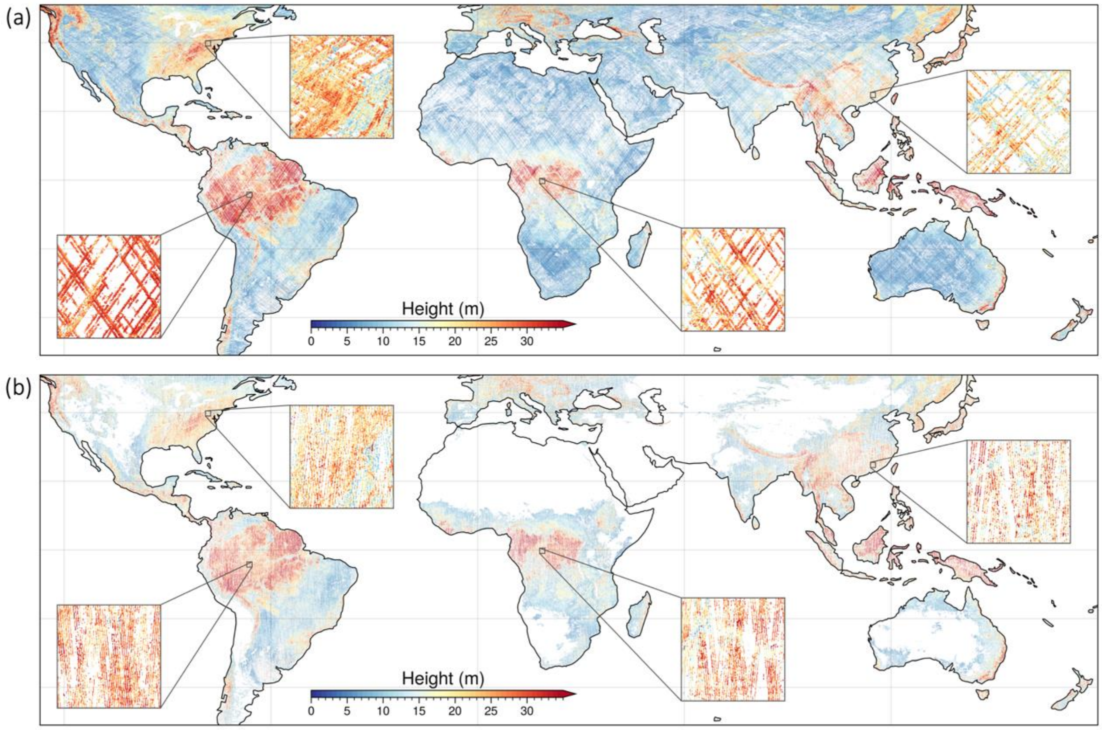


Figure 5.3. Average canopy height at 0.01° by gridding footprint/segment-level observations from GEDI (a) and ICESat-2 (b). The insets highlight fine-scale heterogeneity at selected regions.

By counting area of 0.01° with lidar observations, both missions have sampled about half of the land area within the tree cover gradient identified by the GFC dataset (Figure 5.4). GEDI has higher coverage than ICESat-2 over all gradients of tree cover; the difference is largest at tree cover group 1~10%, probably due to data exclusion in ICESat-2 ATL08. In total, GEDI has sampled about 31.7 million km^2 of 0.01° grids with tree cover above 1%, while ICESat-2 has sampled about 23.8 million km^2. Combining GEDI and ICESat-2 observations together can increase sample coverage. For example, the total sampled area (>1% tree cover) can be increased up to 39.7 million km^2 when GEDI and ICESat-2 data is combined. The coverage increase is more significant at tree cover fractions 1-10%, 10-20%, and 90-100%.

Footprint/segment densities from GEDI and ICESat-2 are shown in Figure 5.5. Generally, GEDI has higher densities than ICESat-2. For example, GEDI has more grids with densities

>60 per grid. About half of 0.01° grids with GEDI data have densities above 20 per grid, while only 20% of grids from ICESat-2 data have the same. Combining GEDI and ICESat-2 observation data increases shot density per grid over crossover grids, potentially improving height heterogeneity at sub-grid scales. For example, combining both datasets results in 18 million grid cells with shot densities above 20, which is 5 million more than GEDI alone and 15 million more than ICESat-2.

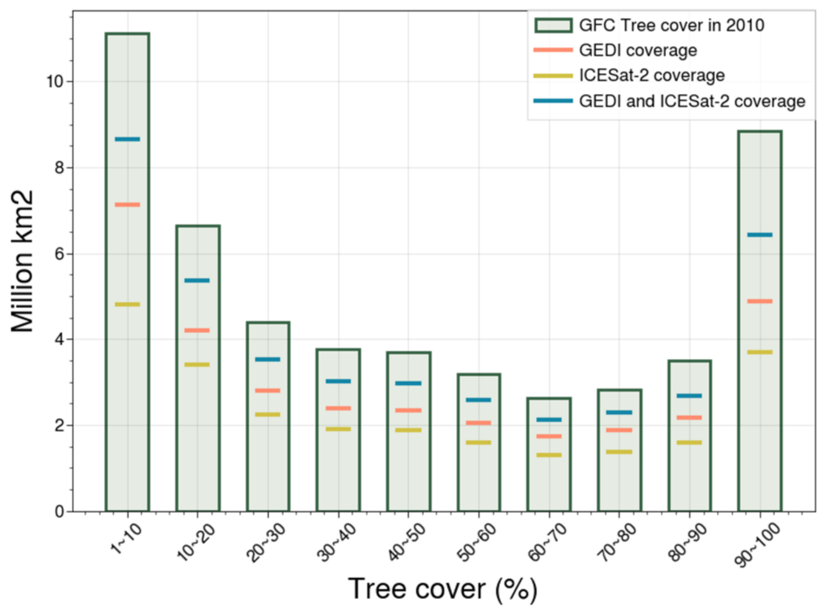


Figure 5.4. Land area sampled by GEDI and ICESat-2. Green bars represent total land area by tree cover groups based on the Global Forest Change dataset in 2010. *Orange and yellow lines represent total area represented by 0.01° grids with at least one footprint/segment observation from GEDI or ICESat-2, respectively. The blue line represents the total area represented by 0.01° grids with observations from both GEDI or ICESat-2.*

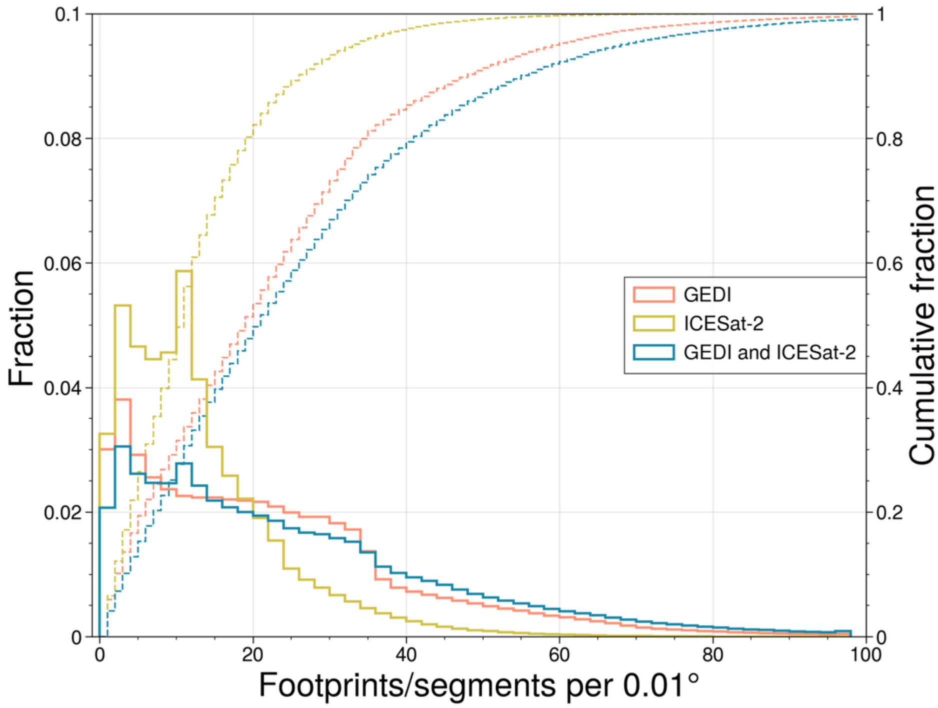


Figure 5.5. Histogram of footprint/segment density of GEDI (orange), ICESat-2 (yellow) and both GEDI and ICESat-2 combined (blue).

Figure 5.6 shows the comparison between GEDI and ICESat-2 average canopy height across 0.01° crossover grids. About 13.8 million grids have at least two shots from either mission. The average height bias between GEDI and ICESat-2 is 0.14 m while the average RMSE is 5.04 m. The majority of crossover grids have average canopy heights under 12 m, and this fraction declines as height increases. This pattern is likely related to coincidence between orbital convergence and latitudinal canopy height gradients. For example, high latitudes have dense tracks from both instruments and generally low canopy heights. It is worth noting that the height difference within crossover grids is not only related to height accuracy at the observation scale, but more likely related to sampling density and which trees areas are being sampled. Each mission could have a different number of shots over the same grid and these shots may measure different areas of trees within the grid.

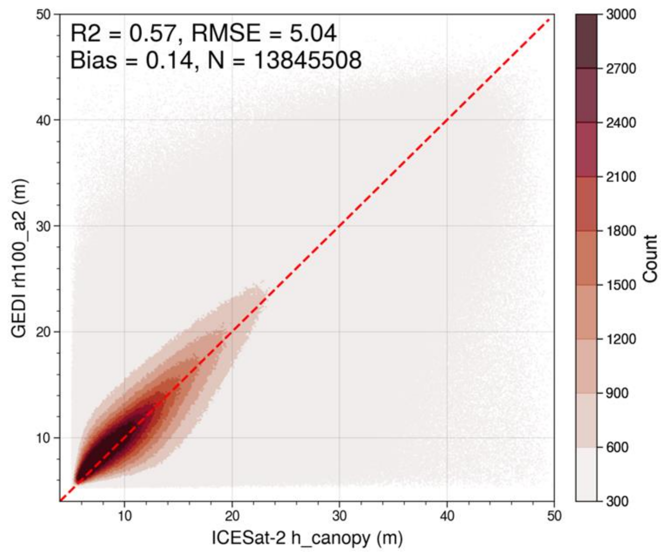


Figure 5.6. Intercomparison of 0.01° canopy height maps between GEDI and ICSat-2 at crossover grids, where there are at least two footprint/segment observations from both instruments.

5.3.2 ED Initialized AGB

ED initialized AGB (Figure 5.7) is generated using the combined canopy height histogram from GEDI and ICESat-2 and tree cover data from GFC. Initialized AGB reveals similar patterns to canopy height, but presents a larger gradient from low to high latitudes. The initialized AGB totals 155.8 Pg C over the pantropical regions (23°N ~ 23°S) and 39.2 Pg C over temperate regions (>23°N or <23°S).

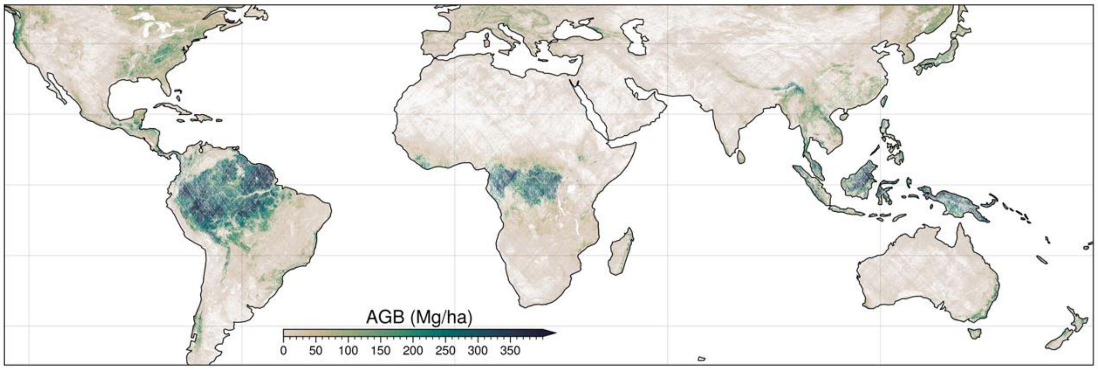


Figure 5.7. ED initialized AGB at 0.01° using the combined gridded canopy height histogram from both GEDI and ICESat-2 and tree canopy cover data from GFC.

Figure 5.8 provides a detailed comparison of initialized AGB from both missions individually and when combined. Although AGB initialized from each mission alone highlights spatial heterogeneity, coverage gaps can still be seen between tracks. With combining the two missions, AGB at high latitudes (Figure 5.8c) reflects nearly wall-to-wall coverage.

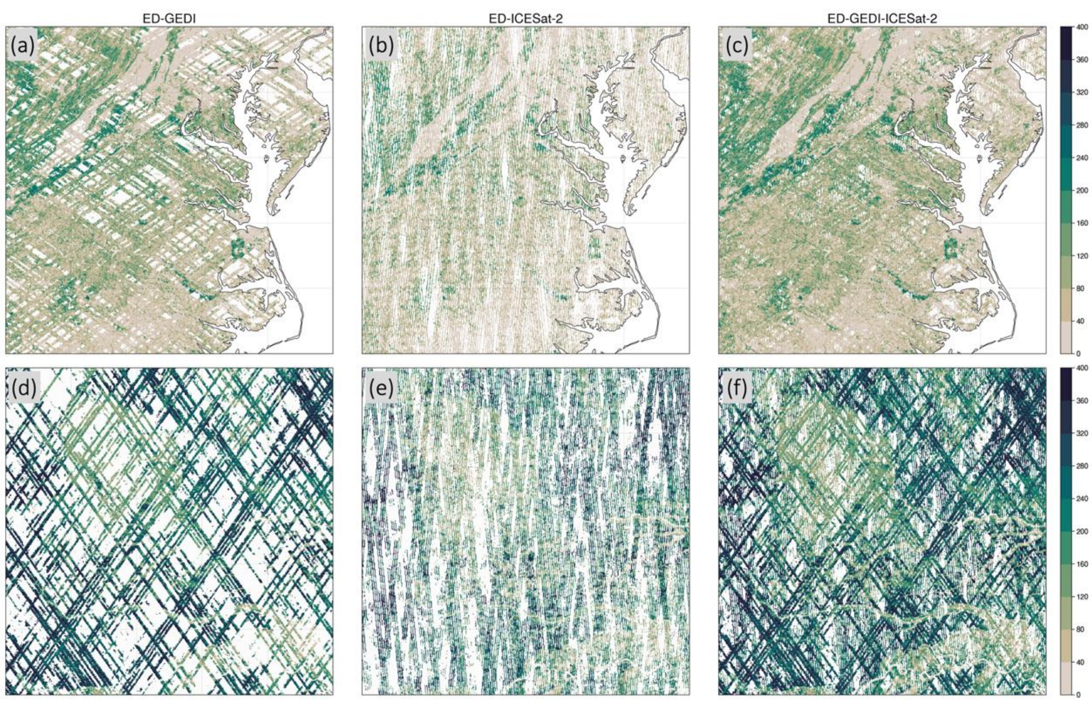


Figure 5.8. ED initialized AGB at 0.01° using gridded canopy height histogram from GEDI and ICESat-2 and tree canopy cover from GFC at eastern US (35°N ~ 40°N, 80°W ~ 75°W)

(top row) and Amazon (3°S ~ 2°N, 70°W ~ 65°W) (bottom row). (a) and (d) use gridded canopy height histogram of GEDI alone, (b) and (e) use the histogram of ICESat-2 alone, (c) and (f) use combined histogram of GEDI and ICESat-2.

Another detailed comparison of deforestation areas shows spatial heterogeneity in initialized AGB. Large areas of forest clearing could be seen in maps of both tree cover (Figure 5.9a) and loss (Figure 5.9b). Forest clearing resulted in forest fragmentation (e.g., change in tree cover), but also in forest structure (Figure 5.8c). Canopy heights from lidar observations reveal significant height variation along forest edges, despite high tree cover (>80%) across these areas. Using these observations of tree cover and canopy height as inputs, ED initialization AGB reveals spatial variation across rainforests, where low AGB appears in deforested areas and along forest edges.

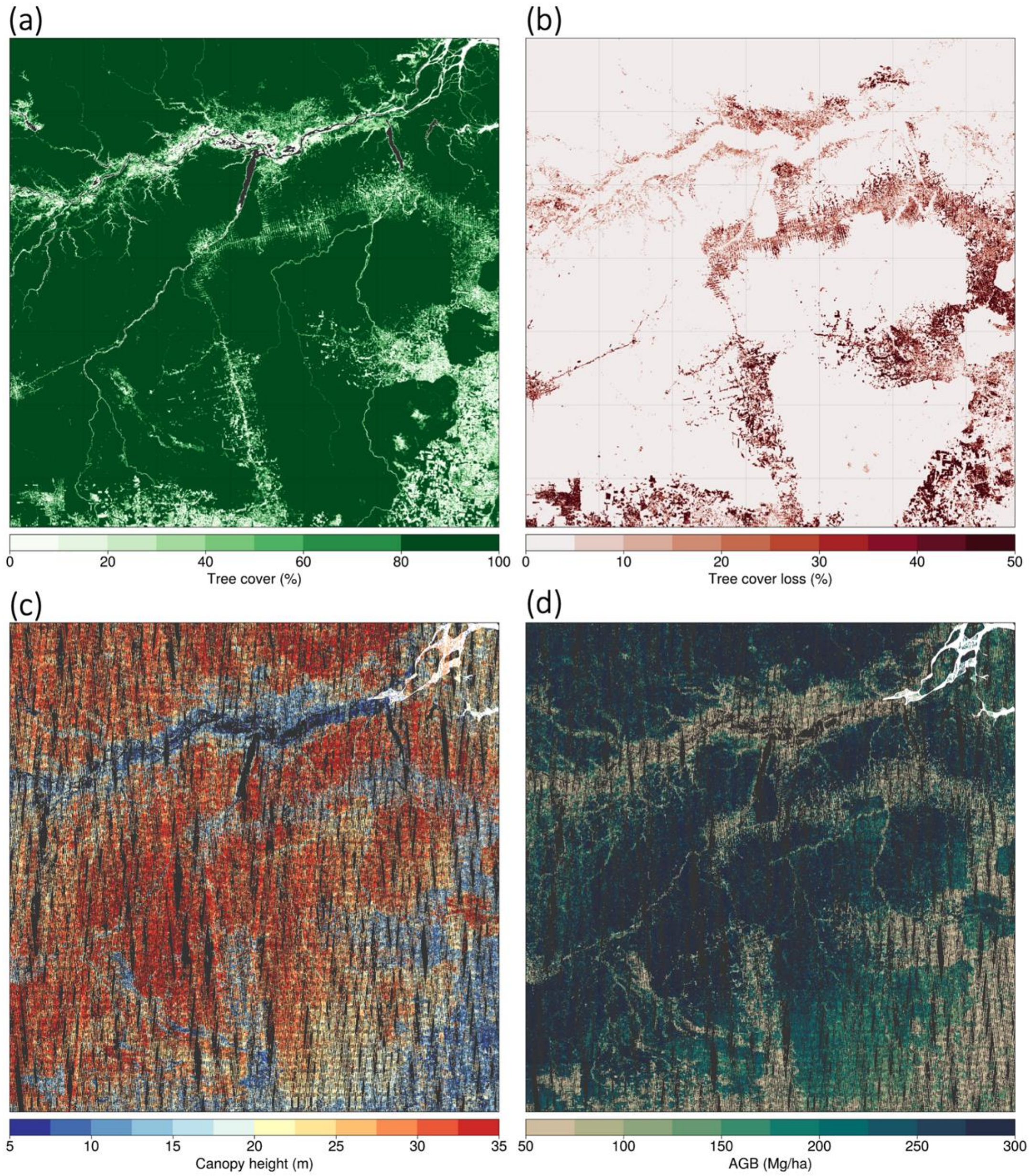


Figure 5.9. Fine-scale details in tree cover (a), tree loss between 2000 and 2010 (b), canopy height from GEDI and ICESat-2 (c) and initialized AGB (d) over a deforested area of the Brazilian Amazon (10°S ~ 0°, 60°W ~ 50°W). Tree loss included here is for identification of causes of low tree cover.

ED initialized AGB is compared across the United States using FIA measurements at the hexagon scale (Menlove and Healey, 2020). ED initialized AGB shows spatially heterogeneity across space, with higher AGB estimates in the Pacific northwest and

Appalachian Mountains. Scatter comparison (Figure 5.10d) suggests high correlation between ED initialized AGB and FIA data, with R^2 about 0.7 and average bias lower than 10 Mg/ha. However, two major negative bias can be seen. For example, the initialized AGB shows lower values in the Pacific northwest, where large redwood trees dominate, and also lower values in the Midwest region, where there is likely woodland not defined as forest in the GFC dataset.

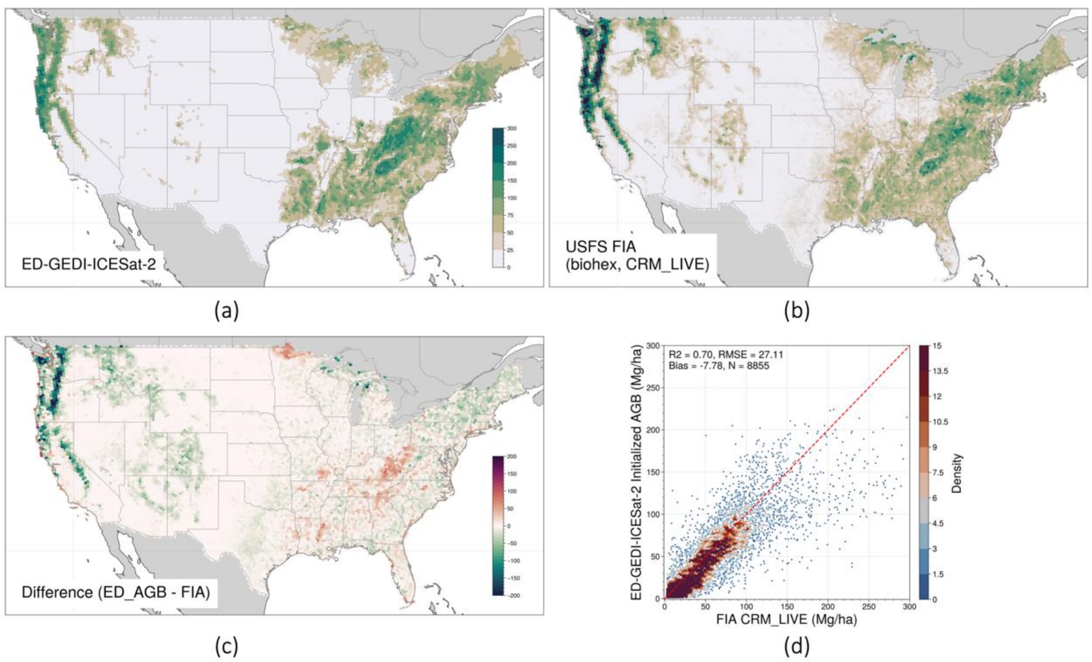


Figure 5.10. Comparison between ED initialized AGB and USFS FIA AGB at the hexagon scale, where (a) is the hexagon-scale average of ED initialized AGB using the combined canopy height histogram of GEDI and ICESat-2; (b) hexagon-scale average AGB from FIA based on the Component Ratio Method allometric equation; (c) AGB difference between (a) and (b); (d) scatter plot between (b) and (a).

5.4 Discussion and Conclusions

We developed a new approach to initialize the ED model at the global scale by combining forest structure observations from GEDI and ICESat-2. This approach produces spatial estimates of AGB at 0.01° resolution over the domain between 51°N and 51°S. Our estimates total 195 Pg C over the study domain. The AGB estimates compare favorably against the USFS FIA dataset across US CONUS. Aided by high-resolution observations of tree cover and vertical canopy structure, our gridded AGB estimates reveal fine-scale spatial heterogeneity, including effects from deforestation and logging.

Our approach integrates state-of-the-art lidar observations into an ecosystem model (i.e., ED) at global scale. This work builds from previous studies at site/region-level that highlight the advantages of model initialization with lidar observation (Hurtt et al., 2004). Linking model with lidar can establish contemporary conditions instead of relying on potential vegetation states. In comparison to other variables from passive remote sensing, lidar canopy height has greater sensitivity to forest succession and carbon status. For example, although other vegetation and leaf area indices exist from many satellite observations and offer wall-to-wall coverage at fine spatial resolutions (e.g., 30 m), these variables have limited sensitivity to forests with low biomass and younger forests that have not yet achieved canopy closure. Linking model with lidar can also increase modeling resolution to 0.01°, a spatial resolution much finer than other ecosystem models and Earth System Models in CMIP5/CMIP6 which are typically initialized with land-use history (LUH1 and LUH2).

Our approach utilizes original lidar observations at fine spatial scales without any spatial extrapolation by other auxiliary datasets, greatly preserving raw observed information. Due to spareness of lidar observations, previous attempts to estimate AGB at fine scale (e.g., 1 km) have required spatial extrapolation of lidar observations to unsampled areas with aid of auxiliary remote sensing variables (e.g., surface reflectance or vegetation indices) (Baccini et al., 2008; Saatchi et al., 2011). However, such spatial extrapolation is subject to uncertainties associated with the weak sensitivity of auxiliary observations to forest structure (Lu, 2005; Mitchard et al., 2013). Thanks to two ongoing spaceborne lidar missions, GEDI and ICESat-

2, and simultaneously collected ground samples, lidar observations of forest structure have being largely enriched across globe. For example, GEDI alone has collected almost 200 million observations over the pan-tropics during its first 3-month on-orbit (Dubayah et al., 2020a); this is a several orders of magnitude change relative to a precursor lidar mission, ICESat-1. These massive observation datasets from GEDI and ICESat-2 now offer opportunities for direct and fine-scale assessments of forest structure and carbon globally. The fine-scale example of deforestation in the Amazon (Figure 5.9) showcases the power of this data to advance models in capturing small-scale impacts of deforestation/degradation on forest structure and AGB.

We also explored the benefits of using GEDI and ICESat-2 data synergistically, taking into account spatial heterogeneity (<0.01°) in height and AGB. Both missions are sample-base measurements, implying they might be limited in their ability to provide wall-to-wall products. Given different track characteristics, combining GEDI and ICESat-2 data could greatly increase data coverage over unsampled areas from either one alone. Combining GEDI and ICESat-2 here does not involve spatial extrapolation with non-lidar auxiliary data, avoiding potential uncertainties with extrapolation approaches. In this study, we also assessed height consistency between two missions. Choosing the best comparable canopy height metric pair, the average bias between GEDI and ICESat-2 at crossover grids is about 0.14 m. More importantly, combining these data greatly increases overall data coverage as well as shot densities on crossover grids. In addition, we chose to use a gridded canopy height histogram (ranging from 5 m to 50 m with bin size of 0.5 m) instead of average canopy height as inputs for initialization. This choice is based on considerable height variation under 0.01° and related nonlinearity issues with upscaling AGB. First, forest canopy height could have large variations at sub-grid scales (Figure 5.2), resulting from disturbance (both natural and anthropogenetic), environmental conditions, species composition, and more. This heterogeneity is more pronounced in highly fragmented or population-intensive areas. Second, AGB is a nonlinear function of canopy height, which suggests that using a single average height over coarse spatial scales would not yield the same result as summing all fine-scale AGB estimates. Therefore, GEDI and ICESat-2 observations at their initial measurement scale (i.e., 25-m footprint for GEDI and 100-m segment for ICESat-2) have been gridded based on a canopy height histogram, preserving original information without aggregation to coarser spatial resolution.

Uncertainties in this study can arise from multiple sources. The first major source is the ED structural uncertainties and its model run for this study. There is inconsistent spatial resolution between the AGB-height lookup table (i.e., 0.5°) and initialization with lidar canopy height (i.e., 0.01°). AGB-height lookup is built up by driving ED at 0.5°, which is determined by the resolution of available model drivers across the study domain. Initialization using AGB-height from a coarse resolution lookup table cannot account for variability in AGB-height relationships at fine spatial scales. This may cause bias in initialized AGB estimates where environmental conditions vary widely within 0.5°, such as across mountainous regions. In addition, uncertainties may also come from height and biomass allometry, and disturbance rate prescribed in the ED model, which may bias estimates in particular ecoregions. For example, evaluation of initialized AGB over US CONUS indicate underestimation in the Pacific Northwest of North America, where there are abundant long-lived and tall tree species (e.g., redwood trees). The current model does not yet characterize these species well. The second major source of uncertainties is associated with initialization inputs from lidar canopy height and tree cover. GEDI and ICESat-2 are sampling-based missions, which means they are not meant to provide wall-to-wall coverage and rely instead on smaller footprint/segment scales (e.g., 25 ~ 100 m). There is likely to be sampling errors in lidar canopy height that propagate errors to AGB estimates. Although this study addresses sub-grid height heterogeneity by using a canopy height histogram, bias may still potentially exist where limited samples cannot fully represent the canopy height of a given population. Such bias may be more pronounced in grids with fewer shots. Furthermore, this study relies on tree cover map to indicate the fraction of each 0.01° grid with trees. Definition differences in forest conditions, or uncertainties in estimating tree cover could propagate uncertainties to the initialized AGB estimates. For example, we found the initialized AGB is lower than forest inventory data within woodland areas where the vegetation is not represented as tree cover in the GFC dataset. A third major source of uncertainty is incomplete evaluation over other regions such as tropics due to paucity of forest inventory data.

Initialization of ecosystem models with remote sensing observations can benefit ecosystem modeling in many ways. Remote sensing observations can improve model representation of forest initial conditions, which could in turn improve subsequent model simulations. Here we show that initialized ecosystem models can serve as helpful tools for attributing and quantifying the terrestrial carbon sink, and assessing responses and feedback between terrestrial ecosystems and climate. Remote sensing observations also allow ecosystem modeling at high-resolutions and the capture of fine-scale heterogeneity in forest structure and carbon storage. This could benefit future studies relating forest structure to ecosystem functioning.

Chapter 6 Conclusions

6.1 Summary of Findings

This research aimed to integrate advances in ecological modeling, remote sensing, and land-use modeling and explore their potential to improve projections of terrestrial carbon dynamics by scaling individual-based ecosystem processes and capturing contemporary forest conditions. To maximize these improvements, this research developed and calibrated a global version of ED and initialized it with land-cover history and lidar remote sensing data. Overall, this work demonstrated that integration of ED and land-cover history can simulate long-term carbon dynamics in both vegetation and soil from the pre-industrial era to present, taking into account the compounding effects of climate change, land-use change, and elevated CO2. However, this integration is limited in capturing spatial heterogeneity at fine-scales (e.g., <10 km), and may involve uncertainties propagated from land-use history estimates. In contrast, integration of ED and lidar remote sensing can improve model estimates of contemporary baseline aboveground carbon and vegetation structure, enable the model to characterize fine-scale spatial heterogeneity (e.g., 90 m – 1 km). However, this integration might be uncertain in estimating soil carbon storage since this approach focuses on aboveground vegetation carbon because it is observable from lidar remote sensing. Below are specific chapter-level findings.

Chapter 2 focused on the translation of land-use history to corresponding land-cover history, which is more relevant for ecosystem models. Choices on translation rules can result in contrasting estimates of contemporary forest area and historical land-use emissions. For example, among translations rules, estimates of global forest area in 2000 ranges from 35.7 ~ 42.7 10^6 km^2 and land-use emissions from 1850-2015 range between 108 and 195 Pg C. This work developed a translation rule for LUH2 implementation within CMIP6 models with three recommended components: (1) completely clear vegetation in land-use changes from primary and secondary land (including both forested and non-forested) to cropland, urban land and managed pasture; (2) completely clear vegetation in land-use changes from primary forest and/or secondary forest to rangeland; and (3) keep vegetation in land-use changes from primary non-forest and/or secondary non-forest to rangeland. With this translation rule, land-cover change history between 850 and 2015 is generated, which then can provide ecosystem models information on where, when, and to what extent forests are disturbed and recovered.

Chapter 3 aimed to develop, calibrate and evaluate the ED model at the global scale. Building from regional scale developments, several refinements are introduced for the model's global performance. With drivers of land-cover change history from Chapter 2, transient CO_2, transient meteorology, the global ED model simulated contemporary vegetation distribution, structure, and carbon stocks and fluxes across different temporal and spatial scales. For example, the ED model reproduced latitudinal patterns of broadleaf and needleleaf trees, resulting from competition between PFTs. The ED model also reproduced positive trends in GPP and NBP, which are evident in satellite observations and atmospheric inversions. Further, the ED model reproduced observed vertical and horizontal canopy height and leaf area. This chapter demonstrated that combining an individual-based ecosystem model with land cover history could improve the characterization of terrestrial carbon dynamics across various temporal and spatial scales.

Chapter 4 focused on high-resolution forest carbon modeling over the RGGI region by leveraging airborne remote sensing observations and the global ED model developed in Chapter 3. This work demonstrated that a global ecosystem model can characterize carbon heterogeneity and dynamics at fine spatial scales with the aid of high-resolution forest structure information from remote sensing. For example, in comparison to Chapter 3, the modeling resolution increased from ~50 km to 90 m. The resulting AGB estimates showed strong spatial heterogeneity at 90 m despite the RGGI region generally sharing similar species compositions and climate conditions. This work also suggests that the future aboveground carbon sequestration gap is much larger than current carbon stocks. Individual contributions to this sequestration gap from existing trees and new afforestation/reforestation varied from grid to grid, depending on current forest structure and cover. This work demonstrates the power of remote sensing in providing estimates of initial forest conditions that advance high-resolution ecosystem modeling at policy relevant scales.

Chapter 5 focused on exploring the potential of spaceborne remote sensing observations to advance ecosystem modeling at the global scale. This work leveraged state-of-the-art lidar observations from GEDI and ICESat-2 missions and initialized the developed global ED model at 0.01°. This work proposed an initialization approach using a canopy height histogram, which accounted for sub-grid height variations. This work found that synergistic usage of data from both missions can greatly increase spatial coverage and observation density per grid. Lidar observations show canopy height variation over dense forest. The initialized ED produced AGB estimates with spatial heterogeneity finer than that of typical global ecosystem models (i.e., ~50 km). This work also demonstrated the potential of spaceborne lidar observations to improve global forest carbon modeling, including projections of global carbon sequestration under future climate change and land-use change scenarios.

6.2 Future Research

This research separately integrated land-use change modeling and lidar remote sensing observation into the ED model. Both integration approaches certainly have their own strengths and limitations. Future research, however, can focus on comparing their results, including of vegetation carbon, soil carbon and resulting short-term and long-term dynamics. It is worth further investigation into the limitations or accuracy of using the land-use history-based approach to estimate contemporary carbon and structure as opposed to the lidar-based approach. Likewise, it is also valuable to investigate the limitations of inferring belowground carbon from lidar observations across different disturbance regimes and land-cover history. More importantly, future research can focus on developing an approach that integrates both approaches together. For example, an integrated approach may continue to initialize vegetation carbon with lidar observations, but initialize soil carbon with land-use history.

Future research can also explore the potential benefits of using these modeling approaches to answer specific questions related to carbon dynamics. The lidar initialized version of global ED could be used as a baseline estimate for projections of future carbon sequestration at a global scale, informing global reforestation commitments. Chapter 4 is an example of such a projection at the regional scale, where the resulting spatially-explicit products have demonstrated their strategic value in studies of carbon pricing and biodiversity corridor mapping (Lamb et al., 2021b; Lamb et al., 2021c). Future work might also focus on quantifying the relative contributions of climate change and CO_2 fertilization to expected forest regrowth and carbon sequestration. In addition, future research could explore a way to retrospectively simulate historical carbon dynamics by back casting from contemporary forests conditions identified by the lidar initialized ED model. On the other hand, the initialized ED with land-cover history could be used to project future carbon dynamics under land-use change scenarios.

Finally, future research can improve the ED model's characterization of belowground dynamics and other PFTs. Currently, the ED model has a relatively simple characterization of soil hydrology (i.e., one-layer bucket model), and evaluations in Chapter 3 suggest a notable discrepancy in soil carbon between reference datasets. Future modeling could include multi-layer soil hydrology and carbon modules which account for carbon and water flows between vertical layers. Future modeling could also include a soil energy budget module that could enable ED to estimate soil temperature without relying on external driver inputs. This development will also allow ED to be driven by commonly used meteorological forcing data that lack of soil temperature estimates. Given the low initialized AGB identified in several locations across the U.S. in Chapter 5, future model development could include additional PFTs that correspond to deciduous needleleaf trees (e.g., larch) and also giant redwood trees.

Appendix A Supplementary material for Chapter 2

Table A.1. Legend translation to produce a common forest canopy cover for various land cover datasets based on (Song et al., 2014). For references see Table. 2.2.

Products	Land cover class	Fraction
GLCC, MODIS LC	Forest (evergreen needleleaf; deciduous needleleaf; evergreen broadleaf; evergreen needleleaf; mixed)	0.80
	Woody savannas	0.45
	Cropland/Natural Vegetation Mosaic	0.25
	Savannas	0.20
	Open shrublands; closed shrublands; grasslands; croplands; urban and build-up; snow and ice; water bodies; permanent wetlands; barren or sparsely vegetated	0
GLC2000	Tree cover (evergreen broadleaved, closed deciduous broadleaved)	0.70
	Tree cover (evergreen needleleaf; deciduous needleleaf; mixed leaf type; regularly flooded fresh or saline)	0.575
	Mosaic: Tree cover/other natural vegetation	0.50
	Tree cover (open deciduous broadleaved)	0.275
	Mosaic: cropland/tree cover/ other natural vegetation	0.25
	Tree cover burnt; shrub cover (evergreen, deciduous); herbaceous cover; sparse herbaceous or sparse shrub cover; regularly flooded shrub and/or herbaceous cover; cultivated and managed areas; mosaic: cropland / Shrub and/or grass cover; bare areas; water bodies; snow and ice; artificial surfaces and associated areas	0
GlobCover	Closed forest (broadleaved deciduous; needle leaved evergreen)	0.70
	Closed to open forest (broadleaved evergreen or semi-deciduous, mixed broadleaved and needle leaved, broadleaved forest regularly flooded)	0.575
	Open broadleaved deciduous forest/woodland; open needle leaved deciduous or evergreen forest;	0.30
	Mosaic vegetation (grassland/shrubland/forest) / cropland; mosaic forest or shrubland / grassland	0.20
	Mosaic grassland / forest or shrubland	0.175
	Mosaic cropland / vegetation (grassland/shrubland/forest)	0.117
	Post-flooding or irrigated croplands (or aquatic); rainfed croplands; closed to open (broadleaved or needle leaved, evergreen or deciduous); closed to open herbaceous vegetation (grassland, savannas or lichens/mosses); sparse vegetation; closed broadleaved forest or shrubland permanently flooded; closed to open grassland or woody vegetation on regularly flooded or waterlogged soil; artificial surfaces and associated areas; bare areas; water bodies; permanent snow and ice	0

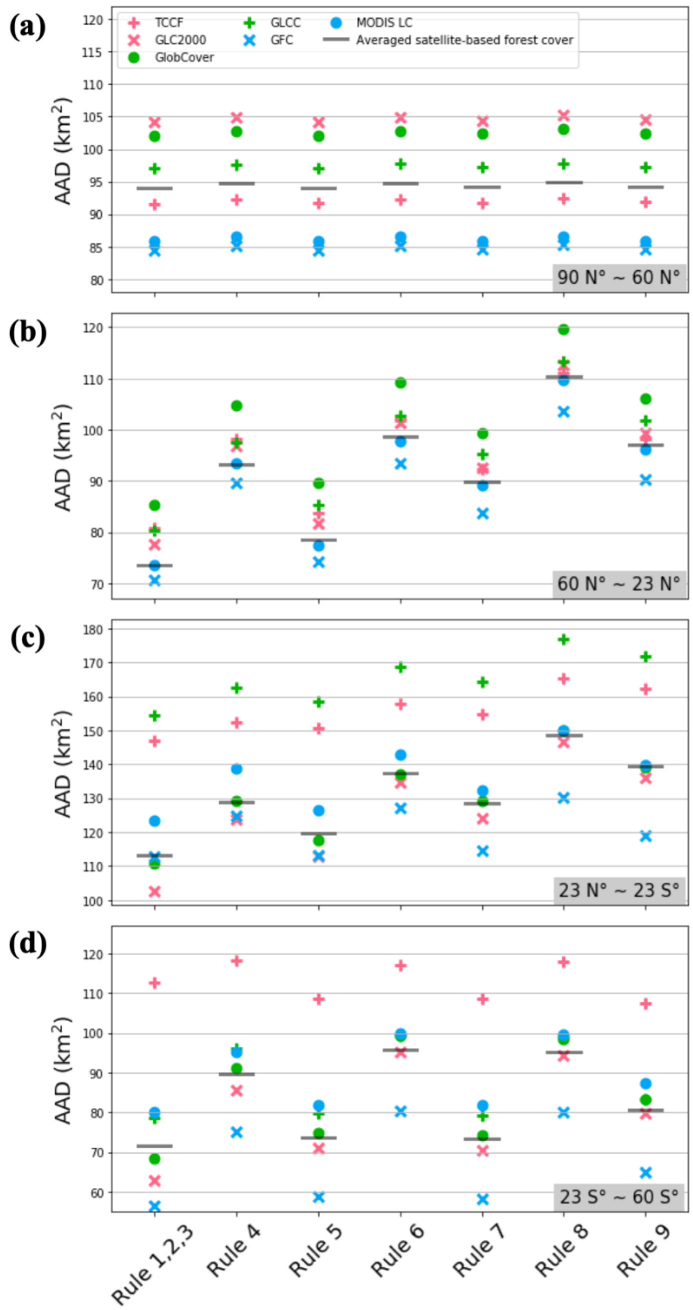


Figure A.1. Regional average of absolute difference in forest area between maps estimated by translation rules, and six satellite-based forest cover maps and the averaged satellite-based forest cover map.

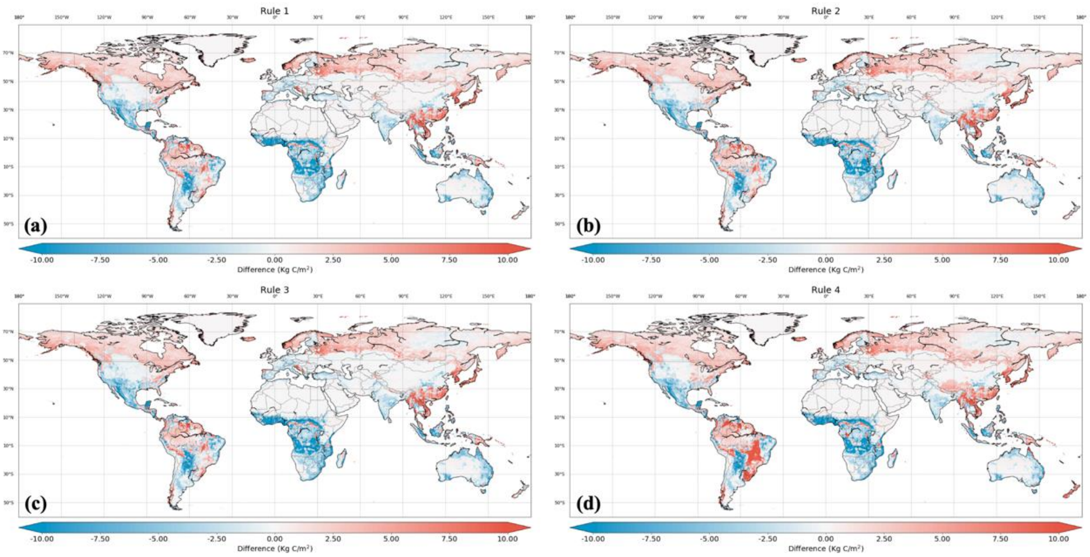


Figure A.2. Global carbon density difference between IPCC biomass Tier-1 (Figure 2.7a) density map and estimates of Rules 1-4 from (a) to (d).

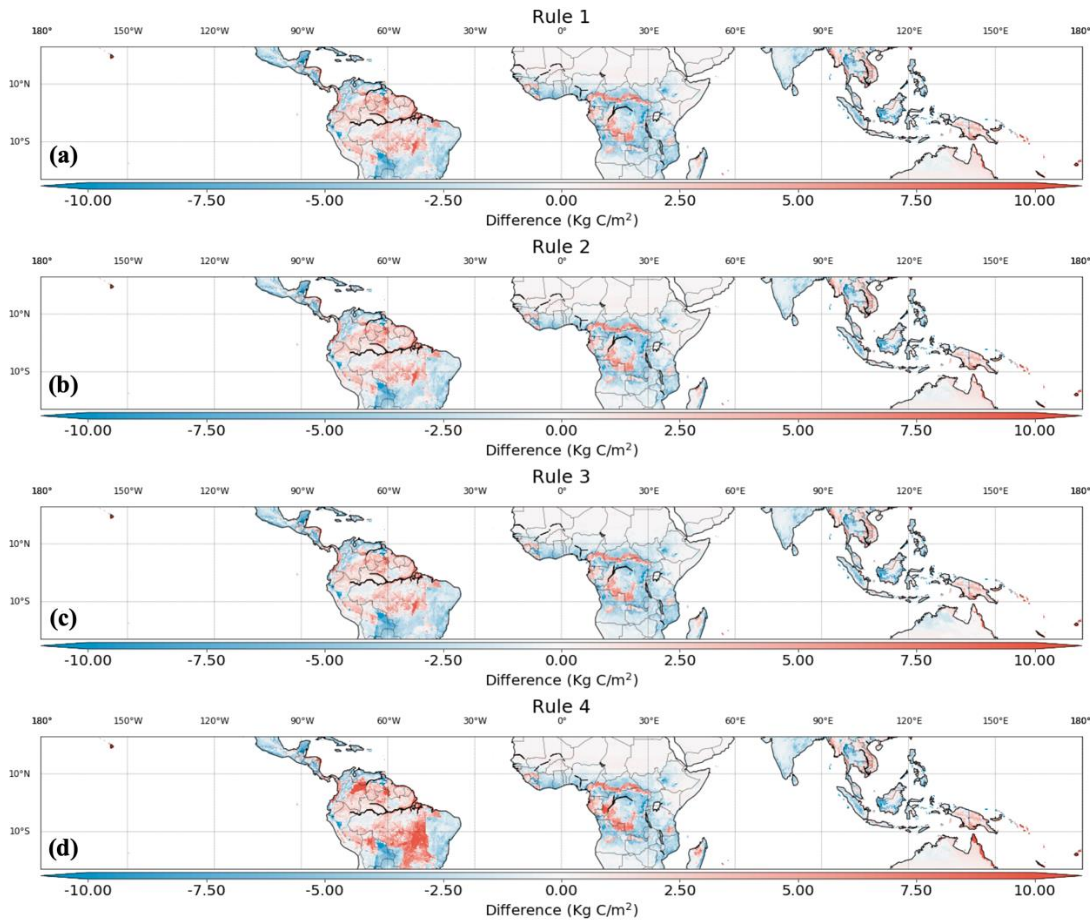


Figure A.3. Global carbon density difference between the Baccini's product (Figure 2.7b) and estimates of Rules 1-4 from (a) to (d).

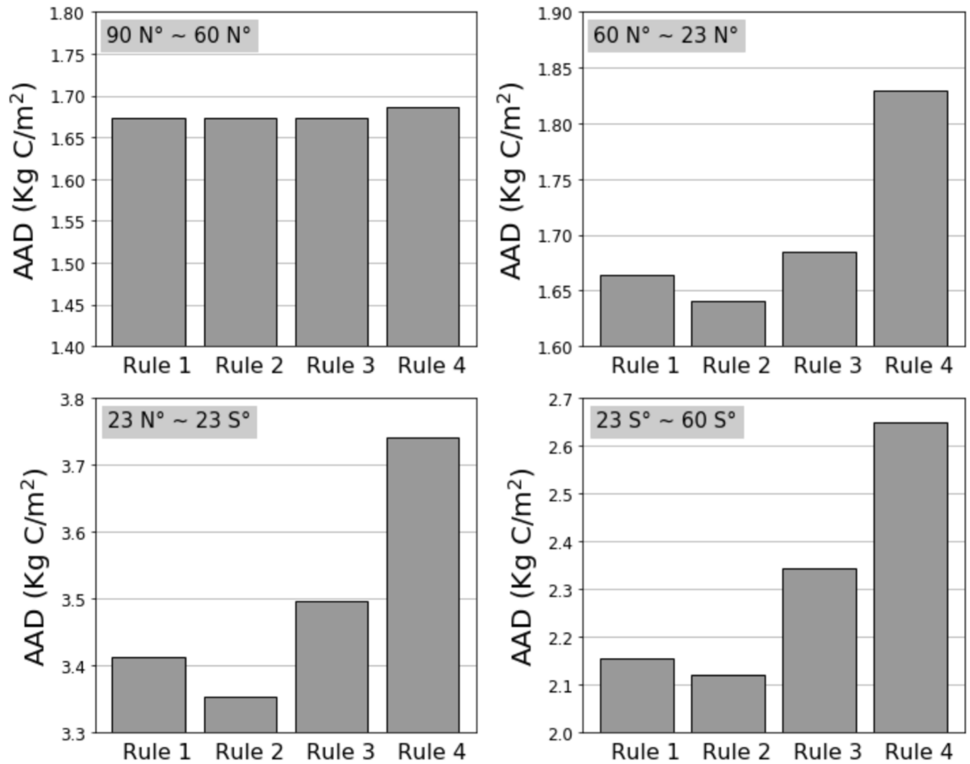


Figure A.4. Average of absolute difference in carbon density between estimations of the Rules 1-4 and the IPCC Tier-1 biomass density map at different latitudinal band zones. 'AR' represents analytical rule.

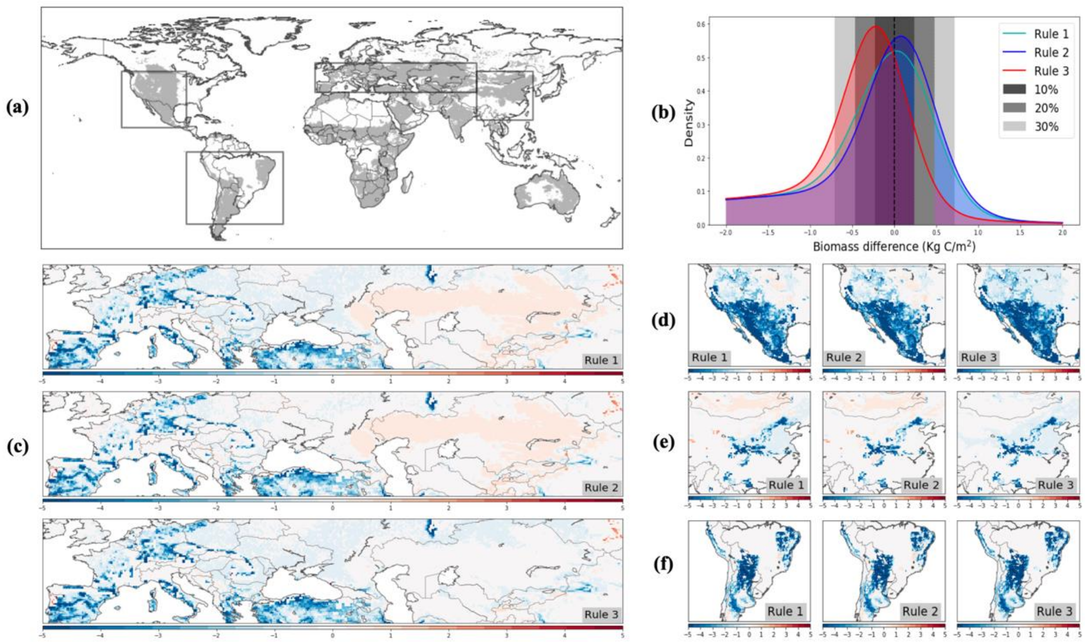


Figure A.5. estimation of Rules 1-3. (a) Shaded regions represent where Rules 1-3 differ in estimates of carbon density; (b) Histogram of carbon density difference of shaded regions in (a), shared bounds present shift range of zero line under three assumed bias levels of the IPCC Tier-1 biomass. (c) – (f) are regional comparison of carbon density difference of Rules 1-3, regions where Rules 1-3 have the same estimate of carbon density are not shown.

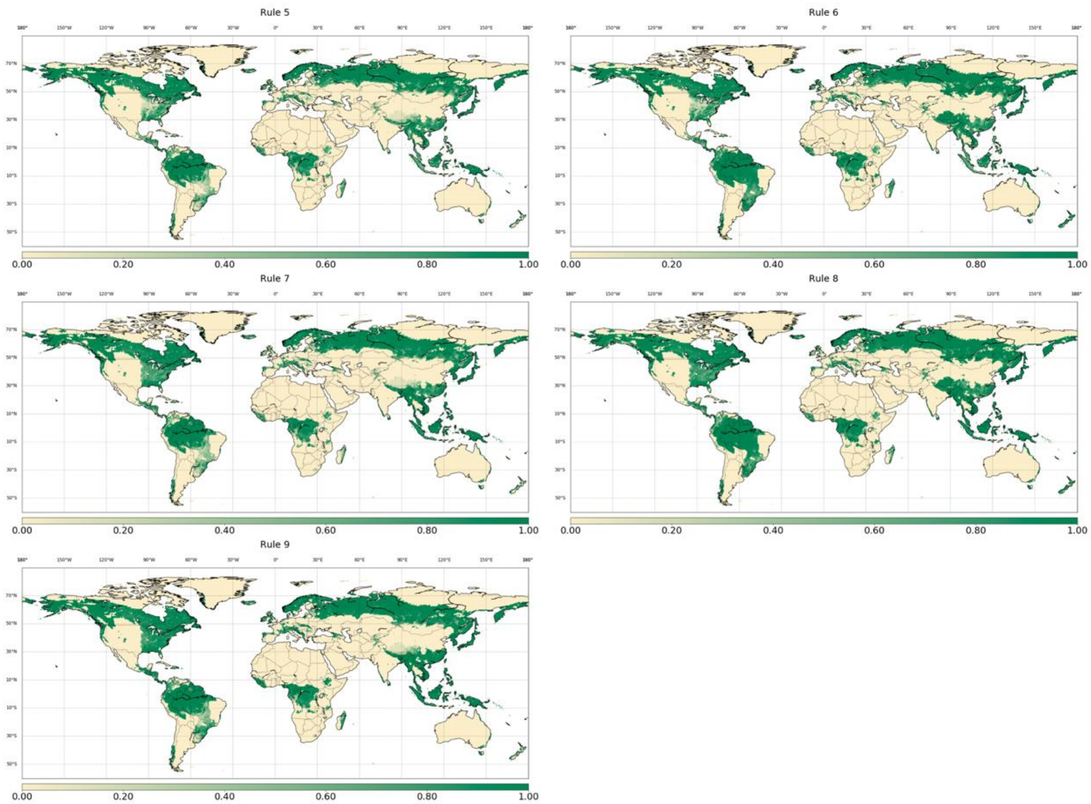


Figure A.6. Forest cover in 2000 from the Rules 5-9 respectively.

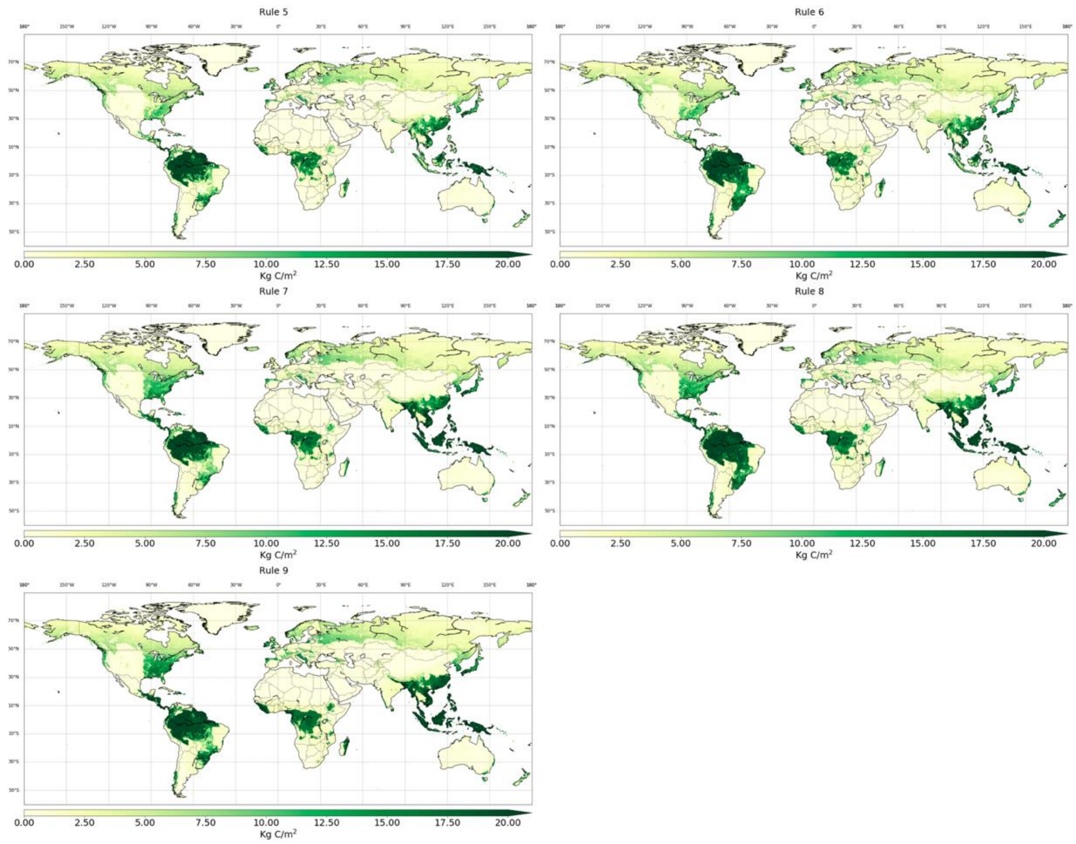


Figure A.7. Global carbon density (above- and below-ground) maps estimated by Rules 5-9 respectively.

Appendix B Supplementary material for Chapter 3

B.1. Plant functional type

Seven major PFTs are defined in ED, where two are grass and shrub type PFTs, namely C3 shrubs and grasses (C3ShG) and C4 shrubs and grasses (C4ShG), three are broadleaf PFTs, namely early-successional broadleaf trees (EaSBT), middle-successional broadleaf trees (MiSBT), and late-successional broadleaf trees (LaSBT), and the final two are needleleaf PFTs, namely northern and southern pines (NSP) and late-successional conifers (LaSC). These PFTs differ primarily in their phenology, leaf traits, allometry, dispersal, and etc. For example, grass and shrub type of PFTs are limited in the maximum height they can attain and have shorter leaf lifespans (less than 1 year). All PFTs are differentiated by their photosynthetic pathways, and C3 and C4 photosynthesis processes are modelled separately (see leaf physiology submodule). Moreover, needleleaf PFTs are characterized by slower leaf and root decay rates than broadleaf PFTs, and also utilize different allometry equations. Here, broadleaf trees are split into early-, mid- and late-successional types, which differ not only in leaf and root decay rate but also in wood density and respective allometry. The empirical relationship between leaf nitrogen content and leaf longevity, and the relationship between specific leaf area and leaf longevity, follow Moorcroft et al., 2001, which follows Reich et al., 1997.

Table B.1. Summary of PFT-dependent parameters. V_{cmax} is used in the leaf physiology submodule; $\rho(\mathbf{x})$, DBH_{max}, a_h, b_h, a_l, b_l, a_s, b_s, $l(\mathbf{x})$, $\alpha_l(\mathbf{x})$ and $\beta_r(\mathbf{x})$ are used in the plant allocation submodule; $m(\mathbf{x})$ is used in reproduction; *phenology*, $T_{crit}(\mathbf{x})$ and $T_{free}(\mathbf{x})$ are used in the leaf phenology and freezing submodule; and $\mu_{DI}(\mathbf{x})$ is used in the mortality submodule. Note that C4ShG is C4 shrubs and grasses, C3ShG is C3 shrubs and grasses, EaSBT is early-successional broadleaf trees, MiSBT is middle-successional broadleaf trees, LaSBT is late-successional broadleaf trees, NSP is northern and southern pines, and LaSC is late-successional conifers. TRO and NTRO are tropical and non-tropical variants of EaSBT, MiSBT, LaSBT.

Parameters	Description	C4ShG	C3ShG	EaSBT		MiSBT		LaSBT		NSP	LaSC
				TRO	NTRO	TRO	NTRO	TRO	NTRO		
V_{cmax}	Maximum rate of Rubisco carboxylation (μ mol m^{-2} s^{-1})	20	80	50	60	45	55	40	50	21	19
$\rho(\mathbf{x})$	Wood density (g cm^{-3})	0.53	0.53	0.53	0.53	0.71	0.71	0.90	0.90	0.70	0.70
DBH_{max}	Corresponding DBH at maximum canopy height (cm)	0.35	0.35	68.31	68.31	68.31	68.31	68.31	68.31	42.09	42.09
a_h	Coefficient of height allometry	-	-	-	-	-	-	-	-	27.14	22.19
b_h	Coefficient of height allometry	-	-	-	-	-	-	-	-	-0.0388	-0.0445
a_l	Coefficient of leaf biomass allometry	-	-	-	-	-	-	-	-	0.024	0.045
b_l	Coefficient of leaf biomass allometry	-	-	-	-	-	-	-	-	1.899	1.683
a_s	Coefficient of structural biomass allometry	-	-	-	-	-	-	-	-	0.147	0.162
b_s	Coefficient of structural biomass allometry	-	-	-	-	-	-	-	-	2.238	2.154
$l(\mathbf{x})$	Specific leaf area (m^2 kg^{-1} C)	22.03	22.03	16.02	28.50	11.64	26.55	9.66	24.42	5.55	5.55
$\alpha_l(\mathbf{x})$	Leaf biomass decay rate (yr^{-1})	2.0	2.0	1.0	3.5	0.5	3.0	0.33	2.5	0.1	0.1
$\alpha_r(\mathbf{x})$	Fine root decay rate (yr^{-1})	2.0	2.0	1.0	0.1	0.5	0.1	0.33	0.1	0.1	0.1
$\beta_r(\mathbf{x})$	Respiration coefficient	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
$m(\mathbf{x})$	Non-local dispersal rate	1.0	1.0	1.0	1.0	0.5	0.5	0.2	0.2	0.78	0.2
phenology	C-cold deciduous; D-drought-deciduous E-evergreen	C, D	C, D	C, D	C, D	C, D	C, D	C, D	C, D	E	E
$T_{crit}(\mathbf{x})$	Temperature threshold triggering leaf drop (°C)	15	5	10	10	10	10	10	10	-	-
$T_{free}(\mathbf{x})$	Temperature threshold of freezing resistance (°C)	-	-	-15	-15	-15	-15	-15	-15	-80	-80
$\mu_{DI}(\mathbf{x})$	Density independent mortality (yr^{-1})	0.081	0.081	0.081	0.032	0.054	0.032	0.025	0.014	0.014	0.014

B.2. Plant allocation submodule

Regardless of PFT type, each individual plant consists of both active tissue (B_a) and structural tissue (B_s). B_a includes leaf biomass (B_l), sapwood biomass (B_{sw}), and fine root biomass (B_r). The biomass in each active tissue component governs plant functioning. For example, leaf biomass determines the number of leaves available for leaf photosynthesis, and the fine root biomass determines the amount of water uptake from soil. Distribution of B_a to B_l, B_{sw} , and B_r is based on ratio factors of $q_l(\mathbf{z},\mathbf{x})$, $q_r(\mathbf{z},\mathbf{x})$ and $q_{sw}(\mathbf{z},\mathbf{x})$, respectively. Assuming B_l and B_r are equal for all PFTs, and the sapwood cross-sectional area is proportional to total leaf area, then $q_l(\mathbf{z},\mathbf{x})$, $q_r(\mathbf{z},\mathbf{x})$ and $q_{sw}(\mathbf{z},\mathbf{x})$ are given by:

$$q_l(\mathbf{z},\mathbf{x}) = \frac{B_l}{B_a} = \frac{1}{2 + 0.00128 l(\mathbf{x}) h} \tag{B.2.1}$$

$$q_r(\mathbf{z},\mathbf{x}) = \frac{B_r}{B_a} = \frac{1}{2 + 0.00128 l(\mathbf{x}) h} \tag{B.2.2}$$

$$q_{sw}(\mathbf{z},\mathbf{x}) = \frac{B_{sw}}{B_a} = \frac{0.00128 l(\mathbf{x}) h}{2 + 0.00128 l(\mathbf{x}) h} \tag{B.2.3}$$

Where $l(\mathbf{x})$ is dependent on PFT-specific leaf area, and h is plant height.

When the plant maintains a positive carbon balance, after taking in account respiration and decay cost from carbon fixation by photosynthesis, the gained carbon will be allocated towards the growth of B_a and B_s. The allocation fraction to B_a defined as $q_a(\boldsymbol{z},\mathbf{x})$ is based on empirical allometry equations, which ensure B_a and B_s stay on allometric trajectory. However, a negative carbon balance in the plant could result in B_a departing from its allometric trajectory as B_a needs to decrease in order to compensate for respiration and decay costs. In this case, subsequent carbon gains will all be allocated to B_a until it resumes its allometry (i.e., $q_a(\boldsymbol{z}, \mathbf{x})$ = 1).

Empirical allometry equations depict the relationship between plant height (h), leaf biomass (B_l) structural tissue (B_s) and Diameter at Breast Height (DBH). For broadleaf PFTs and grass and shrub PFTs, the allometry equations from Moorcroft et al., 2001 are used:

$$h = \begin{cases} 2.34DBH^{0.64} & if\ (DBH \leq\ DBH_{max}) \\ 2.34{DBH_{max}}^{0.64} & if\ (DBH >\ DBH_{max}) \end{cases} \tag{B.2.4}$$

$$B_l = \begin{cases} 0.0419DBH^{1.56}\rho(\mathbf{x})^{0.55} & if\ (DBH \leq\ DBH_{max}) \\ 0.0419{DBH_{max}}^{1.56}\rho(\mathbf{x})^{0.55} & if\ (DBH >\ DBH_{max}) \end{cases} \tag{B.2.5}$$

$$B_s = \begin{cases} 0.069h^{0.572}DBH^{1.94}\rho(\mathbf{x})^{0.931} & if\ (DBH \leq\ DBH_{max}) \\ 0.069{h_{max}}^{0.572}DBH^{1.94}\rho(\mathbf{x})^{0.931} & if\ (DBH >\ DBH_{max}) \end{cases} \quad \text{(B.2.6}$$

Where DBH_{max} is the corresponding DBH when h reaches its max (note that it is not the maximum that DBH the plant can grow), and $\rho(\mathbf{x})$ is PFT-dependent wood density.

For the PFTs of NSP and LaSC, the allometry equations from Albani et al., 2006 are used:

$$h = 1.3 + a_h(1 - e^{b_h DBH}) \tag{B.2.7}$$

$$B_l = \begin{cases} a_l DBH^{b_l} & if\ (DBH \leq\ DBH_{max}) \\ a_l {DBH_{max}}^{b_l} & if\ (DBH \leq\ DBH_{max}) \end{cases} \tag{B.2.8}$$

$$B_s = a_s DBH^{b_s} \tag{B.2.9}$$

Where a_h, b_h, a_l, b_l, a_s and b_s are allometry coefficients.

With ratio q_l from Equation B.2.1 and leaf biomass calculated from Equation B.2.5 or B.2.8, the active tissue biomass on the allometric trajectory is:

$$B_a^{opt} = q_l(\mathbf{z}, \mathbf{x})B_l^* \tag{B.2.10}$$

Thus, when the plant is in positive carbon balance and B_a is not smaller than B_a^{opt}, the allocation fraction of new carbon to B_a is calculated as:

$$q_a(\mathbf{z}, \mathbf{x}) = \frac{\frac{dB_a^{opt}}{dB_s}(B_s)}{1 + \frac{dB_a^{opt}}{dB_s}(B_s)} \tag{B.2.11}$$

B.3. Leaf physiology submodule

The leaf physiology submodule estimates leaf-level photosynthesis and transpiration rates which are key inputs to other submodules (e.g., growth and hydrological submodules). This submodule uses light, CO_2, air temperature, and air humidity as environmental inputs, and generates carbon assimilation and transpiration per leaf area as outputs. Three processes are coupled in this submodule: 1) photosynthesis process, which describes carbon assimilation with consideration of light availability, leaf temperature, air humidity, and CO_2 supply; 2) stomatal conductance process, which describes CO_2 diffusion from ambient air to leaf intercellular space and associated water vapor loss; and 3) leaf energy balance process, which describes the energy budget (i.e., absorbed radiation, emitted thermal radiation, and sensible and latent heat loss) for a leaf and determines leaf temperature.

B.3.1. Photosynthesis process

Photosynthesis process are modelled for C3 and C4 PFTs separately. The Farquhar, von Caemmerer & Berry model (Farquhar et al., 1980) is used to describe the C3 photosynthetic pathway. When limited soil moisture and nutrients are not limited, net photosynthesis rate per unit leaf area is the difference between the gross photosynthesis rate, A, and mitochondrial respiration, R_d. As shown in Equation B.3.1, the gross photosynthesis rate is co-limited by three processes: (1) Rubisco-limited photosynthesis rate (A_c); (2) Light-limited or RuBP regeneration-limited photosynthesis rate (A_j); and (3) Product-limited or triose phosphate utilization-limited photosynthesis rate (A_e).

$$A_n = A - R_d = \min\left(A_c, A_j, A_e\right) - R_d \tag{B.3.1}$$

The Rubisco-limited photosynthesis rate, A_c , is given by:

$$A_c = \frac{V_{cmax}(c_i - \Gamma^*)}{\left[c_i + K_c\left(1 + \frac{o_i}{K_o}\right)\right]} \tag{B.3.2}$$

Where V_{cmax} is the maximum rate of Rubisco carboxylation, c_i and o_i are the intercellular concentrations of CO_2, and O_2, respectively, Γ^* is the CO_2 compensation point, and K_c and K_o are the Michaelis-Menten constants of Rubisco for CO_2 and O_2, respectively. The RuBP regeneration-limited photosynthesis rate A_j is given by:

$$A_j = \frac{J(c_i - \Gamma^*)}{4(c_i + 2\Gamma^*)} \tag{B.3.3}$$

Where J is the electron transport rate and given by:

$$\theta J^2 - (I_{PSII} + J_{max})J + I_{PSII}J_{max} = 0 \tag{B.3.4}$$

$$I_{PSII} = \frac{1-f}{2}\alpha I \tag{B.3.5}$$

$$I = 4.55 \cdot \phi I_{g0} e^{-f_{sha}K_L \int_h^{\infty} L_{ttl}(h)} \xi \tag{B.3.6}$$

In Equation B.3.4, θ is curvature of the light response curve, I_{PSII} is the light utilized in electron transport by photosystem II, and J_{max} is the maximum rate of electron transport. In Equation B.3.5, α is leaf absorbance (set at 0.85), and f is the correction factor for spectral light quality (set at 0.15). In Equation B.3.6, I is incident photosynthetically active radiation (PAR, in unit of $\mu mol\ m^{-2}\ s^{-1}$) at leaf level with height h, I_{g0} is total shortwave radiation at the patch's canopy top, f_{sha} is the degree of shading, scaled from 0 to 1 (set at 0.5), and K_L is light extinction coefficient (set at 0.5). $L_{ttl}(h)$ is cumulative LAI from the canopy top to leaf height, calculated by summing the leaf area of all cohort plants higher than h. ξ is a coefficient representing the proportion of PAR in shortwave radiation (set at 0.5).

The export-limited photosynthesis rate (A_e) is related to rate of triose phosphate utilization (T_p), and it is given by:

$$A_e = 3 \cdot T_p \tag{B.3.7}$$

A model from von Caemmerer et al., 1999 is used to describe C4 photosynthesis. When soil moisture and nutrients are not limited, net photosynthesis rate per unit leaf area is the difference between A and R_d. The gross photosynthesis rate (A) is co-limited by: (1) Enzyme-limited photosynthesis rate (A_c) and (2) Light- and electron transport-limited photosynthesis rate (A_j).

$$A_n = A - R_d = \min(A_c, A_j) - R_d \tag{B.3.8}$$

The enzyme-limited photosynthesis rate (A_c) is given by solving a quadratic equation:

$$a{A_c}^2 + bA_c + c = 0 \tag{B.3.9}$$

Where

$$a = 1 - \frac{\alpha_o}{0.047}\frac{K_c}{K_o} \tag{B.3.10}$$

$$\begin{aligned} b = -\Bigg\{\Bigg(&\left(V_p - R_m + g_{bs}C_m\right) + (V_{cmax} - R_d) \\ &+ g_{bs}K_c\left(1 + \frac{O_m}{K_o}\right)\Bigg) \\ &+ \left(\frac{\alpha_o}{0.047}\left(\gamma_* V_{cmax} + R_d\frac{K_c}{K_o}\right)\right)\Bigg\} \end{aligned} \tag{B.3.11}$$

$$\begin{aligned} c = &(V_{cmax} - R_d)\left(V_p - R_m + g_{bs}C_m\right) \\ &- \left(V_{cmax}g_{bs}\gamma_* O_m + R_d g_{bs}K_c\left(1 + \frac{O_m}{K_o}\right)\right) \end{aligned} \tag{B.3.12}$$

Where α_o in Equation B.3.10 is the fraction of PSII activity in the bundle sheath. In Equation B.3.11 and B.3.12, C_m and O_m are the partial pressure of CO_2 and O_2 in the mesophyll, C_m equals the CO_2 intercellular partial pressure (C_i), if assuming mesophyll conductance, is infinite. g_{bs} is bundle sheath conductance to CO_2, R_m is mesophyll mitochondrial respiration, and γ_* is half of the reciprocal of Rubisco specificity. V_p is the rate of phosphoenolpyruvate (PEP) carboxylation, given by:

$$V_p = min\left\{\left(\frac{C_m V_{pmax}}{C_m + K_p}\right), V_{pr}\right\} \tag{B.3.13}$$

Where V_{pmax} is the maximum PEP carboxylation rate, K_p is the Michaelis-Menten constant for CO_2, and V_{pr} is a constant representing when PEP regeneration is limiting.

The Light- and electron transport-limited photosynthesis rate (A_j) is given by:

$$A_j = min\left\{\left(\frac{xJ}{2} + g_{bs}C_m - 0.5 \cdot R_d\right), \left(\frac{(1-x)J}{3} - R_d\right)\right\} \tag{B.3.14}$$

Where x is a partitioning factor of the electron transport rate. The electron transport rate (J) is estimated using Equation B.3.4-B.3.6, but with J_{max} value of C4 pathway.

Table B.2. Photosynthetic parameters at 25 °C for C3 and C4 pathways and coefficients of E_a, H_a, S_v and Q_{10} to characterize temperature dependency functions.

Parameter	Eqn	Unit	Temperature dependence	Coefficients					
				k_{25}	Q_{10}	E_a (J mol^{-1})	H_a (J mol^{-1})	H_d (J mol^{-1})	S_v (J mol^{-1} K^{-1})
C3 pathway									
Γ^*	3.2	μmol mol^{-1}	A-fun	42.75	-	37,830	-	-	-
K_c	3.2	μbar	A-fun	404.4	-	79,430	-	-	-
K_o	3.2	mbar	A-fun	278.4	-	36,380	-	-	-
V_{cmax}	3.2	μmol mol^{-1}	P-fun	Table B.1	-	-	71,513	200,000	636.29
R_d	3.1	μmol mol^{-1}	P-fun	$0.015V_{cmax}$	-	-	66,400	150650	490
J_{max}	3.4	μmol mol^{-1}	P-fun	$1.54V_{cmax}$	-	-	49,884	200,000	637.2
T_p	3.7	μmol mol^{-1}	P-fun	$0.09V_{cmax}$	-	-	53,100	150650	490
C4 pathway									
K_c	3.10	μbar	A-fun	650	-	-	67,294	-	-
K_o	3.10	mbar	A-fun	450	-	-	36,000	-	-
V_{pr}	3.13	μmol mol^{-1}	Q-fun	80	2.0	-	-	-	-
K_p	3.13	μmol mol^{-1}	Q-fun	80	2.0	-	-	-	-
V_{cmax}	3.11	μmol mol^{-1}	P-fun	Table B.1	-	-	67,294	144,568	472
J_{max}	3.4	μmol mol^{-1}	P-fun	$5V_{cmax}$	-	-	77,900	191,929	627
V_{pmax}	3.13	μmol mol^{-1}	P-fun	$1.4V_{cmax}$	-	-	70,373	117,910	376
R_d	3.12	μmol mol^{-1}	P-fun	$0.01V_{cmax}$	-	-	67,294	144,568	472

Among above equations, variables Γ^*, K_c, K_o, V_{cmax}, V_{pmax} , J_{max}, T_p, R_d, and V_{pr} are temperature dependent. They are described using three types of dependency functions: 1) Arrhenius function (named as A-fun); 2) peak model function (named as P-fun); and (3) Q10 function (named as Q-fun). They are given respectively by:

$$k_T = k_{25}\mathrm{e}^{E_a \frac{T_l - 25}{298(T_l + 273)R}} \tag{B.3.12}$$

$$k_T = k_{25}\mathrm{e}^{H_a \frac{T_l - 25}{298(T_l + 273)R}} \frac{1 + e^{\frac{298 S_v - H_d}{298R}}}{1 + e^{\frac{T_l S_v - H_d}{T_l R}}} \tag{B.3.13}$$

$$k_T = k_{25} Q_{10}^{\frac{T_l - 298}{10}} \tag{B.3.14}$$

Where k_{25} is the base rate of k_T at the reference temperature of 25 °C and T_l is leaf temperature in °C. E_a and H_a are both activation energy, H_d is deactivation energy, S_v is entropy term and Q_{10} is the coefficient representing the proportional change of metabolic rate per 10°C rise, and R is ideal gas constant. The P-fun function is modified from A-fun, and shows the reduction of metabolic rate at high temperatures due to the thermal breakdown of metabolic processes. Table B.2 describes this parameterization based on von Caemmerer 2000; Bernacchi et al., 2001; Massad et al., 2007; Kattge et a; 2007; von Caemmerer et al., 2009.

B.3.2. Stomatal conductance process

The stomatal conductance model governs the exchange rate of CO_2 and water vapor through leaf stomata, determining the leaf intercellular CO_2 concentration and leaf transpiration rates. Here, an empirical model called Ball-Berry-Leuning model (Ball, Woodrow & Berry 1987;

Leuning et al., 1990 and 1995) is used to describe both C3 and C4 photosynthetic pathways, and it is given by:

$$g_{sw} = g_0 + \frac{a_1 A_n}{(c_s - \Gamma^*)\left(1 + \frac{D_s}{D_0}\right)} \qquad \text{(B.3.15)}$$

Where g_{sw} is the stomatal conductance to water vapor, g_0 is g_{sw} at CO_2 compensation point, and a_1 and D_0 are empirical coefficients. D_s and c_s are vapor pressure deficit (VPD) and CO_2 partial pressure at the leaf surface. D_s is estimated as:

$$D_s = e_s(T_l) - e_a \qquad \text{(B.3.16)}$$

Where e_a is the vapor pressure of ambient air, and $e_s(T_l)$ is saturated vapor pressure at leaf temperature T_l.

The boundary layer conductance of g_{bw} to water vapor is estimated by:

$$g_{bw} = 1.4 \cdot 0.147\sqrt{\frac{u}{d}} = 1.4 \cdot 0.147\sqrt{\frac{u}{0.72w}} \qquad \text{(B.3.17)}$$

Where u is wind speed (in unit of m/s), and w is leaf width (m). With the stomatal conductance g_{sw} and the boundary layer conductance g_{bw}, the CO_2 concentration at leaf surface c_s and at leaf intercellular c_i are estimated as:

$$c_s = c_a - \frac{1.4A_n}{g_{bw}} \qquad \text{(B.3.18)}$$

$$c_i = c_s - \frac{1.6A_n}{g_{sw}} \qquad \text{(B.3.19)}$$

Where c_a is the CO_2 concentration of ambient air.

B.3.3. Leaf energy balance

If heat storage and metabolic heat production are assumed to be negligible, the energy budget of a leaf is:

$$R_{abs} - L_{oe} - H - \lambda E_l = 0 \tag{B.3.20}$$

Where R_{abs} is the absorbed shortwave and longwave radiation, L_{oe} is emitted thermal radiation, and H and λE are sensible and latent heat loss, respectively. These equations are given by:

$$L_{oe} = \varepsilon_s \sigma {T_l}^4 \tag{B.3.21}$$

$$H = c_p g_{ha}(T_l - T_a) \tag{B.3.22}$$

$$\lambda E_l = \lambda g_v \frac{e_s(T_l) - e_a}{p_a} \tag{B.3.23}$$

Where ε_s is leaf thermal emissivity, σ is the Stefan-Boltzmann constant, c_p is specific heat capacity of air, T_a is the air temperature, and E is the transpiration rate. g_{ha} and g_v are heat conductance and vapor conductance, respectively, and are given by:

$$g_{ha} = 1.4 \cdot 0.135 \sqrt{\frac{u}{0.72w}} \tag{B.3.24}$$

$$g_v = 0.5 \frac{g_{sw} g_{bw}}{g_{sw} + g_{bw}} \tag{B.3.25}$$

B.3.4. Coupling and solving three processes

The three processes of photosynthesis, stomatal conductance, and leaf energy balance are interdependent. The process of photosynthesis requires leaf temperature (T_l) and leaf intercellular CO_2 concentration (c_i) as inputs, and subsequently offers the net carbon assimilation rate (A_n) as one of its outputs. The stomatal conductance process requires T_l and A_n as inputs, and delivers estimates of c_i and g_{sw} as outputs. The leaf energy balance process requires g_{sw} as an input and in turn provides an estimate of T_l. Therefore, all three processes are solved in numerical iteration. First, A_n is obtained when photosynthesis is initialized by setting T_l and c_i at air temperature T_a and $0.7c_a$, respectively. Second, A_n is then used in the stomatal conductance process to update c_i. Steps one and two are solved using the Newton-Raphson method until the c_i is converged upon. Third, the c_i and g_{sw} from the steps one and two are used in the leaf energy balance process to solve T_l. These three steps are iterated until T_l is converged upon. As a result, the net carbon assimilation rate (A_n) and transpiration rate (E_l) are scaled up to the canopy-level and drive the growth process in other submodules.

B.3.5. Water and nitrogen constraint

The net photosynthesis in Equation B.3.1 and Equation B.3.8 and transpiration rates in Equation B.3.23 are modelled without accounting for stress from soil moisture and nitrogen availability. However, low availability of water and nitrogen could decrease photosynthesis and transpiration by limiting stomatal conductance (g_{sw}), photosynthetic capacity (V_{cmax}) or both. Following Moorcroft et al., 2001, the net photosynthesis rate, $A_n(\mathbf{r}, t, c^*)$, and transpiration rate, $E_l(\mathbf{r}, t, c^*)$, is adjusted for water and nitrogen stress using a simple approach:

$$A_n(\mathbf{r}, t, c^*) = c^* A_n + (1 - c^*) {A_n}^c \qquad \text{(B.3.26)}$$

$$E_l(\mathbf{r}, t, c^*) = c^* E_l + (1 - c^*) {E_l}^c \tag{B.3.27}$$

$$c^* = f_w f_N \tag{B.3.28}$$

Where ${A_n}^c$ and ${E_l}^c$ are net photosynthesis and transpiration when fully constrained, assuming equal to A_n and E_l at zero light input. c^* is the combined stress factor of water limitation f_w and nitrogen limitation f_N. f_w and f_N are calculated based on the ratio of water/nitrogen uptake by fine roots and that demanded by leaves. Fine root uptake is controlled by fine root biomass, the availability of water, and mineralized nitrogen in soil. f_w and f_N are equal to 0 when demand exceeds supply, and they are set to 1 if there is no limitation in supply.

B.4. Leaf phenology and freezing submodule

The total leaf area of a cohort is dynamic, resulting not only from prior carbon balance and allocation but also environmental conditions (i.e., temperature and soil water availability). Three types of dynamic phenology are considered in the model, including evergreens, where the leaves stay year-around; drought-deciduous, where leaves are reduced if soil water drops below a critical threshold (W_{crit}); and cold-deciduous, where leaves are reduced if air temperature is below a PFT-dependent threshold ($T_{crit}(\mathbf{x})$ in Table B.1). When either drought- or cold- deciduous phenology is triggered, leaf biomass B_l is set at zero. A fraction of lost leaf biomass (defined as L_frac and set at 0.5) is relocated to a non-respiring, non-decaying and non-photosynthetic pool called virtual leaf biomass B_{lv}, and the remaining biomass fraction (1- L_frac) is added to the litter pools where the associated carbon and nitrogen will be cycled within the belowground biochemical submodule. The virtual leaf biomass B_{lv} is account for within B_a but does not lead to and photosynthesis and respiration. When both soil water and air temperatures are favourable, leaf biomass B_l recovers instantly to a level depending on remaining B_a and allometry (for further details see the allocation submodule).

For cold-deciduous PFTs, freezing injury will occur if the air temperature continues to drop below the defined PFT-specific threshold of resistance ($T_{free}(\mathbf{x})$ in Table B.1). This freezing reduces virtual leaf biomass by L_frac, which is added to litter pools. Loss of virtual leaf biomass reduces B_a accordingly, in turn affecting the amount of leaf biomass can be recovered when air temperature returns to a favourable level.

B.5. Growth submodule

The growth submodule provides the growth function for $g_a(\mathbf{z}, \mathbf{x}, \bar{r}, t)$ and $g_s(\mathbf{z}, \mathbf{x}, \bar{r}, t)$, as a result of the carbon balance between carbon assimilation and respiration. This submodule follows Moorcroft at al 2001. Plants gain carbon through leaf photosynthesis and lose carbon by respiration and decay of leaves and roots (decay and respiration of sapwood and structural tissues are assumed to be negligible), and devote remaining carbon to production and growth of active and structural tissue. This process is given by:

$$Prod = A(\mathbf{r}, t, c^*)l(\mathbf{x})B_l - R_d l(\mathbf{x})B_l - \beta_r(\mathbf{x})B_r f(T_s) - \alpha_l(\mathbf{x})B_l - \alpha_r(\mathbf{x})B_r \quad \text{(B.5.1)}$$

Where $Prod$ is net carbon production. On the right-hand side, the first term represents total gross carbon fixation by all leaves, the second and third terms represent biomass and temperature dependent respiration of leaves and fine roots, respectively. The last two terms represent decay of leaves and fine roots, respectively, and are only related to biomass. $A(\mathbf{r}, t, c^*)$ and R_d are the gross photosynthesis rate and leaf respiration per unit leaf area with resource $\mathbf{r}$ (light, water, CO_2) and soil water stress (c^*) at time t. $l(\mathbf{x})$ is specific leaf area

(SLA), β_r is the respiration coefficient for fine root, and $f(T_s)$ is the dependence function of respiration on soil temperature (T_s). α_l and α_r are the decay rates of leaves and fine roots, respectively, with values reciprocal to longevity.

The net carbon production $Prod$ can be positive or negative depending on environmental conditions and leaf conditions. This variability results in several cases where carbon is differentially partitioned between the growth of active tissues, structural tissues, and reproduction. When $Prod$ is positive:

$$g_a(\mathbf{z}, \mathbf{x}, \bar{r}, t) = Prod \cdot [1 - rp(\mathbf{x})] \cdot q_a(\mathbf{z}, \mathbf{x}) \tag{B.5.2}$$

$$g_s(\mathbf{z}, \mathbf{x}, \bar{r}, t) = Prod \cdot [1 - rp(\mathbf{x})] \cdot [1 - q_a(\mathbf{z}, \mathbf{x})] \tag{B.5.3}$$

$$RP(\mathbf{z}, \mathbf{x}, \bar{r}, t) = Prod \cdot rp(\mathbf{x}) \tag{B.5.4}$$

Where $rp(\mathbf{x})$ defines the fraction of $Prod$ used for reproduction, $q_a(\mathbf{z}, \mathbf{x})$ represents the fraction of new growth devoted to active tissues B_a (calculated in Equation B.2.11), and $RP(\mathbf{z}, \mathbf{x}, \bar{r}, t)$ is total carbon allocated for new seedlings (see more details in the reproduction submodule). Positive $Prod$ represent situations where a plant's carbon fixation from photosynthesis is sufficient for growth and reproduction, even after deducting carbon losses due to respiration and decay.

In contrast, negative $Prod$ happens when environmental conditions do not favour photosynthesis (e.g., dry air forces leaf stomata closed) or when leaf drop is triggered by soil water stress or low air temperatures. In this case:

$$g_a(\mathbf{z}, \mathbf{x}, \bar{r}, t) = Prod \tag{B.5.5}$$

$$g_s(\mathbf{z}, \mathbf{x}, \bar{r}, t) = 0 \tag{B.5.6}$$

$$RP(\mathbf{z}, \mathbf{x}, \bar{r}, t) = 0 \tag{B.5.7}$$

where all of $Prod$ is used for the plant's active tissue.

B.6. Reproduction submodule

Plants in positive carbon balance maintain enough carbon $RP(\mathbf{z}, \mathbf{x}, \bar{r}, t)$ to reproduce seedings. The fecundity $F(\mathbf{z}, \mathbf{x}, a, t)$ is calculated as:

$$F(\mathbf{z}, \mathbf{x}, a, t) = \frac{RP(\mathbf{z}, \mathrm{x}, \bar{r}, t)}{B_{a0} + B_{s0}}(1 - \lambda_{SD}) \tag{B.6.1}$$

Where B_{a0} and B_{s0} are the initial active and structural biomass of a seedling with functional type **x**, and $1 - \lambda_{SD}$ is the probability of seeding survivorship (λ_{SD} = 0.95). The dead seedlings will be loaded into the soil pools for later carbon and nitrogen decomposition.

Seedling dispersal includes local dispersal, which limits seedlings to the siting patch (i.e., local patch), and non-local dispersal, which distributes seedling to all other patches. Thus, for any patch, it will receive seedlings not only from all plants of different sizes in its own cohorts but also from plants in other patches. Received seedlings will form a new cohort at the local patch, where plant individual density of the new cohort is represented as:

$$n_i(z_0, \mathbf{x}, a, t) \tag{B.6.2}$$
$$= \frac{1}{G_0} \int_0^{\infty} F(\mathbf{z}, \mathbf{x}, a, t)\, n_i(\mathbf{z}, \mathbf{x}, a, t)(1 - m(\mathbf{x}))\, d\mathbf{z}$$
$$+ \frac{1}{G_0} \frac{1}{p_i(a, t)} \int_0^{\infty} \int_0^{\infty} F(\mathbf{z}, \mathbf{x}, a, t)\, n_i(\mathbf{z}, \mathbf{x}, a, t) p_i(a, t) m(\mathbf{x})\, da\, d\mathbf{z}$$

Where $m(\mathbf{x})$ is the PFT-dependent non-local dispersal rate, representing the fraction of plant seedings that will be dispersed to other non-local patches. The first term on the right-hand side of the equation represents seedlings received from all cohorts within the local patch, and the second term represent seedlings from non-local patches.

B.7. Mortality submodule

The plant mortality rate $\mu(\boldsymbol{z}, x, \bar{r}, t)$ includes density-independent $\mu_{DI}(\mathrm{x})$ and density-dependent $\mu_{DD}(z, \mathrm{x}, \bar{r}, t)$ components, where:

$$\mu(\boldsymbol{z}, \mathrm{x}, \bar{r}, t) = \mu_{DI}(\mathrm{x}) + \mu_{DD}(z, \mathrm{x}, \bar{r}, t) \tag{B.7.1}$$

The density-independent $\mu_{DI}(\mathbf{x})$ component is related to disturbance, wood density, and life-history of a PFT, such that $\mu_{DI}(\mathbf{x})$ is the sum of disturbance related $\mu_{DI-DIS}(\mathbf{x})$ and wood-density related $\mu_{DI-\rho}(\mathbf{x})$ components. $\mu_{DI-DIS}(\mathbf{x})$ is set at 0.014 and 0.012 in tropical and non-tropical region, respectively. $\mu_{DI-\rho}(\mathbf{x})$ varies by PFT. For example, in comparison to the late-successional broadleaf PFT, the early- and mid-successional broadleaf PFTs have relatively higher rates of carbon accumulation and lower wood densities, making them susceptible to pathogen attack and to windthrow disturbance. Thus, $\mu_{DI-\rho}(\mathbf{x})$ decreases for early- to mid- and late-successional PFTs. In addition, the tropical variant of the broadleaf

PFTs, has higher $\mu_{DI-\rho}(\mathbf{x})$ than the non-tropical variant. $\mu_{DI}(\mathbf{x})$ for each PFT is shown in Table B.1.

The density-dependent $\mu_{DD}(z, \mathbf{x}, \bar{r}, t)$ component is related to averaged carbon balance over certain time in past, it is calculated as:

$$\mu_{DD}(z, \mathbf{x}, \bar{r}, t) = \frac{10}{1 + e^{20 \frac{\int_{t-\Delta t}^{t} Prod(t)dt}{\int_{t-\Delta t}^{t} Prod_{FS}(t)dt}}} \tag{B.7.2}$$

Where $\int_{t-\Delta t}^{t} Prod(t)dt$ is the cumulative carbon balance of a plant from time $t - \Delta t$ to t, and $\int_{t-\Delta t}^{t} Prod_{FS}(t)\, dt$ is the cumulative carbon balance of the plant under full sun conditions. $\mu_{DD}(z, \mathbf{x}, \bar{r}, t)$ is a nonlinear function of light competition, namely shading from other plants could result in mortality rate.

B.8. Soil biogeochemical submodule

The soil biogeochemical submodule tracks belowground carbon and nitrogen dynamics using a simplified Century model (Parton et al., 1987, 1993). This submodule primarily follows Moorcroft et al., 2001. For each patch, three carbon pools are tracked: structural litter carbon pool $C_1(a, t)$, metabolic litter carbon pool $C_2(a, t)$ and soil slow carbon pool $C_3(a, t)$. By assuming nitrogen is mostly bonded in carbon, nitrogen dynamics have the same three pools as carbon plus a mineralized nitrogen pool which stores nitrogen in plant-available forms (nitrate and ammonium).

Decaying tissues from living plants, and active and structural tissues of dead plants are loaded into structural and metabolic litter carbon pools $C_1(a, t)$ and $C_2(a, t)$. A fraction of both decaying active tissues and dead plant active tissue enter $C_1(a, t)$, with the rest entering

$C_2(a,t)$. Two litter pools decompose the carbon with different decomposition rates; decomposition rates of both two pools depend on defined intrinsic decomposition rates and soil moisture, while the decomposition rate of structural pool is additionally controlled by lignin content in the pool. All decomposed carbon from the metabolic litter pool and part of that from the structural litter pool is lost a part of heterotrophic respiration (RH). The rest of the carbon from the structural litter pool is transported to the slow soil carbon pool, where its decomposed at a relative slower rate. Thus, at time t, change rates of structural, metabolic litter and slow soil carbon pools are given:

$$\frac{dC_1(a,t)}{dt} = C_{1,decay}(r,a,t) + C_{1,dead}(r,a,t) - C_{1,decomp}(r,a,t) \tag{B.8.1}$$

$$\frac{dC_2(a,t)}{dt} = C_{2,decay}(r,a,t) + C_{2,dead}(r,a,t) - C_{2,decomp}(r,a,t) \tag{B.8.2}$$

$$\frac{dC_3(a,t)}{dt} = (1 - r_{stsc})C_{1,decomp}(r,a,t) - C_{3,decomp}(r,a,t) \tag{B.8.3}$$

Where $C_{1,decay}(r,a,t)$ and $C_{1,dead}(r,a,t)$ represent the carbon loaded to the structural litter carbon pool from decaying tissues of living plants, and active and structural tissues from dead plants and seedlings, respectively, $C_{1,decomp}(r,a,t)$ is decomposed carbon from the structural litter carbon pool. $C_{2,decay}(r,a,t)$ and $C_{2,dead}(r,a,t)$ represent carbon loaded into the metabolic litter carbon pool from decaying tissues of living plants, and active and structural tissues from dead plants seedlings, respectively. $C_{2,decomp}(r,a,t)$ is decomposed carbon from the metabolic litter carbon pool. Decomposition rates for the three pools are calculated as:

$$C_{1,decomp}(r,a,t) = A\big(a,t,T_s,W(a,t)\big)K_1e^{-3L_s}C_1(a,t) \tag{B.8.4}$$

$$C_{2,decomp}(r,a,t) = A\big(a,t,T_s,W(a,t)\big)K_2C_2(a,t) \tag{B.8.5}$$

$$C_{3,decomp}(r,a,t) = A\big(a,t,T_s,W(a,t)\big)K_3C_3(a,t) \tag{B.8.6}$$

Where $A\big(a,t,T_s,W(a,t)\big)$ is a combined factor (ranging from 0-1) of soil temperature and moisture, K_1, K_2 and K_3 are constant coefficients, and L_s is the relative fraction of lignin in the structural carbon pool. With Equation B.8.1, B.8.2 and B.8.3, the total heterotrophic respiration at time *t* is:

$$\begin{aligned} R_h(a,t) = r_{stsc}C_{1,decomp}(r,a,t) + C_{2,decomp}(r,a,t) \\ + C_{3,decomp}(r,a,t) \end{aligned} \tag{B.8.7}$$

Nitrogen pools include the structural litter nitrogen pool $N_1(a,t)$, metabolic litter nitrogen pool $N_2(a,t)$, soil slow nitrogen pool $N_3(a,t)$, and mineralized nitrogen pool $N_4(a,t)$. Nitrogen is assumed to largely be bonded with carbon. The carbon to nitrogen ratio is fixed at 150 for the structural litter pool and 10 for the soil slow pool but floating for the metabolic pool depending on the PFT's leaf nitrogen content. Nitrogen dynamics across pools are similar to carbon dynamics, except that the nitrogen attached to carbon lost during heterotrophic respiration is assumed to be mineralized, and subsequently added to the mineralized nitrogen pool $N_4(a,t)$:

$$\begin{aligned} \frac{dN_1(a,t)}{dt} = N_{1,decay}(r,a,t) + N_{1,dead}(r,a,t) \\ - N_{1,immbo}(r,a,t) \end{aligned} \tag{B.8.8}$$

$$\frac{dN_2(a,t)}{dt} = N_{2,decay}(r,a,t) + N_{2,dead}(r,a,t) - N_{2,min}(r,a,t) \tag{B.8.9}$$

$$\frac{dN_3(a,t)}{dt} = N_{1,immbo}(r,a,t) - N_{3,min}(r,a,t) \tag{B.8.10}$$

$$\frac{dN_4(a,t)}{dt} = N_{2,min}(r,a,t) + N_{3,min}(r,a,t) - N_{up}(r,a,t) - N_{lea}(r,a,t) \tag{B.8.11}$$

Where $N_{1,decay}(r,a,t)$ and $N_{1,dead}(r,a,t)$ are nitrogen inputs into the structural litter nitrogen pool from decaying tissues of living, and active and structural tissue from dead plants and seedlings, respectively. $N_{1,immbo}(r,a,t)$ is decomposed nitrogen which will be transported to the soil slow nitrogen pool. $N_{2,decay}(r,a,t)$ and $N_{2,dead}(r,a,t)$ are nitrogen inputs to the metabolic litter nitrogen pool from either the decaying tissues of living, and active and structural tissue from dead plants and seedlings, respectively. $N_{2,min}(r,a,t)$ and $N_{3,min}(r,a,t)$ are mineralized nitrogen from the metabolic litter and soil slow pools, and $N_{up}(r,a,t)$ is nitrogen uptake by plants. $N_{lea}(r,a,t)$ is leached nitrogen, which is assumed to be linearly related to the percolation and runoff rate $perc(a,t)$ which is calculated in hydrology submodule. Nitrogen flows in the above equations are calculated stoichiometrically as a product of the corresponding carbon flow and carbon to nitrogen ratio.

B.9. Hydrology submodule

The hydrology submodule tracks incoming soil water flow from precipitation and snow melt and outgoing flow through percolation, runoff, and evapotranspiration from the soil and plant canopy. At time t, soil water change rate is given by:

$$\frac{dW(a,t)}{dt} = P(a,t) + SM(a,t) - perc(a,t) - E_{soil,canopy}(a,t) - W_{up}(a,t) \quad \text{(B.9.1)}$$

Where $W(a,t)$ is soil water availability, $P(a,t)$, $SM(a,t)$ are incoming water flux from snowmelt, and $perc(t)$ is water loss due to percolation and runoff, $E_{soil,canopy}(a,t)$ is water loss due to evaporation from the soil and canopy, and $W_{up}(a,t)$ is plant water uptake for transpiration.

$W_{up}(a,t)$ equals the total transpiration of all leaves:

$$W_{up}(a,t) = \int_0^\infty E_l(\mathbf{r},t,c^*)l(\mathbf{x})B_l n_i(\mathbf{z},\mathbf{x},a,t)\,d\mathbf{z} \quad \text{(B.9.2)}$$

Where $E_l(\mathbf{r},t,c^*)$ is the leaf transpiration rate per leaf area, given in the leaf physiology submodule.

When the monthly average air temperature drops below the freezing point, precipitation falls as snow to accumulate snowpack; no water is loaded into the soil. When the monthly average air temperature rises above the freezing point, precipitation falls as rain and snowpack start to melt at a rate linearly related to air temperature until depletion; both precipitation and snowmelt are loaded into the soil. The snowmelt and snowpack change rate is given by:

$$\frac{dSP(a,t)}{dt} = P_s(a,t) - SM(a,t) \quad \text{(B.9.3)}$$

$$SM(a,t) = \begin{cases} 0, & T_a < 0°\text{C or } SP(a,t) = 0 \\ T_a k_{melt}, & T_a \geq 0°\text{C and } SP(a,t) > 0 \end{cases} \quad \text{(B.9.4)}$$

Where $SP(a,t)$ is snowpack, $P_s(a,t)$ equals to $P(a,t)$ when air temperature is below the freezing point and otherwise equal to zero. k_{melt} is the coefficient constant of the melting rate, set at $100\ mm\ °\mathrm{C}^{-1}\ month^{-1}$. Snowmelt ceases when cumulated snowpack is depleted.

Percolation and runoff rate $perc(a,t)$ is related to hydraulic conductivity, which is a nonlinear function of soil water availability. This relationship is given as:

$$perc = K_{sat,MvG} S_e(a,t)^{L_{MvG}} (1 - (1 - S_e(a,t)^{\frac{1}{m_{MvG}}})^{m_{MvG}})^2 \tag{B.9.5}$$

$$S_e(a,t) = \frac{\frac{W(a,t)}{d_{soil}} - \theta_{res,MvG}}{\theta_{sat,MvG} - \theta_{res,MvG}} \tag{B.9.6}$$

Where $K_{sat,MvG}$ is saturated hydraulic conductivity, $\theta_{res,MvG}$ and $\theta_{sat,MvG}$ are residual and saturated volumetric water content. $S_e(a,t)$ is effective volumetric saturation, d_{soil} is soil depth (in mm). L_{MvG} and m_{MvG} are Mualem–van Genuchten (MvG) coefficients (van Genuchten, 1980), specified by gridded soil hydraulic data external to ED (e.g., Montzka et al., 2014).

Evaporation from the soil and canopy is estimated using a model developed by Mu et al., 2011, with the sum represented as:

$$E_{soil,canopy}(a,t) = E_{canopy}(a,t) + E_{soil}(a,t) \tag{B.9.7}$$

Both $E_{soil}(a,t)$ and $E_{canopy}(a,t)$ are estimated based on the Penman-Monteith (P-M) equation (Monteith, 1965):

$$\lambda E = \frac{s \cdot R + \frac{\rho \cdot c_p \cdot (e_{sat} - e)}{r_a}}{s + \gamma \cdot \left(1 + \frac{r_s}{r_a}\right)} \tag{B.9.8}$$

Where s is slope of the curve relating saturated water vapor pressure e_{sat} to temperature, R is available energy partitioned between sensible heat, latent heat, and soil heat fluxes, ρ is air density, c_p is the specific heat capacity of air, γ is the psychrometric constant, r_a is aerodynamic resistance, r_s is an effective resistance to evaporation from the land surface. Calculations of r_a and r_s are different for soil and canopy.

Canopy evaporation $E_{canopy}(a,t)$ comes from wet canopy which intercepts precipitation. Based on the P-M equation, $E_{canopy}(a,t)$ is given by:

$$E_{canopy}(a,t) = \frac{1}{\lambda} \frac{\left[s \cdot R_{canopy} + \frac{\rho \cdot c_p \cdot (e_{sat} - e)}{rhrc}\right] \cdot F_c \cdot F_{wet}}{s + \frac{P_a \cdot c_p \cdot rvc}{\lambda \cdot \varepsilon \cdot rhrc}} \tag{B.9.9}$$

Where R_{canopy} is part of R in Equation B.9.8 allocated to canopy, F_c is the patch fraction covered by plants, and F_{wet} is the wet fraction of the land surface, correlated to air relative humidity (Fisher et al., 2008). $rhrc$ and rvc are aerodynamic resistance and wet canopy resistance to evaporation from wet canopy. Calculation of F_c, R_c, F_{wet}, $rhrc$, and rvc can be found in Mu et al., 2011.

Soil evaporation $E_{soil}(a,t)$ consists of potential evaporation from both the saturated soil surface and moist soil surface, thereby $E_{soil}(a,t)$ equals to:

$$E_{soil}(a,t) = E_{wet\_soil}(a,t) + E_{pot\_soil}(a,t)\left(\frac{e_{sat} - e}{100}\right)^{(e_{sat}-e)/200} \quad \text{(B.9.10)}$$

Then $E_{wet\_soil}(a,t)$ and $E_{pot\_soil}(a,t)$ are estimated as:

$$E_{wet\_soil}(a,t) = \frac{1}{\lambda}\frac{\left[s \cdot R_{soil} + \frac{\rho \cdot c_p \cdot (1 - F_c) \cdot (e_{sat} - e)}{ras}\right] \cdot F_{wet}}{s + \frac{\gamma \cdot rtot}{ras}} \quad \text{(B.9.11)}$$

$$E_{pot\_soil}(a,t) = \frac{1}{\lambda}\frac{\left[s \cdot R_{soil} + \frac{\rho \cdot c_p \cdot (1 - F_c) \cdot (e_{sat} - e)}{ras}\right] \cdot (1 - F_{wet})}{s + \frac{\gamma \cdot rtot}{ras}} \quad \text{(B.9.12)}$$

Where R_{soil} is the portion of R in Equation B.9.8 allocated to the soil surface, ras is the aerodynamic resistance at the soil surface, and $rtot$ is the sum of the soil surface resistance and aerodynamic resistance to water vapor transport. Calculation of ras and $rtot$ is related to air temperature, and further details can be found in Mu et al., 2011.

$E_{soil}(a,t)$ and $E_{canopy}(a,t)$ are calculated separately for day and night, using the same equations but different parameter values. The sum of both day and night evaporation is then weighted by the daytime fraction.

B.10. Disturbance and fire submodule

The disturbance submodule describes the impacts of natural disturbance (treefall, hurricane, and fire) on patch and cohort dynamics as well as the associated carbon cycle. Disturbance impact on patch demography has been depicted in the patch dynamic PDE equation, where the second term on the right-hand side denotes changes in the proportion of patch natural disturbance. Currently three types of disturbance are included: treefall, hurricane, and fire. The disturbance rate $\lambda_i(a,t)$ is given by:

$$\lambda_i(a,t) = max\left(\lambda_{treefall} + \lambda_{hurricane}, \lambda_{fire}(a,t)\right) \tag{B.10.1}$$

Where $\lambda_{treefall}$ is set as 0.014 yr^{-1} and 0.012 yr^{-1} for tropical and non-tropical regions, respectively. $\lambda_{hurricane}$ is specified either by an internal parameter or via external data. $\lambda_{fire}(a,t)$ is either calculated in fire submodule or specified by external data.

Disturbance reduces the area of all patches proportionally and then form a new patch, the boundary conditions of area, carbon, nitrogen and water pool for this new patch are represented as:

$$p_i(0,t) = \int_0^\infty \lambda_i(a,t) p_i(a,t)\, da \tag{B.10.2}$$

$$PL_i(0,t) = \int_0^\infty PL_i(a,t) \frac{\lambda_i(a,t) p_i(a,t)}{p_i(0,t)}\, da \tag{B.10.3}$$

Where PL represents each pool of soil carbon, nitrogen and water. As the above two equation shows, the new patch proportionally inherits pools from the source patches.

In addition to area and pool changes, disturbance also removes a fraction of the plants within involved patches. Some plants from the reduced patch area survive the disturbance and are relocated to the new patch, rest of plants die and their carbon and nitrogen are loaded into the soil pools. Individual density of surviving plants is represented as:

$$n_i(\mathbf{z}, \mathbf{x}, 0, t) = \int_0^{\infty} S(\mathbf{x}) n_i(\mathbf{z}, \mathbf{x}, a, t) \lambda_i(a, t)\, da \tag{B.10.4}$$

Where $S(\mathbf{x})$ is survivorship dependent on the disturbance and PFT type. For non-fire related disturbance (i.e., treefall or hurricane), survivorship is differentiated by tree height. Thereby $S(\mathbf{x})$ is given by:

$$S(\mathbf{x}) = \begin{cases} s_{lt}(\mathbf{x}), & h(\mathbf{z}, \mathbf{x}, a, t) < h_{treefall} \\ s_{gt}(\mathbf{x}), & h(\mathbf{z}, \mathbf{x}, a, t) \geq h_{treefall} \end{cases} \tag{B.10.5}$$

Where $h(\mathbf{z}, \mathbf{x}, a, t)$ is the height of a cohort, $h_{treefall}$ is a defined height threshold, and $s_{lt}(\mathbf{x})$ and $s_{gt}(\mathbf{x})$ are the survivorship rate (scaled from 0 to 1) for a plant with a height above $h_{treefall}$ or below it, respectively. Currently, $s_{lt}(\mathbf{x})$ and $s_{gt}(\mathbf{x})$ are the same for all PFTs, (i.e., values are 1.0 and 0.0, respectively), and $h_{treefall}$ is set as 0, meaning all plants will not survive in treefall disturbance.

For fire-related disturbance, survivorship is different for grasses where:

$$S(\mathbf{x}) = \begin{cases} 1.0, & \mathbf{x} = \mathrm{C3ShG\ or\ C4ShG} \\ 0.3, & otherwise. \end{cases} \tag{B.10.6}$$

Total carbon of dead plants involved in disturbance is given by:

$$C_{rem,dis}(t) = \int_0^{\infty} \int_0^{\infty} [B_a(\mathbf{z}, \mathbf{x}, a, t) + B_s(\mathbf{z}, \mathbf{x}, a, t)][1 - S(\mathbf{x})] n_i(\mathbf{z}, \mathbf{x}, a, t) \lambda_i(a, t) d\mathbf{z}\, da \tag{B.10.7}$$

The total carbon of dead plants is partitioned between soil carbon pools and emissions:

$$C_{rem,dis}(t) = \left[1 - C_{rem,dis}(t)\right]C_{rem,dis}(t) + f_{loss}C_{rem,dis}(t) \qquad \text{(B.10.8)}$$

Where the two terms on the right-hand side of the equation represent the carbon partitioned to soil carbon pools and to emissions, respectively. f_{loss} is the fraction of carbon lost as CO_2 emissions. f_{loss} is set as 0.3 for fire-related disturbance (i.e., smoke fraction), 0 for treefall- and hurricane-related disturbance, which means no carbon will lost as emissions.

Fire disturbance rate $\lambda_{fire}(a,t)$ can be either or specified by external burned area data or estimated by fire submodule (described below). Following Hurtt et al., 2001, fire risk is controlled by fuel and ignition rate, thereby $\lambda_{fire}(a,t)$ is given by:

$$\lambda_{fire}(a,t) \ = B_{fuel}(a,t)f_{ignition}(a,t) \qquad \text{(B.10.9)}$$

$$B_{fuel}(a,t) \ = \int_0^\infty \left[B_l(\mathbf{z},\mathbf{x},a,t) + f_{agb}B_{sw}(\mathbf{z},\mathbf{x},a,t) + f_{agb}B_s(\mathbf{z},\mathbf{x},a,t)\right]n_i(\mathbf{z},\mathbf{x},a,t)d\mathbf{z} \qquad \text{(B.10.10)}$$

$$\lambda_{fire}(a,t) \ = \begin{cases} \left(\dfrac{\overline{D}}{30000}\right)^{10}, & P(a,t) < 100\ mm\ month^{-1} \\ 0.0, & otherwise. \end{cases} \qquad \text{(B.10.11)}$$

Where $B_{fuel}(a,t)$ is total aboveground carbon as fuel, f_{agb} is aboveground ratio of structural biomass B_s, which is set as 0.8. $\overline{D}$ is annual average drought index, calculated from rolling monthly estimates of the number of days precipitation is below potential evapotranspiration rate.

B.11. Land use submodule

The land use submodule describes the demographic dynamics of patches and cohorts by tracking the sub-grid heterogeneity associated with different land use types and transitions. A wide range of land use activities are accounted for including deforestation, reforestation, shifting cultivation, and wood harvest. In this submodule, land use activities can alter the demography of patches and cohorts. For example, deforestation for cropland results in area decrease of forest patches and area increase of new cropland patch, and correspondingly resets the age of patches and cohorts. In addition, land use activities alter carbon dynamics, including redistribution of carbon between plant, soil and wood timber product pools, and legacy effects on the carbon balance such as elevated heterotrophic respiration from dead plants and enhanced carbon sequestration from plant regrowth. Currently, the submodule is structured for use of standard land use forcing for CMIP5 and CMIP6 (i.e., the Land Use Harmonization 1 (LUH1) and 2 (LUH2) datasets). These datasets provide historical gridded land use fractions and transitions between land use types on an annual basis.

Four land use types are characterized: primary land, secondary land, cropland, and pasture. Patches are tagged with a particular land use type (i.e., primary (v), secondary (s), cropland (c), and pasture (p)), and labelled with the corresponding subscript of $p_i(a,t)$ in the Equation 3.1 (the core PDE equation for patch dynamic). Transition types among the four land use types are listed in Table B.3, along with their corresponding input variables in LUH1 and LUH2. In this table, $\lambda_{\mathrm{v,c}}$, $\lambda_{\mathrm{v,p}}$, $\lambda_{\mathrm{s,c}}$ and $\lambda_{\mathrm{s,p}}$ represent deforestation, $\lambda_{\mathrm{v,s}}$ and $\lambda_{\mathrm{s,s}}$ represent wood harvest, $\lambda_{\mathrm{c,s}}$ and $\lambda_{\mathrm{p,s}}$ represent reforestation. For each grid cell, patch area is subject to:

$$\int_0^\infty p_i(a,t)\,da = \mathrm{LU}_i(t) \qquad (i = \mathrm{v}, \mathrm{s}, \mathrm{c}\ and\ \mathrm{p}) \tag{B.11.1}$$

Where $\mathrm{LU}_i(t)$ is the area of the land use type i at time t, specified by the external land use change dataset (e.g., LUH1 or LUH2).

Land use transitions drive patch demographic changes by reducing the area and land-use proportion of existing patches, which is described as:

$$\frac{\partial}{\partial t} p_i(a,t) = -\frac{\partial}{\partial a} p_i(a,t) - \lambda_i(a,t) p_i(a,t) - \sum_j \lambda_{j,i}(a,t) p_i(a,t) \quad \text{(B.11.2)}$$

The above equation has been described in section 2.1, governing patch dynamics in terms of ageing and disturbance due to both natural and anthropogenic land use change. The last term on the right-hand size of the equation represents the patch fraction $p_i(a,t)$ that decreases due to a land use transition from current type *i* to new type *j*. Along with this fractional decrease for all involved patches, a new patch with land use type *j* will be formed. The area, carbon, nitrogen and water boundary conditions for this new patch are represented as:

$$p_j(0,t) = \sum_i \int_0^\infty \lambda_{j,i}(a,t) p_i(a,t)\, da \qquad (i,j = \mathrm{v,s,c}\ and\ \mathrm{p}) \quad \text{(B.11.3)}$$

$$PL_j(0,t) = \sum_i PL_i(0,t) \frac{\int_0^\infty \lambda_{j,i}(a,t) p_i(a,t)\, da}{p_j(0,t)} \qquad (i,j = \mathrm{v,s,c}\ and\ \mathrm{p}) \quad \text{(B.11.4)}$$

Where PL represents each pool of soil carbon, nitrogen, and water. The above two equations shows that the new patch inherits pools from the source patches proportionally.

Depending on the specific transition type, land use transitions may also involve plant removal (Table B.3). Plant removal will clear native plants and distribute associated carbon to either wood product pools, or soil litter pools. The carbon from plant removal is partitioned between carbon pools as follows:

$$C_{res,i}(t) = \int_0^\infty \int_0^\infty [B_a(\mathbf{z}, \mathbf{x}, a, t) + B_s(\mathbf{z}, \mathbf{x}, a, t)] n_i(\mathbf{z}, \mathbf{x}, a, t) p_i(a, t) \lambda_{i,j}(a, t) \zeta_{\mathrm{res}}(\mathbf{x}, i) \, d\mathbf{z} \, da \quad (i, j = \mathrm{v, s, c} \; and \; \mathrm{p}) \tag{B.11.5}$$

$$\Delta C_{wood,1yr}(t) = \int_0^\infty \int_0^\infty [B_a(\mathbf{z}, \mathbf{x}, a, t) + B_s(\mathbf{z}, \mathbf{x}, a, t)] n_i(\mathbf{z}, \mathbf{x}, a, t) p_i(a, t) \lambda_{i,j}(a, t) [1 - \zeta_{\mathrm{res}}(\mathbf{x}, i, j)] \eta_{1yr}(\mathbf{x}, i, j) \, d\mathbf{z} \, da \quad (i, j = \mathrm{v, s, c} \; and \; \mathrm{p}) \tag{B.11.6}$$

$$\Delta C_{wood,10yr}(t) = \int_0^\infty \int_0^\infty [B_a(\mathbf{z}, \mathbf{x}, a, t) + B_s(\mathbf{z}, \mathbf{x}, a, t)] n_i(\mathbf{z}, \mathbf{x}, a, t) p_i(a, t) \lambda_{i,j}(a, t) [1 - \zeta_{\mathrm{res}}(\mathbf{x}, i, j)] \eta_{10yr}(\mathbf{x}, i, j) \, d\mathbf{z} \, da \quad (i, j = \mathrm{v, s, c} \; and \; \mathrm{p}) \tag{B.11.7}$$

$$\Delta C_{wood,100yr}(t) = \int_0^\infty \int_0^\infty [B_a(\mathbf{z},\mathbf{x},a,t) + B_s(\mathbf{z},\mathbf{x},a,t)] n_i(\mathbf{z},\mathbf{x},a,t) p_i(a,t) \lambda_{i,j}(a,t) [1 - \zeta_{\mathrm{res}}(\mathbf{x},i,j)] \eta_{100yr}(\mathbf{x},i,j)\, d\mathbf{z}\, da \quad (i,j = \mathrm{v,s,c}\ and\ \mathrm{p}) \tag{B.11.8}$$

Where $C_{res,i}(t)$ is removed carbon that is allocated to soil litter pools. $\Delta C_{wood,1yr}(t)$, $\Delta C_{wood,10yr}(t)$ and $\Delta C_{wood,100yr}(t)$ are removed carbon that is allocated to wood product pools with decay rates of 1-year, 10-year and 100-year, respectively. The coefficient $\zeta_{\mathrm{res}}(\mathbf{x},i,j)$ is the carbon fraction left on-site; $\eta_{1yr}(\mathbf{x},i,j)$, $\eta_{10yr}(\mathbf{x},i,j)$ and $\eta_{100yr}(\mathbf{x},i,j)$ are the relative fractions entering each of the three wood product pools; The four coefficients are differentiated among PFTs and between primary or secondary land (Table B.4), the parameterization is based on Hansis et al., 2015.

In addition to patch dynamics arising from land use transitions, cropland patches are routinely harvested and planted on an annual basis, with planting and harvesting dates specified by an external crop calendar (Sacks et al., 2010). Crop harvesting only leaves a limited number of plants in each patch to ensure reproduction in the following years, removing all other plants. For pasture patches, grazing is routinely implemented to similarly remove a fraction of plants from each pasture patch. The removed carbon from harvesting and grazing are given by:

$$C_{rem,\mathrm{c}}(t) = \int_0^\infty \int_0^\infty [B_a(\mathbf{z},\mathbf{x},a,t) + B_s(\mathbf{z},\mathbf{x},a,t)]\left[n_\mathrm{c}(\mathbf{z},\mathbf{x},a,t) - n_{\mathrm{c,min}}\right] d\mathbf{z}\, da \quad \text{(B.11.9)}$$

$$C_{rem,\mathrm{p}}(t) = \int_0^\infty \int_0^\infty [B_a(\mathbf{z},\mathbf{x},a,t) + B_s(\mathbf{z},\mathbf{x},a,t)] n_\mathrm{p}(\mathbf{z},\mathbf{x},a,t) \lambda_{graz\_inten}\, d\mathbf{z}\, da \quad \text{(B.11.10)}$$

Where $n_{\mathrm{c,min}}$ is the minimum density of crop plants that are retained post-harvest, λ_{graz_inten} is the grazing intensity which specifies the fraction of plants to be removed due to grazing.

The removed carbon is distributed to the product pools and soil carbon pool, the partitioning is given by:

$$C_{rem,\mathrm{c}}(t) = \zeta_{\mathrm{res,c}} C_{rem,\mathrm{c}}(t) + \left(1 - \zeta_{\mathrm{res,c}}\right) C_{rem,\mathrm{c}}(t) \quad \text{(B.11.11)}$$

$$C_{rem,\mathrm{p}}(t) = \zeta_{\mathrm{res,p}} C_{rem,\mathrm{p}}(t) + \left(1 - \zeta_{\mathrm{res,p}}\right) C_{rem,\mathrm{p}}(t) \quad \text{(B.11.12)}$$

In above two equations, the first term on the right-hand side of the equation represents on-site plant residuals on cropland or pasture, respectively, these residuals will be loaded into soil litter pools. The second term represents the removed carbon allocated to the product pools of harvested crop and grazed grass. $\zeta_{\mathrm{res,c}}$ and $\zeta_{\mathrm{res,p}}$ are the on-site fraction coefficients, set at 0.5 for cropland and 0.1 for pasture.

Table B.3. Land use transition types and their corresponding input variables from LUH1 and LUH2. Note crops includes C3 annual crops (c3ann), C4 annual crops (c4ann), C3 perennial crops (c3per), C4 perennial crops (c4per), and C3 nitrogen-fixing crops (c3nfx). All transitions represent clearing type except primary land harvesting ($\lambda_{v,s}$) and secondary land harvesting ($\lambda_{s,s}$). Clearing and harvesting types have different parameterization for plant removal (see Table B.4).

Land use transition	LUH1	LUH2	Plant removal
$\lambda_{v,s}$	*gflvh, gflvh2*	*primf_harv, primn_harv*	Y
$\lambda_{v,c}$	*gflvc*	*primf_to_crops, primn_to_crops*	Y
$\lambda_{v,p}$	*gflvp*	*primf_to_pastr, primn_to_pastr primf_to_range, primn_to_range*	Y
$\lambda_{s,s}$	*gfsh1, gfsh2, gfsh3*	*secyf_harv, secmf_harv, secnf_harv*	Y
$\lambda_{s,c}$	*gfsc*	*secdf_to_crops, secdn_to_crops*	Y
$\lambda_{s,p}$	*gflsp*	*secdf_to_pastr, secdn_to_pastr secdn_to_range, secdn_to_range*	Y
$\lambda_{c,s}$	*gflcs*	*crops_to_secdf, crops_to_secdn*	N
$\lambda_{c,p}$	*gflcp*	*crops_to_pastr, crops_to_range*	N
$\lambda_{p,s}$	*gflps*	*pastr_to_secdf, pastr_to_secdn range_to_secdf, range_to_secdn*	N
$\lambda_{p,c}$	*gflpc*	*pastr_to_crops, range_to_crops*	Y

As Equation B.11.6, B.11.7, B.11.8, B.11.11, and B.11.12 shows, carbon that is partially removed during land use transitions will be allocated to the respective product (e.g., wood, crop, or grass). These pools decay with different rates, for example, crop and grass pools are assumed to decay immediately, and are lost to the atmosphere as land use emissions. However, wood product pools decay slowly over time with a rate following an exponential curve:

$$\frac{dC_{wood,nyr}(a,t)}{dt} = \Delta C_{wood,nyr}(t) + C_{wood,nyr}(a,t)e^{-\tau_{nyr}dt} \qquad \text{(B.11.13)}$$

Where $C_{wood,nyr}$ is the nyr product pool (nyr=1yr, 10yr, or 100yr), $\Delta C_{wood,nyr}(t)$ is newly loaded carbon due to land use transitions, τ_{nyr} is the coefficient governing the decay rate. This rate is currently set at -1.873, 0.187 and 0.018 for the three wood pools ($C_{wood,1yr}$,

$C_{wood,10yr}$ and $C_{wood,100yr}$) respectively, such that three pools reduce to 15% of their respective size within 1 year, 10 years, or 100 years. Decayed carbon from all of three wood product pools contribute to land use emissions.

Table B.4. Parameters for land use transitions involved in plant removals (i.e., Equation B.11.5-8).

Parameters	C4ShG	C3ShG	EaSBT, MiSBT, LaSBT		NSP, LaSC
			TRO	NTRO	
Harvesting on primary land					
$\eta_{1yr}(\mathbf{x}, \mathrm{v}, \mathrm{s})$	1.0	1.0	0.90	0.40	0.40
$\eta_{10yr}(\mathbf{x}, \mathrm{v}, \mathrm{s})$	0.0	0.0	0.04	0.24	0.24
$\eta_{100yr}(\mathbf{x}, \mathrm{v}, \mathrm{s})$	0.0	0.0	0.06	0.36	0.36
$\zeta_{\mathrm{res}}(\mathbf{x}, \mathrm{v}, \mathrm{s})$	0.860	0.780	0.825	0.795	0.870
Harvesting on secondary land					
$\eta_{1yr}(\mathbf{x}, \mathrm{s}, \mathrm{s})$	1.0	1.0	0.90	0.40	0.40
$\eta_{10yr}(\mathbf{x}, \mathrm{s}, \mathrm{s})$	0.0	0.0	0.04	0.24	0.24
$\eta_{100yr}(\mathbf{x}, \mathrm{s}, \mathrm{s})$	0.0	0.0	0.06	0.36	0.36
$\zeta_{\mathrm{res}}(\mathbf{x}, \mathrm{s}, \mathrm{s})$	0.810	0.700	0.750	0.725	0.820
Clearing					
$\eta_{1yr}(\mathbf{x}, i, j)$	1.0	1.0	0.59	0.59	0.59
$\eta_{10yr}(\mathbf{x}, i, j)$	0.0	0.0	0.41	0.31	0.31
$\eta_{100yr}(\mathbf{x}, i, j)$	0.0	0.0	0.00	0.10	0.10
$\zeta_{\mathrm{res}}(\mathbf{x}, i, j)$	0.50	0.50	0.33	0.33	0.33

Appendix C Supplementary material for Chapter 4

C.1. ED initialization and projection methods

This study generally follows the initialization and projection approach used in Hurtt *et al* 2019, but proposes a modification to the initialization method (here defined as weighting-based initialization method) to improve AGB estimates where ED-modelled canopy height saturates. The workflow of initialization and projection process is illustrated in figure C.1(a) and comprises the following steps:

1. ED model run: ED model was run for 500 years with meteorological drivers, CO_2 and soil properties (as described in section 2.3.1) to create an AGB-Height lookup table. The lookup table stored a 500-year time series of ED-modelled AGB and corresponding canopy height for each land cell.

2. ED initialization: this step combined the lookup table from Step 1 with lidar canopy height and NAIP tree canopy cover to obtain contemporary AGB over the entire land cell (hereafter referred as ED initialized AGB). The NAIP tree canopy cover map was used to determine the relative fraction of land cell that is covered by trees (i.e., the forested fraction). The AGB of the forested fraction (hereafter referred as ED-indexed AGB) was obtained by indexing the AGB-Height lookup table with lidar canopy height; the AGB over the non-tree/non-forested fraction of the land cell is assumed to be zero. Figure C.1(b) illustrates the indexing process. The blue and yellow lines respectively represent the 500-year time series of ED-modelled canopy height and AGB for a given site. The AGB and height grow together until simulation year 300, at which point AGB continues growing but canopy height saturates. When indexing lidar canopy height for different land cells along the ED-modelled canopy height curve, two cases emerge. The initialization approach to each case differs as follows:

    1) In Case I, the recorded lidar canopy height of the land cell is below saturated ED canopy height (i.e., below the dotted line at 23m). Consequently, the resulting ED-indexed AGB is taken from the AGB time-series where ED-modelled canopy height most closely matches lidar canopy height.

2) In Case II, the recorded lidar canopy height exceeds ED canopy height (i.e., where ED height saturates) and thus cannot be found in the ED-modelled canopy height time-series. Differing from the mid-point method used in Hurtt et al., 2019, which uses the AGB at the middle simulation year between lower-bound and upper bound AGB, this study proposed a weight-based method to obtain the ED-indexed AGB by weighting lower-bound and upper-bound AGB. Lower-bound and upper-bound ED AGB are defined as the minimum and maximum AGB possible where ED-modelled canopy height has saturated. The weight of the lower-bound is proportional to the difference between maximum lidar canopy height and the lidar canopy height (i.e., $\frac{\Delta h_2 - \Delta h_1}{\Delta h_2}$), while the weight of upper-bound AGB is proportional to the difference between lidar canopy height and the saturated ED canopy height (i.e., $\frac{\Delta h_1}{\Delta h_2}$). Maximum lidar height is defined as the 95th percentile of lidar canopy height for a given spatial domain (e.g., RGGI region). The closer the recorded lidar canopy height is to the saturated ED canopy height, the more the lower-bound AGB contributes to the ED-indexed AGB.

3. ED projection: this step combined the lookup table from Step 1 and ED-indexed AGB from Step 2 to generate a 500-year AGB growth time-series towards maximum AGB. For each land cell, annual AGB estimates are the area-weighted sum of continued growth over the tree covered fraction and regrowth over the non-tree/non-forested fraction. Continued growth is obtained by sub-setting the lookup table between the indexed time in Step 2 and the 500th year; regrowth is obtained by using the entire AGB time-series of the lookup table, which starts with zero biomass in year 1. Tree/non-tree fractions were determined from NAIP tree canopy cover. Any

proportion of the land area currently covered by impervious surface, open water or herbaceous wetland, as identified via NLCD 2011, was excluded from potential regrowth. Several additional metrics are defined as the carbon sequestration potential (CSP), 95% of the maximum AGB over the final 50 years of the lookup table, and the carbon sequestration potential gap (CSPG), the difference between CSP and ED initialized AGB.

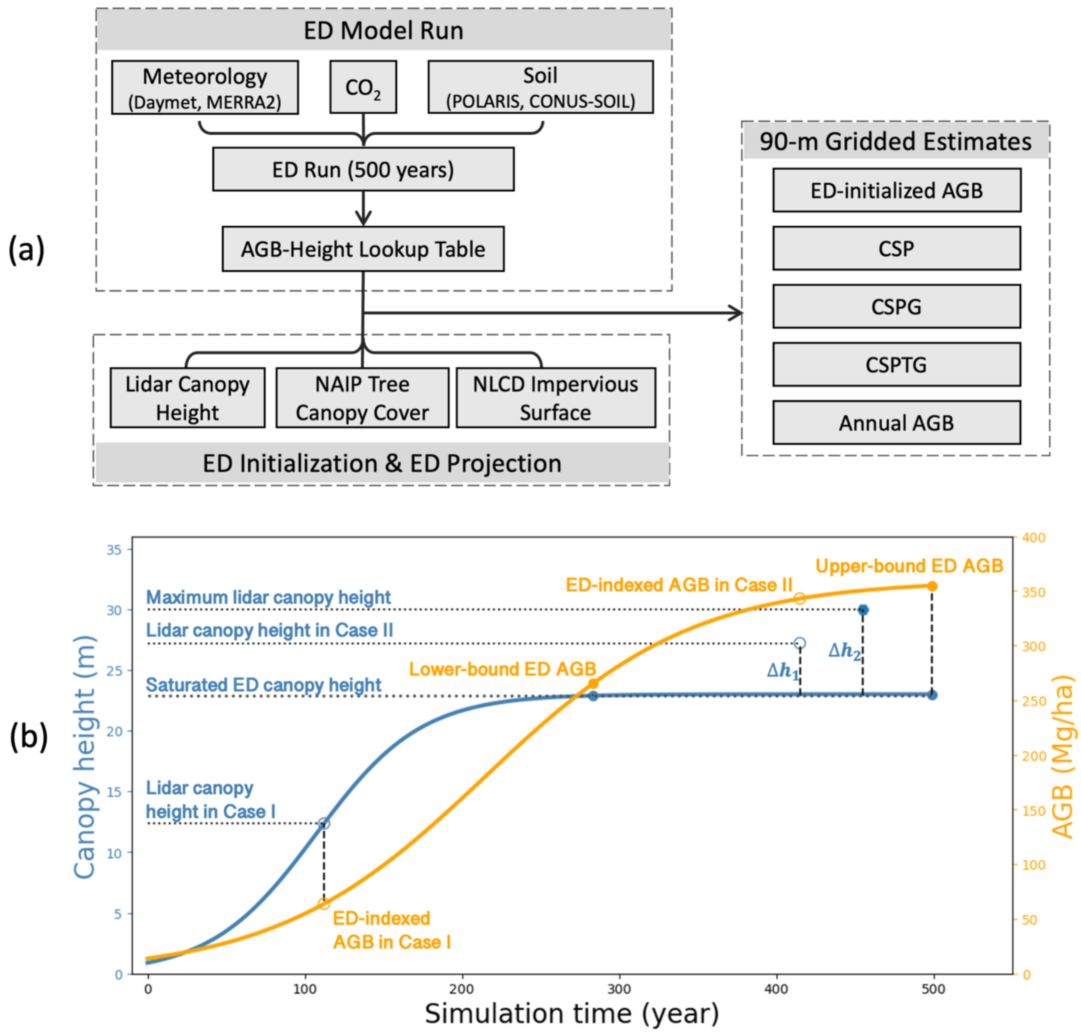


Figure C.1. Illustration of ED initialization and projection workflow (a) and the indexing of ED-modelled AGB-Height lookup table with lidar canopy height in (b).

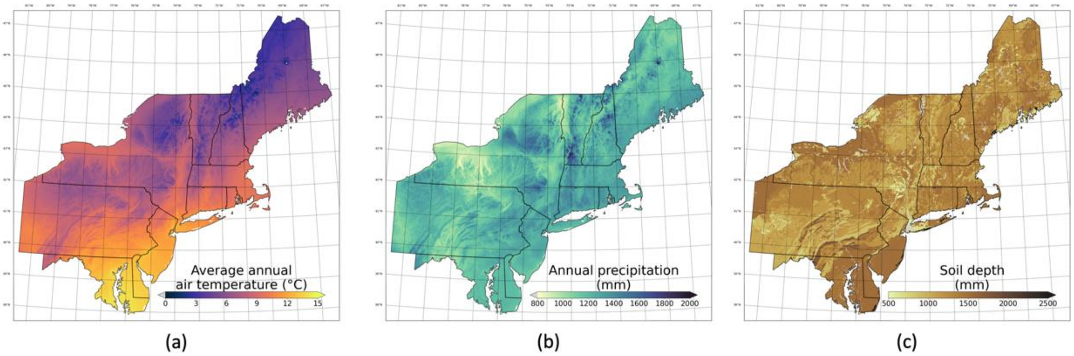


Figure C.2. Examples of ED input drivers of average annual air temperature (a) and annual precipitation (b) from Daymet and soil depth from CONUS-PSU (c).

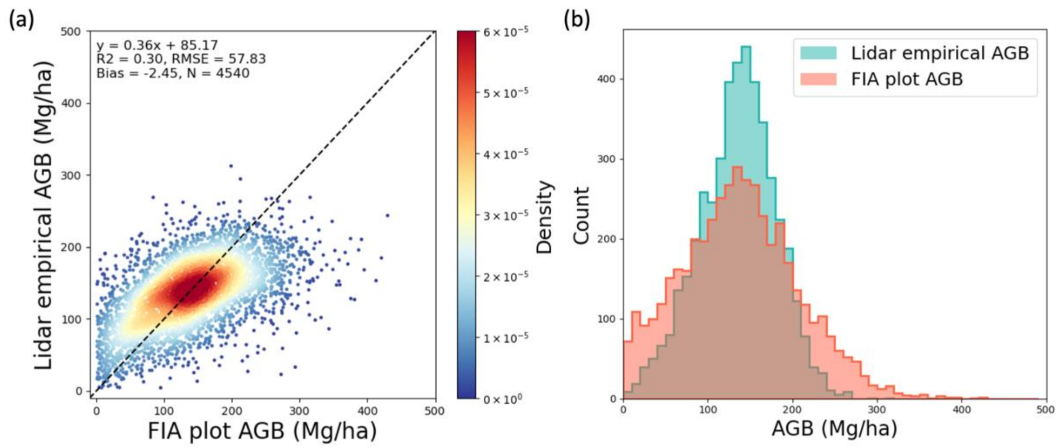


Figure C.3. Comparison of lidar empirical AGB to FIA plot AGB with a density scatter plot (a) and histogram (b).

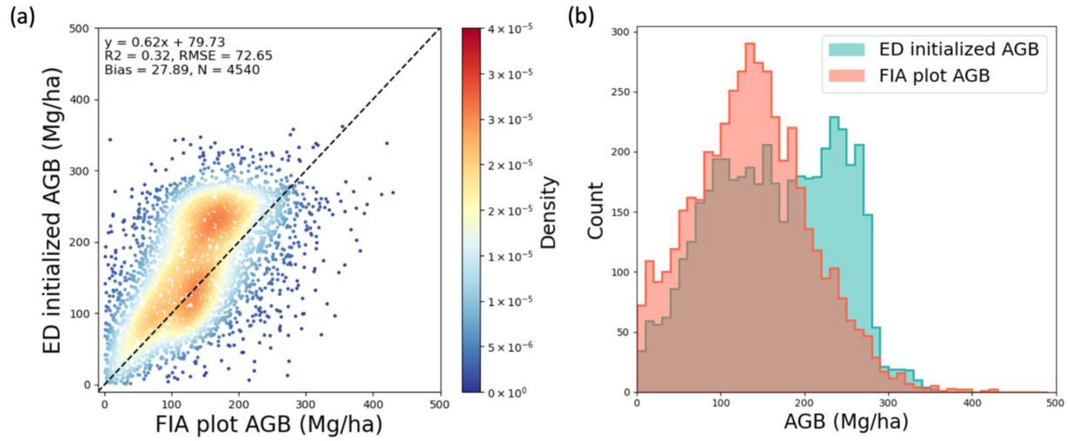


Figure C.4. Comparison of ED initialized AGB, using the mid-point initialization method, to FIA plots AGB with a density scatter plot (a) and histogram (b).

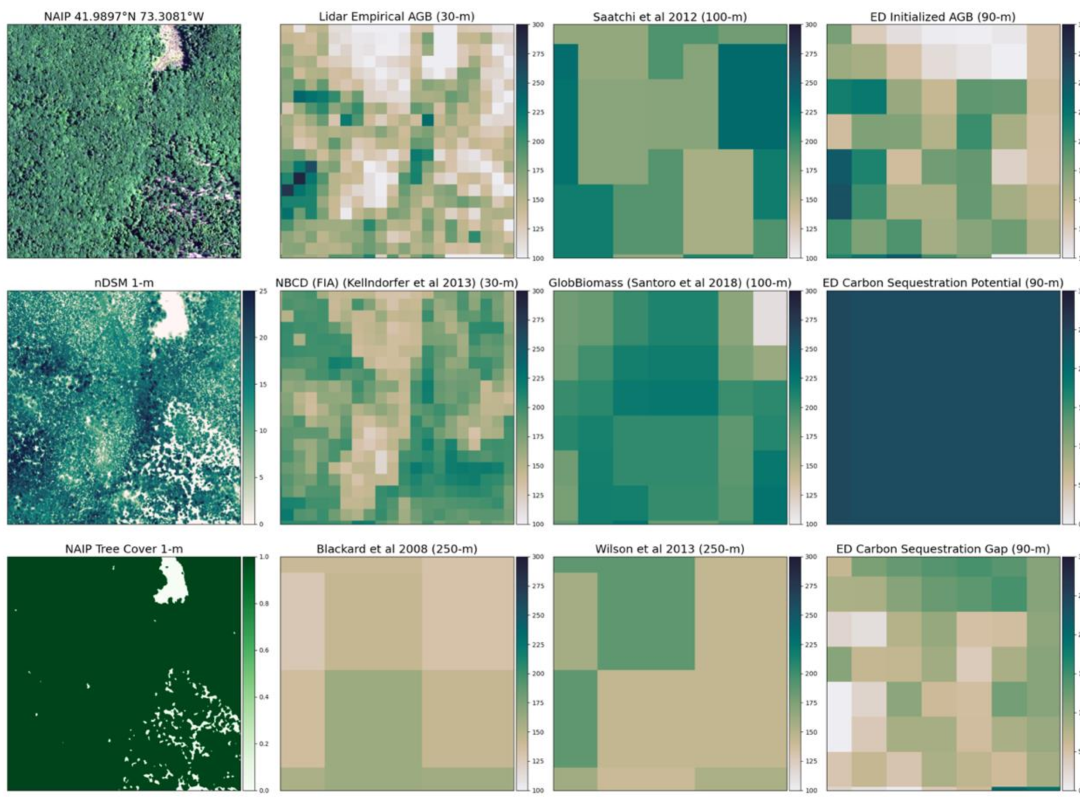


Figure C.5. Fine-scale maps of a forested area in Connecticut (41.9879 °N, 73.3081°W) using NAIP aerial imagery at 1-m, lidar canopy height at 1-m, NAIP tree cover classification at 1-m, lidar empirical AGB and NBCD at 30-m, AGB of Blackard et al 2008 at 250-m, AGB of Saatchi *et al* 2012 at 100-m, GlobBiomass at 90-m, AGB of Wilson *et al* 2013 and ED initialized AGB, carbon sequestration potential and the carbon sequestration potential gap at 90-m.

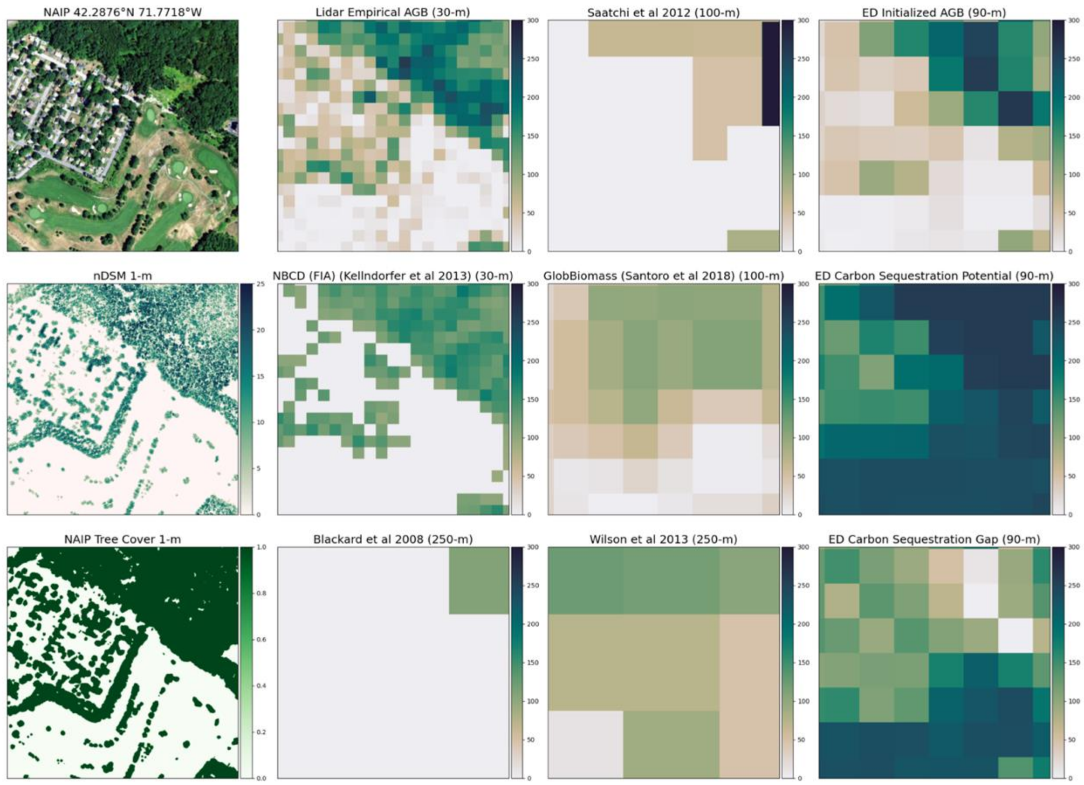


Figure C.6. As in figure C.5 but for a residential area located in the state of Massachusetts (41.2876°N, 71.7718°W).

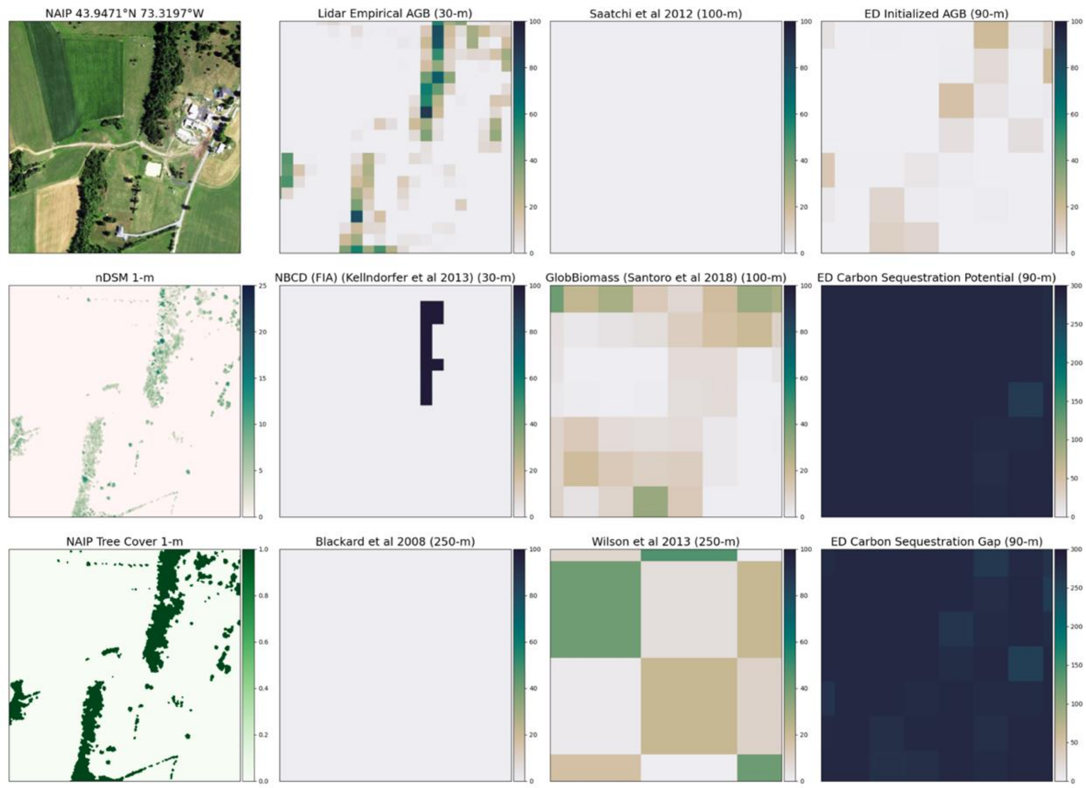


Figure C.7. As in figure C.5 but for an agricultural area located in the state of Vermont (43.9471°N, 73.3197°W).

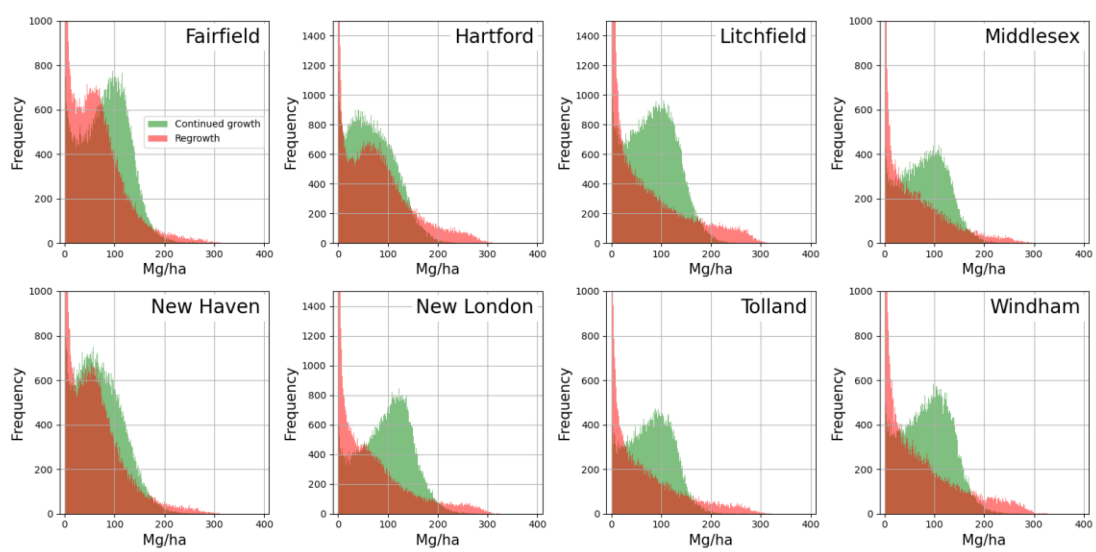


Figure C.8. CSPG over areas with continued growth (green) vs that over regrowth (red) for all counties and county-equivalents in Connecticut.

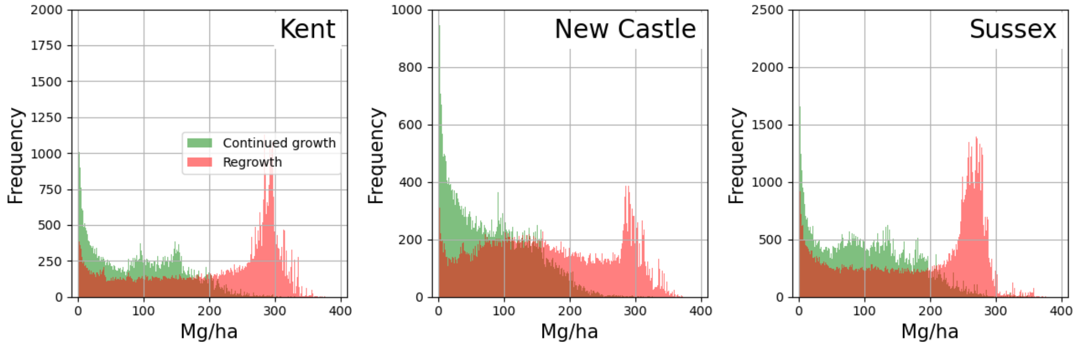


Figure C.9. As in Figure C.8 but for counties in Delaware.

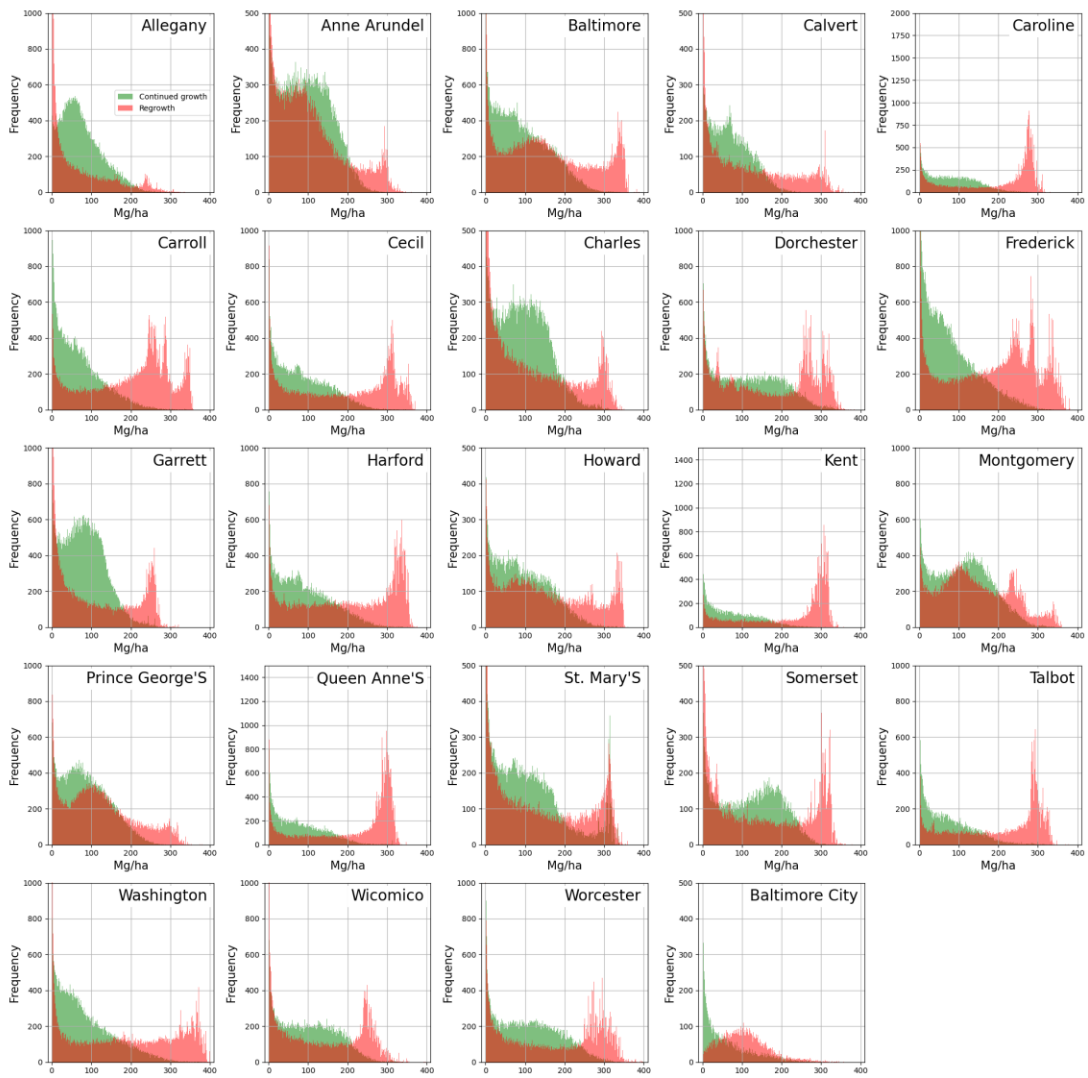


Figure C.10. As in Figure C.8 but for counties in Maryland.

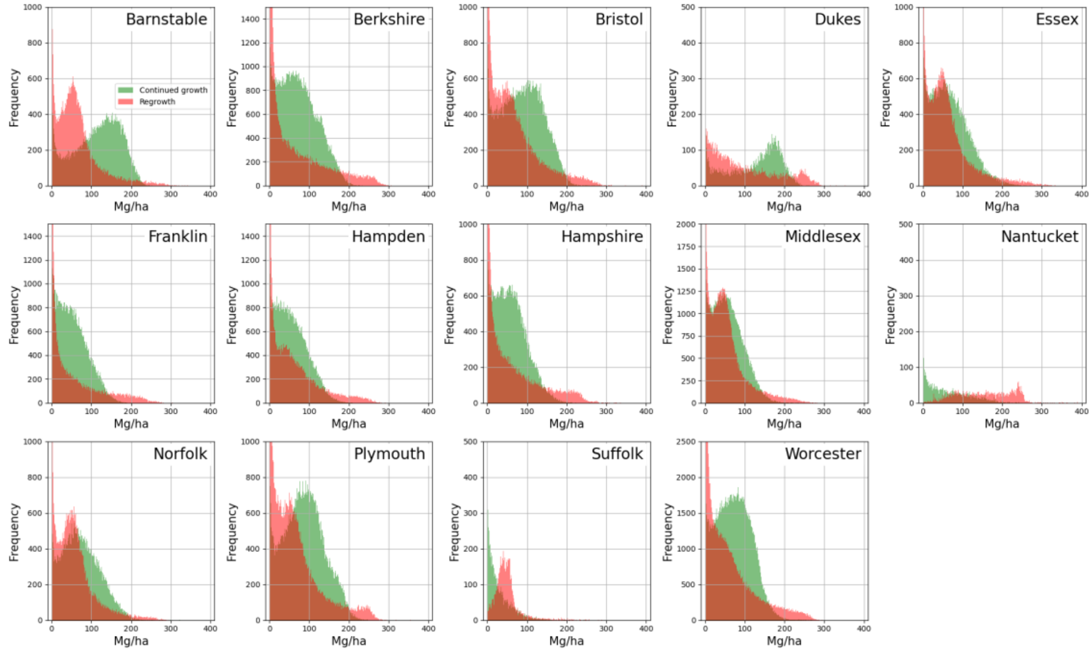


Figure C.11. As in Figure C.8 but for counties in Massachusetts.

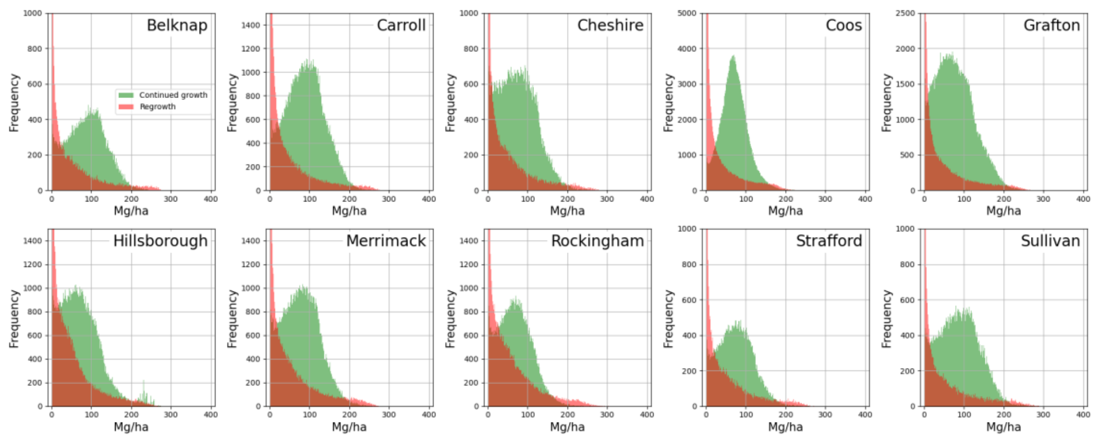


Figure C.12. As in Figure C.8 but for counties in New Hampshire.

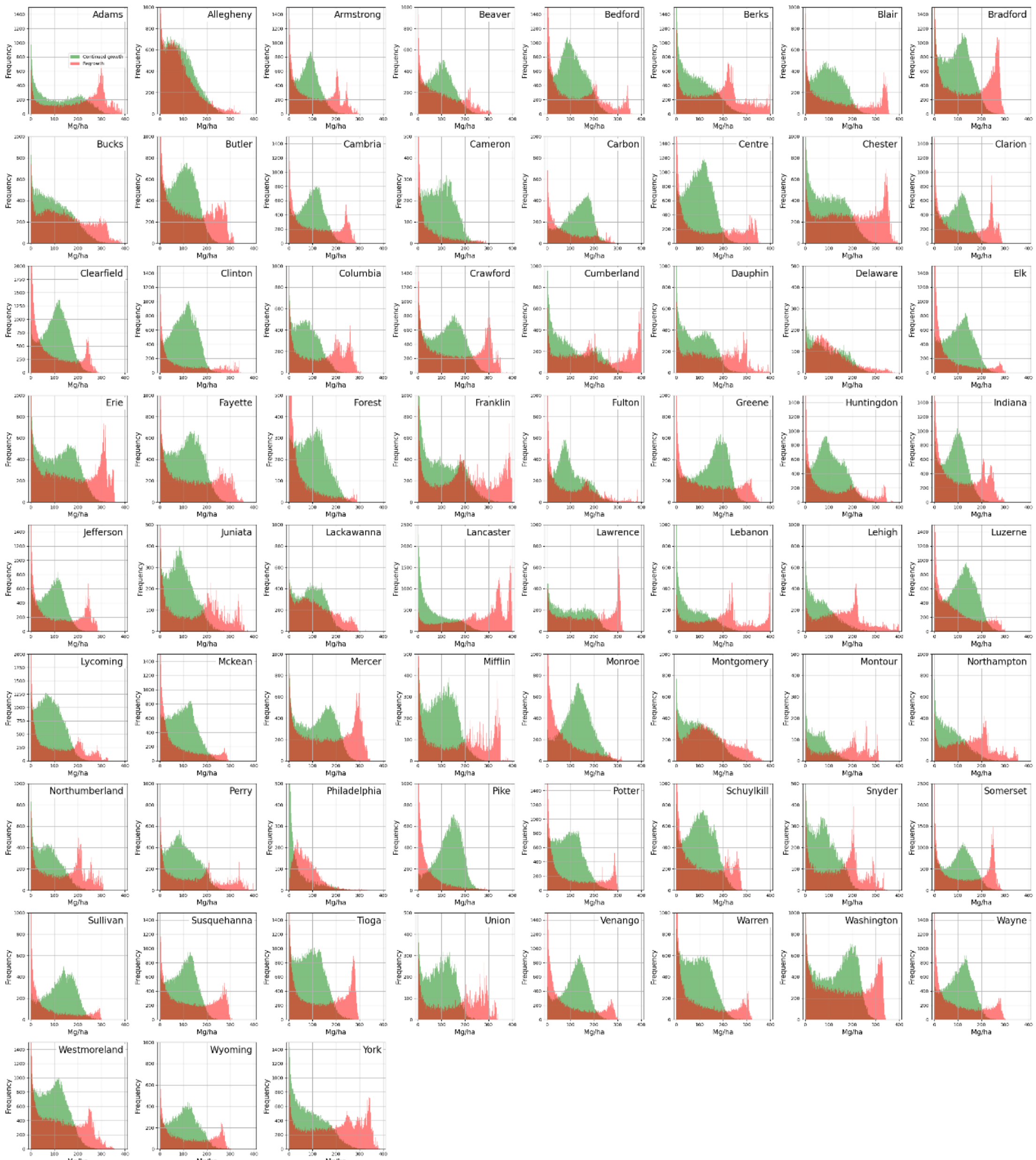


Figure C.13. As in Figure C.8 but for counties in Pennsylvania.

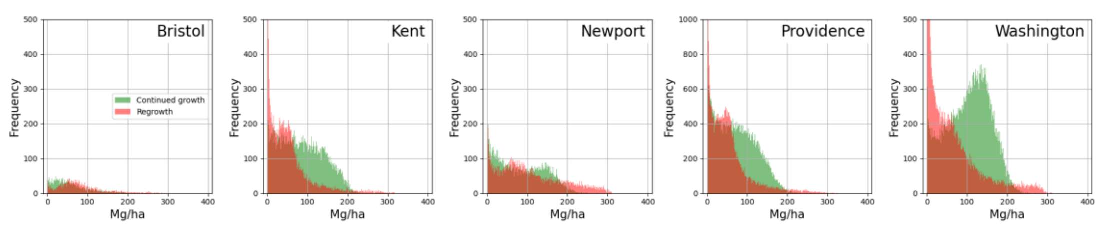


Figure C.14. As in Figure C.8 but for counties in Rhode Island.

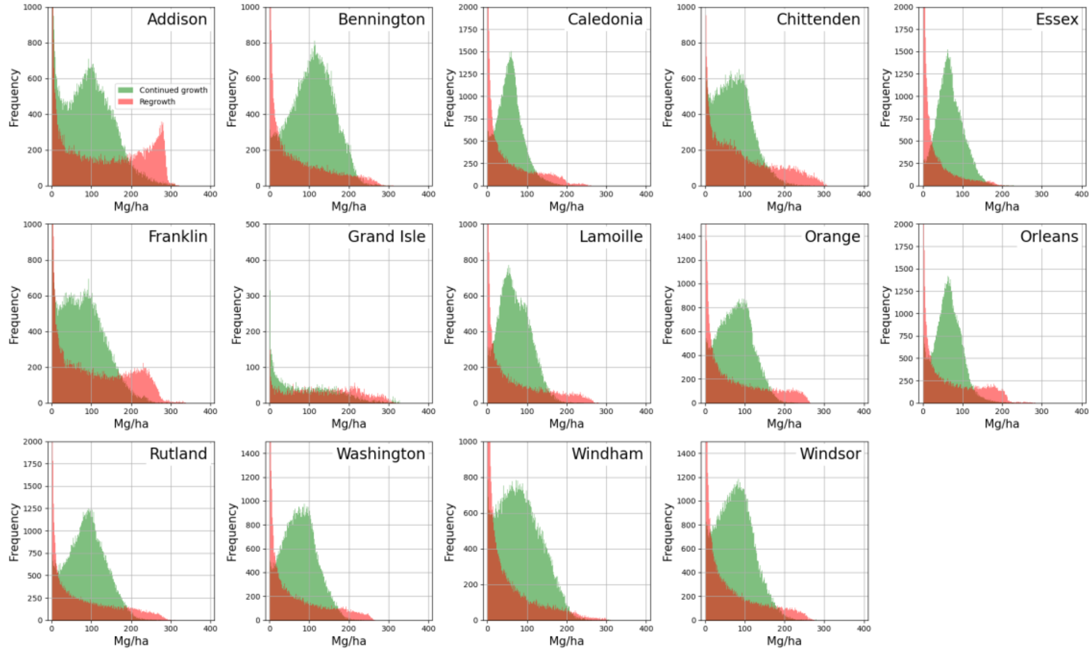


Figure C.15. As in Figure C.8 but for counties in Vermont.

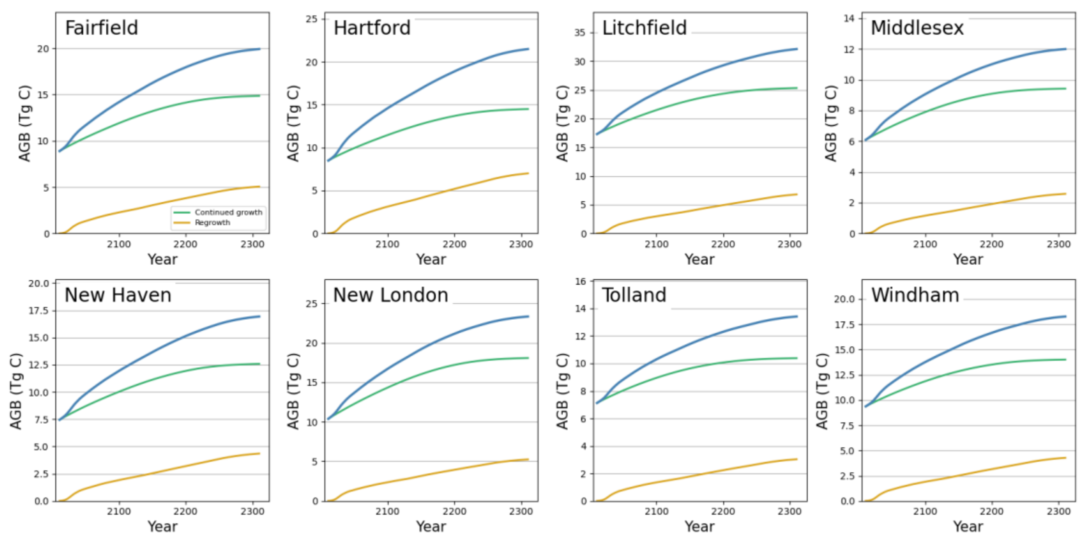


Figure C.16. Carbon sequestration time-series for all counties and county-equivalents in Connecticut. Contribution by contemporary tree and non-tree are colored in blue and orange respectively.

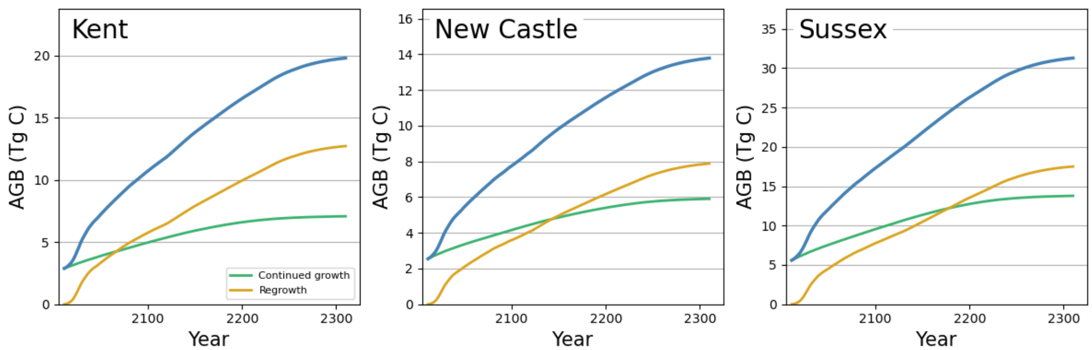


Figure C.17. As in Figure C.16 but for counties in Delaware.

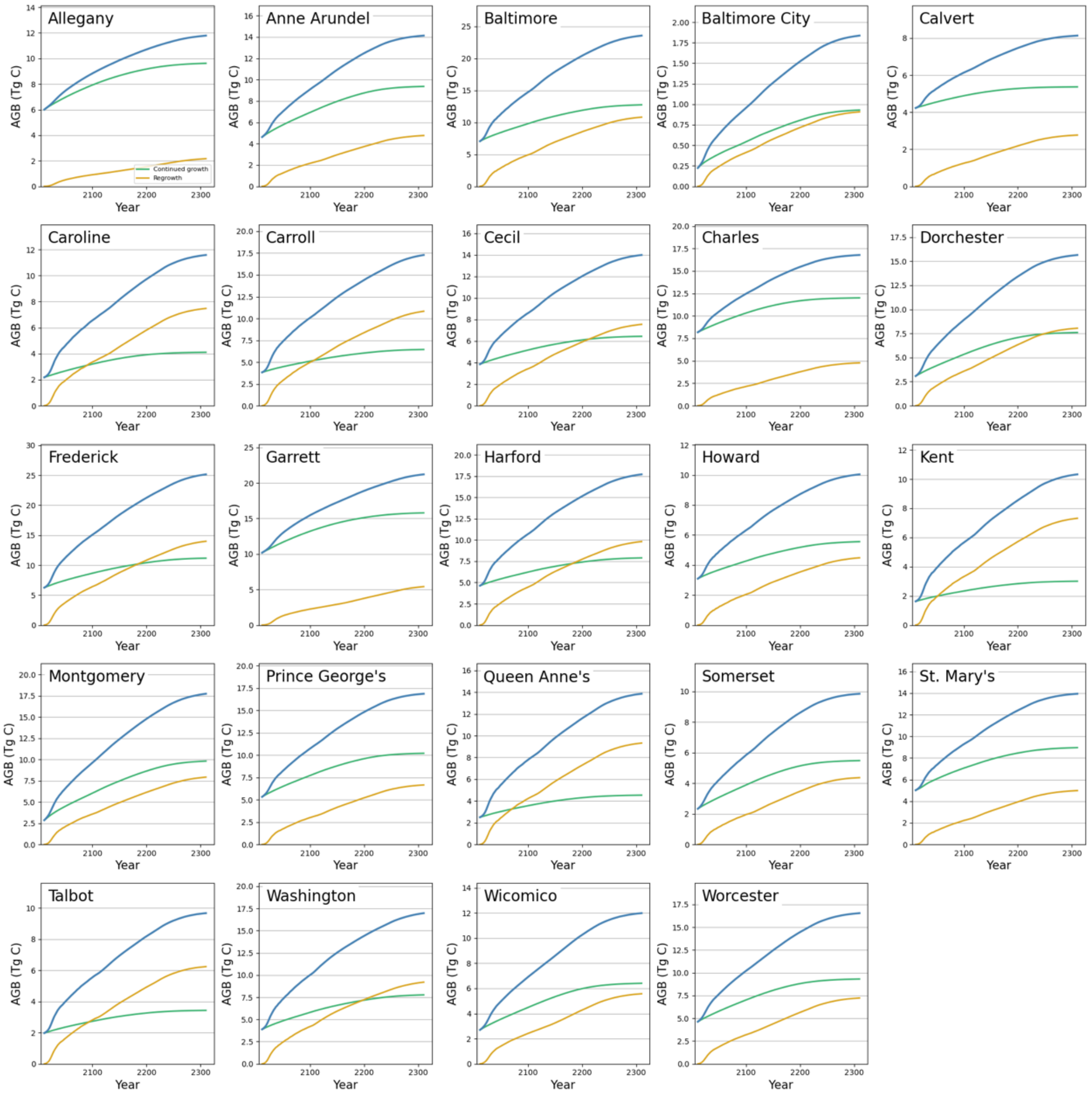


Figure C.18. As in Figure C.16 but for counties in Maryland.

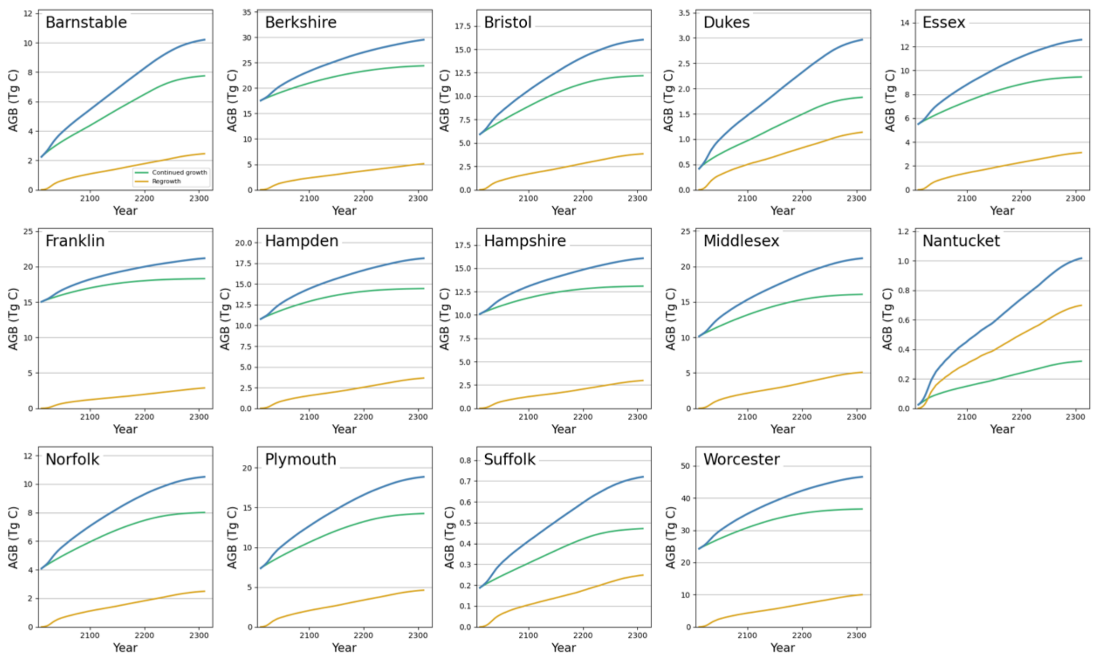


Figure C.19. As in Figure C.16 but for counties in Massachusetts.

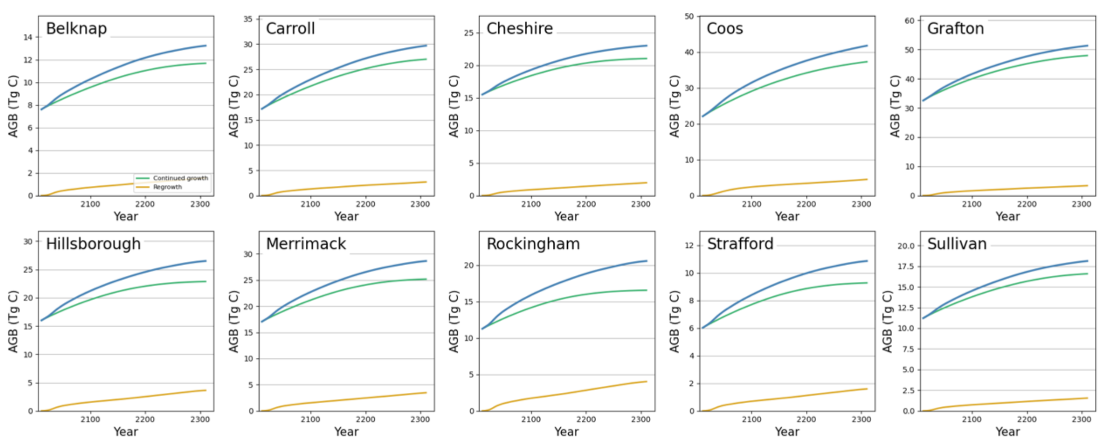
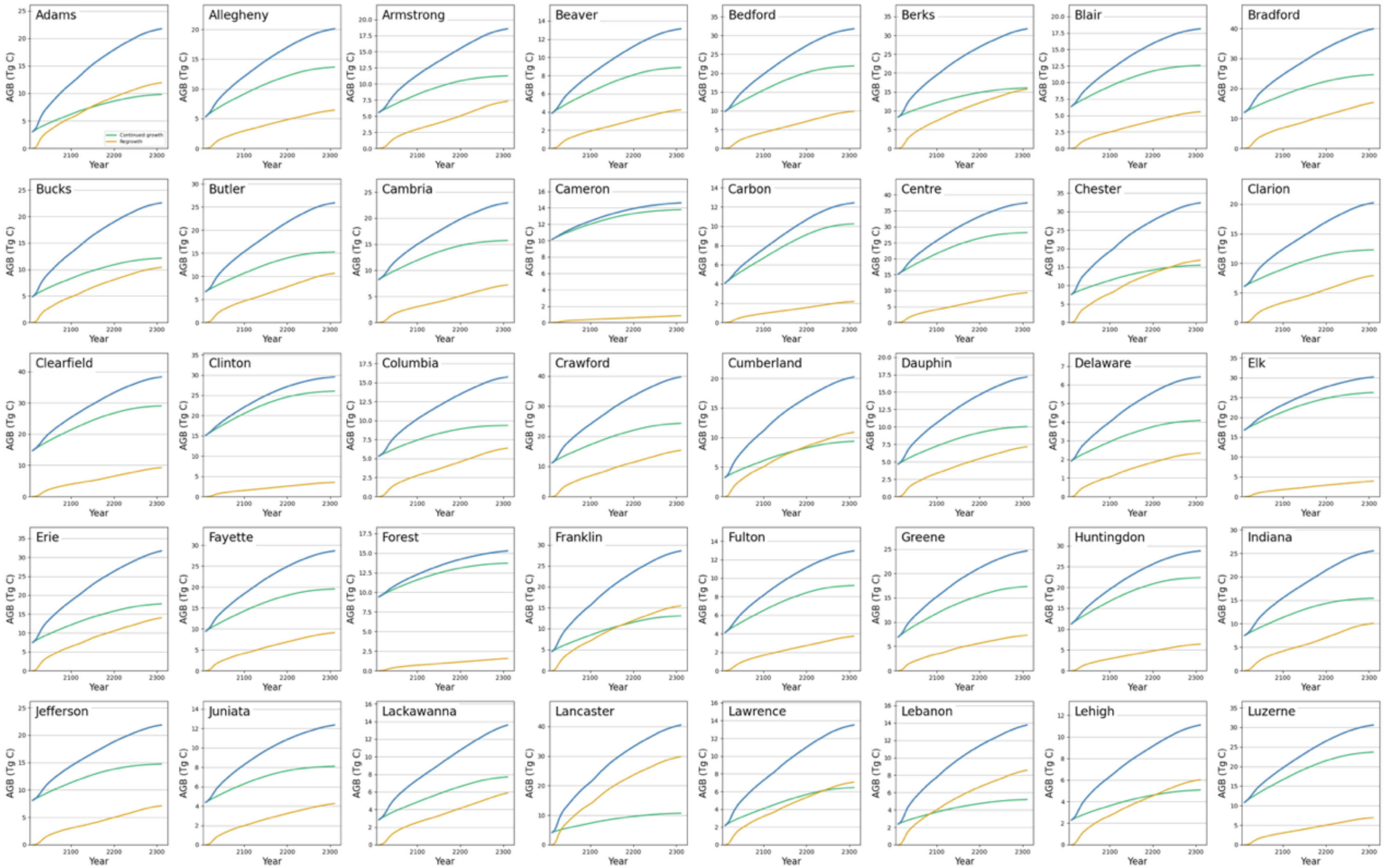


Figure C.20. As in Figure C.16 but for counties in New Hampshire.


Adams
Allegheny
Armstrong
Beaver
Bedford
Berks
Blair
Bradford
Bucks
Butler
Cambria
Cameron
Carbon
Centre
Chester
Clarion
Clearfield
Clinton
Columbia
Crawford
Cumberland
Dauphin
Delaware
Elk
Erie
Fayette
Forest
Franklin
Fulton
Greene
Huntingdon
Indiana
Jefferson
Juniata
Lackawanna
Lancaster
Lawrence
Lebanon
Lehigh
Luzerne
AGB (Tg C)
Year

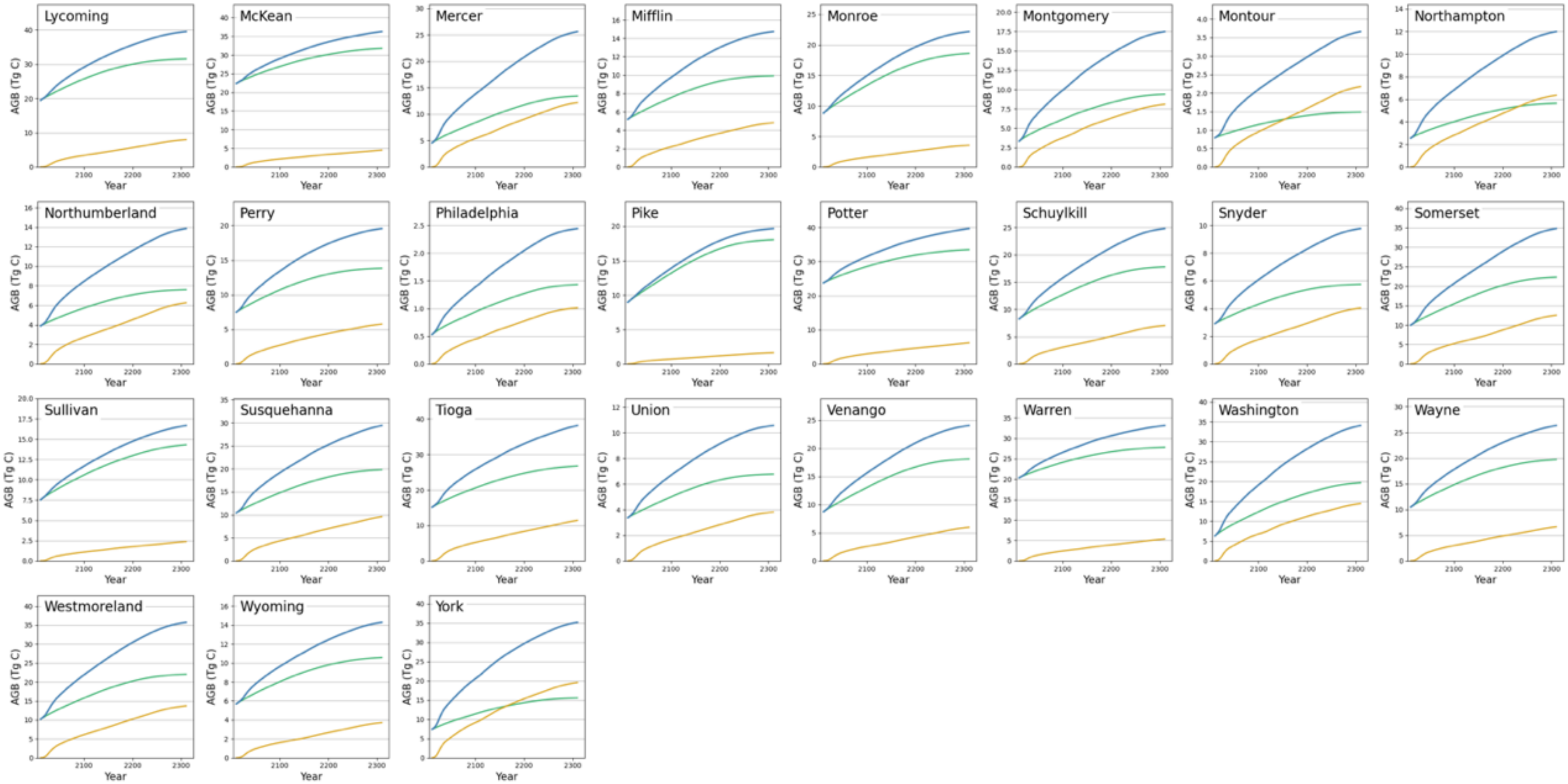


Figure C.21. As in Figure C.16 but for counties in Pennsylvania.

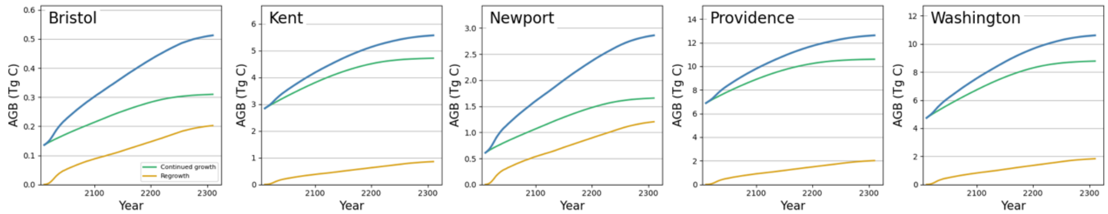


Figure C.22. As in Figure C.16 but for counties in Rhode Island.

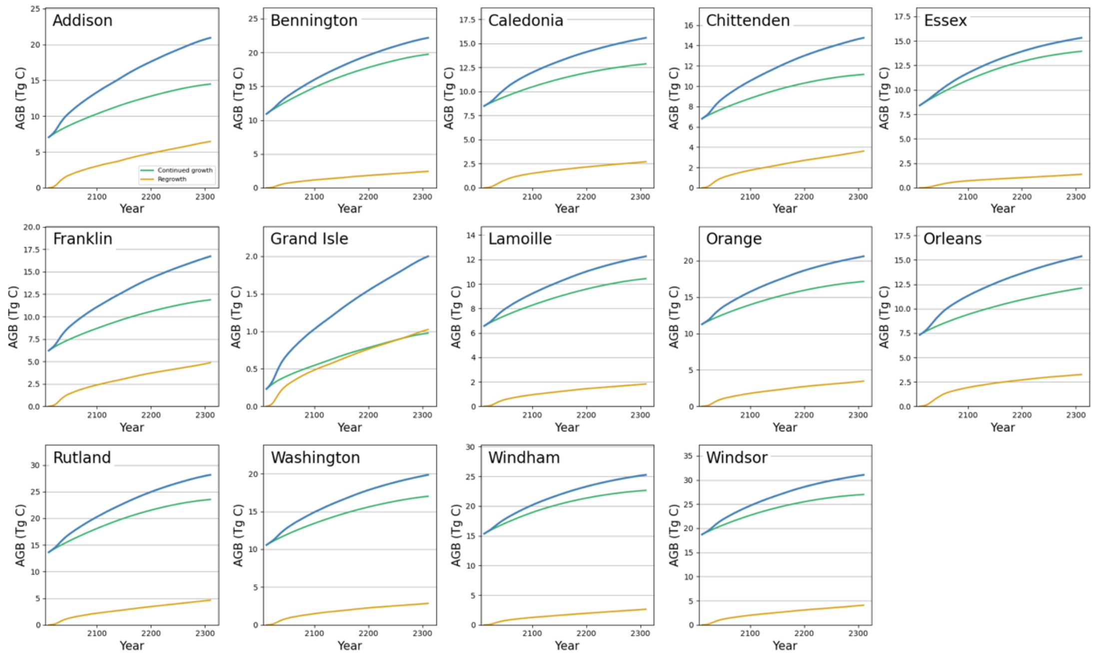


Figure C.23. As in Figure C.16 but for counties in Vermont.

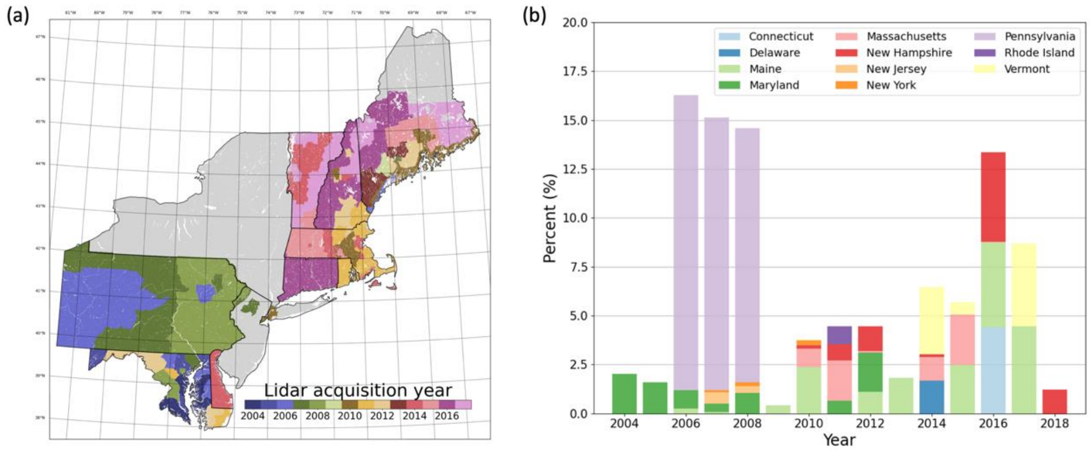


Figure C.24. Lidar canopy height acquisition year map (a) and histogram (b).

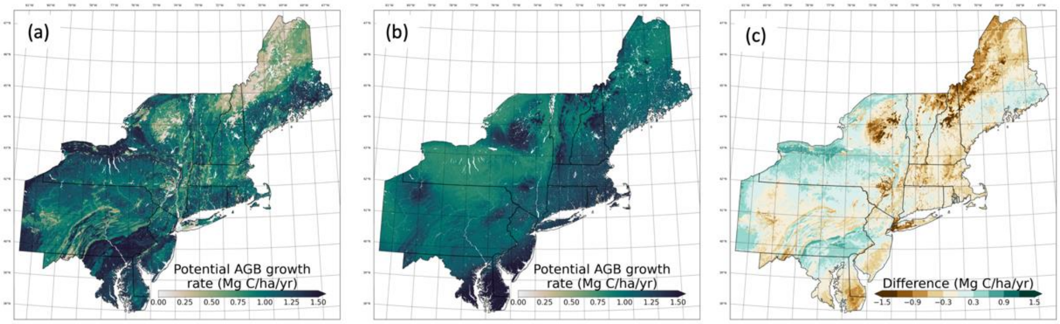


Figure C.25. RGGI region maps of average potential AGB growth rate for the first 30 years of natural forest regrowth from this study (a), from Cook-Patton et al 2020 (b) and the absolute difference between this study and Cook-Patton et al 2020 (c).

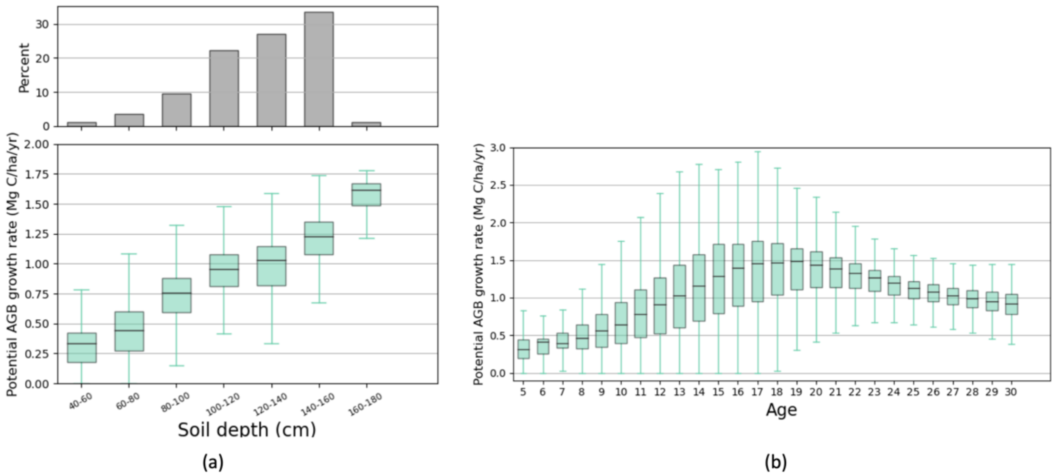


Figure C.26. Stratification of average AGB growth rate (figure C.19b) by soil depth (figure C.2c) for the first 30 years of natural forest regrowth from this study. Annual AGB growth rate as function of stand age between 5 and 30 years (b).

Table C.1. County-level ED initialized AGB and projected carbon stocks (Tg C) over continued growth and new regrowth areas of Connecticut, in 2020, 2030, 2040, 2050, respectively. Total land area is reported in km^2.

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Fairfield	1024.32	357.59	8.92	9.26	9.64	10.00	10.35	0.14	0.67	1.09	1.33
Hartford	1064.36	530.50	8.52	8.87	9.25	9.61	9.95	0.16	0.84	1.47	1.83
Litchfield	1810.56	504.44	17.35	17.82	18.35	18.86	19.34	0.15	0.77	1.39	1.78
Middlesex	689.06	195.85	6.11	6.31	6.54	6.76	6.97	0.07	0.33	0.55	0.68
New Haven	952.22	340.57	7.47	7.77	8.09	8.40	8.70	0.11	0.53	0.90	1.13
New London	1244.38	372.25	10.41	10.86	11.36	11.83	12.28	0.15	0.72	1.15	1.40
Tolland	749.33	223.09	7.14	7.36	7.59	7.81	8.03	0.07	0.37	0.63	0.80
Windham	983.61	309.39	9.38	9.67	9.99	10.29	10.58	0.11	0.53	0.90	1.12

Table C.2. As in Table C.1 but for counties in Delaware.

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Kent	461.10	845.23	2.90	3.14	3.41	3.66	3.89	0.39	1.80	2.77	3.36
New Castle	369.29	509.02	2.56	2.78	2.99	3.18	3.36	0.26	1.15	1.73	2.11
Sussex	972.02	1263.72	5.61	6.12	6.64	7.11	7.55	0.46	2.27	3.70	4.52

Table C.3. As in Table C.1 but for counties in Maryland.

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Allegany	849.70	197.96	6.01	6.26	6.51	6.74	6.96	0.04	0.19	0.37	0.51
Anne Arundel	629.23	326.28	4.64	4.94	5.24	5.52	5.78	0.15	0.68	1.05	1.26
Baltimore	743.82	646.06	7.08	7.44	7.81	8.14	8.44	0.40	1.64	2.36	2.90
Baltimore City	56.44	57.29	0.22	0.28	0.32	0.36	0.39	0.03	0.14	0.20	0.24
Calvert	340.28	179.91	4.24	4.31	4.40	4.48	4.56	0.09	0.40	0.60	0.73
Caroline	287.05	532.19	2.19	2.31	2.44	2.56	2.68	0.20	1.00	1.60	1.94
Carroll	408.44	716.26	3.83	4.00	4.18	4.34	4.48	0.34	1.55	2.38	2.93
Cecil	398.02	471.13	3.89	4.03	4.21	4.36	4.51	0.26	1.12	1.66	2.02
Charles	766.56	309.75	8.20	8.45	8.72	8.98	9.22	0.15	0.68	1.04	1.25
Dorchester	482.24	542.06	3.14	3.37	3.66	3.92	4.17	0.25	1.12	1.72	2.09
Frederick	724.15	934.78	6.27	6.57	6.91	7.21	7.48	0.44	1.98	3.07	3.77
Garrett	1205.78	429.66	10.21	10.56	10.95	11.33	11.69	0.09	0.48	0.98	1.33
Harford	462.08	587.43	4.64	4.85	5.07	5.26	5.44	0.36	1.49	2.15	2.63
Howard	331.32	274.09	3.10	3.25	3.41	3.55	3.68	0.16	0.67	0.98	1.21
Kent	198.67	487.99	1.65	1.72	1.82	1.90	1.98	0.23	1.05	1.61	1.94
Montgomery	627.59	530.08	2.86	3.33	3.77	4.15	4.50	0.25	1.11	1.70	2.11
Prince George's	646.11	430.27	5.34	5.65	5.96	6.24	6.51	0.21	0.96	1.46	1.76
Queen Anne's	295.74	627.07	2.52	2.64	2.79	2.92	3.04	0.29	1.33	2.04	2.46
Somerset	349.31	282.34	2.35	2.53	2.72	2.91	3.08	0.17	0.64	0.94	1.14
St. Mary's	562.25	323.35	5.05	5.24	5.57	5.84	6.07	0.16	0.71	1.07	1.29

Table C.3 (continued)

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Talbot	222.93	423.89	2.00	2.08	2.18	2.27	2.36	0.20	0.89	1.36	1.64
Washington	527.88	551.40	3.91	4.15	4.41	4.64	4.86	0.35	1.38	2.01	2.47
Wicomico	458.81	420.85	2.71	2.93	3.15	3.36	3.56	0.15	0.70	1.15	1.42
Worcester	612.39	482.55	4.65	4.93	5.24	5.52	5.79	0.25	1.01	1.55	1.87

Table C.4. As in Table C.1 but for counties in Massachusetts.

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Barnstable	585.82	192.87	2.27	2.54	2.82	3.08	3.31	0.06	0.29	0.51	0.64
Berkshire	1909.64	424.38	17.59	17.99	18.43	18.85	19.25	0.11	0.54	1.04	1.38
Bristol	920.58	303.62	5.94	6.28	6.64	6.99	7.33	0.10	0.47	0.80	1.00
Dukes	133.74	85.15	0.42	0.51	0.59	0.65	0.71	0.03	0.14	0.24	0.30
Essex	725.08	244.77	5.56	5.75	5.99	6.22	6.44	0.07	0.37	0.65	0.83
Franklin	1474.65	259.06	15.08	15.30	15.56	15.81	16.04	0.05	0.24	0.51	0.70
Hampden	1149.92	315.67	10.81	11.06	11.34	11.60	11.85	0.06	0.32	0.65	0.88
Hampshire	1049.25	260.10	10.13	10.33	10.56	10.78	10.98	0.05	0.25	0.53	0.72
Middlesex	1310.63	442.67	10.20	10.54	10.92	11.29	11.64	0.09	0.45	0.92	1.24
Nantucket	25.15	55.41	0.03	0.05	0.07	0.09	0.10	0.02	0.08	0.14	0.18
Norfolk	619.44	200.59	4.10	4.32	4.55	4.77	4.98	0.05	0.28	0.51	0.65
Plymouth	1080.53	365.08	7.41	7.78	8.19	8.57	8.94	0.11	0.56	0.97	1.22
Suffolk	42.10	23.23	0.19	0.20	0.22	0.23	0.24	0.00	0.02	0.05	0.06
Worcester	2773.27	810.84	24.35	25.07	25.89	26.69	27.44	0.19	0.98	1.90	2.54

Table C.5. As in Table C.1 but for counties in New Hampshire.

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Belknap	894.62	132.15	7.64	7.85	8.09	8.32	8.55	0.03	0.17	0.33	0.43
Carroll	2138.53	238.32	17.22	17.74	18.31	18.85	19.37	0.06	0.30	0.58	0.78
Cheshire	1601.28	168.06	15.56	15.88	16.26	16.61	16.95	0.04	0.20	0.40	0.53
Coos	3974.24	600.55	22.10	22.94	23.84	24.68	25.49	0.08	0.36	0.80	1.21
Grafton	4031.34	322.01	32.60	33.52	34.49	35.40	36.25	0.07	0.35	0.71	0.96
Hillsborough	1816.71	324.13	16.06	16.46	16.93	17.37	17.80	0.06	0.31	0.66	0.90
Merrimack	1971.02	300.57	17.11	17.56	18.08	18.58	19.06	0.07	0.34	0.68	0.92
Rockingham	1305.31	334.36	11.33	11.63	12.00	12.35	12.69	0.08	0.39	0.77	1.02
Strafford	753.63	140.76	6.05	6.23	6.44	6.64	6.84	0.03	0.14	0.30	0.41
Sullivan	1246.32	132.07	11.23	11.54	11.87	12.18	12.48	0.03	0.17	0.33	0.44

Table C.6. As in Table C.1 but for counties in Pennsylvania.

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Adams	580.10	745.43	3.06	3.49	3.93	4.30	4.64	0.39	1.79	2.68	3.25
Allegheny	1038.33	500.11	5.34	5.87	6.38	6.83	7.26	0.15	0.78	1.34	1.70
Armstrong	976.03	658.67	5.60	5.92	6.27	6.60	6.92	0.11	0.56	1.20	1.67
Beaver	707.80	351.86	3.88	4.18	4.48	4.76	5.02	0.09	0.45	0.85	1.10
Bedford	1799.38	810.12	9.84	10.52	11.22	11.88	12.52	0.21	1.00	1.85	2.43
Berks	1032.23	1071.91	8.39	8.89	9.39	9.84	10.27	0.48	2.15	3.41	4.26
Blair	951.75	385.13	6.39	6.76	7.14	7.49	7.84	0.15	0.71	1.16	1.46
Bradford	1760.80	1168.88	12.06	12.77	13.50	14.18	14.82	0.33	1.70	3.15	4.10
Bucks	747.35	664.22	4.90	5.38	5.84	6.25	6.63	0.33	1.51	2.31	2.79
Butler	1156.89	819.90	6.70	7.20	7.70	8.17	8.61	0.22	1.14	2.09	2.71
Cambria	1178.43	568.42	8.26	8.70	9.16	9.59	10.01	0.13	0.67	1.34	1.80
Cameron	961.68	62.49	10.17	10.39	10.63	10.86	11.07	0.02	0.09	0.17	0.22
Carbon	771.07	179.47	4.10	4.43	4.76	5.06	5.35	0.04	0.21	0.41	0.55
Centre	2178.18	658.77	15.17	15.94	16.73	17.49	18.22	0.23	1.10	1.88	2.41
Chester	879.40	979.82	7.61	8.13	8.63	9.08	9.50	0.62	2.60	3.73	4.58
Clarion	930.48	631.40	6.16	6.50	6.87	7.21	7.54	0.14	0.74	1.47	1.98
Clearfield	2194.84	747.09	14.75	15.56	16.41	17.20	17.95	0.17	0.87	1.74	2.33
Clinton	2005.24	266.09	15.13	15.76	16.42	17.06	17.68	0.08	0.40	0.71	0.91
Columbia	737.71	526.07	5.33	5.58	5.85	6.11	6.35	0.12	0.64	1.22	1.60
Crawford	1544.49	1017.41	11.24	12.00	12.78	13.48	14.14	0.41	2.02	3.33	4.11
Cumberland	606.62	694.93	3.26	3.65	4.02	4.35	4.66	0.39	1.56	2.37	2.93

Table C.6 (continued)

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Dauphin	746.11	524.35	4.68	5.02	5.36	5.66	5.95	0.19	0.94	1.55	1.94
Delaware	239.08	144.48	1.92	2.06	2.19	2.30	2.42	0.08	0.34	0.50	0.62
Elk	1813.86	290.79	16.87	17.43	18.01	18.56	19.07	0.09	0.46	0.83	1.06
Erie	1070.92	890.68	7.50	8.11	8.71	9.25	9.77	0.42	1.97	3.09	3.81
Fayette	1357.75	619.96	9.42	10.03	10.65	11.21	11.76	0.24	1.19	1.94	2.41
Forest	982.77	118.09	9.43	9.70	9.98	10.25	10.50	0.03	0.17	0.31	0.41
Franklin	914.82	1039.72	4.62	5.15	5.65	6.11	6.54	0.52	2.15	3.33	4.14
Fulton	789.40	343.13	4.13	4.43	4.73	5.01	5.27	0.08	0.37	0.70	0.94
Greene	1054.31	460.44	6.93	7.57	8.19	8.77	9.31	0.22	1.07	1.64	1.99
Huntingdon	1760.44	507.65	11.26	11.91	12.58	13.22	13.83	0.15	0.73	1.29	1.66
Indiana	1246.29	862.15	7.52	7.97	8.44	8.89	9.33	0.16	0.83	1.74	2.38
Jefferson	1111.46	573.68	8.11	8.50	8.90	9.29	9.66	0.13	0.66	1.33	1.78
Juniata	626.71	324.64	4.39	4.61	4.85	5.08	5.29	0.11	0.54	0.90	1.15
Lackawanna	586.86	485.28	2.84	3.11	3.39	3.63	3.85	0.11	0.57	1.11	1.47
Lancaster	627.50	1714.38	4.19	4.63	5.04	5.41	5.75	1.13	4.59	6.65	8.06
Lawrence	426.81	470.80	2.19	2.46	2.71	2.93	3.14	0.19	0.96	1.54	1.88
Lebanon	353.24	553.50	2.40	2.58	2.76	2.92	3.07	0.29	1.21	1.88	2.34
Lehigh	354.46	438.61	2.29	2.48	2.66	2.82	2.97	0.17	0.73	1.22	1.56
Luzerne	1692.51	522.54	10.85	11.57	12.31	12.99	13.64	0.15	0.76	1.39	1.79
Lycoming	2503.80	679.73	19.53	20.27	21.05	21.79	22.51	0.15	0.75	1.45	1.94
McKean	2173.52	340.93	22.38	22.97	23.57	24.12	24.64	0.10	0.54	0.99	1.26
Mercer	871.32	810.24	4.58	5.10	5.61	6.06	6.48	0.31	1.57	2.60	3.19
Mifflin	722.72	331.22	5.21	5.48	5.77	6.05	6.32	0.14	0.65	1.04	1.29

Table C.6 (continued)

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Monroe	1287.07	262.49	8.85	9.40	9.97	10.50	11.01	0.08	0.42	0.74	0.94
Montgomery	562.72	506.08	3.36	3.76	4.14	4.47	4.78	0.26	1.20	1.81	2.19
Montour	130.50	180.17	0.80	0.84	0.89	0.93	0.97	0.05	0.23	0.42	0.55
Northampton	403.04	478.58	2.57	2.77	2.97	3.14	3.30	0.16	0.74	1.27	1.64
Northumberland	625.89	536.73	3.92	4.13	4.37	4.58	4.79	0.12	0.62	1.20	1.57
Perry	1027.61	424.94	7.50	7.88	8.28	8.66	9.03	0.17	0.76	1.24	1.57
Philadelphia	92.95	74.60	0.53	0.59	0.65	0.70	0.74	0.03	0.13	0.21	0.26
Pike	1257.45	121.94	9.00	9.46	9.96	10.45	10.92	0.04	0.19	0.33	0.42
Potter	2322.98	468.05	23.76	24.38	25.01	25.60	26.15	0.15	0.76	1.37	1.75
Schuylkill	1396.60	580.81	8.27	8.84	9.41	9.93	10.43	0.13	0.69	1.36	1.78
Snyder	467.94	343.52	2.92	3.09	3.26	3.42	3.58	0.08	0.41	0.76	1.01
Somerset	1691.91	1012.58	9.99	10.65	11.35	11.99	12.60	0.21	1.10	2.26	3.06
Sullivan	984.53	175.04	7.55	7.93	8.32	8.68	9.02	0.06	0.29	0.52	0.65
Susquehanna	1390.37	716.92	10.46	10.98	11.53	12.05	12.54	0.22	1.12	2.02	2.59
Tioga	2001.70	892.26	15.16	15.86	16.57	17.23	17.85	0.25	1.29	2.39	3.11
Union	521.35	288.36	3.39	3.57	3.77	3.96	4.14	0.09	0.45	0.78	0.98
Venango	1297.43	439.55	8.76	9.26	9.79	10.29	10.77	0.12	0.64	1.18	1.53
Warren	1893.46	380.19	20.39	20.85	21.34	21.80	22.23	0.13	0.66	1.16	1.46
Washington	1211.72	915.12	6.30	7.16	7.96	8.67	9.34	0.43	2.11	3.25	3.91
Wayne	1377.26	492.24	10.52	11.02	11.57	12.08	12.57	0.15	0.77	1.40	1.79
Westmoreland	1605.74	1008.20	10.22	10.93	11.65	12.31	12.93	0.31	1.61	2.81	3.58
Wyoming	756.80	280.39	5.71	5.98	6.28	6.56	6.82	0.08	0.40	0.74	0.96
York	954.86	1233.72	7.40	7.95	8.48	8.97	9.42	0.65	2.92	4.36	5.35

Table C.7. As in Table C.1 but for counties in Rhode Island.

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Bristol	24.14	16.16	0.14	0.15	0.16	0.17	0.17	0.01	0.03	0.04	0.05
Kent	327.66	64.87	2.85	2.95	3.07	3.19	3.30	0.02	0.11	0.18	0.22
Newport	116.86	86.01	0.61	0.67	0.73	0.78	0.83	0.04	0.16	0.26	0.32
Providence	757.72	155.60	6.91	7.13	7.38	7.62	7.86	0.05	0.25	0.42	0.53
Washington	603.40	131.74	4.75	4.97	5.22	5.45	5.68	0.05	0.25	0.39	0.49

Table C.8. As in Table C.1 but for counties in Vermont.

County	Continued growth area	Regrowth area	ED initialized AGB	Continued growth				Regrowth			
				2020	2030	2040	2050	2020	2030	2040	2050
Addison	1175.56	508.74	7.07	7.51	7.96	8.36	8.73	0.16	0.79	1.41	1.79
Bennington	1493.19	202.86	10.93	11.42	11.92	12.39	12.84	0.05	0.28	0.53	0.69
Caledonia	1318.58	333.75	8.51	8.76	9.01	9.25	9.48	0.05	0.25	0.56	0.82
Chittenden	936.48	317.19	6.82	7.07	7.33	7.57	7.80	0.08	0.42	0.78	1.02
Essex	1510.37	195.73	8.42	8.72	9.05	9.36	9.66	0.02	0.10	0.22	0.35
Franklin	1019.18	433.45	6.24	6.56	6.90	7.21	7.48	0.11	0.58	1.09	1.42
Grand Isle	81.61	90.58	0.24	0.29	0.33	0.37	0.41	0.03	0.13	0.23	0.29
Lamoille	966.86	179.00	6.58	6.80	7.02	7.22	7.42	0.04	0.21	0.41	0.56
Orange	1385.70	316.28	11.29	11.62	11.97	12.30	12.61	0.07	0.40	0.79	1.06
Orleans	1282.41	400.01	7.34	7.61	7.90	8.15	8.39	0.07	0.35	0.76	1.10
Rutland	1922.42	399.92	13.64	14.19	14.76	15.29	15.80	0.10	0.50	0.97	1.28
Washington	1484.56	276.09	10.58	10.94	11.32	11.67	12.00	0.06	0.32	0.65	0.88
Windham	1759.43	230.02	15.38	15.82	16.28	16.72	17.13	0.05	0.28	0.55	0.73
Windsor	2115.04	364.78	18.73	19.22	19.73	20.22	20.68	0.08	0.44	0.88	1.18

Bibliography

Achard, F., DeFries, R., Eva, H., Hansen, M., Mayaux, P., & Stibig, H.-J. (2007). Pan-tropical monitoring of deforestation. Environmental Research Letters, 2(4), 045022. https://doi.org/10.1088/1748-9326/2/4/045022

Ahlström, A., Schurgers, G., Arneth, A., & Smith, B. (2012). Robustness and uncertainty in terrestrial ecosystem carbon response to CMIP5 climate change projections. Environmental Research Letters, 7(4), 044008.

Albani, M., Medvigy, D., Hurtt, G. C., & Moorcroft, P. R. (2006). The contributions of land-use change, CO2 fertilization, and climate variability to the Eastern US carbon sink. Global Change Biology, 12(12), 2370–2390. https://doi.org/10.1111/j.1365-2486.2006.01254.x

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., & Zhao, M. (2015). Spatiotemporal patterns of terrestrial gross primary production: A review. Reviews of Geophysics, 53(3), 785–818. https://doi.org/10.1002/2015RG000483

Antonarakis, Alexander S., Saatchi, S. S., Chazdon, R. L., & Moorcroft, P. R. (2011). Using Lidar and Radar measurements to constrain predictions of forest ecosystem structure and function. Ecological Applications, 21(4), 1120–1137. https://doi.org/10.1890/10-0274.1

Antonarakis, A. S., Munger, J. W., & Moorcroft, P. R. (2014). Imaging spectroscopy- and lidar-derived estimates of canopy composition and structure to improve predictions of forest carbon fluxes and ecosystem dynamics. Geophysical Research Letters, 41(7), 2535–2542. https://doi.org/10.1002/2013GL058373

Avitabile, V., Herold, M., Heuvelink, G. B., Lewis, S. L., Phillips, O. L., Asner, G. P., Armston, J., Ashton, P. S., Banin, L., & Bayol, N. (2016). An integrated pan-tropical biomass map using multiple reference datasets. Global Change Biology, 22(4), 1406–1420.

Baccini, A, Laporte, N., Goetz, S. J., Sun, M., & Dong, H. (2008). A first map of tropical Africa's above-ground biomass derived from satellite imagery. Environmental Research Letters, 3(4), 045011. https://doi.org/10.1088/1748-9326/3/4/045011

Baccini, A, Goetz, S. J., Walker, W., Laporte, N., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P., Dubayah, R., & Friedl, M. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nature Climate Change, 2(3), 182–185.

Baker, I. T., Prihodko, L., Denning, A. S., Goulden, M., Miller, S., & Rocha, H. R. da. (2008). Seasonal drought stress in the Amazon: Reconciling models and observations. Journal of Geophysical Research: Biogeosciences, 113(G1). https://doi.org/10.1029/2007JG000644

Bartholomé, E., & Belward, A. S. (2005). GLC2000: A new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing, 26(9), 1959–1977.

Bechtold, W. A., & Patterson, P. L. (2005). The enhanced forest inventory and analysis program—National sampling design and estimation procedures (Vol. 80). USDA Forest Service, Southern Research Station.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., … Papale, D. (2010). Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate. Science, 329(5993), 834. https://doi.org/10.1126/science.1184984

Bernacchi, C. J., Singsaas, E. L., Pimentel, C., Jr, A. R. P., & Long, S. P. (2001). Improved temperature response functions for models of Rubisco-limited photosynthesis. Plant, Cell & Environment, 24(2), 253–259. https://doi.org/10.1111/j.1365-3040.2001.00668.x

Besnard, S., Carvalhais, N., Arain, M. A., Black, A., De Bruin, S., Buchmann, N., Cescatti, A., Chen, J., Clevers, J. G., & Desai, A. R. (2018). Quantifying the effect of forest age in annual net forest carbon balance. Environmental Research Letters, 13(12), 124018.

Betts, R. A. (2001). Biogeophysical impacts of land use on present-day climate: Near-surface temperature change and radiative forcing. Atmospheric Science Letters, 2(1–4), 39–51.

Bicheron, P., Amberg, V., Bourg, L., Petit, D., Huc, M., Miras, B., Brockmann, C., Delwart, S., Ranéra, F., & Hagolle, O. (2008). Geolocation assessment of 300 m resolution MERIS Globcover ortho-rectified products. Proceedings of MERIS/AATSR Colloque, Frascati.

Birdsey, R., Pregitzer, K., & Lucier, A. (2006). Forest Carbon Management in the United States. Journal of Environmental Quality, 35(4), 1461–1469. https://doi.org/10.2134/jeq2005.0162

Blackard, J. A., Finco, M. V., Helmer, E. H., Holden, G. R., Hoppus, M. L., Jacobs, D. M., Lister, A. J., Moisen, G. G., Nelson, M. D., Riemann, R., Ruefenacht, B., Salajanu, D., Weyermann, D. L., Winterberger, K. C., Brandeis, T. J., Czaplewski, R. L., McRoberts, R. E., Patterson, P. L., & Tymcio, R. P. (2008). Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. Remote Sensing of Environment, 112(4), 1658–1677. https://doi.org/10.1016/j.rse.2007.08.021

Bonan, G. B. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. Science, 320(5882), 1444–1449.

Botkin, D. B., Janak, J. F., & Wallis, J. R. (1972). Some ecological consequences of a computer model of forest growth. The Journal of Ecology, 849–872.

Brovkin, V., Sitch, S., Von Bloh, W., Claussen, M., Bauer, E., & Cramer, W. (2004). Role of land cover changes for atmospheric CO2 increase and climate change during the last

150 years. Global Change Biology, 10(8), 1253–1266. https://doi.org/10.1111/j.1365-2486.2004.00812.x

Brovkin, V., Claussen, M., Driesschaert, E., Fichefet, T., Kicklighter, D., Loutre, M.-F., Matthews, H. D., Ramankutty, N., Schaeffer, M., & Sokolov, A. (2006). Biogeophysical effects of historical land cover changes simulated by six Earth system models of intermediate complexity. Climate Dynamics, 26(6), 587–600.

Bugmann, H. K. M., & Solomon, A. M. (2000). EXPLAINING FOREST COMPOSITION AND BIOMASS ACROSS MULTIPLE BIOGEOGRAPHICAL REGIONS. Ecological Applications, 10(1), 95–114. https://doi.org/10.1890/1051-0761(2000)010[0095:EFCABA]2.0.CO;2

Canadell, J. G., Kirschbaum, M. U. F., Kurz, W. A., Sanz, M.-J., Schlamadinger, B., & Yamagata, Y. (2007). Factoring out natural and indirect human effects on terrestrial carbon sources and sinks. Options for Including Agriculture and Forestry Activities in a Post-2012 International Climate Agreement, 10(4), 370–384. https://doi.org/10.1016/j.envsci.2007.01.009

Canadell, J. G., & Raupach, M. R. (2008). Managing Forests for Climate Change Mitigation. Science, 320(5882), 1456. https://doi.org/10.1126/science.1155458

Carvalhais, N., Reichstein, M., Ciais, P., Collatz, G. J., Mahecha, M. D., Montagnani, L., Papale, D., Rambal, S., & Seixas, J. (2010). Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints. Global Change Biology, 16(10), 2813–2829.

Chaney, N. W., Wood, E. F., McBratney, A. B., Hempel, J. W., Nauman, T. W., Brungard, C. W., & Odgers, N. P. (2016). POLARIS: A 30-meter probabilistic soil series map of the contiguous United States. Geoderma, 274, 54–67. https://doi.org/10.1016/j.geoderma.2016.03.025

Chapman, M., Walker, W. S., Cook-Patton, S. C., Ellis, P. W., Farina, M., Griscom, B. W., & Baccini, A. (2020). Large climate mitigation potential from adding trees to agricultural lands. Global Change Biology, 26(8), 4357–4365. https://doi.org/10.1111/gcb.15121

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., & Ciais, P. (2005). Inferring CO2 sources and sinks from satellite observations: Method and application to TOVS data. Journal of Geophysical Research: Atmospheres, 110(D24). https://doi.org/10.1029/2005JD006390

Chini, LP, Hurtt, G., Klein Goldewijk, K., Frolking, S., Shevliakova, E., Thornton, P., & Fisk, J. (2012). Addressing the pasture anomaly: How uncertainty in historical pasture data leads to divergence of atmospheric CO2 in Earth System Models. 2012, GC11D-1021.

Chini, Louise, Hurtt, G., Sahajpal, R., Frolking, S., Klein Goldewijk, K., Sitch, S., Ganzenmüller, R., Ma, L., Ott, L., & Pongratz, J. (2021). Land-Use Harmonization Datasets for Annual Global Carbon Budgets. Earth System Science Data Discussions, 1–27.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., & Heimann, M. (2014). Carbon and other biogeochemical cycles. In Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (pp. 465–570). Cambridge University Press.

Claussen, M., Brovkin, V., & Ganopolski, A. (2001). Biogeophysical versus biogeochemical feedbacks of large-scale land cover change. Geophysical Research Letters, 28(6), 1011–1014.

Cook-Patton, S. C., Leavitt, S. M., Gibbs, D., Harris, N. L., Lister, K., Anderson-Teixeira, K. J., Briggs, R. D., Chazdon, R. L., Crowther, T. W., Ellis, P. W., Griscom, H. P., Herrmann, V., Holl, K. D., Houghton, R. A., Larrosa, C., Lomax, G., Lucas, R., Madsen, P., Malhi, Y., … Griscom, B. W. (2020). Mapping carbon accumulation potential from global natural forest regrowth. Nature, 585(7826), 545–550. https://doi.org/10.1038/s41586-020-2686-x

Cramer, W., Bondeau, A., Woodward, F. I., Prentice, I. C., Betts, R. A., Brovkin, V., Cox, P. M., Fisher, V., Foley, J. A., Friend, A. D., Kucharik, C., Lomas, M. R., Ramankutty, N., Sitch, S., Smith, B., White, A., & Young-Molling, C. (2001). Global response of terrestrial ecosystem structure and function to CO2 and climate change: Results from six dynamic global vegetation models. Global Change Biology, 7(4), 357–373. https://doi.org/10.1046/j.1365-2486.2001.00383.x

Crowther, T. W., Todd-Brown, K. E. O., Rowe, C. W., Wieder, W. R., Carey, J. C., Machmuller, M. B., Snoek, B. L., Fang, S., Zhou, G., Allison, S. D., Blair, J. M., Bridgham, S. D., Burton, A. J., Carrillo, Y., Reich, P. B., Clark, J. S., Classen, A. T., Dijkstra, F. A., Elberling, B., … Bradford, M. A. (2016). Quantifying global soil carbon losses in response to warming. Nature, 540(7631), 104–108. https://doi.org/10.1038/nature20150

Defries, R. S., Hansen, M. C., Townshend, J. R., Janetos, A., & Loveland, T. R. (2000). A new global 1-km dataset of percentage tree cover derived from remote sensing. Global Change Biology, 6(2), 247–254.

DeFries, R. S., Houghton, R. A., Hansen, M. C., Field, C. B., Skole, D., & Townshend, J. (2002). Carbon emissions from tropical deforestation and regrowth based on satellite observations for the 1980s and 1990s. Proceedings of the National Academy of Sciences, 99(22), 14256–14261.

DeHayes, D. H. (1992). Winter Injury and Developmental Cold Tolerance of Red Spruce. In C. Eagar & M. B. Adams (Eds.), Ecology and Decline of Red Spruce in the Eastern United States (pp. 295–337). Springer. https://doi.org/10.1007/978-1-4612-2906-3_8

Dewar, R. C. (1991). Analytical model of carbon storage in the trees, soils, and wood products of managed forests. Tree Physiology, 8(3), 239–258.

Depro, B. M., Murray, B. C., Alig, R. J., & Shanks, A. (2008). Public land, timber harvests, and climate mitigation: Quantifying carbon sequestration potential on U.S. public timberlands. Forest Ecology and Management, 255(3), 1122–1134. https://doi.org/10.1016/j.foreco.2007.10.036

Domke, G. M., Oswalt, S. N., Walters, B. F., & Morin, R. S. (2020a). Tree planting has the potential to increase carbon sequestration capacity of forests in the United States. Proceedings of the National Academy of Sciences, 117(40), 24649–24651. https://doi.org/10.1073/pnas.2010840117

Domke, G. M., Walters, B. F., Nowak, D. J., Smith, J., Ogle, S. M., Coulston, J. W., & Wirth, T. C. (2020b). Greenhouse gas emissions and removals from forest land, woodlands, and urban trees in the United States, 1990-2018. Resource Update FS-227. Madison, WI: U.S. Department of Agriculture, Forest Service, Northern Research Station. 5 p., 227, 1–5. https://doi.org/10.2737/FS-RU-227

Drake, J. B., Dubayah, R. O., Clark, D. B., Knox, R. G., Blair, J. B., Hofton, M. A., Chazdon, R. L., Weishampel, J. F., & Prince, S. (2002). Estimation of tropical forest structural characteristics using large-footprint lidar. Remote Sensing of Environment, 79(2), 305–319. https://doi.org/10.1016/S0034-4257(01)00281-4

Dubayah, R. O., & Drake, J. B. (2000). Lidar Remote Sensing for Forestry. Journal of Forestry, 98(6), 44–46. https://doi.org/10.1093/jof/98.6.44

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P. L., Qi, W., & Silva, C. (2020a). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Science of Remote Sensing, 1, 100002. https://doi.org/10.1016/j.srs.2020.100002

Dubayah, R., Hofton, M., Blair, J., Armston, J., Tang, H., & Luthcke, S. (2020b). GEDI L2A Elevation and Height Metrics Data Global Footprint Level V001 [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/GEDI/GEDI02_A.001

Dubayah, R., Tang, H., Armston, J., Luthcke, S., Hofton, M., & Blair, J. (2020c). GEDI L2B Canopy Cover and Vertical Profile Metrics Data Global Footprint Level V001 [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/GEDI/GEDI02_B.001

Erb, K.-H., Kastner, T., Luyssaert, S., Houghton, R. A., Kuemmerle, T., Olofsson, P., & Haberl, H. (2013). Bias in the attribution of forest carbon sinks. Nature Climate Change, 3(10), 854–856. https://doi.org/10.1038/nclimate2004

Erb, K.-H., Kastner, T., Plutzar, C., Bais, A. L. S., Carvalhais, N., Fetzel, T., Gingrich, S., Haberl, H., Lauk, C., & Niedertscheider, M. (2018). Unexpectedly large impact of forest management and grazing on global vegetation biomass. Nature, 553(7686), 73–76.

ESA. (2017). Land Cover CCI Product User Guide Version 2. Tech. Rep. Available at: maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf

Fadrique, B., Báez, S., Duque, Á., Malizia, A., Blundo, C., Carilla, J., Osinaga-Acosta, O., Malizia, L., Silman, M., & Farfán-Ríos, W. (2018). Widespread but heterogeneous responses of Andean forests to climate change. Nature, 564(7735), 207–212.

FAO. (2015). Global Forest Resources Assessment 2015. Food and Agriculture Organization of the United Nations.

Fargione, J. E., Bassett, S., Boucher, T., Bridgham, S. D., Conant, R. T., Cook-Patton, S. C., Ellis, P. W., Falcucci, A., Fourqurean, J. W., Gopalakrishna, T., Gu, H., Henderson, B., Hurteau, M. D., Kroeger, K. D., Kroeger, T., Lark, T. J., Leavitt, S. M., Lomax, G., McDonald, R. I., … Griscom, B. W. (2018). Natural climate solutions for the United States. Science Advances, 4(11), eaat1869. https://doi.org/10.1126/sciadv.aat1869

Farjon, A., & Filer, D. (2013). An Atlas of the World's Conifers: An Analysis of their Distribution, Biogeography, Diversity and Conservation Status. In An Atlas of the World's Conifers. Brill. https://brill.com/view/title/20587

Farquhar, G. D., von Caemmerer, S., & Berry, J. A. (1980). A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. Planta, 149(1), 78–90. https://doi.org/10.1007/BF00386231

Feddema, J. J., Oleson, K. W., Bonan, G. B., Mearns, L. O., Buja, L. E., Meehl, G. A., & Washington, W. M. (2005). The importance of land-cover change in simulating future climates. Science, 310(5754), 1674–1678.

Finzi, A. C., Norby, R. J., Calfapietra, C., Gallet-Budynek, A., Gielen, B., Holmes, W. E., Hoosbeek, M. R., Iversen, C. M., Jackson, R. B., & Kubiske, M. E. (2007). Increases in nitrogen uptake rather than nitrogen-use efficiency support higher rates of temperate forest productivity under elevated CO2. Proceedings of the National Academy of Sciences, 104(35), 14014–14019.

Fischer, R., Bohn, F., de Paula, M. D., Dislich, C., Groeneveld, J., Gutiérrez, A. G., Kazmierczak, M., Knapp, N., Lehmann, S., & Paulick, S. (2016). Lessons learned from applying a forest gap model to understand ecosystem and carbon dynamics of complex tropical forests. Ecological Modelling, 326, 124–133.

Fisher, J. B., Tu, K. P., & Baldocchi, D. D. (2008). Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. Remote Sensing of Environment, 112(3), 901–919.

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

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

Fisk, J. P., Hurtt, G. C., Chambers, J. Q., Zeng, H., Dolan, K. A., & Negrón-Juárez, R. I. (2013). The impacts of tropical cyclones on the net carbon balance of eastern US forests (1851–2000). Environmental Research Letters, 8(4), 045017. https://doi.org/10.1088/1748-9326/8/4/045017

Fisk, J. P. (2015). Net effects of disturbance: Spatial, temporal, and soci-etal dimensions of forest disturbance and recovery on terrestrial car-bon balance [Ph.D thesis]. University of Maryland.

Flanagan, S. A., Hurtt, G. C., Fisk, J. P., Sahajpal, R., Zhao, M., Dubayah, R., Hansen, M. C., Sullivan, J. H., & Collatz, G. J. (2019). Potential Transient Response of Terrestrial Vegetation and Carbon in Northern North America from Climate Change. Climate, 7(9), 113. https://doi.org/10.3390/cli7090113

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114(1), 168–182.

Friedl, M. A., Gray, J. M., Melaas, E. K., Richardson, A. D., Hufkens, K., Keenan, T. F., Bailey, A., & O'Keefe, J. (2014). A tale of two springs: Using recent climate anomalies to characterize the sensitivity of temperate forest phenology to climate change. Environmental Research Letters, 9(5), 054006.

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., & Knutti, R. (2014). Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks. Journal of Climate, 27(2), 511–526.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., … Zaehle, S. (2019). Global Carbon Budget 2019. Earth System Science Data, 11(4), 1783–1838. https://doi.org/10.5194/essd-11-1783-2019

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., & Sitch, S. (2020). Global carbon budget 2020. Earth System Science Data, 12(4), 3269–3340.

Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers, R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., … Woodward, F. I. (2014). Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2. Proceedings of the National Academy of Sciences, 111(9), 3280–3285. https://doi.org/10.1073/pnas.1222477110

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., … Zhao, B. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). Journal of Climate, 30(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Gibbs, H. K., Brown, S., Niles, J. O., & Foley, J. A. (2007). Monitoring and estimating tropical forest carbon stocks: Making REDD a reality. Environmental Research Letters, 2(4), 045023. https://doi.org/10.1088/1748-9326/2/4/045023

Goldstein, A., Turner, W. R., Spawn, S. A., Anderson-Teixeira, K. J., Cook-Patton, S., Fargione, J., Gibbs, H. K., Griscom, B., Hewson, J. H., Howard, J. F., Ledezma, J. C., Page, S., Koh, L. P., Rockström, J., Sanderman, J., & Hole, D. G. (2020). Protecting irrecoverable carbon in Earth's ecosystems. Nature Climate Change, 10(4), 287–295. https://doi.org/10.1038/s41558-020-0738-8

Goward, S. N., Huang, C., Zhao, F., Schleeweis, K., Rishmawi, K., Lindsey, M., Dungan, J. L., & Michaelis, A. R. (2016). NACP NAFD Project: Forest Disturbance History from Landsat, 1986-2010. ORNL DAAC. https://doi.org/10.3334/ORNLDAAC/1290

Griscom, B. W., Adams, J., Ellis, P. W., Houghton, R. A., Lomax, G., Miteva, D. A., Schlesinger, W. H., Shoch, D., Siikamäki, J. V., Smith, P., Woodbury, P., Zganjar, C., Blackman, A., Campari, J., Conant, R. T., Delgado, C., Elias, P., Gopalakrishna, T., Hamsik, M. R., … Fargione, J. (2017). Natural climate solutions. Proceedings of the National Academy of Sciences, 114(44), 11645–11650. https://doi.org/10.1073/pnas.1710465114

Gu, L., Hanson, P. J., Post, W. M., Kaiser, D. P., Yang, B., Nemani, R., Pallardy, S. G., & Meyers, T. (2008). The 2007 Eastern US Spring Freeze: Increased Cold Damage in a Warming World? BioScience, 58(3), 253–262. https://doi.org/10.1641/B580311

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., Frankenberg, C., Huete, A. R., Zarco-Tejada, P., Lee, J.-E., Moran, M. S., Ponce-Campos, G., Beer, C., Camps-Valls, G., Buchmann, N., Gianelle, D., Klumpp, K., Cescatti, A., Baker, J. M., & Griffis, T. J. (2014). Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. Proceedings of the National Academy of Sciences, 111(14), E1327–E1333. https://doi.org/10.1073/pnas.1320008111

Guo, L. B., & Gifford, R. M. (2002). Soil carbon stocks and land use change: A meta analysis. Global Change Biology, 8(4), 345–360.

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., Lovejoy, T. E., Sexton, J. O., Austin, M. P., Collins, C. D., Cook, W. M., Damschen, E. I., Ewers, R. M., Foster, B. L., Jenkins, C. N., King, A. J., Laurance, W. F., Levey, D. J., Margules, C. R., … Townshend, J. R. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. Science Advances, 1(2), e1500052. https://doi.org/10.1126/sciadv.1500052

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. Science, 342(6160), 850–853. https://doi.org/10.1126/science.1244693

Hansis, E., Davis, S. J., & Pongratz, J. (2015). Relevance of methodological choices for accounting of land use change carbon fluxes. Global Biogeochemical Cycles, 29(8), 1230–1246. https://doi.org/10.1002/2014GB004997

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. PLOS ONE, 12(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Hickler, T., Smith, B., Prentice, I. C., Mjöfors, K., Miller, P., Arneth, A., & Sykes, M. T. (2008). CO2 fertilization in temperate FACE experiments not representative of boreal and tropical forests. Global Change Biology, 14(7), 1531–1542.

Houghton, R. (1999). The annual net flux of carbon to the atmosphere from changes in land use 1850–1990. Tellus B, 51(2), 298–313.

Houghton, Richard A. (2003). Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850–2000. Tellus B: Chemical and Physical Meteorology, 55(2), 378–390.

Houghton, R. A., Hall, F., & Goetz, S. J. (2009). Importance of biomass in the global carbon cycle. Journal of Geophysical Research: Biogeosciences, 114(G2). https://doi.org/10.1029/2009JG000935

Houghton, Richard A. (2010). How well do we know the flux of CO2 from land-use change? Tellus B: Chemical and Physical Meteorology, 62(5), 337–351.

Houghton, Richard A, House, J. I., Pongratz, J., Van Der Werf, G. R., DeFries, R. S., Hansen, M. C., Quéré, C. L., & Ramankutty, N. (2012). Carbon emissions from land use and land-cover change. Biogeosciences, 9(12), 5125–5142.

Houghton, Richard A, & Nassikas, A. A. (2017). Global and regional fluxes of carbon from land use and land cover change 1850–2015. Global Biogeochemical Cycles, 31(3), 456–472.

Huang, W., Swatantran, A., Johnson, K., Duncanson, L., Tang, H., O'Neil Dunne, J., Hurtt, G., & Dubayah, R. (2015). Local discrepancies in continental scale biomass maps: A case study over forested and non-forested landscapes in Maryland, USA. Carbon Balance and Management, 10(1), 19. https://doi.org/10.1186/s13021-015-0030-9

Huang, W., Dolan, K., Swatantran, A., Johnson, K., Tang, H., O'Neil-Dunne, J., Dubayah, R., & Hurtt, G. (2019). High-resolution mapping of aboveground biomass for forest carbon monitoring system in the Tri-State region of Maryland, Pennsylvania and Delaware, USA. Environmental Research Letters, 14(9), 095002. https://doi.org/10.1088/1748-9326/ab2917

Huntingford, C., Fisher, R. A., Mercado, L., Booth, B. B. B., Sitch, S., Harris, P. P., Cox, P. M., Jones, C. D., Betts, R. A., Malhi, Y., Harris, G. R., Collins, M., & Moorcroft, P. (2008). Towards quantifying uncertainty in predictions of Amazon 'dieback.' Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1498), 1857–1864. https://doi.org/10.1098/rstb.2007.0028

Hurtt, G. C., Moorcroft, P. R., Pacala, S. W., & Levin, S. A. (1998). Terrestrial models and global change: Challenges for the future. Global Change Biology, 4(5), 581–590. https://doi.org/10.1046/j.1365-2486.1998.t01-1-00203.x

Hurtt, G. C., Pacala, S. W., Moorcroft, P. R., Caspersen, J., Shevliakova, E., Houghton, R. A., & Moore, B. (2002). Projecting the future of the U.S. carbon sink. Proceedings of the National Academy of Sciences, 99(3), 1389–1394. https://doi.org/10.1073/pnas.012249999

Hurtt, G. C., Dubayah, R., Drake, J., Moorcroft, P. R., Pacala, S. W., Blair, J. B., & Fearon, M. G. (2004). Beyond Potential Vegetation: Combining Lidar Data and a Height-Structured Model for Carbon Studies. Ecological Applications, 14(3), 873–883. https://doi.org/10.1890/02-5317

Hurtt, G. C., Frolking, S., Fearon, M., Moore, B., Shevliakova, E., Malyshev, S., Pacala, S. W., & Houghton, R. (2006). The underpinnings of land-use history: Three centuries of global gridded land-use transitions, wood-harvest activity, and resulting secondary lands. Global Change Biology, 12(7), 1208–1229.

Hurtt, G. C., Fisk, J., Thomas, R. Q., Dubayah, R., Moorcroft, P. R., & Shugart, H. H. (2010). Linking models and data on vegetation structure. Journal of Geophysical Research: Biogeosciences, 115(G2). https://doi.org/10.1029/2009JG000937

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R., Feddema, J., Fischer, G., Fisk, J., Hibbard, K., Houghton, R., & Janetos, A. (2011). Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. Climatic Change, 109(1), 117–161.

Hurtt, G. C., Wickland, D., Jucks, K., Bowman, K., Brown, M., Duren, R., Hagen, S., & Verdy, A. (2014). NASA CARBON MONITORING SYSTEM: Prototype Monitoring, Reporting, and Verification. 1–37.

Hurtt, G. C., Thomas, R. Q., Fisk, J. P., Dubayah, R. O., & Sheldon, S. L. (2016). The Impact of Fine-Scale Disturbances on the Predictability of Vegetation Dynamics and Carbon Flux. PLOS ONE, 11(4), e0152883. https://doi.org/10.1371/journal.pone.0152883

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J., Fisk, J., Fujimori, S., Goldewijk, K. K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J., Krisztin, T., Lawrence, D., Lawrence, P., … Zhang, X. (2019a). Harmonization of Global Land Use Change and Management for the Period 850-2015. Earth System Grid Federation. https://doi.org/10.22033/ESGF/input4MIPs.10454

Hurtt, G. C., Zhao, M., Sahajpal, R., Armstrong, A., Birdsey, R., Campbell, E., Dolan, K., Dubayah, R., Fisk, J. P., Flanagan, S., Huang, C., Huang, W., Johnson, K., Lamb, R., Ma, L., Marks, R., O'Leary, D., O'Neil-Dunne, J., Swatantran, A., & Tang, H. (2019b). Beyond MRV: High-resolution forest carbon modeling for climate mitigation planning over Maryland, USA. Environmental Research Letters, 14(4), 045013. https://doi.org/10.1088/1748-9326/ab0bbe

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., & Klein Goldewijk, K. (2020). Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. Geoscientific Model Development, 13(11), 5425–5464.

Jeong, S.-J., & Medvigy, D. (2014). Macroscale prediction of autumn leaf coloration throughout the continental United States. Global Ecology and Biogeography, 23(11), 1245–1254. https://doi.org/10.1111/geb.12206

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., & Xian, G. (2013). A comprehensive change detection method for updating the National Land Cover Database to circa 2011. Remote Sensing of Environment, 132, 159–175. https://doi.org/10.1016/j.rse.2013.01.012

Joiner, J., Yoshida, Y., Zhang, Y., Duveiller, G., Jung, M., Lyapustin, A., Wang, Y., & Tucker, C. J. (2018). Estimation of Terrestrial Global Gross Primary Production (GPP) with Satellite Data-Driven Models and Eddy Covariance Flux Data. Remote Sensing, 10(9), 1346. https://doi.org/10.3390/rs10091346

Jones, C., Robertson, E., Arora, V., Friedlingstein, P., Shevliakova, E., Bopp, L., Brovkin, V., Hajima, T., Kato, E., & Kawamiya, M. (2013). Twenty-first-century compatible CO2 emissions and airborne fraction simulated by CMIP5 earth system models under four representative concentration pathways. Journal of Climate, 26(13), 4398–4413.

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S., Bodesheim, P., Carvalhais, N., Chevallier, F., Gans, F., Goll, D. S., Haverd, V., Köhler, P., Ichii, K., Jain, A. K., Liu, J., Lombardozzi, D., … Reichstein, M. (2020). Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the FLUXCOM approach. Biogeosciences, 17(5), 1343–1365. https://doi.org/10.5194/bg-17-1343-2020

Kaplan, J. O., Krumhardt, K. M., & Zimmermann, N. (2009). The prehistoric and preindustrial deforestation of Europe. Quaternary Science Reviews, 28(27–28), 3016–3034.

Kattge, J., & Knorr, W. (2007). Temperature acclimation in a biochemical model of photosynthesis: A reanalysis of data from 36 species. Plant, Cell & Environment, 30(9), 1176–1190. https://doi.org/10.1111/j.1365-3040.2007.01690.x

Keeling, R. F. (2008). Recording Earth's Vital Signs. Science, 319(5871), 1771. https://doi.org/10.1126/science.1156761

Keenan, Trevor F, Prentice, I. C., Canadell, J. G., Williams, C. A., Wang, H., Raupach, M., & Collatz, G. J. (2016). Recent pause in the growth rate of atmospheric CO 2 due to enhanced terrestrial carbon uptake. Nature Communications, 7(1), 1–10.

Keenan, T. F., & Williams, C. A. (2018). The Terrestrial Carbon Sink. Annual Review of Environment and Resources, 43(1), 219–243. https://doi.org/10.1146/annurev-environ-102017-030204

Kellndorfer, J., Walker, W., Kirsch, K., Fiske, G., BISHOP, J., Lapoint, L., Hoppus, M., & Westfall, J. (2013). NACP Aboveground Biomass and Carbon Baseline Data, V.2 (NBCD 2000), U.S.A., 2000. https://doi.org/10.3334/ORNLDAAC/1161

Kim, Y., Knox, R. G., Longo, M., Medvigy, D., Hutyra, L. R., Pyle, E. H., Wofsy, S. C., Bras, R. L., & Moorcroft, P. R. (2012). Seasonal carbon dynamics and water fluxes in an Amazon rainforest. Global Change Biology, 18(4), 1322–1334. https://doi.org/10.1111/j.1365-2486.2011.02629.x

Klein Goldewijk, K., Beusen, A., Doelman, J., & Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene–HYDE 3.2. Earth System Science Data, 9(2), 927–953.

Köhler, P., & Huth, A. (1998). The effects of tree species grouping in tropical rainforest modelling: Simulations with the individual-based model FORMIND. Ecological Modelling, 109(3), 301–321.

Körner, C. (2017). A matter of tree longevity. Science, 355(6321), 130. https://doi.org/10.1126/science.aal2449

Kucharik, C. J., Foley, J. A., Delire, C., Fisher, V. A., Coe, M. T., Lenters, J. D., Young-Molling, C., Ramankutty, N., Norman, J. M., & Gower, S. T. (2000). Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure. Global Biogeochemical Cycles, 14(3), 795–825.

Lamarche, C., Santoro, M., Bontemps, S., d'Andrimont, R., Radoux, J., Giustarini, L., Brockmann, C., Wevers, J., Defourny, P., & Arino, O. (2017). Compilation and validation of SAR and optical data products for a complete and global map of inland/ocean water tailored to the climate modeling community. Remote Sensing, 9(1), 36.

Lamb, R. L., Hurtt, G., Boudreau, T. J., Campbell, E., Carlo, E. S., Chu, H.-H., de Mooy, J., Dubayah, R. O., Gonsalves, D., & Guy, M. (2021a). Context and future directions for integrating forest carbon into sub-national climate mitigation planning in the RGGI region of the US. Environmental Research Letters.

Lamb, R., Ma, L., Sahajpal, R., Edmonds, J., Hultman, N., Dubayah, R., Kennedy, J.,& Hurtt, G. (2021b). Geospatial assessment of the economic opportunity for reforestation in Maryland (USA). Environmental Research Letters. In Review.

Lamb, R., Hurtt, G., Spivy, A., Ma, L., Campbell, E., Dubayah, R., Goetz, S., Hoffman Delett, C., Jantz, P., Panday, F.M., & Tang, H. (2021c). High-resolution geospatial framework for future forest carbon storage and protected area expansion in Maryland (USA). Environmental Research Letters. In Prep.

Lark, T. J., Spawn, S. A., Bougie, M., & Gibbs, H. K. (2020). Cropland expansion in the United States produces marginal yields at high costs to wildlife. Nature Communications, 11(1), 4295. https://doi.org/10.1038/s41467-020-18045-z

Law, B. E., Turner, D., Campbell, J., Sun, O., Van Tuyl, S., Ritts, W., & Cohen, W. (2004). Disturbance and climate effects on carbon stocks and fluxes across Western Oregon USA. Global Change Biology, 10(9), 1429–1444.

Lawrence, D. M., Hurtt, G. C., Arneth, A., Brovkin, V., Calvin, K. V., Jones, A. D., Jones, C. D., Lawrence, P. J., Noblet-Ducoudré, N. de, & Pongratz, J. (2016). The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6: Rationale and experimental design. Geoscientific Model Development, 9(9), 2973–2998.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., & Canadell, J. G. (2018). Global carbon budget 2018. Earth System Science Data, 10(4), 2141–2194.

Lee, J.-E., Frankenberg, C., van der Tol, C., Berry, J. A., Guanter, L., Boyce, C. K., Fisher, J. B., Morrow, E., Worden, J. R., Asefi, S., Badgley, G., & Saatchi, S. (2013). Forest productivity and water stress in Amazonia: Observations from GOSAT chlorophyll fluorescence. Proceedings of the Royal Society B: Biological Sciences, 280(1761), 20130171. https://doi.org/10.1098/rspb.2013.0171

Levine, N. M., Zhang, K., Longo, M., Baccini, A., Phillips, O. L., Lewis, S. L., Alvarez-Dávila, E., Andrade, A. C. S. de, Brienen, R. J. W., Erwin, T. L., Feldpausch, T. R., Mendoza, A. L. M., Vargas, P. N., Prieto, A., Silva-Espejo, J. E., Malhi, Y., & Moorcroft, P. R. (2016). Ecosystem heterogeneity determines the ecological resilience of the Amazon to climate change. Proceedings of the National Academy of Sciences, 113(3), 793–797. https://doi.org/10.1073/pnas.1511344112

Lippke, B., Oneil, E., Harrison, R., Skog, K., Gustavsson, L., & Sathre, R. (2011). Life cycle impacts of forest management and wood utilization on carbon mitigation: Knowns and unknowns. Carbon Management, 2(3), 303–333. https://doi.org/10.4155/cmt.11.24

Longo, Marcos, Knox, R. G., Levine, N. M., Alves, L. F., Bonal, D., Camargo, P. B., Fitzjarrald, D. R., Hayek, M. N., Restrepo-Coupe, N., & Saleska, S. R. (2018). Ecosystem heterogeneity and diversity mitigate Amazon forest resilience to frequent extreme droughts. New Phytologist, 219(3), 914–931.

Longo, M., Knox, R. G., Medvigy, D. M., Levine, N. M., Dietze, M. C., Kim, Y., Swann, A. L. S., Zhang, K., Rollinson, C. R., Bras, R. L., Wofsy, S. C., & Moorcroft, P. R. (2019a). The biophysics, ecology, and biogeochemistry of functionally diverse, vertically and horizontally heterogeneous ecosystems: The Ecosystem Demography model, version 2.2 – Part 1: Model description. Geoscientific Model Development, 12(10), 4309–4346. https://doi.org/10.5194/gmd-12-4309-2019

Longo, M., Knox, R. G., Levine, N. M., Swann, A. L. S., Medvigy, D. M., Dietze, M. C., Kim, Y., Zhang, K., Bonal, D., Burban, B., Camargo, P. B., Hayek, M. N., Saleska, S. R., da Silva, R., Bras, R. L., Wofsy, S. C., & Moorcroft, P. R. (2019b). The biophysics, ecology, and biogeochemistry of functionally diverse, vertically and horizontally heterogeneous ecosystems: The Ecosystem Demography model, version 2.2 – Part 2: Model evaluation for tropical South America. Geoscientific Model Development, 12(10), 4347–4374. https://doi.org/10.5194/gmd-12-4347-2019

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing, 21(6–7), 1303–1330.

Lu, D. (2005). Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. International Journal of Remote Sensing, 26(12), 2509–2525.

Ma, L., Hurtt, G. C., Chini, L. P., Sahajpal, R., Pongratz, J., Frolking, S., Stehfest, E., Klein Goldewijk, K., O'Leary, D., & Doelman, J. C. (2020a). Global rules for translating land-use change (LUH2) to land-cover change for CMIP6 using GLM2. Geoscientific Model Development, 13(7), 3203–3220.

Ma, L., Hurtt, G., Ott, L., Sahajpal, R., Fisk, J., Flanagan, S., Poulter, B., Liang, S., Sullivan, J., & Dubayah, R. (2020b, December 21). Global Ecosystem Demography Model (ED-global v1.0): Development, Calibration and Evaluation for NASA's Global Ecosystem Dynamics Investigation (GEDI) (world) [Poster]. Earth and Space Science Open Archive; Earth and Space Science Open Archive. https://doi.org/10.1002/essoar.10505486.1

Ma, L., Hurtt, G., Tang, H., Lamb, R. L., Campbell, E., Dubayah, R. O., Guy, M., Huang, W., Lister, A., & Lu, J. (2021). High-resolution forest carbon modeling for climate mitigation planning over the RGGI region, USA. Environmental Research Letters. https://doi.org/10.1088/1748-9326/abe4f4

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., … Zwally, J. (2017). The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation. Remote Sensing of Environment, 190, 260–273. https://doi.org/10.1016/j.rse.2016.12.029

Massad, R.-S., Tuzet, A., & Bethenod, O. (2007). The effect of temperature on C4-type leaf photosynthesis parameters. Plant, Cell & Environment, 30(9), 1191–1204. https://doi.org/10.1111/j.1365-3040.2007.01691.x

McDowell, N. G., Allen, C. D., Anderson-Teixeira, K., Aukema, B. H., Bond-Lamberty, B., Chini, L., Clark, J. S., Dietze, M., Grossiord, C., Hanbury-Brown, A., Hurtt, G. C., Jackson, R. B., Johnson, D. J., Kueppers, L., Lichstein, J. W., Ogle, K., Poulter, B., Pugh, T. A. M., Seidl, R., … Xu, C. (2020). Pervasive shifts in forest dynamics in a changing world. Science, 368(6494). https://doi.org/10.1126/science.aaz9463

Medvigy, D. (2006). The State of the Regional Carbon Cycle: Results from a Constrained Coupled Ecosystem-atmosphere Model [Ph.D thesis, Harvard University]. https://books.google.com/books?id=ZJJwYgEACAAJ

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: Biogeosciences, 114(G1). https://doi.org/10.1029/2008JG000812

Medvigy, David, & Moorcroft, P. R. (2012). Predicting ecosystem dynamics at regional scales: An evaluation of a terrestrial biosphere model for the forests of northeastern North America. Philosophical Transactions of the Royal Society B: Biological Sciences, 367(1586), 222–235. https://doi.org/10.1098/rstb.2011.0253

Menlove, J., & Healey, S. P. (2020). A Comprehensive Forest Biomass Dataset for the USA Allows Customized Validation of Remotely Sensed Biomass Estimates. Remote Sensing, 12(24), 4141.

Menzel, A., & Fabian, P. (1999). Growing season extended in Europe. Nature, 397(6721), 659–659. https://doi.org/10.1038/17709

Mercado, L. M., Bellouin, N., Sitch, S., Boucher, O., Huntingford, C., Wild, M., & Cox, P. M. (2009). Impact of changes in diffuse radiation on the global land carbon sink. Nature, 458(7241), 1014–1017.

Miller, A. D., Dietze, M. C., DeLucia, E. H., & Anderson-Teixeira, K. J. (2016). Alteration of forest succession and carbon cycling under elevated CO2. Global Change Biology, 22(1), 351–363. https://doi.org/10.1111/gcb.13077

Miller, D. A., & White, R. A. (1998). A Conterminous United States Multilayer Soil Characteristics Dataset for Regional Climate and Hydrology Modeling. Earth Interactions, 2(2), 1–26. https://doi.org/10.1175/1087-3562(1998)002<0001:ACUSMS>2.3.CO;2

Mitchard, E. T., Saatchi, S. S., Baccini, A., Asner, G. P., Goetz, S. J., Harris, N. L., & Brown, S. (2013). Uncertainty in the spatial distribution of tropical forest biomass: A comparison of pan-tropical maps. Carbon Balance and Management, 8(1), 1–13.

Monteith, J. (1965). Evaporation and environment. Symposia of the Society for Experimental Biology, 19, 205–234. PubMed.

Montzka, C., Herbst, M., Weihermüller, L., Verhoef, A., & Vereecken, H. (2017). A global data set of soil hydraulic properties and sub-grid variability of soil water retention and hydraulic conductivity curves. Earth System Science Data, 9(2), 529–543.

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

Murray-Tortarolo, G., Anav, A., Friedlingstein, P., Sitch, S., Piao, S., Zhu, Z., Poulter, B., Zaehle, S., Ahlström, A., Lomas, M., Levis, S., Viovy, N., & Zeng, N. (2013). Evaluation of Land Surface Models in Reproducing Satellite-Derived LAI over the High-Latitude Northern Hemisphere. Part I: Uncoupled DGVMs. Remote Sensing, 5(10), 4819–4838. https://doi.org/10.3390/rs5104819

Myneni, R., Knyazikhin, Y., & Park. (2015). MCD15A3H MODIS/Terra+Aqua Leaf Area Index/FPAR 4-day L4 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MCD15A3H.006

Nave, L. E., Vance, E. D., Swanston, C. W., & Curtis, P. S. (2010). Harvest impacts on soil carbon storage in temperate forests. Forest Ecology and Management, 259(5), 857–866.

Neuenschwander, A., & Pitts, K. (2019). The ATL08 land and vegetation product for the ICESat-2 Mission. Remote Sensing of Environment, 221, 247–259.

Neuenschwander, A. L., Popescu, S. C., Nelson, R. F., Harding, D., Pitts, K. L., & Robbins, J. (2020). ATLAS/ICESat-2 L3A Land and Vegetation Height, version 3 [Data set]. NASA National Snow and Ice Data Center DAAC. https://doi.org/10.5067/ATLAS/ATL08.003

Norby, R. J., & Zak, D. R. (2011). Ecological Lessons from Free-Air CO2 Enrichment (FACE) Experiments. Annual Review of Ecology, Evolution, and Systematics, 42(1), 181–203. https://doi.org/10.1146/annurev-ecolsys-102209-144647

O’Neil-Dunne, J., MacFaden, S., & Royar, A. (2014a). A Versatile, Production-Oriented Approach to High-Resolution Tree-Canopy Mapping in Urban and Suburban Landscapes Using GEOBIA and Data Fusion. Remote Sensing, 6(12), 12837–12865. https://doi.org/10.3390/rs61212837

O’Neil-Dunne, J., MacFaden, S., Royar, A., Reis, M., Dubayah, R., & Swatantran, A. (2014b). An object-based approach to statewide land cover mapping. Proceedings of ASPRS 2014 Annual Conference, 23–28.

Ordway, E. M., & Asner, G. P. (2020). Carbon declines along tropical forest edges correspond to heterogeneous effects on canopy structure and function. Proceedings of the National Academy of Sciences, 117(14), 7863–7870. https://doi.org/10.1073/pnas.1914420117

O’Sullivan, M., Smith, W. K., Sitch, S., Friedlingstein, P., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Kautz, M., Lombardozzi, D., Nabel, J. E. M. S., Tian, H., Vuichard, N., Wiltshire, A., Zhu, D., & Buermann, W. (2020). Climate-Driven Variability and Trends in Plant Productivity Over Recent Decades Based on Three Global Products. Global Biogeochemical Cycles, 34(12), e2020GB006613. https://doi.org/10.1029/2020GB006613

Ott, L., Hurtt, G. C., Randerson, J. T., Chatterjee, A., Chen, Y., Chini, L. P., Davis, S. J., Hubacek, K., Lee, E., MA, L., Poulter, B., Rousseaux, C. S., Sun, L., Woodard, D., & Zeng, F. (2018). Toward integrated seasonal predictions of land and ocean carbon flux: Lessons from the 2015-16 El Nino. AGU Fall Meeting Abstracts, 51. http://adsabs.harvard.edu/abs/2018AGUFM.B51E1990O

Pacala, S. W., Canham, C. D., Saponara, J., Silander Jr, J. A., Kobe, R. K., & Ribbens, E. (1996). Forest models defined by field measurements: Estimation, error analysis and dynamics. Ecological Monographs, 66(1), 1–43.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S., & Hayes, D. (2011). A Large and

Persistent Carbon Sink in the World's Forests. Science, 333(6045), 988–993. https://doi.org/10.1126/science.1201609

Pan, Yude, Birdsey, R. A., Phillips, O. L., & Jackson, R. B. (2013). The structure, distribution, and biomass of the world's forests. Annual Review of Ecology, Evolution, and Systematics, 44, 593–622.

Pappas, C., Fatichi, S., Rimkus, S., Burlando, P., & Huber, M. O. (2015). The role of local-scale heterogeneities in terrestrial ecosystem modeling. Journal of Geophysical Research: Biogeosciences, 120(2), 341–360. https://doi.org/10.1002/2014JG002735

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., & Tans, P. P. (2007). An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. Proceedings of the National Academy of Sciences, 104(48), 18925. https://doi.org/10.1073/pnas.0708986104

Peylin, P., Law, R., Gurney, K., Chevallier, F., Jacobson, A., Maki, T., Niwa, Y., Patra, P., Peters, W., & Rayner, P. (2013). Global atmospheric carbon budget: Results from an ensemble of atmospheric CO 2 inversions. Biogeosciences, 10(10), 6699–6720.

Piao, S., Ciais, P., Friedlingstein, P., Peylin, P., Reichstein, M., Luyssaert, S., Margolis, H., Fang, J., Barr, A., Chen, A., Grelle, A., Hollinger, D. Y., Laurila, T., Lindroth, A., Richardson, A. D., & Vesala, T. (2008). Net carbon dioxide losses of northern ecosystems in response to autumn warming. Nature, 451(7174), 49–52. https://doi.org/10.1038/nature06444

Piao, S., Wang, X., Park, T., Chen, C., Lian, X., He, Y., Bjerke, J. W., Chen, A., Ciais, P., Tømmervik, H., Nemani, R. R., & Myneni, R. B. (2020). Characteristics, drivers and feedbacks of global greening. Nature Reviews Earth & Environment, 1(1), 14–27. https://doi.org/10.1038/s43017-019-0001-x

Pielke Sr., R. A., Pitman, A., Niyogi, D., Mahmood, R., McAlpine, C., Hossain, F., Goldewijk, K. K., Nair, U., Betts, R., Fall, S., Reichstein, M., Kabat, P., & de Noblet, N. (2011). Land use/land cover changes and climate: Modeling analysis and observational evidence. WIREs Climate Change, 2(6), 828–850. https://doi.org/10.1002/wcc.144

Pongratz, J., Reick, C., Raddatz, T., & Claussen, M. (2008). A reconstruction of global agricultural areas and land cover for the last millennium. Global Biogeochemical Cycles, 22(3).

Pongratz, J., Reick, C., Raddatz, T., & Claussen, M. (2009). Effects of anthropogenic land cover change on the carbon cycle of the last millennium. Global Biogeochemical Cycles, 23(4).

Pongratz, J., Reick, C., Raddatz, T., & Claussen, M. (2010). Biogeophysical versus biogeochemical climate response to historical anthropogenic land cover change. Geophysical Research Letters, 37(8).

Pongratz, J., Reick, C. H., Houghton, R., & House, J. (2014). Terminology as a key uncertainty in net land use and land cover change carbon flux estimates. Earth System Dynamics, 5(1), 177–195.

Pongratz, J., Dolman, H., Don, A., Erb, K., Fuchs, R., Herold, M., Jones, C., Kuemmerle, T., Luyssaert, S., & Meyfroidt, P. (2018). Models meet data: Challenges and opportunities in implementing land management in Earth system models. Global Change Biology, 24(4), 1470–1487.

Post, W. M., & Kwon, K. C. (2000). Soil carbon sequestration and land-use change: Processes and potential. Global Change Biology, 6(3), 317–327.

Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., & Peylin, P. (2015). Plant functional type classification for earth system models: Results from the European Space Agency's Land Cover Climate Change Initiative. Geoscientific Model Development, 8(7), 2315–2328. https://doi.org/10.5194/gmd-8-2315-2015

Poulter, Benjamin, Heyder, U., & Cramer, W. (2009). Modeling the Sensitivity of the Seasonal Cycle of GPP to Dynamic LAI and Soil Depths in Tropical Rainforests. Ecosystems, 12(4), 517–533. https://doi.org/10.1007/s10021-009-9238-4

Pregitzer, K. S., & Euskirchen, E. S. (2004). Carbon cycling and storage in world forests: Biome patterns related to forest age. Global Change Biology, 10(12), 2052–2077.

Prentice, I. Colin, Bondeau, A., Cramer, W., Harrison, S. P., Hickler, T., Lucht, W., Sitch, S., Smith, B., & Sykes, M. T. (2007). Dynamic Global Vegetation Modeling: Quantifying Terrestrial Ecosystem Responses to Large-Scale Environmental Change. In J. G. Canadell, D. E. Pataki, & L. F. Pitelka (Eds.), Terrestrial Ecosystems in a Changing World (pp. 175–192). Springer. https://doi.org/10.1007/978-3-540-32730-1_15

Prentice, Iain Colin, & Cowling, S. A. (2013). Dynamic global vegetation models. Encyclopedia of Biodiversity, 670–689. https://doi.org/10.1016/B978-0-12-384719-5.00412-3

Pugh, T. A., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V., & Calle, L. (2019). Role of forest regrowth in global carbon sink dynamics. Proceedings of the National Academy of Sciences, 116(10), 4382–4387.

Purves, D., & Pacala, S. (2008). Predictive Models of Forest Dynamics. Science, 320(5882), 1452–1453. https://doi.org/10.1126/science.1155359

Quesada, B., Arneth, A., Robertson, E., & de Noblet-Ducoudré, N. (2018). Potential strong contribution of future anthropogenic land-use and land-cover change to the terrestrial carbon cycle. Environmental Research Letters, 13(6), 064023.

Quillet, A. Q., Peng, C. P., & Garneau, M. G. (2010). Toward dynamic global vegetation models for simulating vegetation–climate interactions and feedbacks: Recent developments, limitations, and future challenges. Environmental Reviews. https://doi.org/10.1139/A10-016

Ramankutty, N., & Foley, J. A. (1999). Estimating historical changes in global land cover: Croplands from 1700 to 1992. Global Biogeochemical Cycles, 13(4), 997–1027.

Randerson, J., Van Der Werf, G., Giglio, L., Collatz, G., & Kasibhatla, P. (2015). Global Fire Emissions Database, Version 4.1 (GFEDv4). ORNL DAAC.

Reich, P. B., Walters, M. B., & Ellsworth, D. S. (1997). From tropics to tundra: Global convergence in plant functioning. Proceedings of the National Academy of Sciences, 94(25), 13730. https://doi.org/10.1073/pnas.94.25.13730

Reick, C., Raddatz, T., Pongratz, J., & Claussen, M. (2010). Contribution of anthropogenic land cover change emissions to pre-industrial atmospheric CO2. Tellus B: Chemical and Physical Meteorology, 62(5), 329–336.

Reinmann, A. B., Smith, I. A., Thompson, J. R., & Hutyra, L. R. (2020). Urbanization and fragmentation mediate temperate forest carbon cycle response to climate. Environmental Research Letters, 15(11), 114036. https://doi.org/10.1088/1748-9326/abbf16

Rhemtulla, J. M., Mladenoff, D. J., & Clayton, M. K. (2009). Historical forest baselines reveal potential for continued carbon sequestration. Proceedings of the National Academy of Sciences, 106(15), 6082–6087. https://doi.org/10.1073/pnas.0810076106

Rödenbeck, C., Le Quéré, C., Heimann, M., & Keeling, R. F. (2008). Interannual variability in oceanic biogeochemical processes inferred by inversion of atmospheric O2/N2 and CO2 data. Tellus B: Chemical and Physical Meteorology, 60(5), 685–705. https://doi.org/10.1111/j.1600-0889.2008.00375.x

Roxburgh, S. H., Wood, S. W., Mackey, B. G., Woldendorp, G., & Gibbons, P. (2006). Assessing the carbon sequestration potential of managed forests: A case study from temperate Australia. Journal of Applied Ecology, 43(6), 1149–1159. https://doi.org/10.1111/j.1365-2664.2006.01221.x

Ruesch, A., & Gibbs, H. K. (2008). New IPCC Tier-1 global biomass carbon map for the year 2000. Available Online from the Carbon Dioxide Information Analysis Center [Http://Cdiac. Ornl. Gov], Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., & Morel, A. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. Proceedings of the National Academy of Sciences, 108(24), 9899–9904. https://doi.org/10.1073/pnas.1019576108

Saatchi, S., Yu, Y., Fore, A., Nuemann, M., Chapman, B., Nguyen, Nemani, R., Ganguly, S., Zhang, G., Votava, P., Masek, J., Tucker, C., Hall, F., Nelson, R., Cook, B., Birdsey, R., Healey, S., Johnson, K., Dubayah, R., & Lefsky, M. (2012). CMS-Multi-scale forest biomass maps based on multiple remote sensing inputs. CMS Biomass Pilot Project. https://nacp-files.nacarbon.org/biomass_pilot/JPL_ARC_Maps/Continental_US/

Sacks, W. J., Deryng, D., Foley, J. A., & Ramankutty, N. (2010). Crop planting dates: An analysis of global patterns. Global Ecology and Biogeography, 19(5), 607–620.

Sakai, A., & Weiser, C. J. (1973). Freezing Resistance of Trees in North America with Reference to Tree Regions. Ecology, 54(1), 118–126. https://doi.org/10.2307/1934380

Sakai, Akira, & Larcher, W. (2012). Frost Survival of Plants: Responses and Adaptation to Freezing Stress. Springer Science & Business Media.

Santoro, M., Beaudoin, A., Beer, C., Cartus, O., Fransson, J. E., Hall, R. J., Pathe, C., Schmullius, C., Schepaschenko, D., & Shvidenko, A. (2015). Forest growing stock volume of the northern hemisphere: Spatially explicit estimates for 2010 derived from Envisat ASAR. Remote Sensing of Environment, 168, 316–334.

Santoro, M., Cartus, O., Mermoz, S., Bouvet, A., Le Toan, T., Carvalhais, N., Rozendaal, D., Herold, M., Avitabile, V., Quegan, S., Carreiras, J., Rauste, Y., Balzter, H., Schmullius, C., & Seifert, F. M. (2018). GlobBiomass—Global datasets of forest biomass (p. 174 data points) [Text/tab-separated-values]. PANGAEA - Data Publisher for Earth & Environmental Science. https://doi.org/10.1594/PANGAEA.894711

Scheiter, S., Langan, L., & Higgins, S. I. (2013). Next-generation dynamic global vegetation models: Learning from community ecology. New Phytologist, 198(3), 957–969. https://doi.org/10.1111/nph.12210

Schimel, D., Pavlick, R., Fisher, J. B., Asner, G. P., Saatchi, S., Townsend, P., Miller, C., Frankenberg, C., Hibbard, K., & Cox, P. (2015). Observing terrestrial ecosystems and the carbon cycle from space. Global Change Biology, 21(5), 1762–1776. https://doi.org/10.1111/gcb.12822

Schimel, D., Stephens, B. B., & Fisher, J. B. (2015). Effect of increasing CO2 on the terrestrial carbon cycle. Proceedings of the National Academy of Sciences, 112(2), 436–441.

Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G., Wild, J., Ascoli, D., Petr, M., Honkaniemi, J., Lexer, M. J., Trotsiuk, V., Mairota, P., Svoboda, M., Fabrika, M., Nagel, T. A., & Reyer, C. P. O. (2017). Forest disturbances under climate change. Nature Climate Change, 7(6), 395–402. https://doi.org/10.1038/nclimate3303

Sexton, J. O., Noojipady, P., Song, X.-P., Feng, M., Song, D.-X., Kim, D.-H., Anand, A., Huang, C., Channan, S., & Pimm, S. L. (2016). Conservation policy and the measurement of forests. Nature Climate Change, 6(2), 192–196.

Shevliakova, E., Pacala, S. W., Malyshev, S., Hurtt, G. C., Milly, P. C. D., Caspersen, J. P., Sentman, L. T., Fisk, J. P., Wirth, C., & Crevoisier, C. (2009). Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink. Global Biogeochemical Cycles, 23(2). https://doi.org/10.1029/2007GB003176

Shevliakova, E., Stouffer, R. J., Malyshev, S., Krasting, J. P., Hurtt, G. C., & Pacala, S. W. (2013). Historical warming reduced due to enhanced land carbon uptake. Proceedings of the National Academy of Sciences, 110(42), 16730–16735.

Shugart, H. H., & West, D. (1977). Development of an Appalachian deciduous forest succession model and its application to assessment of the impact of the chestnut blight.

Shugart, H. H., Wang, B., Fischer, R., Ma, J., Fang, J., Yan, X., Huth, A., & Armstrong, A. H. (2018). Gap models and their individual-based relatives in the assessment of the consequences of global change. Environmental Research Letters, 13(3), 033001. https://doi.org/10.1088/1748-9326/aaaacc

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. Journal of Geophysical Research: Biogeosciences, 116(G4). https://doi.org/10.1029/2011JG001708

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., & Sykes, M. T. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Global Change Biology, 9(2), 161–185.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C. D., Prentice, I. C., & Woodward, F. I. (2008). Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). Global Change Biology, 14(9), 2015–2039. https://doi.org/10.1111/j.1365-2486.2008.01626.x

Smith, L. J., & Torn, M. S. (2013). Ecological limits to terrestrial biological carbon dioxide removal. Climatic Change, 118(1), 89–103.

Smith, W. B. (2002). Forest inventory and analysis: A national inventory and monitoring program. Environmental Pollution, 116, S233–S242.

Song, X.-P., Huang, C., Feng, M., Sexton, J. O., Channan, S., & Townshend, J. R. (2014). Integrating global land cover products for improved forest cover characterization: An application in North America. International Journal of Digital Earth, 7(9), 709–724.

Spawn, S. A., Sullivan, C. C., Lark, T. J., & Gibbs, H. K. (2020). Harmonized global maps of above and belowground biomass carbon density in the year 2010. Scientific Data, 7(1), 112. https://doi.org/10.1038/s41597-020-0444-4

Stocker, B., Strassmann, K., & Joos, F. (2011). Sensitivity of Holocene atmospheric CO 2 and the modern carbon budget to early human land use: Analyses with a process-based model. Biogeosciences, 8(1), 69–88.

Tang, H., Dubayah, R., Swatantran, A., Hofton, M., Sheldon, S., Clark, D. B., & Blair, B. (2012). Retrieval of vertical LAI profiles over tropical rain forests using waveform lidar at La Selva, Costa Rica. Remote Sensing of Environment, 124, 242–250.

Tang, H., Ma, L., Lister, A., O'Neill-Dunne, J., Lu, J., Lamb, R. L., Dubayah, R., & Hurtt, G. (2021). High-resolution forest carbon mapping for climate mitigation baselines over the RGGI region, USA. Environmental Research Letters, 16(3), 035011. https://doi.org/10.1088/1748-9326/abd2ef

Tharammal, T., Bala, G., Devaraju, N., & Nemani, R. (2019). A review of the major drivers of the terrestrial carbon uptake: Model-based assessments, consensus, and uncertainties. Environmental Research Letters, 14(9), 093005.

Thomas, R. Q., Hurtt, G. C., Dubayah, R., & Schilz, M. H. (2008). Using lidar data and a height-structured ecosystem model to estimate forest carbon stocks and fluxes over mountainous terrain. Canadian Journal of Remote Sensing, 34(sup2), S351–S363. https://doi.org/10.5589/m08-036

Thomas, R. Q., Canham, C. D., Weathers, K. C., & Goodale, C. L. (2010). Increased tree carbon storage in response to nitrogen deposition in the US. Nature Geoscience, 3(1), 13–17.

Thornton, M. M., Thornton, P. E., Wei, Y., Mayer, B. W., Cook, R. B., & Vose, R. S. (2016). Daymet: Monthly Climate Summaries on a 1-km Grid for North America, Version 3. https://doi.org/10.3334/ORNLDAAC/1345

Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. a. G., & Allison, S. D. (2013). Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. Biogeosciences, 10(3), 1717–1736. https://doi.org/10.5194/bg-10-1717-2013

Townsend, A. R., Asner, G. P., & Cleveland, C. C. (2008). The biogeochemical heterogeneity of tropical forests. Trends in Ecology & Evolution, 23(8), 424–431.

Turner, D. P., Guzy, M., Lefsky, M. A., Tuyl, S. V., Sun, O., Daly, C., & Law, B. E. (2003). Effects of land use and fine-scale environmental heterogeneity on net ecosystem production over a temperate coniferous forest landscape. Tellus B: Chemical and Physical Meteorology, 55(2), 657–668. https://doi.org/10.3402/tellusb.v55i2.16756

Urban, D., Bonan, G., Smith, T., & Shugart, H. (1991). Spatial applications of gap models. Forest Ecology and Management, 42(1–2), 95–110.

U.S. Environmental Protection Agency. (2019). State CO2 Emissions from Fossil Fuel Combustion. Web. https://www.epa.gov/statelocalenergy/state-co2-emissions-fossil-fuel-combustion

U.S. Geological Survey, 2019, 3D Elevation Program 1-Meter Resolution Digital Elevation Model (published 20200606), accessed October 23, 2019 at URL https://www.usgs.gov/core-science-systems/ngp/3dep/data-tools

van der Laan-Luijkx, I. T., van der Velde, I. R., van der Veen, E., Tsuruta, A., Stanislawska, K., Babenhauserheide, A., Zhang, H. F., Liu, Y., He, W., Chen, H., Masarie, K. A., Krol, M. C., & Peters, W. (2017). The CarbonTracker Data Assimilation Shell (CTDAS) v1.0: Implementation and global carbon balance 2001–2015. Geoscientific Model Development, 10(7), 2785–2800. https://doi.org/10.5194/gmd-10-2785-2017

Verger, A., Baret, F., & Weiss, M. (2014). Near Real-Time Vegetation Monitoring at Global Scale. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(8), 3473–3481. https://doi.org/10.1109/JSTARS.2014.2328632

Vieira, S., de Camargo, P. B., Selhorst, D., Da Silva, R., Hutyra, L., Chambers, J. Q., Brown, I. F., Higuchi, N., Dos Santos, J., & Wofsy, S. C. (2004). Forest structure and carbon dynamics in Amazonian tropical rain forests. Oecologia, 140(3), 468–479.

Vitasse, Y., Lenz, A., & Körner, C. (2014). The interaction between freezing tolerance and phenology in temperate deciduous trees. Frontiers in Plant Science, 5. https://doi.org/10.3389/fpls.2014.00541

Von Caemmerer, Susanne, & Furbank, R. T. (1999). Modeling C4 photosynthesis. C4 Plant Biology, 173–211.

Von Caemmerer, Susanna. (2000). Biochemical models of leaf photosynthesis. Csiro publishing.

von Caemmerer, S., Farquhar, G., & Berry, J. (2009). Biochemical Model of C3 Photosynthesis. In A. Laisk, L. Nedbal, & Govindjee (Eds.), Photosynthesis in silico: Understanding Complexity from Molecules to Ecosystems (pp. 209–230). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9237-4_9

Walker, A. P., De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Iversen, C. M., Asao, S., Guenet, B., Harper, A., Hickler, T., Hungate, B. A., Jain, A. K., Luo, Y., Lu, X., Lu, M., Luus, K., Megonigal, J. P., Oren, R., Ryan, E., Shu, S., … Norby, R. J. (2019).

Decadal biomass increment in early secondary succession woody ecosystems is increased by CO2 enrichment. Nature Communications, 10(1), 454. https://doi.org/10.1038/s41467-019-08348-1

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M., Hoegh-Guldberg, O., & Bairlein, F. (2002). Ecological responses to recent climate change. Nature, 416(6879), 389–395. https://doi.org/10.1038/416389a

Wei, Y., Liu, S., Huntzinger, D. N., Michalak, A. M., Viovy, N., Post, W. M., Schwalm, C. R., Schaefer, K., Jacobson, A. R., & Lu, C. (2014). The North American carbon program multi-scale synthesis and terrestrial model intercomparison project–Part 2: Environmental driver data. Geoscientific Model Development, 7(6), 2875–2893.

Westerling, A. L., Hidalgo, H. G., Cayan, D. R., & Swetnam, T. W. (2006). Warming and Earlier Spring Increase Western U.S. Forest Wildfire Activity. Science, 313(5789), 940. https://doi.org/10.1126/science.1128834

White, M. A., Running, S. W., & Thornton, P. E. (1999). The impact of growing-season length variability on carbon assimilation and evapotranspiration over 88 years in the eastern US deciduous forest. International Journal of Biometeorology, 42(3), 139–145. https://doi.org/10.1007/s004840050097

Wieder, W. R., Boehnert, J., Bonan, G. B., & Langseth, M. (2014). Regridded Harmonized World Soil Database v1.2. ORNL DAAC. https://doi.org/10.3334/ORNLDAAC/1247

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., & Kattge, J. (2009). Improving land surface models with FLUXNET data. Biogeosciences, 6(7), 1341–1359.

Williams, C. A., Collatz, G. J., Masek, J., & Goward, S. N. (2012). Carbon consequences of forest disturbance and recovery across the conterminous United States. Global Biogeochemical Cycles, 26(1).

Wilson, B. T., Woodall, C. W., & Griffith, D. M. (2013). Imputing forest carbon stock estimates from inventory plots to a nationally continuous coverage. Carbon Balance and Management, 8(1), 1. https://doi.org/10.1186/1750-0680-8-1

Xian, G. Z., Homer, C. G., Dewitz, J., Fry, J., Hossain, N., & Wickham, J. (2011). Change of impervious surface area between 2001 and 2006 in the conterminous United States. Photogrammetric Engineering and Remote Sensing, 77(8), 758–762. USGS Publications Warehouse.

Xiao, J., Chevallier, F., Gomez, C., Guanter, L., Hicke, J. A., Huete, A. R., Ichii, K., Ni, W., Pang, Y., & Rahman, A. F. (2019). Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. Remote Sensing of Environment, 233, 111383.

Xu, X., Medvigy, D., Powers, J. S., Becknell, J. M., & Guan, K. (2016). Diversity in plant hydraulic traits explains seasonal and inter-annual variations of vegetation dynamics in seasonally dry tropical forests. New Phytologist, 212(1), 80–95. https://doi.org/10.1111/nph.14009

Yang, X., Tang, J., Mustard, J. F., Lee, J.-E., Rossini, M., Joiner, J., Munger, J. W., Kornfeld, A., & Richardson, A. D. (2015). Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest. Geophysical Research Letters, 42(8), 2977–2987. https://doi.org/10.1002/2015GL063201

Zhang, K., de Almeida Castanho, A. D., Galbraith, D. R., Moghim, S., Levine, N. M., Bras, R. L., Coe, M. T., Costa, M. H., Malhi, Y., Longo, M., Knox, R. G., McKnight, S., Wang, J., & Moorcroft, P. R. (2015). The fate of Amazonian ecosystems over the coming century arising from changes in climate, atmospheric CO2, and land use. Global Change Biology, 21(7), 2569–2587. https://doi.org/10.1111/gcb.12903

Zhang, Y., Joiner, J., Alemohammad, S. H., Zhou, S., & Gentine, P. (2018). A global spatially contiguous solar-induced fluorescence (CSIF) dataset using neural networks. Biogeosciences, 15(19), 5779–5800.

Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., Ciais, P., Sitch, S., Friedlingstein, P., Arneth, A., Cao, C., Cheng, L., Kato, E., Koven, C., Li, Y., Lian, X., Liu, Y., Liu, R., Mao, J., … Zeng, N. (2016). Greening of the Earth and its drivers. Nature Climate Change, 6(8), 791–795. https://doi.org/10.1038/nclimate3004

Zomer, R. J., Neufeldt, H., Xu, J., Ahrends, A., Bossio, D., Trabucco, A., van Noordwijk, M., & Wang, M. (2016). Global Tree Cover and Biomass Carbon on Agricultural Land: The contribution of agroforestry to global and national carbon budgets. Scientific Reports, 6(1), 29987. https://doi.org/10.1038/srep29987