



New Applications of Forest Models in Combination with Remote Sensing

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt
Universität Osnabrück
Fachbereich Mathematik/Informatik

von Hans Henniger
geboren am 09.11.1989 in Wernigerode

September 5, 2023

Bibliographische Beschreibung

Hans Henniger

New Applications of Forest Models in combination with Remote Sensing

Universität Osnabrück, Fachbereich Mathematik/Informatik

Dissertation (in englischer Sprache), 2023

166 Seiten, 60 Abbildungen, 11 Tabellen, 290 Literaturzitate

Zusammenfassung

Wälder bedecken 31% der globalen Landoberfläche. Sie spielen eine wichtige Rolle für den globalen Kohlenstoffkreislauf, da sie als Kohlenstoffspeicher fungieren und zum globalen Kohlenstoffkreislauf beitragen. Waldökosysteme sind Umweltveränderungen ausgesetzt, die auf die Klimaerwärmung und die damit verbundenen Veränderungsprozesse wie Dürren, Hitzewellen, Brände, Stürme oder Schädlingsbefall zurückzuführen sind, aber auch auf Entwaldung und Fragmentierung. Diese Veränderungsprozesse finden schneller, häufiger und immer gleichzeitiger statt.

Waldmodelle tragen zum Verständnis der Wälder und der Dynamik der Ökosysteme unter sich ändernden Umweltbedingungen bei. Mit der zunehmenden Verfügbarkeit von Fernerkundungsdaten und der steigenden Rechenleistung ergeben sich neue Möglichkeiten für die Anwendung von Waldmodellen. Dies bedeutet auch, dass Waldmodelle angepasst, erweitert oder auf neue Weise angewandt werden müssen, um diese neuen Bedingungen voll ausnutzen zu können.

In dieser Arbeit werden innovative und neuartige Anwendungen von Waldmodellen vorgestellt, die dazu beitragen könnten, von diesen neuen Möglichkeiten zu profitieren. Die erste Studie dieser Arbeit (Kapitel 2) etabliert eine neue Art der Nutzung von Waldmodellen, indem sie den Forest Factory Ansatz erweitert und ihn für Wälder in verschiedenen Biomen anwendbar macht. Dieser Ansatz erlaubt es, Wälder unter Verwendung der Architektur und der Prozesse von Waldmodellen zu generieren (hier verwenden wir das individualbasierte gap modell FORMIND). In dieser Studie werden 700.000 Waldbestände in sieben verschiedenen Ökoregionen mit Hilfe der Forest Factory 2.0 generiert. Im Gegensatz zur traditionellen Untersuchung der Entwicklung einzelner Waldbestände über die Zeit, haben wir die Forest Factory 2.0 als Werkzeug eingesetzt, um durch die Analyse des Zustandsraums der Wälder Wissen über diese zu gewinnen. Wir haben eine strukturelle Sensitivitätsanalyse durchgeführt, um die Beziehungen zwischen den strukturellen Eigenschaften und der Biomasse, der Produktivität sowie der species evenness von Wäldern zu vergleichen. Wir analysieren den Zustandsraum von Wäldern in verschiedenen Biomen und demonstrieren das Potenzial dieses Ansatzes für die theoretische Ökologie. Mit der Forest Factory 2.0 können Forscher*innen virtuelle Wälder für ihre Bedürfnisse erzeugen oder die Open-Source-Walddaten nutzen, um ein digitales Walduniversum von Waldzuständen zu analysieren.

Die zweite Studie in dieser Arbeit (Kapitel 3) gibt Einblicke, wie Fernerkundungsmessungen in Waldmodelle integriert werden können. Es handelt sich um einen neuen Ansatz, der die Berechnung der hyperspektralen Strahlungsreflexion von Wäldern ermöglicht. Er verwendet das mehrschichtige Strahlungstransfermodell mScope und das individualen basierte Waldmodell FORMIND. Diese Arbeit bietet einen forward model-

ling Ansatz, um die Waldreflexion mit den Waldeigenschaften in Beziehung zu setzen. Mit diesem Werkzeug ist es möglich, eine große Anzahl von Waldbeständen und ihre entsprechende reflektierte Strahlung (im sichtbaren und nahen Infrarotbereich) zu analysieren. Dies eröffnet die Möglichkeit zu verstehen, wie Waldreflexion mit Sukzession und anderen Waldbedingungen zusammenhängt.

Um die Vorteile der zunehmenden Zahl von Fernerkundungsmessungen zu nutzen und Synergieeffekte mit Waldmodellen zu erzielen, wäre es sinnvoll, das Design von Satellitenmissionen mit den Fähigkeiten von Waldmodellen abzustimmen. Die letzte Studie dieser Arbeit (Kapitel 4) befasst sich mit der Verwendung der erwarteten Biomasseverteilungen, die von der bevorstehenden RADAR BIOMASS P-Band-Satellitenmission (Start im Jahr 2024 durch die Europäische Weltraumorganisation) bereitgestellt werden, um die Produktivität tropischer Wälder vorherzusagen. Die Ergebnisse zeigen eine hohe Korrelation zwischen der Schätzung der Produktivität und der Biomasseverteilung bei einer räumlichen Auflösung von 4 ha und 1 ha. Eine höhere vertikale Auflösung führt im Allgemeinen zu besseren Vorhersagen für die Produktivität (GPP, NPP). Außerdem zeigen die Ergebnisse den Einfluss der räumlichen Auflösung mit Unterschieden zwischen gestörten und ausgewachsenen Wäldern. Der vorgestellte Ansatz bietet eine Reihe von Innovationen: (i) die Verwendung von erwarteten Fernerkundungsmessungen, (ii) die Verwendung von individuenbasierten Waldmodellen für Vorstudien zu Satellitenmissionen, (iii) die Verwendung von RADAR-Satellitenmessungen für die Vorhersage von Produktivität und carbon turnover times und (iv) die Untersuchung der Vorhersagequalität für verschiedene Waldtypen. Die erzielten Ergebnisse unterstreichen den Wert der bevorstehenden BIOMASS-Satellitenmission und verdeutlichen das Potenzial der Ableitung von Waldproduktivität aus Informationen über die Waldstruktur.

Die in dieser Arbeit vorgestellten Studien bieten eine Grundlage für künftige Anwendungen. Darüber hinaus zeigen sie Beispiele für erste Anwendungen, wie die neu entwickelten und neuartigen Methoden eingesetzt werden können, um die Beziehung zwischen Waldstruktur und Produktivität zu untersuchen (Kapitel 2 und 4), um virtuelle Fernerkundungsmessungen für eine Fallstudie in Finnland zu untersuchen (Kapitel 3) und schließlich um die Produktivität für typische tropische Wälder mit Hilfe von Radar-Fernerkundungsdaten zu schätzen (Kapitel 4). Die vorgestellten Anwendungen können letztlich zu einem besseren Verständnis von Waldökosystemen beitragen.

Schlagworte

Waldmodell, Strahlungstransfer, Vegetationsindizes, individuenbasiert, Waldreflexion, machine learning, Fernerkundung, RADAR, Waldgenerator, Ökosystemfunktionen, Produktivität, Waldbiomasse, forest factory

Bibliographic description

Hans Henniger

New Applications of Forest Models in combination with Remote Sensing

phd thesis (in English), 2023

166 pages, 60 figures, 11 tables, 290 references

Abstract

Forests cover 31% of the global land surface. They play a major role for the global carbon cycle because of their role as carbon storage and their contributions to global carbon fluxes. Forest ecosystems are exposed to environmental changes due to climate warming and related change processes as droughts, heat waves, fires, storms or pest outbreaks, but also deforestation and fragmentation which accelerating and occurring more and more simultaneously.

Forest models contribute to the understanding of forests and the dynamics of ecosystems under changing environmental conditions. With the increasing availability of remote sensing data and increasing computing power, new opportunities are emerging for the application of forest models. This also means that forest models need to be adapted, extended or applied in new ways to take full advantage of these new conditions.

This thesis presents innovative and novel applications of forest models, which could help to profit from such new opportunities. The first study of this thesis (Chapter 2) establish a new way of using forest models by extending the forest factory approach by and make it applicable for forests in different biomes. This approach allows to generate forests using the architecture and processes of forest models (here we use the individual-based gap model FORMIND). In this study, 700,000 forest stands in seven different ecoregions are generated by using the Forest Factory 2.0. In contrast to the tradition of investigating the development of individual forest stands over time, we used the Forest Factory 2.0 as a tool to gain knowledge about forests by analyzing the state space of forests. We conducted a structural sensitivity analysis to compare the relationships between structural properties and biomass, productivity, as well as (tree) species evenness of forests. We analyze the state space of forests in different biomes and demonstrate the potential of this approach for theoretical ecology. With the Forest Factory 2.0, researchers can generate virtual forests for their needs or use the open-source forest data to analyze a digital forest universe of forest states.

The second study in this thesis (Chapter 3) provides insights into how remote sensing measurements can be incorporated in forest models. It is about a new approach which enable the calculation of hyperspectral reflectance of forests. It uses the multi-layer radiative transfer model mScope and the individual-based forest model FORMIND. This work provides a forward modeling approach for relating forest reflectance to forest characteristics. With this tool, it is possible to analyze a large set of forest stands and their corresponding reflected radiance (in the visible and near infrared range). This opens up the possibility to understand how forest reflectance is related to succession and different forest conditions.

In order to take advantage of the increasing number of remote sensing measurements and to achieve synergy effects with forest models, it would be useful to align the de-

sign of satellite missions with the capabilities of forest models. The last study of the thesis (Chapter 4) is about using expected biomass distributions provided by the upcoming RADAR BIOMASS P-band satellite mission (launching in 2024 by the European Space Agency) to predict the productivity of tropical forests. The results show a high correlation for estimating productivity with a biomass distribution at a spatial resolution of 4 ha and 1 ha. Increased vertical resolution leads generally to better predictions for productivity (GPP, NPP). Further, the results demonstrate the influence of spatial resolution with differences between disturbed and mature forests. The presented approach offers a number of innovations: (i) the use of expected remote sensing measurements, (ii) the use of individual-based forest models for preliminary studies of satellite missions, (iii) the use of RADAR satellite measurements for the prediction of productivity and carbon turnover times and (iv) the exploration of the prediction quality for different forest types. The obtained results emphasize the value of the forthcoming BIOMASS satellite mission and highlight the potential of deriving estimates for forest productivity from information on forest structure.

The studies presented in the thesis are providing a basis for future applications. In addition, they show first applications of how these newly developed and novel methods can be used to investigate the relationship between forest structure and productivity (Chapter 2 and 4), to explore virtual remote sensing measurements for a case study in Finland (Chapter 3), and finally to estimate productivity for typical tropical forests using RADAR remote sensing data (Chapter 4). The applications presented may ultimately contribute to a better understanding of forest ecosystems.

Key words

forest model, radiative transfer, vegetation indices, individual-based, forest reflectance, machine learning, remote sensing, RADAR, forest generator, ecosystem functions, productivity, forest biomass, forest factory

Danksagung

Zu allererst möchte ich mich bei meinen Betreuern Andreas Huth und Friedrich Bohn bedanken. Ich denke, nur wenige Doktorand*innen können sich über eine so enge und stetige Betreuung freuen, ohne die es nicht möglich gewesen wäre, in der Zeitspanne die Arbeit zu schreiben und meine Motivation aufrecht zu erhalten. Auch wenn es manchmal sehr intensiv war, haben wir es immer geschafft konstruktiv, ehrlich und respektvoll zu sein, wodurch viele tolle Ideen in Gemeinschaft entstanden sind und wir voneinander lernen konnten.

Zusätzlich möchte ich dem FORMIND-Team und den Menschen im Department für Ökosystemanalyse danken. Anregende Diskussionen, morgendliche Pläusche, Mittagessen in der Sonne, Veröffentlichungssekt und letztendlich den Anfangsimpuls und die Möglichkeit diese Arbeit zu schreiben, verdanke ich euch - vor allem Karin, Friedrich und Andreas. Bewunderung für Ricos stets liebenswerte sowie verständnisvolle Art und auch für Samuels und Michaels Hilfsbereitschaft.

Ich möchte mich vor allem auch bei meinen Freundinnen und meiner WG bedanken. Lauren, Ebba, Rebecca, Lena, Lisi, Henrike, Alex, Lori, Hanna, Zora, Lena, Marie und Sophia - ihr habt euch viele meiner Strapazen angehört und mich immer unterstützt. Unzählige Bibtage wart ihr an meiner Seite, Putzplanvernachlässigung und Telefonunlust habt ihr mir nachgesehen. Ich danke euch für eure Nähe und Fürsorge.

Contents

List of Figures	XI
List of Tables	XV
List of Abbreviations	XVI
1 Introduction	1
1.1 Overview	1
1.2 History of forest modelling	2
1.3 Remote Sensing	5
1.4 Applications of forest models	8
1.5 Objectives of this thesis	11
2 Creating virtual forests around the globe and analysing their state space, Henniger et al. 2023	15
2.1 Abstract	15
2.2 Introduction	15
2.3 Methodical concept	18
2.4 Results	25
2.5 Discussion	31
3 A New Approach Combining a Multilayer Radiative Transfer Model with an Individual-Based Forest Model: Application to Boreal Forests in Finland, Henniger et al. 2023	36
3.1 Abstract	36
3.2 Introduction	36
3.3 Materials and Methods	39
3.4 Results	46
3.5 Discussion	49
3.6 Conclusions	51
4 A new Approach to derive Productivity of Tropical Forests using Radar Remote Sensing Measurement	53
4.1 Abstract	53
4.2 Introduction	53
4.3 Methods	55
4.4 Results	58
4.5 Discussion	64

4.6 Conclusion	69
5 Conclusion	70
5.1 Main results, limitations and potentials	70
5.2 Outlook	78
References	83
A Appendix Chapter 2	111
B Appendix Chapter 3	122
C Appendix Chapter 4	137

List of Figures

1	Main sensor types used by environmental satellites	6
2	Interdependencies of FORMIND, Forest Factory 2.0 and the R package	19
3	Concept of the Forest Factory 2.0	20
4	Analysis of the state space of generated forests	26
5	Relationship between structural properties (basal area, maximum height) and functional characteristics	28
6	Relationship between structural properties (LAI, height heterogeneity) and functional characteristics	29
7	Comparison of biomass, aboveground productivity and species evenness derived from forests with a similar state space	31
8	Visualization of the forest inventory of the 28 forest stands in Finland	40
9	Different concepts of forest representation	44
10	Reflectance spectra for detailed and simplified forest representation by using layers with a size of 10 m	47
11	Reflectance spectra for detailed and simplified forest representation by using standard layers with a size of 0.5 m	47
12	Comparison of vegetation indices	48
13	Simulated forest dynamics over time	56
14	Example for the assumed distribution of biomass over height	57
15	Relationship between biomass and GPP, NPP and τ (carbon turnover time)	59
16	Relationship between the biomass in the height layer 40 – 60 m and GPP, NPP, carbon turnover time and total biomass over time	60
17	Comparison of predicted NPP derived from the boosted regression tree and reference NPP derived from a forest model	61
18	Comparison of predicted GPP derived from the boosted regression tree and reference GPP derived from a forest model	62
19	Comparison of correlation between the estimated and the reference GPP, NPP and carbon turnover time	63
20	Comparison of correlation between the estimated and the reference GPP, NPP and carbon turnover time for different types of forests	64
A.1	Comparison of Forest Factory by Bohn & Huth 2017 and Forest Factory 2.0	113
A.2	Relationship between biomass and above-ground productivity for analyzed ecoregions	114

A.3	Relationship between structural properties and above-ground wood productivity1	114
A.4	Relationship between structural properties and above-ground wood productivity2	115
A.5	Relationship between structural properties and functional characteristics of all forest stands1	115
A.6	Maximum values of functional characteristics of forest stands	116
A.7	Standard deviation of functional characteristics of forest stands	116
A.8	Relationship between structural properties and functional characteristics of all forest stands2	117
A.9	Maximum values of functional characteristics of forest stands	117
A.10	Standard deviation of functional characteristics of forest stands	118
A.11	Comparison of biomass from different regions derived from forests with a similar state space	119
A.12	Comparison of AWP from different regions derived from forests with a similar state space	120
A.13	Comparison of AWP from different regions derived from forests with a similar state space	121
B.1	Study map of the 28 forest stands in the region North Karelia	122
B.2	Soil reflection	123
B.3	Analysis of allometries in FORMIND	123
B.4	Concept of height layers	123
B.5	LAI Profiles for each forest stand	125
B.6	Comparison of simulated reflectance spectra with Sentinel measurements assuming a simple forest representation using different descriptions of the vertical forest structure	126
B.7	Comparison of simulated reflectance spectra with Sentinel measurements assuming a detailed forest representation using different descriptions of the vertical forest structure	127
B.8	Comparison of simulated reflectance spectra with Sentinel measurements assuming a spectra-averaged forest representation using different descriptions of the vertical forest structure	127
B.9	Comparison of simulated reflectance spectra with Sentinel measurements assuming different types of forest representations and using 10 m height layers	128

B.10 Comparison of simulated reflectance spectra with Sentinel measurements assuming different types of forest representations and using 0.5 m height layers	128
B.11 Comparison of simulated reflectance spectra assuming different types of forest representations and using 0.5 m height layers	129
B.12 Comparison of simulated reflectance spectra and measured reflectance spectra using a distance index	129
B.13 Comparison of simulated and measured reflectance of 28 forest stands and different forest representations	130
B.14 Comparison of reflectance spectra and additional information of forest stand 15 (classified as outlier) with forest stand 5	131
B.15 Comparison of reflectance spectra and additional information of forest stands 7 and 8 (classified as outliers) with forest stand 11	132
B.16 Comparison of reflectance spectra and additional information of forest stand 19 (classified as outlier) with forest stand 22	132
B.17 Comparison of reflectance spectra and additional information of forest stand 18 (classified as outlier) with forest stand 23	133
B.18 Relationship between LAI (field data) and NDVI (Sentinel-2 measurements) of 28 Finland forest stands	134
B.19 Comparison of vegetation indices for 28 forest stands in Finland . . .	135
C.1 Comparison of predicted NPP derived from the boosted regression tree and reference NPP derived from simulations using the forest model FORMIND	137
C.2 Comparison of predicted GPP derived from the boosted regression tree and reference GPP derived from simulations using the forest model FORMIND	138
C.3 Comparison of predicted carbon turnover time (τ) derived from the boosted regression tree and reference τ derived from simulations using the forest model FORMIND	139
C.4 Comparison of predicted NPP derived from the boosted regression tree and reference NPP derived from simulations using the forest model FORMIND	140
C.5 Comparison of predicted GPP derived from the boosted regression tree and reference GPP derived from simulations using the forest model FORMIND	141

C.6	Comparison of predicted carbon turnover time (τ) derived from the boosted regression tree and reference τ derived from simulations using the forest model FORMIND	142
C.7	Comparison of correlation between the estimated and the reference GPP, NPP and carbon turnover time (τ)	143
C.8	Comparison of correlation between the estimated and the reference GPP, NPP and carbon turnover time (τ) for different types of forests .	143

List of Tables

1	Properties of selected global open data satellite missions	8
2	Improvements of the Forest Factory 2.0 in comparison to the Forest Factory (Bohn and Huth 2017)	22
3	Overview of the ecoregions, parameterizations and climate used for this study	24
A.1	Description of the Data product of 700,000 forests stands from 7 different ecoregions	112
B.1	Leaf parameters	122
B.2	Attributes of the forest stands from Finland used for this study	124
B.3	Spectral configuration of the 10 Sentinel-2A bands used in this study .	125
B.4	Analysis of vegetation indices for different forest representations . . .	136
C.1	Parameters of the derived Boosted regression trees (BRTs) for the whole forest data (0 - 320 years of simulation)	144
C.2	Parameters of the derived Boosted regression trees (BRTs) for disturbed forests (0 - 160 years of simulation)	145
C.3	Parameters of the derived Boosted regression trees (BRTs) for mature forests (160 - 320 years of simulation)	146

List of Abbreviations

- AWP** Above-ground Wood Productivity
- BRT** Boosted Regression Tree
- EnMAP** Environmental Mapping and Analysis Program
- EVI** Enhanced Vegetation Index
- GPP** Gross Primary Production
- DGVM** Dynamic Global Vegetation Model
- IBM** Individual Based Model
- RTM** Radiative Transfer Model
- kNDVI** kernel Normalized Difference Vegetation Index
- LAI** Leaf Area Index
- LiDAR** Light Detection And Ranging
- LPJ** Lund-Jena-Potsdam Model
- MSI** Moisture Stress Index
- NDMI** Normalized Difference Moisture Index
- NDVI** Normalized Difference Vegetation Index
- NPP** Net Primary Productivity
- RADAR** Radio Detection And Ranging
- SIF** Solar-induced Chlorophyll Fluorescence

1 Introduction

1.1 Overview

Forests cover 31% of the global land surface (FAO 2020; Keenan et al. 2015). They play a major role for the global carbon cycle because of their function as carbon storage and their contributions to global carbon fluxes (Bonan 2008; Grace et al. 2014). Forests are important for sustaining biodiversity and provide habitat for 70% of all terrestrial animal species (Gibson et al. 2011; Myers et al. 2000; Pimm et al. 2014). Further, forests exhibit a diversity of spatial structures and change their structure due to natural succession, management or disturbances (Pan et al. 2013).

While forest ecosystems are exposed to environmental change, like all complex adaptive systems, they have a certain capacity to cope with it. However, if these change processes occur too frequently, on too large spatial scales or/and with too high intensity, the adaptive capacity of the forests may be exceeded. Global change processes such as climate change and associated impacts such as droughts, heat waves, fires, storms or pest outbreaks, but also deforestation (Stocker 2014) and fragmentation (FAO 2022; R. Fischer 2021; Taubert et al. 2018) are accelerating and occurring more and more simultaneously. As a result, the dynamics of the forests would change as well as their tree species composition and structure. Therefore, forests appear to be under increasing pressure (McDowell et al. 2020), affecting forest biodiversity in general as well as the diversity of functions provided by forests. This shows the urgency of sustaining their functioning, understanding and enhancing their adaptive capacity, and appropriately adjusting their management.

Determining the impacts of changing drivers on forests development is difficult (McDowell et al. 2020). However, there is evidence from individual published studies of drivers and their effects on plant communities, and new modeling and observational efforts are now providing a more complete picture of disturbance and forest demography (Hartmann et al. 2018; Pugh et al. 2019; Stovall et al. 2019). Prerequisite for addressing these challenges, however, is a sound understanding of structure-function relationships, esp. between the properties of forests (species-compositional and spatial-structural) and their functions (e.g. carbon flux and storage). Unfortunately, there is no universal relationship between stand structural characteristics and forest function; rather, this relationship is dependent on environmental conditions, biotic interactions, stand age, and disturbance intensity within a given forest ecosystem (Ali 2019; Ali & Mattsson 2017; Paquette & Messier 2011; van der Sande et al. 2017; Yuan et al. 2018; Y. Zhang & Chen 2015).

Filling knowledge gaps (Courchamp et al. 2015; Franklin et al. 2020) is one goal for

modern ecological science, where the perspectives of community ecology (organismal aspects, diversity of species and structure) and ecosystem ecology (matter and energy flux aspects, biogeochemical cycles) are inseparable. A mechanistic understanding of the functioning of ecosystems can only be gained if these two perspectives are adequately linked to one another (Loreau 2010). However, although there are many empirical studies based on forest inventories, the number of available samples and plots or the lack of focus in monitoring makes it difficult to get a sufficiently complete picture (Lindenmayer & Likens 2009; Lindenmayer et al. 2011). It remains particularly demanding to answer questions that deal with continental to global scales, or that relate to short time periods and thus require recent data (for current applications). In addition, there is a wide variety of environmental factors that may influence forest characteristics (species composition and structure) and structure-function relationships. It is a challenge to capture the inherent spatial heterogeneity of environmental conditions and how they change in response to projected changing processes, especially among different biomes. Many of the variables relevant to exploring these relationships are often not fully covered by existing inventories and measurements.

Parts of the mentioned challenges may be overcome by using remote sensing data, but relating and condensing these often large-scaled data to the local or individual scale remains difficult (Ma et al. 2020). However, an increasing amount of field and remote sensing measurements of various spatial and temporal extents and resolutions are available. It remains demanding to integrate these data into a coherent picture (Chave 2013; Estes et al. 2018; Levin 1992; Maréchaux et al. 2021) and forest models can help to gain understanding of relationships between changing environment and forest development as well as to bridge between different spatial-temporal scales. A collaborative effort is needed in which forest models, remote sensing measurements and new statistical methods must be intertwined. The presented thesis is a contribution to this collaboration.

1.2 History of forest modelling

Forest modelling, in particular forest gap models and global vegetation models, has been an important part of ecological research since the 1970s and has been the subject of a number of reviews in recent decades (e.g., Bugmann 2001; Bugmann et al. 1996; Dale et al. 1985; Fisher et al. 2018; Liu & Ashton 1995; Maréchaux et al. 2021; Peng 2000; Quillet et al. 2010; Shugart 1998; Shugart et al. 1984; Shugart & Smith 1996; Shugart & West 1980; Urban & Shugart 1992). In this part of the introduction I follow the argumentation of Bugmann 2001; Maréchaux et al. 2021; Peng 2000; Shugart & Smith 1996 and Quillet et al. 2010.

For a long time, there has been a high level of fascination among scientists and foresters for describing, understanding and predicting long-term dynamics of forest ecosystems (Clements 1916; Gleason 1926; Trudgill 2007; Watt 1923; Whittaker 1953). The first predictions about the behavior of the forest come from forestry. In order to have a steady supply of timber, foresters need to know how many trees can be harvested. These thinning rates to maintain forest yield are related to the growth of trees in a stand, and subsequently to the growth of an individual tree. Empirical observations of thinning and growth of a large collection of standard trees in forest stands can be condensed into these rates through statistical analysis. This upscaling from the growth of a standard tree has been and continues to be used in empirical forestry in so called yield tables, the bulwark of forecasting for foresters (Shugart et al. 2018). Nevertheless, it was difficult to develop methods and quantitative theories for forest dynamics because of insufficient data availability for sometimes long-lasting processes, environmental influences, interactions, and the lack of tools to synthesize the data.

In the 1960s and 1970s, the advance of digital computer technology enabled scientists to develop novel mathematical models of forest ecosystems (Newnham & Smith 1964; Siccama et al. 1969). As a result, the first forest models capable of representing small forest patches were developed. They were designed to reproduce observations and results from small forests observed in nature. Only data for small forest patches were available and a certain homogeneity of the forest was necessary to formulate the model. The first models of this type were primarily designed to provide a quantitative method for forecasting, mainly using information from yield tables and other data sources. A more complex type of model is the individual-based forest model, in which establishment, mortality, and growth of each individual tree is simulated as a function of abiotic (climate and soil) and biotic (competition) factors. The first such model was the JABOWA model (Botkin et al. 1970, 1972, 1973), which was developed for forests in New England and which served as a parent model for many other individual forest models later. The JABOWA model used the gap approach, which is a widely used approach in combination with individual-based models (Shugart & West 1980). It simulates the development of a forest on a large number of small independent forest patches, each patch having the size of a large canopy (e.g. 20 m x 20 m).

By the 2000s, a wide range of gap models had been developed to investigate boreal forests to tropical forests and even extended to grasslands (Coffin & Lauenroth 1989, 1990; Taubert et al. 2012), alpine tundra (Humphries et al. 1996), and savannas (Menant et al. 1990). This type of model, have since become one of the most widely used model types in ecology (Botkin et al. 1972; Shugart et al. 2018; Shugart & WEST 1977). By 2016, there were a combined 6.500 publications citing 12 different early

forest gap models (Shugart et al. 2018).

One successful applied forest model is the individual-based gap model FORMIND (Köhler & Huth 1998), which has first be used to simulate tropical forests (R. Fischer et al. 2016; Rödig et al. 2017) but is also used for forests in temperate regions (Bohn et al. 2014) as well as grasslands (Taubert et al. 2012). It has been in use continuously for 30 years in 28 locations around the globe. The model considers four main process groups: growth of single trees, mortality, recruitment, and competition (e.g. for light and space). It is working on a spatial resolution of 20 m x 20 m and normally on a yearly time step (also daily time step is used like in Holtmann et al. 2021). Unlike other gap models, it is able to use plant functional types or specific tree species, and it includes gaps caused by tree fall (which is an important disturbance regime especially in unmanaged forests like those in the tropics, Robert 2003). FORMIND is also used for large scale simulations (Paulick et al. 2017; Rödig et al. 2019) e.g. forest-wide carbon balances in the Amazon.

Some phenomena cannot be studied in small forest areas, such as competition between different vegetation types, so other models are needed. Such type of forest models are global vegetation models and later dynamic global vegetation models (DGVMs, Friend et al. 1997; Smith 2001) which were developed in the 1990s, with links to forest gap models in terms of fundamental model assumptions (Bugmann & Seidl 2022). DGVMs have their origins in four research fields: plant geography, biogeochemistry, vegetation dynamics, and biophysics (Prentice et al. 2007). They originally focused on bio-geochemical processes as the exchange of carbon and water between vegetation and atmosphere at the global scale (e.g. Bonan et al. 2003; Krinner et al. 2005) on a low spatial (e.g. 0.5° longitude/latitude in the Biome Mode by Prentice et al. 1992) and high temporal resolution (hourly to daily). This comes with the cost of a realistic representation of forest diversity, competition, and structure (Feeley et al. 2007; Hurtt et al. 1998; Moorcroft et al. 2001). This means that these models were originally developed to simulate competition between different types of vegetation and not between individual trees, which led to the concept of using pools (e.g. for biomass) to summarize and represent vegetation. In addition, DGVMs use plant functional types to aggregate functionally similar species to represent functional properties at the biome scale.

One of the first DGVM is the Lund-Potsdam-Jena Model (Sitch et al. 2003; Smith 2001) which incorporates process-based, large-scale representations of terrestrial vegetation dynamics and land-atmosphere carbon and water exchanges in a modular framework. LPJ explicitly considers key ecosystem processes such as photosynthesis, carbon allocation, mortality, resource competition, fire disturbance and soil heterotrophic respiration in a modelled area/grid cell of 100 – 2500 km². It includes competition be-

tween vegetation in different grid cells, but not competition between individual species or individual trees within a grid cell. LPJ is constantly used and developed over the last 20 years. It has been applied e.g. to simulate global carbon and CH₄ emissions and uptake by different ecosystems (Prentice et al. 2000; Spahni et al. 2011; Wania et al. 2010), to model dynamic interactions and feedbacks between the terrestrial biosphere and the water cycle (Gerten et al. 2004), to investigate the role of fire disturbance for global vegetation dynamics (Pfeiffer et al. 2013; Thonicke et al. 2001) or to investigate the effects of soil freezing and thawing on vegetation carbon density (Beer et al. 2007). Looking back at the history of forest modelling, different types of models have been developed for different purposes, such as for regional analysis or for the study of global trends. However, it is not only the purpose that has driven the different developments, but also limitations such as data availability or computational costs have led to the need for new types of models. Due to the relatedness of individual-based models and dynamic global vegetation models the boundaries between them are blurred. An example for that is LPJ-GUESS, a combination of LPJ with the general ecosystem simulator (Guess). It is able to simulate individuals and cohorts of PFTs (Plant functional types) in a patch of 0.1 ha (Smith 2001). It is capable of simulating competition between individual trees or cohorts, but differs from individual-based forest models in that the simulation is done in such detail only for a sample of patches and then extrapolated to the entire simulation area. Previous regional applications of LPJ-GUESS explored vegetation dynamics and plant biogeography in forests in parts of the U.S. (Emmett et al. 2021; Hickler et al. 2004; Tang et al. 2012) and Europe (Hickler et al. 2012; Koca et al. 2006; Morales et al. 2007; Smith 2001; Smith et al. 2008). Another example for implementing interaction between individuals is the development of individual-based global dynamic vegetation models such as SEIB-DGVM (Sato et al. 2007). At the same time, the applications of individual-based models are expanding from local to regional scales (e.g. entire Amazon rainforest, Rödiger et al. 2019).

1.3 Remote Sensing

To monitor the condition of forests, foresters and ecologists have long relied on measuring forest inventories. Collecting inventories is time-consuming. In addition, for e.g., in tropical forests, national forest inventories are often missing. Due to the interaction of electromagnetic radiation with different components of forests (e.g. chemical composition of leaves, leaves, branches and trunks) there is another approach to monitor forests. It is based on remote sensing observations, which provide relevant data for large areas. Different techniques are used for remote sensing applications (Figure 1),

resulting in a variety of capabilities and satellite products.

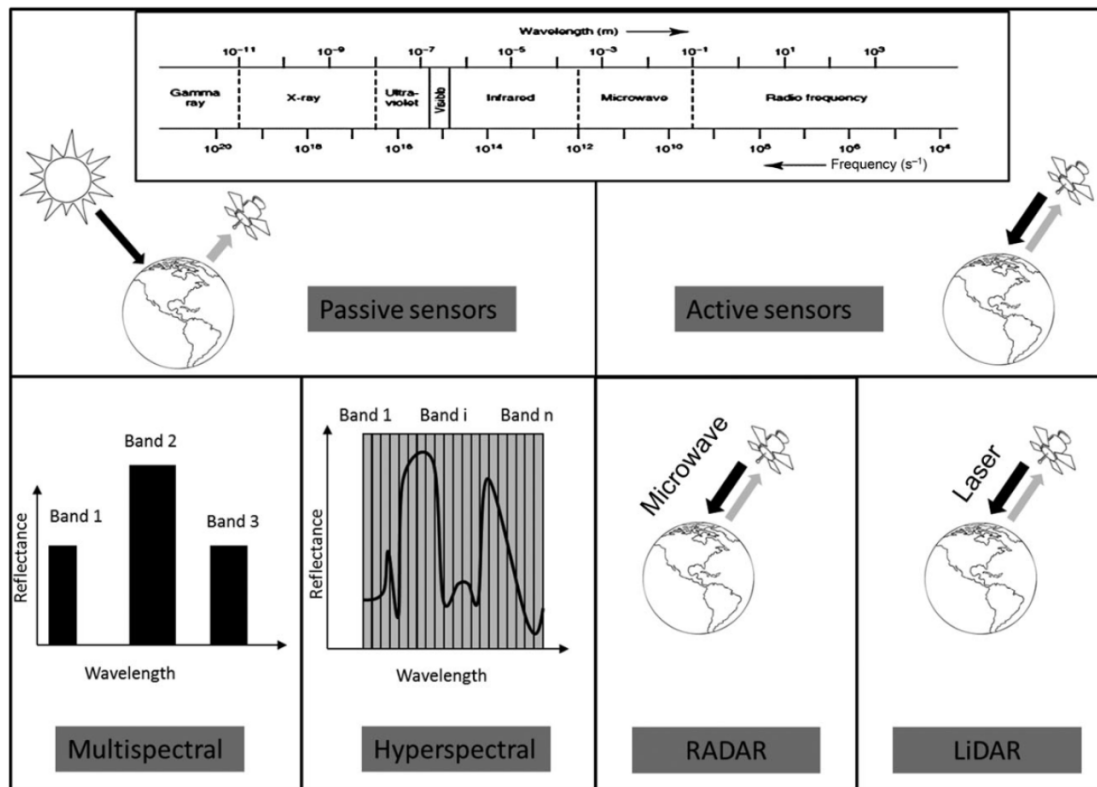


Figure 1. Main sensor types used by environmental satellites. Figure from Pettorelli et al. 2014.

These techniques result in different types of remote sensing measurements: multispectral and hyperspectral measurements, LiDAR and RADAR. Passive sensors (hyper- and multispectral) measure radiation emitted or reflected by the Earth. Reflected sunlight of the surface is the most common source of radiation measured by these sensors. Multispectral imagery is produced by sensors that measure reflected energy within several specific bands of the electromagnetic spectrum. Hyperspectral sensors, on the other hand, measure reflected energy in narrower and more numerous bands than multispectral sensors. This provides a relatively continuous measurement of a portion of the electromagnetic spectrum. The disadvantage of these passive sensors is that they can only measure during the day when the sun is shining and there are no clouds. Active sensors (active RADAR and LiDAR) emit an electromagnetic radiation (pulse) and later measure the energy bounced back to them. Radio Detection And Ranging (RADAR) sensors on-board satellites use longer wavelengths (1 – 10 cm) and are side looking (off-nadir), while Light Detection And Ranging (LiDAR) sensors emit laser pulses (usually at 1064 nm, but other wavelengths are possible) and are nadir viewing

(description of figure from Pettorelli et al. 2014).

RADAR remote sensing has the advantage to observe forests, regardless atmospheric conditions (such as clouds) through its longer wavelength. RADAR and LiDAR measurements can provide information about the vegetation structure (e.g. tree trunks, tree height) which enables the estimation of e.g. biomass. As de Paula 2017 summarizes, RADAR sensors are particularly useful in measuring canopy height, which is an important component of forest structure (Neumann et al. 2009). Note that there are also passive RADAR measurements, which are possible because of the emission of long-wave radiation from the Earth's surface, and are used, for example, to estimate carbon loss from forest degradation (Qin et al. 2021). LiDAR, another active remote sensing concept, typically uses laser pulses to estimate tree heights, and provide predictions of forest structural parameters (e.g. Zolkos et al. 2013). However, LiDAR measurements for the observation of vegetation were a long time limited to aerial platforms, which entail high cost and limited areas (Guimarães et al. 2020; Su et al. 2016). The NASA has successfully launched the Global Ecosystems Dynamics Investigation Lidar (GEDI) mission in 2018 and installed a LiDAR module at the International Space Station (ISS). The resulting measurements are used e.g. for global forest canopy height maps with the integration of Landsat data (Potapov et al. 2021) or for analysis of forest disturbances and biomass changes in Italy (Francini et al. 2022 also in combination with Landsat measurements).

The amount of data is significantly raising with more and more Earth-observing satellite missions launched in the last ten years (Guanter et al. 2015; Zeng et al. 2022). Most important impact of the application of earth observation satellite data on scientific research have the multispectral satellite missions Sentinel, Landsat and MODIS (properties of these satellite missions can be found in Table 1) followed by Gaofen/GF and WorldView (analysis of Remote Sensing Impact Factor by Radočaj et al. 2020; Zhao et al. 2022). Zhao et al. 2022 analyzed that the number of publications from Landsat and MODIS account for 40.2% and 31.7% of all investigated publications and have increased rapidly. Sentinel is a relatively new satellite mission (launched in 2015) that has the most potential but will take time to accumulate.

These well-known and high impact satellite missions are producing mainly multispectral measurements (except Sentinel-1). Multispectral data can be used to derive so called vegetation indices. These indices (an overview by Montero et al. 2023; Zeng et al. 2022) such as Normalized Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI) can be used to correlate several vegetation processes (e.g. Leaf area of forests Q. Wang et al. 2005 or estimates of forest productivity Sims et al. 2006).

Table 1

Properties of selected global open data satellite missions Handbook & Tools 2015; Missions 2016; Vermote et al. 2011. Landsat 9 has similar characteristics as Landsat 8 but its sensors includes also two thermal infrared bands (Lulla et al. 2021).

Properties	Sentinel-2	Landsat 7	Landsat 8	MODIS
Spatial resolution (m)	10, 20, 60	(15), 30, 60	(15), 30, 100	250, 500, 1000
Temporal resolution/ revisit time (days)	2–3	16	16	1–2
Spectral resolution	13 bands	8 bands	11 bands	25 bands
Wavelength range (nm)	442–2186	450–12500	433–12500	459–2155
Supported study area scale	local, national	regional, national	regional, national	regional, global

Remote sensing observations offer the opportunity to gain a better understanding of forests with respect to their structure and dynamics. Satellite measurements vary in their resolution and coverage. Thus, for global observations, there is a trade-off between the spatial and temporal resolution of satellite (e.g., Landsat, Sentinel) and airborne products. The combined methods of remote sensing and field observations offers the opportunity to gain a better understanding of forests with respect to their structure and dynamics. However, the ecological interpretation of remote sensing observations of forests is challenging, and in many cases still in development. For example, there are problems of the estimation of LAI and productivity out of NDVI values, because of saturation effects (A. R. Huete et al. 1997). Additionally, reflected signals from the top of the canopy is influenced by the understory of the vegetation and soil, which make it difficult to estimate forest attributes out of the reflectance (Eriksson et al. 2006). In the detection of forest degradation, the capacity of remote sensing applications is limited at this stage, particularly in the detection of degradation on a small scale and at a local level, such as that caused by low-intensity selective logging in specific locations (Ellis et al. 2017; Y. Gao et al. 2020; Hernández-Gómez et al. 2019).

The advances in remote sensing in the last decades especially the availability of data and products are resulting in a greater democratization of remote sensing. In this context it means the opportunity for everyone to participate in supporting forest management and conservation activities in parts of the world where environmental problems are most pressing (Lechner et al. 2020).

1.4 Applications of forest models

Over the past 50 years, the complexity of forest models has increased substantially. This partly reflects enhanced ecological knowledge and strongly increasing computing power, but partly also the desire to develop models that more realistically represent

natural processes (Bugmann & Seidl 2022). There is a long tradition of using forest models to analyze the impact of climate change on forest structure, species composition, and biogeochemical cycles (e.g., Bugmann & Fischlin 2002; Bugmann & Seidl 2022; Kienast 1991; Solomon 1986). Forest gap models were also used to investigate questions related to biodiversity (Shugart et al. 2018), for e.g. in a case study the forest dynamics of a tropical rainforest with more than 400 tree species was simulated (Köhler & Huth 2004). In addition to that, logging modules were included in forest gap models and they have been used e.g. to investigate different logging scenarios to explore the long-term impacts on forest carbon stocks (Brazhnik et al. 2017; Huth et al. 2004; Köhler & Huth 2004). Also, the impact of other types of disturbances have been investigated with forest gap models like landslides, forest wildfires and windstorms (e.g. Brazhnik et al. 2017; Brazhnik & Shugart 2017; Dislich & Huth 2012; Doyle 1981; R. Fischer 2021; Gutiérrez & Huth 2012; Keane et al. 1996; Shugart & Noble 1981; Shuman et al. 2017).

Dynamic global vegetation models have been used e.g. to investigate the changes in the atmospheric concentration of CO_2 (Brovkin et al. 2002; Joos et al. 2004), the greening trend in northern high latitudes (Lucht et al. 2002; New et al. 2000) or the effect of CO_2 and climate warming on productivity (Cao & Woodward 1998; Cramer et al. 2001; Kicklighter et al. 1999; Prentice et al. 2007; Schaphoff et al. 2006). Additionally, they have been applied to ecosystems undergoing disturbance events such as fire, drought, elevated CO_2 , land-use change, and insect defoliation (Medvigy et al. 2012; Miller et al. 2016; Trugman et al. 2016; K. Zhang et al. 2015).

Prentice & Cowling 2013 critically reflect that DGVMs applications to questions of biophysics (land–atmosphere interactions) and biogeochemistry (carbon and nitrogen cycling) have tended to dominate over applications to questions of community composition and biodiversity. These are important perspectives to be included in ecosystem ecology (Loreau 2010). Nowadays there is a general acceptance of the need to include cohort models (which are representing several individual trees with the same attributes) to “bridge” between the small scale and large scale (Argles et al. 2022; Franklin et al. 2020). There are several studies making progress in this regard (Argles et al. 2020; Burton et al. 2019, overview in Fisher et al. 2018), but there are still challenges, such as the implementation of plant demography within the explicit modelling of plant sizes instead of patch age (Argles et al. 2022; Maréchaux et al. 2021; Melton & Arora 2016; Sakschewski et al. 2015).

Important technical developments have revolutionized the forest modelling field and improved prediction of forest ecosystem dynamics at large scales (Blanco et al. 2020; Shugart et al. 2015, 2020):

- technical innovations in airborne and satellite remote sensing instruments (Belward & Skøien 2015) and significantly increase of available data (Guanter et al. 2015)
- development of autonomous continuous measurement devices for soil, vegetation, and atmospheric variables (Guimarães et al. 2020; Sethi et al. 2022)
- development of new statistical methods (Tredennick et al. 2021)
- increased computing power, which is permitting continental scale implementation of individual-based forest models as well as DGVMs at high spatial and temporal scales (Argles et al. 2022; Blanco & Lo 2023).

Recent applications of forest models have taken advantage of these new possibilities, and in the following section I will focus on applications of individual-based models. As Maréchaux et al. 2021 pointed out, forest models are able to conduct virtual experiments (Fyllas et al. 2017; Morin et al. 2018; Schmitt et al. 2020), as well as to test assumptions about ecological processes (Maris et al. 2018; Mouquet et al. 2015). They can also be used to identify potential for improvement and possible gaps in ecological knowledge, and to guide the design of further field experiments (Medlyn et al. 2016; Norby et al. 2016). Due to increasing confidence in forest models' ability to reflect the true behavior of the system (e.g., due to high agreement in model comparisons - Bugmann et al. 2019; Cramer et al. 2001; Fisher et al. 2018), they are increasingly being used to investigate important relationships between different forest characteristics (e.g., Bohn & Huth 2017; Bohn et al. 2018; Köhler & Huth 2010; Maréchaux & Chave 2017; Morin et al. 2020, 2011; Rödig et al. 2018; Sakschewski et al. 2016; Schmitt et al. 2020).

Studies which use forest models are more and more incorporating remote sensing measurements which are providing increased potential for large-scale measurements of forest attributes (Shugart et al. 2015). This is especially true for direct measurements of forest height and predictions of biodiversity and carbon stocks (Van der Sande et al. 2017). For example, individual-based gap models are used to derive information from remote sensing measurements about the heterogeneity of forest structure (R. Fischer et al. 2019) or estimate carbon dynamics of forests with the use of new allometric models parameterized with tree crown data derived out of airborne measurements (Jucker et al. 2017). Forest models can be used to derive relationships for the interpretation of remote sensing measurements, e.g. between forest height and carbon stock (Köhler & Huth 2010), or to validate correlations for these relationships (Minh et al. 2013). It is also possible to use remote sensing and gap models to identify patterns across scales

from micro to global scale (Shugart et al. 2020). In this study, the predictions of a gap model are compared with data from micrometeorological eddy-covariance towers and then scaled-up to produce maps of global patterns of evapotranspiration, net primary production, gross primary production and respiration. The fusion of high-resolution data, such as those derived from airborne laser scanning, with individual-based forest modelling can additionally offer insight into how plant size contributes to large-scale biogeochemical processes (F. J. Fischer et al. 2019). Here individual-based models can serve as data integrators for dynamic global vegetation models. Forest models in combination with space-borne LiDAR measurements can be used to improve the carbon estimates for large regions like the whole amazon (Rödig et al. 2017).

Moreover, forest models can be used to create virtual remote sensing data by combining them with forward modelling of LiDAR to directly link measurements to forests (Knapp et al. 2018). New methods use forest models not only to generate virtual remote sensing data, they also generate virtual forests. Such a generator was developed by Bohn & Huth 2017, the so-called ‘forest factory approach’. In one of their studies they generated virtual forest stands that possibly could exist in Central Europe. This multivariate dataset enabled a multidimensional investigation of the relationships between structural properties, plant diversity and productivity (Bohn & Huth 2017; Bohn et al. 2018). This promising approach has shown on the basis of simple mechanisms that over a broad range of forest stands, several forest properties (biodiversity and structure) have to be considered to understand forest productivity. The forest factory approach establishes a new way to analyze forests which does not require simulating forests over long periods of time. Instead the focus of the analysis is on the state space of the forests (described by structure properties, like forest height, basal area or others).

As Blanco & Lo 2023 pointed out “we are at an exciting moment, in which the development of new statistical and measurement techniques is finally creating opportunities for developing true inter-scale models, from individuals to regions and beyond”. Additionally, “the need to better understand ecological process is also more important than ever as climate, biodiversity, and land-use changes move forest ecology of the Earth to novel conditions”.

1.5 Objectives of this thesis

In the introduction, I emphasized the importance of a deeper understanding of forest ecosystems because of the changing environment and our dependence on forest ecosystem functions and services. I explained the origins of forest modeling and the needs

and limitations under which different types of forest models have been developed. Individual-based forest models face the challenge of applying their well-established, detailed implementation of mechanistic processes to large scales. DGVMs, on the other hand, attempt to implement processes such as forest demography in more detail. New opportunities, such as increasing remote sensing measurements and computing power, are bringing both families of forest models closer together. But not only have the historical limitations changed, so have the purpose of forest models, and today we face new challenges and uncertainties related to climate warming. Some of the most urgent and important questions today include understanding and quantifying the global carbon cycle and the role of forest ecosystems and vegetation in it (Cabon et al. 2022; Harris et al. 2021; Ruehr et al. 2023). Related to this is the importance of quantifying land use change (Song et al. 2018) and to track which forests, are under limiting environmental factors or disturbances (Curtis et al. 2018; Taubert et al. 2018). Particularly important in the context of land use change is understanding the role of forest management, deforestation and degradation (Roebroek et al. 2023), especially in the Amazon (Albert et al. 2023). Given the new possibilities of improved computational resources and the availability of new methodologies, the opportunities to apply forest models are growing, but forest models need to be adapted, extended or applied in new ways to take full advantage of these conditions.

The first study of this thesis (Chapter 2) establish such a new way of using forest models by extending the forest factory approach by Bohn & Huth 2017 and make it applicable for forests in different biomes. This approach allows to generate forests using the architecture and processes of forest models (here we use the individual-based gap model FORMIND). In this study, 700,000 forest stands in seven different ecoregions are generated by using the Forest Factory 2.0. The main objectives of this study are:

1. Analyzing the structural state space of virtual forests in different ecoregions and compare them with each other
2. Comparing the relationships between the structural properties of forests and (i) biomass (as a proxy for the carbon stock), (ii) above-ground wood production AWP (as a proxy for the carbon flux), and (iii) species evenness (as an example for a biodiversity index).

In contrast to the tradition of investigating the development of individual forest stands over time, the Forest Factory 2.0 is used as a tool to gain knowledge about forests by analyzing the state space of forests. This analysis shows the potential of the presented approach for a wide range of research questions.

Remote sensing data from satellites measuring hyperspectral and multispectral reflectance on high resolution is more and more available (1.3 and 1.4). These data could provide information on forest characteristics related to structure and state. Subsequently, these data need to be analyzed using tools that can relate reflected radiance to forest dynamics at small scales. One possibility for such tools are individual-based forest models. The second study in this thesis is about a new combined approach which enable the calculation of hyperspectral reflectance of forests (Chapter 3). It uses the multi-layer radiative transfer model mScope (Yang et al. 2017) and the individual-based forest model FORMIND. The objectives of the second study are:

4. Enlarging the application field of mScope and investigate the calculated reflectance spectra of boreal forests using forests in Finland as an example
5. Investigation of how the concept of forest representation (simple or detailed structure) influence the calculated reflectance spectrum of forests
6. Comparing the simulation output with Sentinel-2 data by calculating typical vegetation indices of the investigated forests

This work provides a forward modelling approach for relating forest reflectance of radiation to forest characteristics. With this tool, it is possible to analyze a large set of forest stands (for example in combination with the first study in Chapter 2) with corresponding reflectances (in the visible and infrared range). This opens up the possibility to understand how reflectance of radiation in forests is related to succession and different forest conditions.

In order to take advantage of the increasing number of remote sensing measurements and to achieve synergy effects with forest models, it would be useful to align the design of satellite missions with the capabilities of forest models.

The last study of the thesis is about using expected biomass measurements provided by the upcoming RADAR BIOMASS satellite mission (P-band, launching in 2024 by the European Space Agency) to predict the productivity of tropical forests (Chapter 4). The main objectives are:

7. Investigating the relationship between horizontal and vertical distribution of above-ground biomass and carbon dynamics in forests (GPP, NPP and carbon turnover time) by using a boosted regression tree
8. Examining the robustness of the research results by applying the approach to different forest types (disturbed and mature forests)

This study presents a novel approach to estimate forest productivity by combining P-band RADAR remote sensing measurements, machine learning and an individual-based forest model. The results highlight the potential of deriving estimates for forest productivity from information on forest structure and gives information about the quality of future satellite products.

This thesis presents innovative and novel applications of forest models. It provides insights into how remote sensing measurements can be incorporated into ecological research and how path dependencies of forest simulations over time can be overcome by introducing a globally applicable forest generator. It shows first applications how these approaches can be used to investigate the relationship between forest structure and productivity (Chapter 2 and 4), to generate virtual remote sensing measurements for a case study (Chapter 3), and finally to estimate productivity for typical tropical forests using RADAR remote sensing measurements (Chapter 4). The applications presented may ultimately contribute to a better understanding of forest ecosystems.

2 Creating virtual forests around the globe and analysing their state space, Henniger et al. 2023

2.1 Abstract

Forests, as one of the most important carbon sinks on earth, are more and more under stress by environmental changes. The dynamics of forests and consequently their functions in general, begin to change. We therefore present a recent model development, called "Forest Factory 2.0", which generates various virtual forest stands for different biomes on earth. This approach allows to generate forests using the architecture and processes of forest models (here we use the individual-based gap model FORMIND). Using Forest Factory 2.0, we generated 700,000 forest stands in seven different ecoregions. In contrast to the tradition of investigating the development of individual forest stands over time, we used the Forest Factory 2.0 as a tool to gain knowledge about forests by analyzing the state space of forests. We conducted a structural sensitivity analysis to compare the relationships between structural properties and biomass, productivity, as well as species evenness of forests. In this study we analyze the state space of forests in different biomes and demonstrate the potential of this approach for theoretical ecology.

2.2 Introduction

Forests cover 25% of the global land surface (Gibson et al. 2011). They play a major role for the global carbon cycle because of their function as carbon storage and their contributions to global carbon fluxes (Bonan 2008; Grace et al. 2014). Forests are important for sustaining biodiversity and provide habitat for 70% of all animal species (Gibson et al. 2011; Myers et al. 2000; Pimm et al. 2014). Further, forests exhibit a diversity of spatial structure and change their structure due to natural succession, management or disturbances (Pan et al. 2013).

While forest ecosystems are exposed to environmental change, like all complex adaptive systems, they have a certain capacity to cope with it. However, if these change processes occur too frequently, on too large spatial scales, with too high intensity, the adaptive capacity of the forests may be exceeded. Global change processes such as climate change and related effects such as drought, heat waves, fire, storms or pest outbreaks (Stocker 2014), but also deforestation (Stocker 2014) and fragmentation (FAO 2022; R. Fischer 2021; Taubert et al. 2018) are accelerating and occur simultaneously. As a result, the dynamics of the forests would change as well as their tree species composition and structure. Therefore, forests appear to be under increasing pressure

(McDowell et al. 2020), affecting forest biodiversity in general as well as the diversity of functions provided by forests. This shows the urgency of sustaining their functioning, understanding and enhancing their adaptive capacity, and appropriately adapting their management.

Prerequisite for addressing these challenges, however, is a sound understanding of structure-function relationships, esp. between the properties of forests (species-compositional and spatial-structural) and their functions (e.g. carbon flux and storage). To analyze forests ecosystems the perspectives of community ecology (organismal aspects, diversity of species and structure) and ecosystem ecology (matter and energy flux aspects, biogeochemical cycles) are not separable. A mechanistic understanding of the functioning of ecosystems can only be gained if these two perspectives are adequately linked to one another (Loreau 2010). However, though there are many empirical studies based on data on forest inventories, the number of available samples and plots or the lack of focus in monitoring make it difficult to create a sufficiently complete picture (Lindenmayer & Likens 2009; Lindenmayer et al. 2011). This challenge may be overcome by using remote sensing data, but relating and condensing this large scaled data to the local or individual scale remains a challenge (Ma et al. 2019). Another challenge is to capture the inherent spatial heterogeneity of environmental conditions and how they change in response to projected changing processes, especially among different biomes. Thus, there is a huge variety in the environmental factors which are supposed to influence forest properties (species composition and structure) and the shape of the structure-function relationship. The needed relevant variables are mostly not fully covered by the existing inventories and datasets.

Forest models can help to bridge the gap between multiscale field data and processes enabling a multivariate view of forests. Nevertheless, different types of models have different application fields. For example, global vegetation models have a focus on large spatial scales and time scales, whereas individual-based models focus on smaller scales, as they consider processes at tree level and can thus also analyze structural dynamics (Maréchaux et al. 2021). Thus, individual-based models are particularly suitable for considering ecosystem dynamics as an emergent outcome from the interaction of processes at individual level. This allows the identification of structural properties and functional characteristics of forests at different spatial scales as they emerge from the assumed environmental conditions. This also opens up the opportunity for correlative analyses of the structure-function relationship (Rödiger et al. 2018; Thurner et al. 2017). However, the causal relationships underlying them are not yet satisfyingly understood.

We introduce a new way of sensitivity analysis. The variation of parameter values or

the comparison of different scenarios is a prominent way of sensitivity analysis to gain causal understanding of relationships. We perform sensitivity analysis not by varying parameters, but by analyzing millions of initial states to gain understanding of the relationship between forest functions and forest structure for different biomes. Performing this way of sensitivity analysis is not common so far and methods for this are rare. Examples for such a powerful application are the use of landscape generators in the context of impact assessments of land use scenarios (Engel et al. 2012; Langhammer et al. 2019) and the use of weather generators in the frame of climate impact analyses (Friend 1998; Kumagai et al. 2004).

For forests, such a generator was developed by Bohn & Huth 2017, the so-called ‘forest factory approach’. In one of their studies they generated virtual forest stands that possibly could exist in Central Europe. This multivariate dataset enabled a multidimensional investigation of the relationships between structural properties, plant diversity and productivity (Bohn & Huth 2017; Bohn et al. 2018). This promising approach has shown on the basis of simple mechanisms that over a broad range of forest stands, several forest properties (biodiversity and structure) have to be considered to understand forest productivity. The forest factory approach establishes a new way to analyze forests which does not require simulating forests over long periods of time. Instead the focus of the analysis is on the state space of the forests (described by structure properties). Due to the regional limitations of the forest factory by Bohn and Huth (focus on European forests), it offers potential for further research. To realize the potential and to analyze a causal relationship between their structural -, diversity -, and productivity relationships for different biomes, a further development and extension of the forest factory approach is necessary.

In this study we present a novel software tool - the Forest Factory 2.0 - which creates millions of virtual forest stands, covering various species compositions and structural properties for different biomes.

Additionally, we provide a data product generated with the software tool to demonstrate the potential of this approach for systematic mechanistic analyses of structure-function relationships across biomes. The data product contains in total 700,000 forest stands including 12 forest properties. These forest stands consist of over 11 million individual trees with over 20 tree properties.

In this study, we show examples of ecological analysis based on the generated forests. First, we compare the state space (based on four structural properties) of forests between seven regions derived for different biomes. Second, we compare the relationships between the four structural properties of forests and (i) biomass (as a proxy for the carbon stock), (ii) aboveground wood production AWP (as a proxy for the carbon

flux), and (iii) species evenness (as an example for a biodiversity index). With the analysis, we want to show the potential of the presented approach for a wide range of research questions.

2.3 Methodical concept

With the Forest Factory 2.0 we have developed a software tool that makes it possible to create virtual natural forests that could exist in nature. The Forest Factory 2.0 allows through its algorithm a fast and generic generation of forests. In contrast to forest simulations, the forest factory approach does not consider and simulate forests over a long period of time. It generates various forests describing different states of succession, as well as management and disturbed forest stands also for different species mixtures. The forests can be generated for different regions of the world. In this study we produced 700,000 forest stands in total for seven different ecoregions. The background knowledge, i.e. the information and processes for the generation of forests, is provided by forest inventories and studies which are represented in the parameterizations of forest models (here we use the forest model FORMIND). A large number of ecological properties can be calculated for each generated forest, which allows a detailed analysis of the relationships between forest properties. Comparison of forest stands for different ecoregions is made possible by using the same algorithm for each forest stand generated.

2.3.1 Forest Factory 2.0

For processes such as competition and productivity, the Forest Factory 2.0 uses the individual- and process-based forest model FORMIND. This forest model allows the simulation of species rich forests and also considers the complex age structure of their tree community. FORMIND has been extensively tested and applied to tropical forests (R. Fischer et al. 2014; Gutiérrez & Huth 2012; Huth & Ditzer 2001; Kammesheidt et al. 2001; Köhler et al. 2003; Köhler & Huth 2004, 2007; Rödiger et al. 2019; Rüger et al. 2008), temperate forests (Bohn et al. 2014; Bruening et al. 2021; Rüger et al. 2007) and grasslands (Taubert et al. 2012). It is an individual-based model which means that the growth of every single tree is simulated. The model considers four main process groups: growth of single trees (increment of tree biomass, stem diameter and height), mortality, recruitment, and competition (e.g. for light and space). FORMIND is also used for large scale simulations (Paulick et al. 2017; Rödiger et al. 2018) e.g. forest-wide carbon balances in the Amazon. The Forest Factory 2.0, is implemented as an

independent module of FORMIND in C++ language and uses processes of the forest model FORMIND (like competition for light and allometries). The processes of the forest model can be modified independently of the Forest Factory 2.0. It is possible to combine the Forest Factory 2.0 with other forest models.

The methodology of the Forest Factory 2.0 follows the Forest Factory (Bohn and Huth, 2017), that generated forest stands for the temperate zone and was implemented in the language R. In this paper, we introduce a new version of the Forest Factory that includes important new components and extensions that make it applicable on a global scale. We also provide an R package that facilitates the use of Forest Factory 2.0 (Fig. 2). This package allows analysis of the data product that we publish or that users generate themselves.

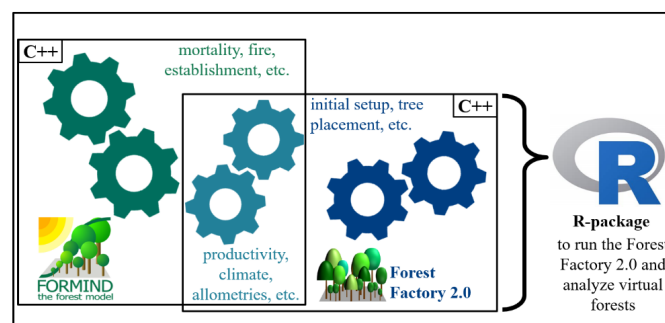


Figure 2. Interdependencies of FORMIND, Forest Factory 2.0 and the R package. The Forest Factory use processes of the forest model FORMIND. The R package (wrapper of C++ code) helps to run the Forest Factory 2.0 and process the generated forest stands to a data product. It also prevents some features for the analysis of the data product.

The Forest Factory 2.0 can produce a large number of virtual forest stands (20 m x 20 m base area and funnel shape) for each available parameterization, which is representing an ecoregion. Every tree in the generated forest stand must have a positive productivity (gross primary production > respiration). In FORMIND a negative productivity causes the dying of trees. To calculate the productivity, we calculate the biomass increment of every placed tree over one year, which results from the different ecoregion-specific parameterizations (e.g. climate). To create forests for an ecoregion the Forest Factory needs information on climate conditions and a parameter set which consists of species-specific parameters e.g. concerning the tree geometry, productivity and species pool (see Section 2.3.3 for details) which are representative for an ecoregion.

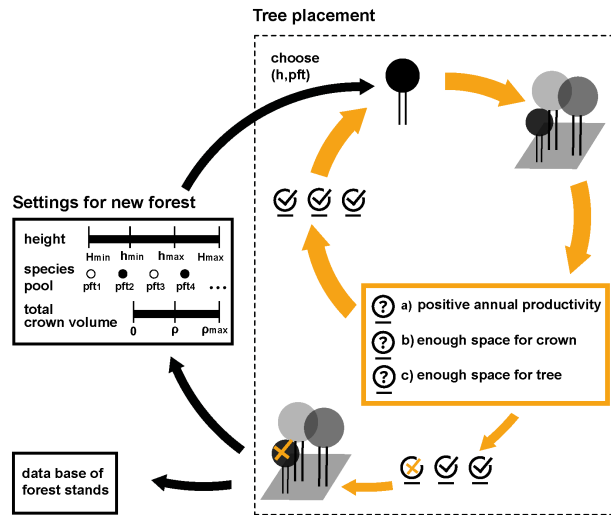


Figure 3. Concept of the Forest Factory 2.0. For each forest, Forest Factory 2.0 pre-selects a minimum and maximum height of trees, a group of species and a maximum total crown volume (sum of crown volume of all trees). Each tree is determined by a height (random from height distribution) and a species (random from the species pool). A new tree is added to the forest stand until the new tree has no positive productivity or space. Then the tree is deleted, the forest stand is saved in the data base and a new forest stand is generated.

As an initial information, which is valid for all generated forest stands, the Forest Factory 2.0 assumes a minimum and maximum height of the trees H_{min} and H_{max} , an overall maximum total crown volume ρ_{max} and an initial species pool. The overall maximum total crown volume ρ_{max} is the maximum sum of crown volume of all trees valid for every forest stand. The species pool is defined by the parameterization of each ecoregion and each species/plant functional type is representing a species or group of species with similar functional and morphological characteristics. This initial information is required to start the Forest Factory 2.0 (Fig. 3).

Once started, the Forest Factory 2.0 pre-selects for each forest stand a minimum and maximum height of trees h_{min} and h_{max} (from the initial H_{min} and H_{max}), a maximum total crown volume ρ and a group of plant functional types. The pre-selection for h_{min} , h_{max} and the maximum total crown volume ρ is done by random assuming uniform distributions (the boundaries are $[H_{min}, H_{max}]$ and $[0, \rho_{max}]$). The pre-selection of h_{min} and h_{max} for every forest stands also allows the generation of even aged forests. The pre-selection of the species pool for each forest stand is done by random assuming a uniform distribution to select the number of species (more details in Appendix A).

After the pre-selection for the forest stand is done, one tree after another is planted. The explicit position of a tree in the forest stand is not important due to the spatially

implicit approach of forest gap models, where the position is randomly chosen at the end of the tree placement procedure. A tree height for the tree to be planted is selected from a predefined height distribution ($X \sim Exp(-0.05), h = X | X \in [h_{min}, h_{max}]$). The selected species pool is used to determine randomly the species type of a new tree (each species has an equal probability). For each tree these two attributes (height and species type) are selected and are used to calculate other attributes of the tree. Attributes are derived from processes and the parameter input of the used forest model (here we use FORMIND for different forest biomes). For the tree placement, it is checked if: a) each tree has a positive productivity, b) there is still space for the canopy of this tree (in each height layer, all tree crowns together must not exceed the boundaries of the forest stand) and c) the maximum total crown volume is not exceeded (we allow a certain maximum density in three-dimensional space: the maximum total crown volume). The selection rules for tree height and tree species is the same for each tree. If b) or c) is violated, the tree will not be considered, the tree placement for this forest is terminated and the created forest stand is saved in a database. If a) is violated and the calculated productivity over one year is negative, an attempt is made to replace the tree with a tree of a different species (with the same height and out of the selected species pool for this forest stand). If the tree has now a positive productivity, it is placed, if not, the tree placement for this forest is terminated (and the forest stand is saved). Every time a new tree is placed the annual productivity of all previous planted trees have to be recalculated (e.g. due to the change of light availability). If one or more trees have a negative productivity the algorithm try to replace them with tree(s) of another species and if this doesn't work the tree placement for the forest is terminated. The generation of a new forest stand starts.

In this way, the Forest Factory can be used to generate millions of forests for different ecoregions and climates (by considering input parameterizations). The forests describe different states of succession (e.g. by differentiate H_{min} and H_{max}), as well as managed (e.g. even aged forests by the selection of H_{min} and H_{max} values with a small difference) or disturbed forest stands (e.g. by selecting a low overall maximum total crown volume ρ_{max}) including different species mixtures. The goal is to generate as many potential forest states as possible. For specific analyses of e.g. even-aged forests or late-successional forests, the virtual forests must be filtered according to the desired attributes.

For the derived forest stands a large number of properties and characteristics can be calculated by using the methods of the forest model e.g. for leaf area, diameter increment, LAI per height layer, size distribution, biomass, maintenance respiration, gross primary production (GPP), net ecosystem carbon exchange (NEE). Since we simulate

the productivity of each forest over only one year, we here do not focus on temporal evolution, but on states and benefit from the knowledge contained in widely applied and long-established forest models (here FORMIND).

The Forest Factory 2.0 enables the possibility for a coupling with other forest models. The coupling setup would run an iterative process. The Forest Factory provides tree and forest stand information to the corresponding forest model. The calculation of productivity and tree attributes (e.g. due to allometry) takes place in the forest model, and is reported back to the Forest Factory.

2.3.2 Forest Factory (Bohn & Huth 2017) vs. Forest Factory 2.0

In this Section, we will explain the main differences between the Forest Factory by Bohn and Huth (2017) and the Forest Factory 2.0 and show how we have significantly extended the approach.

One important advantage of ForestFactory 2.0 compared to the Forest Factory by Bohn and Huth (Table 2) is a significant speed increase (3 million forest stands per hour, 30 times faster), which allows the creation of a huge number of forest stands. Further, it is now possible to create forests for all regions of the world for which parameter sets are available (here for the forest model FORMIND).

Table 2

Improvements of the Forest Factory 2.0 in comparison to the Forest Factory (Bohn and Huth 2017).

Forest Factory (Bohn, Huth 2017)	Forest Factory 2.0
10,000 forest stands per hour (standard notebook)	3 mio forest stands per hour (standard notebook)
programming language R	C++ and integration in the actual forest model (here FORMIND)
only temperate forests	forests in different biomes (ecoregions)
15 pre-defined stem diameter	one continuous height distribution for tree placement
only stem diameters up to 0.5 m planting trees until they are nonproductive	no restrictions for stem diameter replacing non-productive trees (by other species)
algorithm produces clustered sampling	more equally distributed sampling
	open source code and open data product

As the Forest Factory 2.0 is a part of the FORMIND model repository, functional model improvements are automatically available for the forest factory. This allows the Forest Factory 2.0 to use recently developed sub-modules of the forest model. For ex-

ample, lidar waveforms or light reflectance spectra can be calculated for the generated forest stands. It is also possible to use the generated forests directly as input for simulations to analyze future development of these forests.

Compared to the Forest Factory by Bohn and Huth, in the Forest Factory 2.0 the input parameters were reduced. There is only one function for tree height distribution to derive stem diameter as input (in the old version there were 15 fixed stem diameter distributions). This increases the flexibility and the possibilities for analysis. With the new Forest Factory 2.0, we can investigate the state space of the forests more evenly, i.e., different characteristics of the forest structure occur with sufficiently equal abundance (Appendix Fig. A1).

We provide an open source R + Python package and a data product of forest stands to enable accessibility to a wide range of users. The R package (git.ufz.de/angermue-/forestfactory) represents an interface which makes it possible to operate with the Forest Factory 2.0 from the R platform. An overview of the forest dataset is given in Appendix (Table A1).

2.3.3 Study sites

The parameterizations (representing ecoregions in Table 3) represent the synthesis of information of many field measurements and inventories, not only concerning the species-specific allometric tree attributes but also concerning tree growth and productivity. Due to this we use for all ecoregions the same kind of information only with different values. The parameterizations can therefore be interpreted as a kind of recipe with always the same ingredients, in different quantities. The cooking process - the algorithm of the Forest Factory 2 - works for all parameterizations according to the same principle. The used parameterizations belong to different forest stands in different regions and we decided to use the names of the ecoregions in the paper to make clear where the investigated forests are located. For the generation of temperate forests in Germany we use a daily based climate data set of the Hainich National Park (Thuringia, Germany) for the year 2007. For the other regions we used reduced climate information which is described in the Appendix (Section 3).

2.3.4 Simulation and analysis

In this study we present results for forest stands in seven different ecoregions (see Table 3). The ecoregions consist of two temperate regions (one in North America, one in Europe) and five tropical regions (two in South America, one in central America, one in Africa, one in Asia). For each region we generated 100,000 forest stands with the Forest Factory 2.0 (initial parameters: $H_{min} = 5$ m, $H_{max} = 65$ m and the

Table 3

Overview of the ecoregions, parameterizations and climate used for this study.

Name in Paper	Biome short description	Number of pfts	Paper
Amazon	entire tropical forest in the Amazon using plant functional types	3	Rödiger et al. 2017
Panama	tropical lowland rainforest on Barro Colorado Island	4	Knapp et al. 2018
Germany	temperate forest in central Europe	8	Bohn et al. 2014
US	temperate forest within the Northeast US	9	Bruening et al. 2021
Ecuador	tropical evergreen montane rain forest in southern Ecuador	7	Dislich et al. 2009
Malaysia	Southeastern Asian tropical rainforest (North Borneo, Malaysia)	4	update von Köhler & Huth 2004
Tanzania	tropical submontane and lower montane rainforest at Mt. Kilimanjaro	6	R. Fischer et al. 2015

overall maximum total crown volume $\rho_{max} = 0.78$). Each region provides an initial species pool. We analyzed all forest stands for structural attributes (basal area, LAI, height heterogeneity, maximum height) and functional characteristics (above-ground wood productivity AWP, aboveground biomass and species evenness as an indicator for biodiversity. Species evenness is calculated by the Shannon Equitability Index (Heip 1974; Peet 1975). The Shannon Index (Shannon 1948) is normalized by the logarithm of the maximum number of species (we treat pfts as species here).

In a first step we explored under which structural conditions forests can exist in different ecoregions. For this, we use a state space approach. This space is determined here by four structural variables: maximum height, basal area, height heterogeneity, and LAI.

In Section 2.4.1 we investigated this state space of forests (mentioned above), by using diagrams (Fig. 4) similar to the classical diagrams of Whittaker, in which he analyzed the relation between climate (average annual temperature and precipitation) and vegetation types (Whittaker et al. 1970). Instead of climatic attributes we investigate here four structural properties (two in each Figure). We analyzed maximum height of trees (this corresponds to the forest height) and basal area which are typical properties to describe the structure of forests. Additionally, we investigated the role of tree height variability (here by using the standard deviation of the tree heights which we define as height heterogeneity) and leaf area index. Each generated forest stand can be represented as a point in the state space by a combination of these structural properties. We generated 700,000 forest stands, each representing a possible state, resulting in

700,000 points in the state space (100,000 for each ecoregion). To analyze the state space of the generated forest stands with positive productivity (Section 2.4.1), we examined the area which is covered by 100,000 forest stands of the same ecoregion by calculating the envelope around the points (each point represents one forest stand). These envelopes are calculated with the R package Concaveman (which uses convex hulls with concavity, for details please see Appendix Section 4).

To investigate the relationship between different forest properties and characteristics (Section 2.4.2) we derived heatmaps (Fig. 5, 6) where the x and y axis describe structural properties and the color describe functional characteristics: biomass (carbon stock), AWP (carbon flow) and evenness (biodiversity). The maps are rastered so one cell contains information of several forest stands with the same structural properties. The shown value for a cell represents the mean value over these forest stands. We also derived maximum value and standard deviation for these analyses (shown in the Appendix).

To allow direct comparison of forests between the seven ecoregions, we examined forests by their functions (biomass, AWP, evenness) that are similar in all four structural properties (Fig. 7). For these similar structured forests, we calculated the mean value of their functional characteristics and compared them in a 1:1 graph for three different regions. Additionally, the regression line and the adjusted R^2 were calculated. For all analysis we considered only forest stands with a basal area under 100. In the Appendix the analysis of the maximum values and the standard deviation of the biomass, AWP and species evenness have been added (see Appendix Figs. A7-A11).

2.4 Results

2.4.1 Analysis of forest structure in different ecoregions

In the first step, we are looking at the structural characteristics of forest stands for the different ecoregions created by the Forest Factory 2.0 (Fig. 4) by calculating the basal area, maximum height, height heterogeneity and LAI for each forest stand. The analysis in Fig. 4 shows which combinations of maximum height/basal area and height heterogeneity/ LAI lead to forests with positive productivity. Forests with properties outside the envelope line, don't have positive productivity. We observe mostly similarly-shaped envelopes with different sizes for the different ecoregions (represented by the different colors).

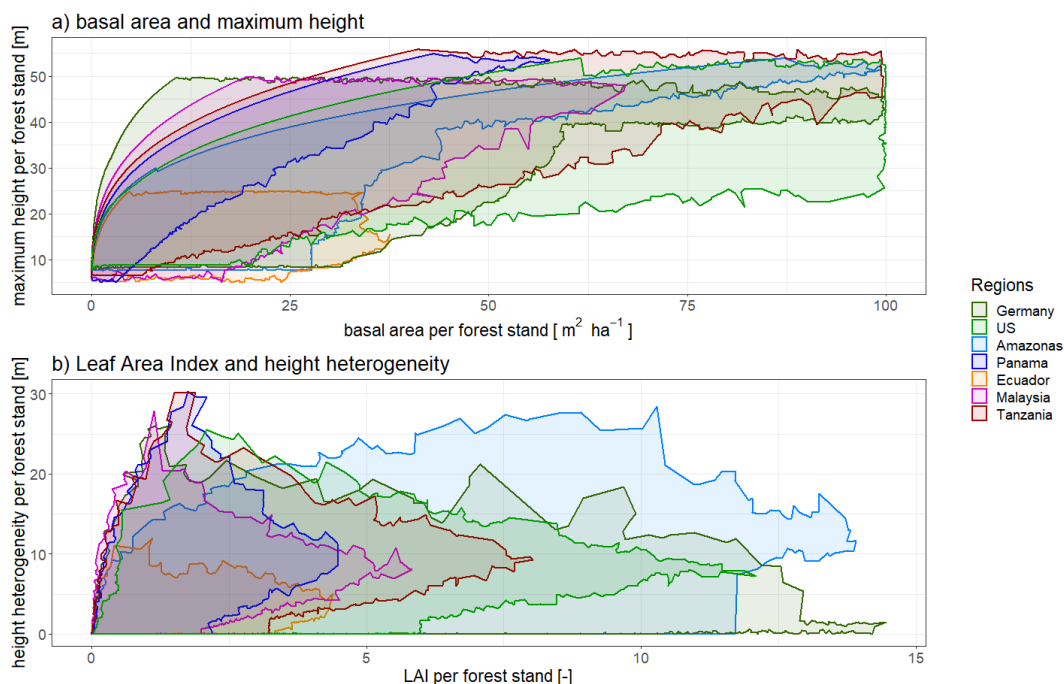


Figure 4. Analysis of the state space of generated forests. We examined the area which is covered by 100,000 forest stands of the same ecoregion by calculating the envelope around the points (for more details please see Appendix). Each generated forest stand can be represented as a point in the state space by a combination of the structural properties: a) maximum height and basal area; b) tree height heterogeneity and LAI. Different colors are indicating different ecoregions.

The analysis shows typical limitations of forest stands. Forest stands with a high basal area and low or moderate maximum height (empty area at the right bottom in Fig. 4a) do not occur. Large trees have large crowns. This tree allometries in combination with limited space restrict the abundance of these trees and also influences the resulting basal area (empty area at top left). The physiological and species-specific allometric interactions result in a typical shape in the state space that curves to the right.

Our forest stands can also be analyzed in a different state space, consisting of the LAI and the height heterogeneity. In most ecoregions, the largest values for tree height heterogeneity occur for forest stands with low LAI values, while the highest LAI values occur in forests with low to moderate tree height heterogeneity (Fig. 4b). As expected, the Brazilian Amazon has a large diversity of forest stands, and the shape of the envelope is quite different compared to other regions, e.g., without a peak at the top left (high height heterogeneity, low LAI).

The smallest area within the envelopes in both Figures (4a + b) is found for mountain forests of Ecuador (low maximum tree height).

We also investigated the frequency distributions of the forest properties of the forests

within the illustrated areas (Appendix Fig. A1).

2.4.2 Relationship between forest structure and ecosystem functions in different ecoregions

The Forest Factory 2.0 allows us also to analyze how structural properties (maximum height and basal area) affect functional characteristics (biomass, above-ground wood productivity (AWP), and evenness in the species composition (normalized Shannon Index as proxy for biodiversity). The biomass-related plots (Fig. 5: a, b, c) reveal a structure-function relationship that is quite similar for all investigated regions. Biomass is largely determined by the basal area and maximum height.

The analysis of the German forest stands (Fig. 5a) shows some interesting details for forests with high biomass and high basal area (top right area). Forests with a lower maximum height (40 m - 45 m) have on average a higher biomass than forests with a larger maximum height (> 45 m). With the Forest Factory, it is possible to analyze each individual tree of the corresponding forest stands in this area. The result is that all of these forest stands consist of trees of the species *Picea abies*. This is the tree species with the largest maximum height in the analysis for this region, but it has a low wood density, which leads to a lower forest biomass.

In all three regions, we observe that AWP increases with the basal area and decreases with maximum height, while shape and strength of the combined effects are region-specific. Also, the range of AWP values differs due to climate variations between the temperate (Germany, Fig. 5d) and the tropic regions (Fig. 5: e, f), which leads to lower AWP values for the German forest stands. Nevertheless, we observe that for Germany (Fig. 5a) and the Amazon (Fig. 5b), forest stands with high AWP have a high maximum height and basal area. In all three ecoregions, there occur also forest stands with high AWP values that have only moderate basal area and height.

Concerning the evenness of species, forest stands (Fig. 5: g, h, i) show a similar structure-function relationship for the Amazon (Fig. 5h) and Tanzania (Fig. 5i). Species evenness is increasing with basal area but decreasing with maximum height. For the temperate forests in Germany (Fig. 5g), the situation is more intricate. Forest stands between low and medium maximum height (0 m - 35 m) and with medium basal area have high evenness values.

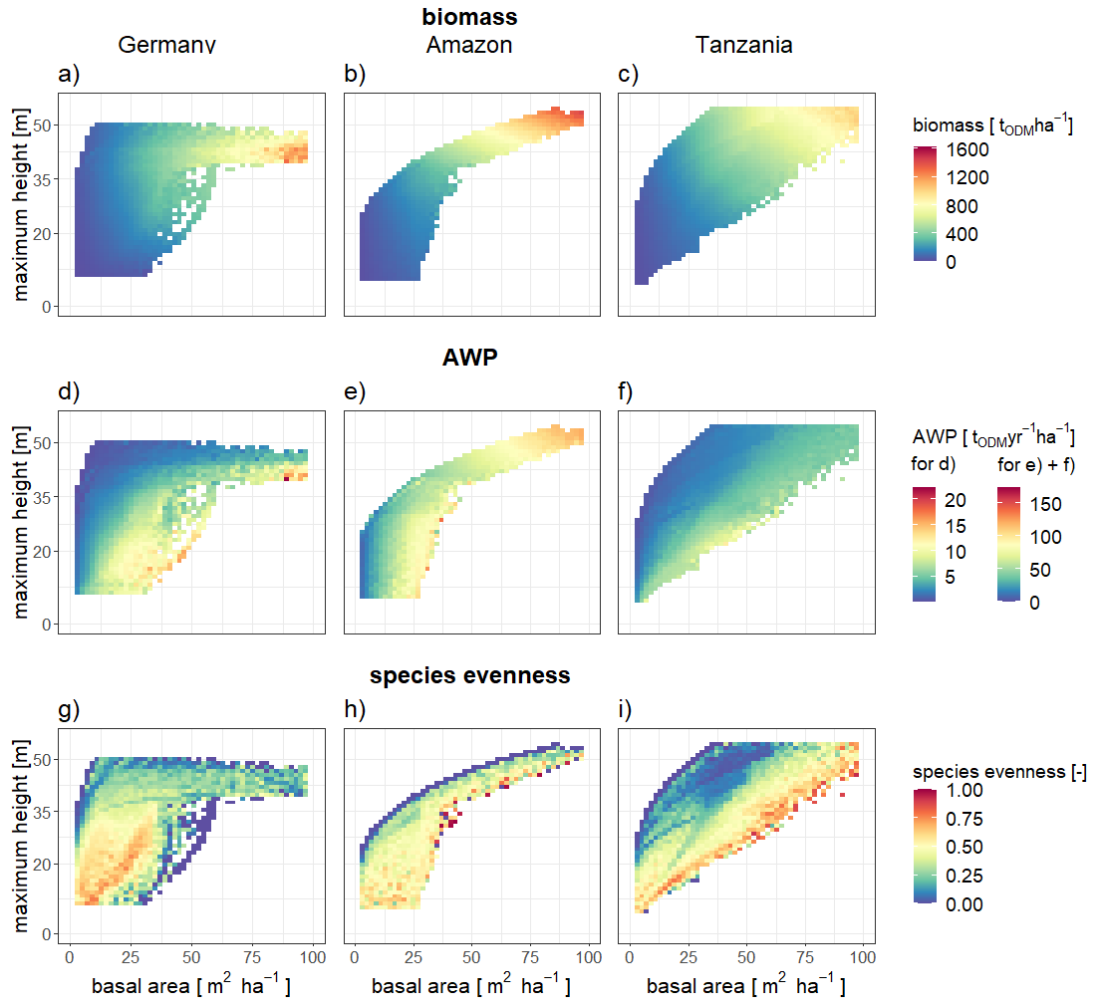


Figure 5. Relationship between structural properties (basal area and maximum height per forest stand) and biomass (a-c), above-ground productivity (d-f) and species evenness (g-i) for three selected ecoregions (Germany, Amazon, Tanzania). The color of each cell in the graph represents the mean value of the investigated property of all forest stands within one cell. Note that the AWP axes are scaled differently between the ecoregions (d-f). All other ecoregions, maximum value and standard deviation per property you can find in the Appendix (Figs. A5-A7). See Appendix Fig. A3 for Figures d) - f) with common AWP legend.

In a second step, we analyzed how two other structural properties (here: height heterogeneity and LAI) affect the functional characteristics of forest stands for the three investigated ecoregions (Fig. 6).

In all cases, forest stands with large biomass values (Fig. 6: a, b, c) can only be found if the LAI is high. Additionally, in the Amazon and Tanzania, these forest stands also need height heterogeneity values above 8 m.

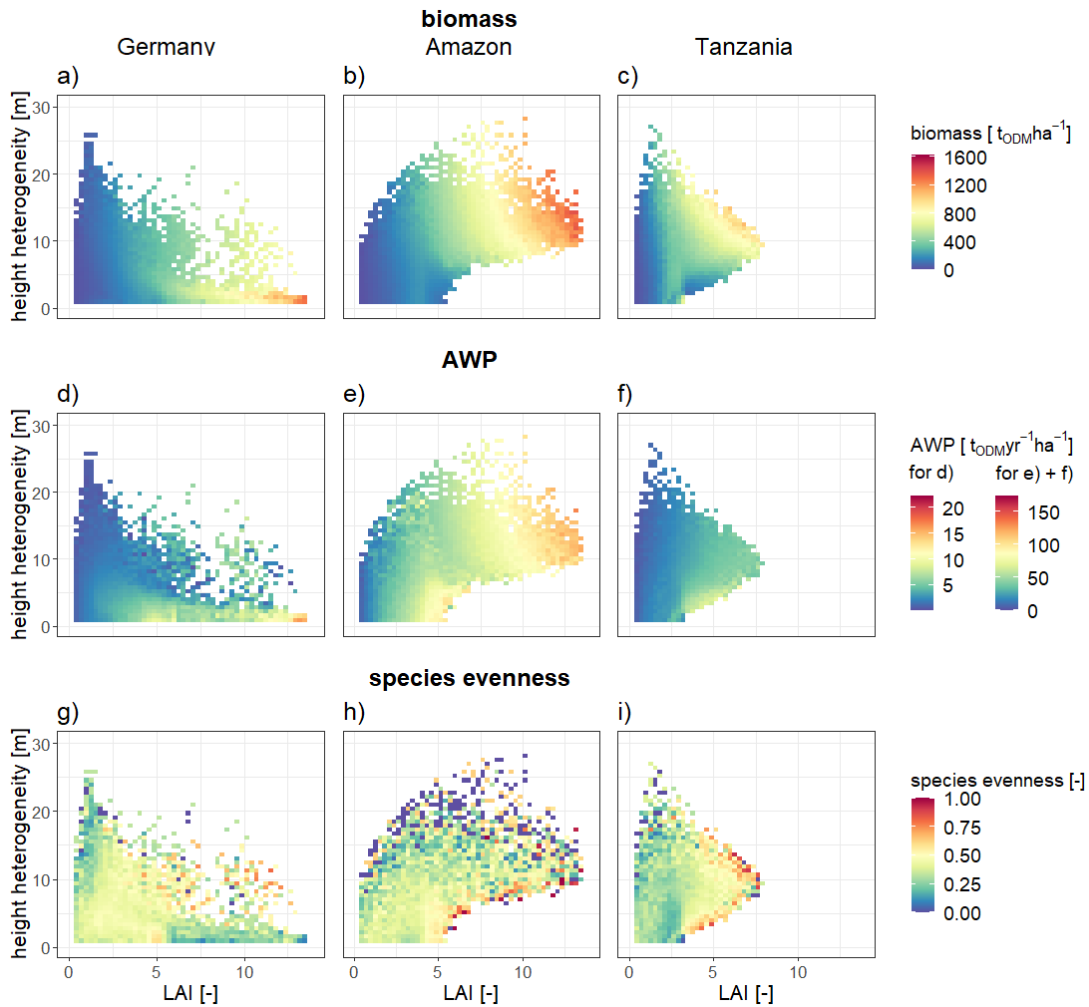


Figure 6. Relationship between structural properties (LAI and tree height heterogeneity) and biomass (a-c), above-ground productivity (d-f) and evenness (g-i) for the generated forests for three selected ecoregions (Germany, Amazon, Tanzania). The color of each cell in the graph represents the mean value of the investigated property of all forest stands within one cell. Note that the AWP axes are scaled differently between the ecoregions (d-f). All other ecoregions, maximum value and standard deviation per property you can find in the Appendix (Figs. A8-A10). See Appendix Fig. A4 for Figures d) - f) with common AWP legend.

We also analyzed forest productivity (Fig. 6: d, e, f). Forest stands with high AWP values have one pattern in common. High productivity goes along with low height heterogeneity and medium LAI in the Amazon and Tanzania ($4-7 \text{ t}_{odm}\text{yr}^{-1}\text{ha}^{-1}$) and high LAI in Germany ($5-13 \text{ t}_{odm}\text{yr}^{-1}\text{ha}^{-1}$). Lower AWP values of forest stands in Germany can be explained by the shorter vegetation period. In contrast to the other ecoregions for the forests in the Amazon, we also observe highly productive forests with large LAI (>10) and high height heterogeneity (>10 m), analogous to the biomass.

For Germany, it is remarkable that we observe only a few forest stands with medium

height heterogeneity (6 m - 10 m) and medium LAI (2–10). Forest stands with these properties have low AWP values. These forests are next to an area in the state space without forest stands (white area), possibly due to the negative productivity of trees. For species evenness (Fig. 6: g, h, i), we got no clear trends (like in the first and second row). Above a certain LAI (>3), we find Tanzanian forests with high species evenness (0.5 - 1). The highest evenness values (>0.7) can be found at the outer edge of the envelope.

The analysis of biomass, AWP and species evenness in German forests (Fig. 6: a, d, g) shows that forests with height heterogeneity smaller than 3 m and LAI larger than 6 have on average a lower evenness, besides all these forests have a high biomass and a high productivity. Results for the standard deviation and maximal biomass values, productivity and species evenness can be found in the Appendix (in Fig. 6, we analyzed mean values; for details, see methods and Appendix Figs. A8–A10).

2.4.3 Comparison of structure function relationships for different ecoregions

In the previous Sections, we examined structure-function relationships for different ecoregions. Here, we directly compare the structure-function relationships for three ecoregions (Amazonian, German and Tanzanian forests) to explore how generally the derived relationships apply (for comparisons for all ecoregions see Appendix Figs. A11-A13). Specifically, we compare mean biomasses (blue points in Fig. 7: a, b, c), mean AWP (red points in Fig. 7: d, e, f) and mean species evennesses (green points in Fig. 7: g, h, i) of forest stands that have similar states (according to the four structural properties used in the Figures above) but are from different regions. We show them in 1:1 graphs. We consider forest states as similar if they have similar maximal height, height heterogeneity, LAI and basal area (details in Section 2.3.4).

We observe a strong correlation for the biomass (high R^2 value). The biomass of forest stands with similar properties are not identical (not on the 1:1 line).

For the AWP (Fig. 7: d, e, f), we see a good correlation. The four structural dimensions are sufficient to find relations between AWP for different regions but less effective than between the biomass.

We see no correlation in the evenness relationships for the different regions (Fig. 7: g, h, i). That indicates that we may need more information in addition to structural properties to get a better correlation.

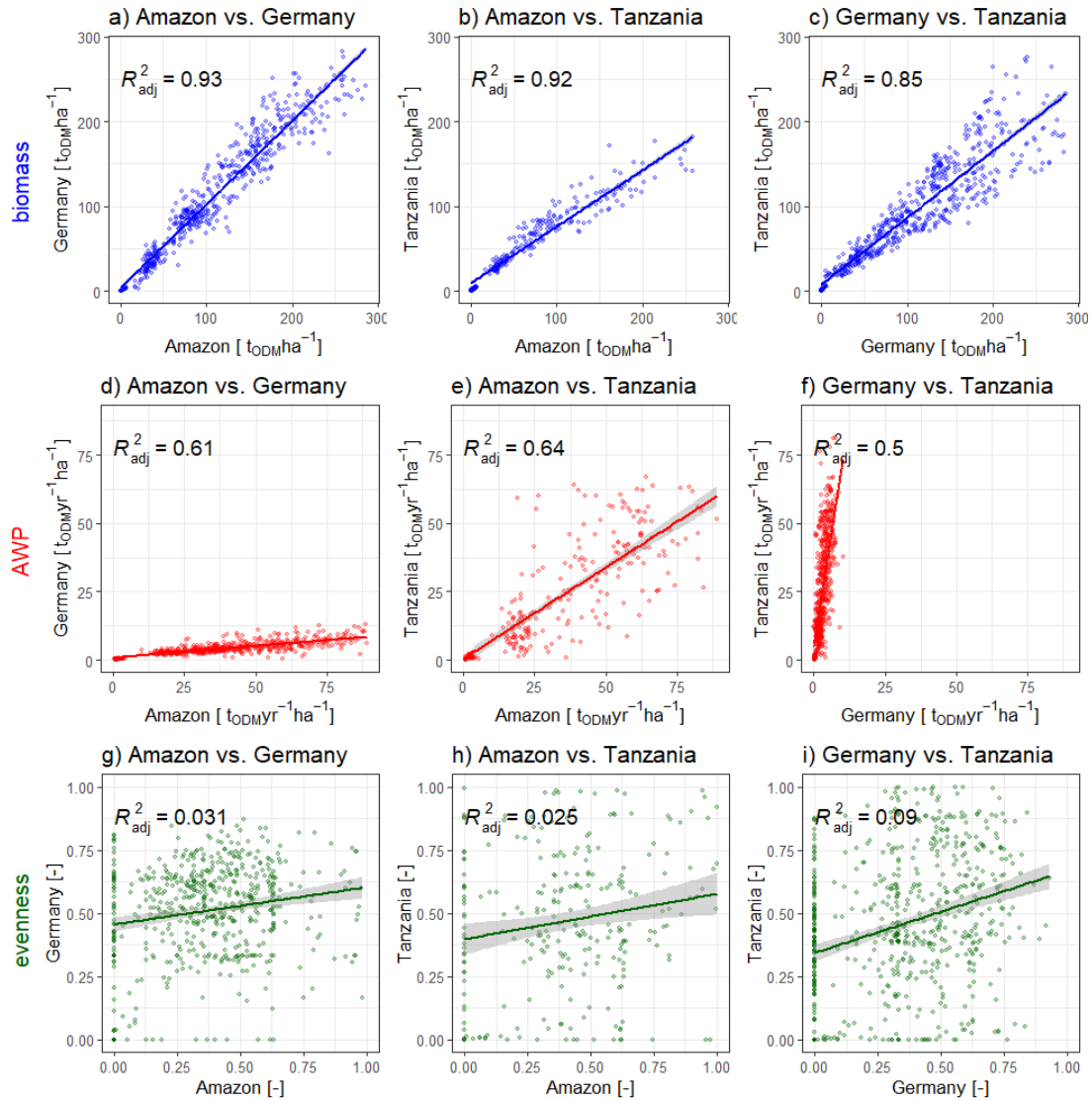


Figure 7. Comparison of biomass, aboveground productivity and species evenness derived from forests with a similar state space (by a 2% quantile of the four structural properties LAI, basal area, tree height heterogeneity and maximum tree height). Each graph compares forest stands out of two ecoregions. We show the pairwise comparisons for three illustrative ecoregions. Each point represents the mean values of the investigated functional characteristics.

2.5 Discussion

In this paper, we explored the Forest Factory 2.0 a new open source software tool to simulate and analyze forests from different biomes on earth. We demonstrated several benefits of the approach and provide insights into how this method can increase our knowledge on structure-function relationships of forests and overall forest functioning. Breaking with the tradition of investigating the development of individual forest

stands over time, we used the Forest Factory 2.0 as a tool to gain knowledge about forests by analyzing the state space of forests, resulting from species pool and environmental factors.

The simple algorithm of the Forest Factory allows comparison of a large number of forest stands from different biomes (3 million forest stands per hour) generated with the same process-driven architecture. This also provides a causal understanding of forest structure-function relationships (as we showed in Section 3). In this manuscript, we present a method to investigate the relationship between structure (maximum height, basal area, LAI, height heterogeneity) and productivity (biomass and AWP) or biodiversity (species evenness) of forests. With the Forest Factory 2.0, it is also possible to analyze other forest properties, such as diameter increment or net ecosystem exchange. Additionally, it enables us to explore the role of other more complex structural characteristics like stem size distribution or height-layer specific information. With this systematic approach we could investigate the causes of the differences and similarities of forest stands e.g. why forests with similar structure show different or similar biomass or productivity values. This could allow us to calculate transfer functions for structure-function relationships of forests from one ecoregion to another (outlined in Section 2.4.3). Here, we generated forest datasets for seven forest regions to illustrate the approach. It is also possible to use other parameterizations from other forest models to generate forests for additional regions.

With this approach it is not only possible to create forests that already exist but also could occur. Using the Forest Factory 2.0 to create forest states beyond the currently existing ones provides a fuller understanding of forests beyond the constraints of empirical data such as national forest inventories or remote sensing data. Some of these forest states may be due to current changes in disturbance regimes or management, and for some forest states it may not even be clear which successional or disturbance pathways will lead to them. In addition to the promising research area of realistic forest selection, it is also interesting to study forests that have almost no or even negative productivity. This analysis can be used to identify stressed forest stands (in forest inventories) or generally describe and understand the state space of stressed forests. This might help to detect potential regime shifts and to explore adaptive capacities of forests and forest ecosystems. As seen in Fig. 6d, there are forests with medium height heterogeneity and LAI that have low productivity. These forests are next to an area in the state space without forest stands (white area). White areas may indicate that forest stands in this area of the state space have negative productivity. Such forests are not generated by the algorithm due to the productivity condition. Further analysis could reveal if this white area represents a transition from forests with low positive to

forests with negative productivity. This area in the state space could give information on the limits of coping capacity of forest stands. If this were the case, these forests could change from being a carbon sink to being a carbon source. In this study we presented how the Forest Factory 2.0 can be used to study region-specific patterns and the ecological mechanisms behind them. Every forest stand consists of many individual trees that are modeled by the selected forest model (here FORMIND). For each single tree, additional information is available (data product of the Forest Factory 2.0). This allows the analysis of specific forest attributes by analyzing the productivity or other properties of each tree in the forest, hence yielding a deeper understanding of forest dynamics. The Forest Factory 2.0 also offers the possibility for jointly addressing research questions from community ecology (organismal aspects, diversity of species and structure) to ecosystem ecology (matter and energy flux aspects, biogeochemical cycles) (Loreau 2010).

For making Forest Factory 2.0 easier to use for different user groups it might be useful to generate forests with only certain tree species for user groups that want to generate lidar data with Forest Factory 2.0. At the moment, forests with certain tree species can of course be sorted out of the data product or generated by changing the parameterization. Possible tree species should be selected during the initialization of Forest Factory 2.0. Perhaps users only want to study multilayer forests, so it would be interesting to allow other height distributions that make these forests more likely (e.g., bimodal height distributions), even if they already exist in the data product. Another direction would be to allow different spatial resolutions for Forest Factory 2.0 if users want to create larger contiguous forests without filtering and rearranging the ones already generated. An additional extension could be to allow different mechanisms for tree placement. It would be possible to remove trees in the virtual forest stands to mimic interventions. Also, we could implement mechanisms which guarantee a denser packing of forest stands and may widen the envelopes in Fig. 3. Nevertheless, the presented envelopes show that we can already cover a broad range of different forest structures with the current approach.

Furthermore, the coupling of the Forest Factory with other modules of FORMIND allows us to explore additional properties and characteristics of the generated forest, for example to derive typical remote sensing data and indexes based on radiative transfer models. For instance, Bruening et al. 2021 use the Forest Factory 2.0 to explore the relationship between lidar profiles and aboveground biomass. It is also possible to combine radiative transfer models with the Forest Factory 2.0 to generate reflection spectra for a huge number of forest stands (Henniger, Huth, et al. 2023). Virtual forests are also used in studies by the remote sensing community (Frazer et al. 2011,

2005; Widlowski et al. 2015). In addition to the typical remote sensing forest variables (point clouds, lidar profiles), the generated forests allow the calculation of additional properties (basal area, LAI, AWP, net ecosystem exchange) also at the tree level. Thus, the presented approach can help to downscale the satellite-imagery-based data and to translate the remote sensing measurements available for large areas to the level of individual trees.

It is also possible to combine the Forest Factory 2.0 approach with other forest models. The new approach of looking at forests in terms of states rather than simulations over time, along with Forest Factory 2.0's free coupling possibility, offers a promising path to compare forest models and learn more about their capabilities and limitations. Specifically, it opens up the possibility of using different forest models to generate different databases of forest stands, as shown in this study with FORMIND, and then analyzing these comparatively using the methods presented. In addition to the possibility of combining the Forest Factory 2.0 with other forest models, the Forest Factory 2.0 is also an additional test for parameterizations. We can analyze forests that cannot be created by the forest succession for which the parameterization was made. These forests may be possible under different environmental conditions (like climate change) or due to disturbances (e.g. fallen trees).

Another possible application is the use of generated forest stands to initialize models simulating forest development (for different forest models) over a longer period of time. With this application it is possible e.g. to analyze the further behavior of these forests under climate change or management scenarios (natural extinction processes or implementation of new species). Again, the advantage is that we can simulate forests with states beyond those that currently exist and gain information that we cannot obtain from inventory or remote sensing observations. With forest models we can analyze the development of these forest stands which allows new ways of analysis. For example, we can explore forest states that are more resilient to climate change and should be pursued in forest management.

The presented way of analyzing forests in a digital universe of processes and mechanisms also offers new possibilities for data scientists. The freely available datasets of generated forest stands can be used to train artificial intelligence models (AI) that estimates additional forest/tree attributes from just a few attributes of forest stands. The resulting relationships could be used to gain a deeper understanding at the level of individual trees from large-scale remote sensing observations. In addition, all relationships shown in the graphs and the data product could be condensed into equations with symbolic regression AIs.

With the Forest Factory 2.0, researchers can generate virtual forests for their needs or

use the open-source forest data to analyze a digital forest universe of forest states.

3 A New Approach Combining a Multilayer Radiative Transfer Model with an Individual-Based Forest Model: Application to Boreal Forests in Finland

3.1 Abstract

To understand forest dynamics under today's changing environmental conditions, it is important to analyze the state of forests at large scales. Forest inventories are not available for all regions, so it is important to use other additional methods, e.g., remote sensing observations. Increasingly, remotely sensed data based on optical instruments and airborne LiDAR are becoming widely available for forests. There is great potential in analyzing these measurements and gaining an understanding of forest states. In this work, we combine the new-generation radiative transfer model mScope with the individual-based forest model FORMIND to generate reflectance spectra for forests. Combining the two models allows us to account for species diversity at different height layers in the forest. We compare the generated reflectances for forest stands in Finland, in the region of North Karelia, with Sentinel-2 measurements. We investigate which level of forest representation gives the best results and explore the influence of different calculation methods of mean leaf parameters. For the majority of the forest stands, we generated good reflectances with all levels of forest representation compared to the measured reflectance. Good correlations were also found for the vegetation indices (especially NDVI with $R^2 = 0.62$). This work provides a forward modeling approach for relating forest reflectance to forest characteristics. With this tool, it is possible to analyze a large set of forest stands with corresponding reflectances. This opens up the possibility to understand how reflectance is related to succession and different forest conditions.

3.2 Introduction

Forests play a major role in the terrestrial component of the global carbon cycle. They account for about 55% of the global above-ground carbon stock (Pan et al. 2011) and represent approximately 40% of the global terrestrial carbon sink (Ciais et al. 2014; Malhi 2010). Forests shape the surface of the Earth by comprising 31% of the land area (FAO 2022) and they influence the energy balance by reflecting and absorbing sunlight. They are important for sustaining biodiversity and provide habitat for 70% of all faunal species (Gibson et al. 2011; Myers et al. 2000; Pimm et al. 2014). Forests exhibit a diversity of spatial structures that can be dynamic due to natural succession,

management or disturbances (Pan et al. 2013).

To monitor the state of forests, the conventional standard practice for foresters and ecologists alike has long been the measurement of forest inventories. Collecting inventories is time-consuming. However, in tropical forests, national forest inventories are often missing. Another approach to monitor forests is based on remote sensing observations, which provide relevant data at large scales. The amount of data is significantly raising with more and more Earth-observing satellite missions launched in the last ten years (Guanter et al. 2015). The spatial and temporal remote sensing observations offer the opportunity to gain a better understanding of forests with respect to their structure and dynamics. Satellite measurements vary in their resolution and coverage. Thus, for global observations, there is a trade-off between the spatial and temporal resolution of satellite (e.g., Landsat, Sentinel) and airborne products. The combined methods of remote sensing and field observations offers the opportunity to gain a better understanding of forests with respect to their structure and dynamics. However, the ecological interpretation of remote sensing observations of forests is challenging, and in many cases still in development.

One way to obtain information from remote sensing measurements concerning target vegetation variables (e.g., Leaf Area Index (LAI), species composition, productivity) is to use models that link the measured remote sensing measurements to the vegetation. Vegetation models have been successfully applied to study change in forests for nearly four decades, many of which differ in their applications. As one example, dynamic global vegetation models (e.g., ED by Moorcroft et al. 2001 and CLM4 by Lawrence et al. 2011), were initially developed to represent the interaction between vegetation and the global carbon cycle as stand-alone simulation models, but also to represent vegetation dynamics in the context of Earth system models, or alongside atmospheric (general circulation models), oceanic and cryospheric modeling frameworks (Maréchaux et al. 2021). These models focus on large-scale applications and they rely on simplifications to reduce complexity and computational demand (e.g., individual species simplified to plant functional types). They do not offer information at the individual tree level. For the analysis of forests in forestry and ecology, there has been a long tradition (Shugart et al. 2018) of using individual forest models (e.g., FORMIND by Köhler & Huth 1998 and LPJ-GUESS by Smith 2001). FORMIND is able to represent the ecosystem dynamics of the forest by simulating each individual tree in a forest (forest gap model). FORMIND allows for the simulation of species-rich forests and also considers the size and age structure of the simulated tree community. At the same time, with increasing computing capacity, there is an opportunity to use these models to simulate large forest areas. Due to the simulation of single trees, they

are also able to consider the heterogeneity of forest structure and dynamics.

An important component in vegetation models is solar irradiance and the competition for light between plants. One simple way to calculate the light climate is based on Lambert–Beer’s law, which is often used by forest models. It describes the decreasing intensity of radiation as it passes through a medium (e.g., tree crowns), depending on the composition of the medium and the height of the layer. Radiative transfer models (RTMs) calculate the light climate in the forests in a more detailed way. They simulate the reflectance, interception, absorption and transmission of light through a canopy. Radiative transfer is influenced, e.g., by the amount of leaves, their characteristics (i.e., amount of chlorophyll and carotenoids, water content), the angle of the leaves struck by light and the angle between the leaves and the Sun. All these parameters are combined by coupled differential equations and allow for the calculation of reflectance of a forest for light of different wavelengths (between 300 nm and 2500 nm, depending on the model) including the reflectance, absorption and transmission of the leaves. Some RTMs are able to provide results for multiple canopy layers, whereas others assume a homogeneous canopy. RTMs are able to simulate the reflectance of the canopy, as it is measured by satellites. Canopy radiative transfer is one of the primary and long-relied-upon mechanisms by which models relate vegetation properties to surface reflectance as captured by remote sensing (Sellers 1985), as radiative transfer in combination with vegetation can be modeled at different levels of complexity. The representation of the vegetation for which the radiative transfer is calculated can range from a simple homogeneous to a detailed and heterogeneous 3D representation of the vegetation structure. The complexity of the solution of radiative transfer problems also varies (Kokhanovsky et al. 2013; Kuusk 2018) from numerical Monte Carlo ray tracing approaches (e.g., Brazhnik & Shugart 2017; Deutschmann et al. 2011) to analytical solutions using, e.g., four stream technology (e.g., Verhoef et al. 2007).

Some of the global vegetation models are coupled with simple RTMs to calculate reflectance for a wavelength from 300 to 2500 nm. The two-stream approximation is used to calculate radiative transfer in CLM4.5 (Bonan et al. 2011), ED2 (Medvigy et al. 2009) and CLM(SPA) (Bonan et al. 2014). Mostly, these models only use a few plant functional types and a low number of canopy layers.

With the new generation of RTMs (such as DART by Gastellu-Etchegorry et al. 2017 and mScope by Yang et al. 2017), it is possible to consider heterogeneous vegetation. The more complex the structure of the vegetation, the more computationally intensive the simulation of light reflectance and the interaction with the vegetation. The same applies to the simulation of vegetation on a global level. As mentioned, global vegetation models must make strong simplifications in order to be able to simulate large areas

in appropriate timespans. Individual-based models describe forest structure in a more detailed way, but they are difficult to apply on a global scale due to the computational requirements. Nevertheless, they endorse the fundamental premise that the structure of forests represents an important factor for ecosystem dynamics that is lost in more aggregated modeling approaches (Shugart et al. 2018).

Individual-based forest models in combination with the new generation of RTMs are therefore a promising approach to consider the complexity of forest structure and species. Their combination will aid in the development of a mechanistic understanding of the linkage between forest reflectance and forest properties such as structure and species diversity. The challenge is to develop an approach which is sensitive to forest structure and species diversity within the current, but ever-increasing, computational constraints both in simulating vegetation and radiative transfer, in order to allow for the analysis of huge forest simulations. Such a tool can also be used to gain a more general understanding of the relationships between reflectance and vegetation properties.

Here, we present an approach by coupling the new-generation RTM mScope with the individual-based forest model FORMIND. We enlarge the application field of mScope and investigate the calculated reflectance spectra of Boreal forests using forests in Finland as an example. Comparing the simulation output with Sentinel-2 data allows us to answer the following questions: How does the concept of forest representation (homogeneous or heterogeneous structure) influence the reflectance spectrum? Can the approach reproduce the variety of reflectance spectra in Finland? Furthermore, how well can we calculate the vegetation indices of the forests with this approach?

3.3 Materials and Methods

For coupling the individual-based forest model FORMIND and the radiative transfer model RTM mScope, we implemented mScope (in an adapted version of Yang et al. 2017) as an additional process in the forest model FORMIND. By using inventories for forest stands in Finland and the forest model, we were able to reconstruct these forests. In combination with the RTM, it was possible to calculate reflectance spectra for the visible and near-infrared range. We then compared the simulated reflectance with measured reflectance spectra from remote sensing observations (Sentinel-2).

To analyze possible applications, different levels of the forest complexity were analyzed and their influence on the reflection spectra was investigated. In addition, several vegetation indices were calculated and analyzed.

3.3.1 Study Site

For this study, we investigated 28 Boreal forest stands in Finland in the region of North Karelia, which are located in an area of about $150 \text{ km} \times 150 \text{ km}$ (see Figure B.1). The inventory data were collected for the FunDivEUROPE project (<http://project.fundiv-europe.eu>, Baeten et al. 2013) in summer (August) in 2012 and again in 2017. Each of the 28 inventory plots had a size of $30 \text{ m} \times 30 \text{ m}$ (Figure 8). The forest inventory contains information on species type, tree positions (x- and y-coordinates) and stem diameters at breast height from all trees. Information about the understory (e.g., shrubs, grasses, mosses) is not provided by the inventory. Based on stem diameter and tree species, other important forest attributes, such as tree height, crown diameter and leaf area index (LAI) are calculated by the forest model FORMIND. The investigated forest stands include as main species *Picea Abies* (Norway spruce), *Pinus Sylvestris* (Baltic pine), *Betula Pendula* (silver birch) and *Betula Pubescens* (downy birch). Information about species richness and evenness, biomass, basal area and LAI can be found in Table B.2 and Figure B.5.

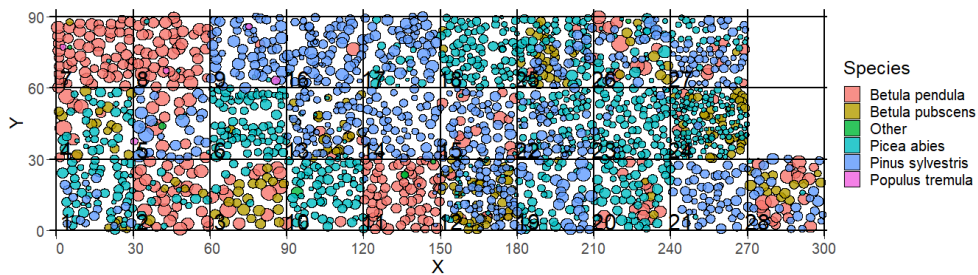


Figure 8. Visualization of the forest inventory of the 28 forest stands in Finland (reconstruction of 2015). Each circle represents a tree and its location in the plot (x and y coordinates). The color of the circles represents the species of a tree and the size of the circle represents its crown diameter. The number in the squares indicates the number of the forest stand and corresponds to the numbering in the FORMIND simulation. The forest stands are shown side by side but are originally distributed over an area of $150 \text{ km} \times 150 \text{ km}$ (a map is shown in Appendix Figure B.1).

For all forest stands, tree size was measured in 2012 and 2017. Here, we took the mean stem diameter (at breast height) of the measured stem diameter values of 2012 and 2017 as a proxy for the stem diameter in the year 2015 (same year as the analysis of Sentinel-2 data Ma et al. 2019) and used these values for the forest reconstruction with FORMIND. We then compared the calculated reflectance spectra with remote sensing observation using atmosphere-corrected Sentinel-2 measurements (Ma et al. 2019) from August 2015. For the simulation of the reflectance spectra, information on observation geometries (Sun and observer, in terms of zenith and azimuth) for each forest stand was provided by Ma et al. 2019.

3.3.2 The Individual-Based Forest Model FORMIND

For the simulation of the 28 forest stands, we used the individual- and process-based forest model FORMIND, which belongs to the model family of individual-based forest gap models. This means that the growth of every single tree is simulated and that individual trees interact with each other. Additionally, FORMIND allows for the simulation of forests with different tree species and also considers the size structure of the tree community. FORMIND can be used for small-scale simulations as well as large-scale simulations (Paulick et al. 2017; Rödiger et al. 2018), e.g., in the Amazon.

The model includes four main process groups: recruitment, mortality, competition (e.g., for light and space) and growth of each individual tree (increment of tree biomass, stem diameter and height). For our investigations, we implemented the RTM mScope as an additional process in FORMIND (in an adapted version in C++).

The stem position (x- and y-coordinate), species information and the diameter at breast height were used as input information in FORMIND. Via different allometry formulas, FORMIND calculates tree height, crown diameter and LAI. This also depends on a set of species-specific parameters and allometry equations. FORMIND has been extensively tested and applied to tropical forests (R. Fischer et al. 2014; Gutiérrez & Huth 2012; Huth & Ditzer 2001; Kammesheidt et al. 2001; Köhler et al. 2003; Köhler & Huth 2004, 2007; Rödiger et al. 2018; Rüger et al. 2008), temperate forests (Bohn et al. 2014; Bruening et al. 2021; Rüger et al. 2007), grasslands (Taubert et al. 2012) and boreal forests (Reyer et al. 2020). The parameterization of Bohn et al. 2014 includes all tree species of the investigated forest stands (North Karelia, Finland) and is used for our simulations on a 30 m × 30 m scale.

3.3.3 Coupling mScope with FORMIND

mScope is an RTM which, on the one hand, can handle several canopy layers, and on the other hand, has a short computation time. For this study, we coupled mScope (Yang et al. 2017) with FORMIND (as a part of the FORMIND code). It is based on Scope (Soil Canopy Observation of Photochemistry and Energy fluxes, van der Tol et al. 2009). The Scope model is a vertical, one-dimensional, integrated radiative transfer and energy balance model, which simulates short-wave reflectance spectra (400–2500 nm) and the fluorescence of homogeneous vegetation. In its original version, it combines two basic RTMs: Fluspect (Vilfan et al. 2016) (on the base of PROSPECT, Féret et al. 2017) for calculations of reflectance, transmittance and fluorescence at leaf level and SAIL-based models (Scattering by Arbitrary Inclined Leaves,

Verhoef 1984) for calculating the radiative transfer in the canopy. Compared to Scope, mScope has multiple layers to include the variation in the distribution of leaves, which enables the representation and simulation of heterogeneous vegetation.

We use mScope (Yang et al. 2017) to simulate the reflectance spectra of forest stands. For calculations at leaf level (reflectance, transmittance and fluorescence), our mScope version uses the model PROSPECT-D (Féret et al. 2017). At canopy level (radiative transfer), a modified version of Scope is used.

For the parameterization of the leaf model, the following attributes are used:

- Leaf structure (number of internal leaf layers [layer]);
- The amount of pigments in the leaf (chlorophyll a and b [$\mu\text{g cm}^{-2}$], carotenoids [$\mu\text{g cm}^{-2}$], anthocyanins [$\mu\text{g cm}^{-2}$], senescent pigments [fraction]);
- Dry matter [g cm^{-2}] and leaf water content [g cm^{-2}];
- Traits describing vegetation structure as the mean and bi-modality of the leaf inclination distribution function, LAI [$\text{m}^2 \text{m}^{-2}$], canopy height [m].

The parameters for the different species were taken from the "CABO 2018-2019 Leaf-Level Spectra Data set" by Kothari et al. 2023 and can be found in Table B.1. These values are generalized values (measurements from Finland were not available). Due to physiological similarities, the species *Betula Pendula* and *Betula Pubescens* are combined to one species group called *Betula* (birches). Additional information that is used is soil reflectance spectra (see Figure B.2) and atmospheric constants, which are taken from Yang et al. 2017.

3.3.4 Representations of Different Levels of Forest Complexity (Heterogeneous Structure)

Using the individual-based approach in forest modeling, it is possible to simulate and describe forest structure at fine scales, which allows for the heterogeneity of a forest to be considered. Individual-based forest models (here, FORMIND) make it possible to gain tree- and forest-specific properties for each forest patch (e.g., $30 \text{ m} \times 30 \text{ m}$) in different height layers (each height layer has a thickness/size Δh) from the bottom/soil up to the top of the canopy.

One important property to calculate radiative transfer is the LAI. FORMIND enables the calculation of LAI distributions for each tree over height (the above-described height layers). In order to determine the species composition, we used the LAI fraction of a species as a measure of its abundance. MScope uses a fixed number of height

layers (in the version of Ma et al. 2019: 60 layers). In our modified version, we use a fixed layer height $\Delta h = 10$ m for a low height resolution and later $\Delta h = 0.5$ m for a high resolution (see Figure B.4). We analyze forests up to forest heights of 50 m. Thus, we use 5 or 100 height layers, respectively, for our calculations. Depending on the structure of the forest, the leaves are located in different height layers. Height layers without leaves do not contribute to the reflectance spectra.

To calculate leaf reflectance and transmittance (using the leaf model PROSPECT-D), the RTM utilized information from the forest model for each layer, which included a leaf parameterization containing leaf properties for each layer. Additionally, the distribution of the orientation of leaves was considered—it was assumed to be spherical for all species—but it is also possible to choose other distributions. MScope also includes observation geometry (Sun and satellite, azimuth and zenith). Vegetation information from the reconstructed simulated forest, which is provided by the forest model, could be processed in different ways and then be transferred to the radiative transfer model. In this paper, we analyze three cases, each resulting in a different representation of the vegetation. The processing differs according to the LAI and according to the species composition (Figure 9).

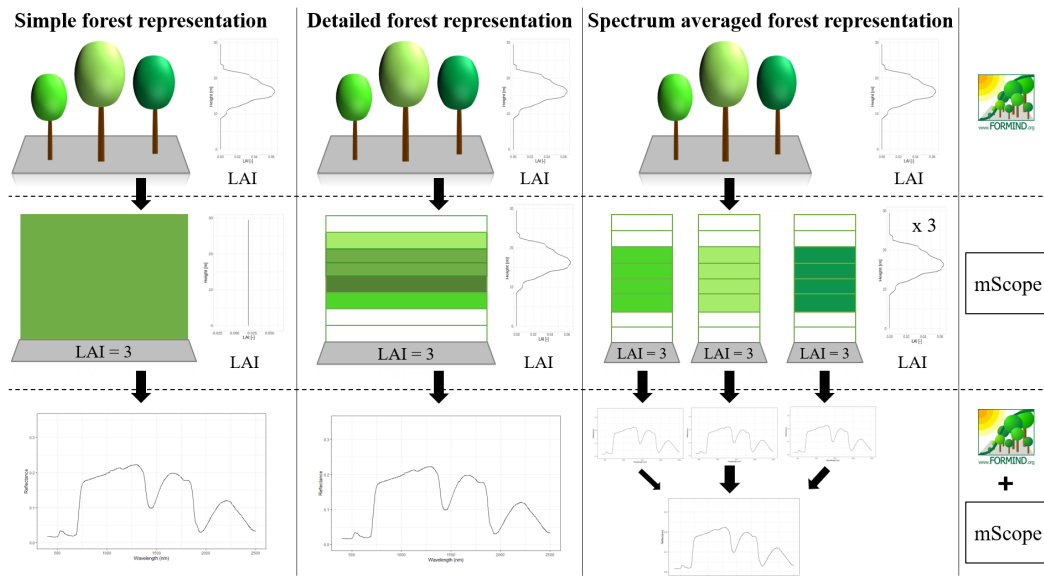


Figure 9. Different concepts of forest representation. Visualization of the different representations of a sample forest during the simulation (first column: simple forest representation, second column: detailed forest representation, third column: spectra-averaged forest representation). We see how, under the given concept, the forest is represented in FORMIND (first row), how it is simulated in mScope (second row) and how the output is built (third row). The sample forest has 3 different species (represented by the different colors). The concepts of representation are described in detail in the text below.

1. Simple forest representation

The simplified forest representation only uses reduced information of the forest. It assumes the same mixture of species and the same LAI for each height layer of the forest stand. The leaf parameterization is calculated by averaging the leaf attributes of the occurring species (weighted by LAI, as a measure of abundance). The LAI of the forest stand is equally distributed among all layers.

2. Detailed forest representation

The detailed representation of the forest assigns to each height layer different mixtures of species and different LAIs. The leaf parameterization for each layer is calculated by averaging the leaf attributes of the occurring species weighted by LAI in the height layer, as a measure of abundance. For each layer of the forest, the calculated LAI of the reconstructed forest stand will be used.

3. Spectra-averaged representation

In this case, the forest is divided into different "sub-forests". In each sub-forest stand, we maintain the total number of trees and the structure of the main forest stand. However, we assume that all trees in a sub-forest stand are of only one

species. Thus, there are as many sub-forests as there are tree species. For each layer, the calculated LAI of the reconstructed forest stand is used. For each of these single-species sub-forests, the reflectance spectra are calculated using the species-specific leaf parameters. The final reflectance spectrum is determined by averaging the species-specific spectra weighted by LAI fraction, as a measure of abundance.

The processed Sentinel-2 observations (Ma et al. 2019) include reflectance values for only 10 wavebands. MScope calculates radiative transfer for wavelengths in the range from 400 nm to 2500 nm (with a resolution of 1 nm). For better comparability with the simulated reflectance profiles, we averaged the simulated reflectance values for the different Sentinel-2A bands (e.g., Sentinel Band 704 nm has a range of 15 nm, so we averaged 15 reflectance values; for more information on bands see Appendix Table B.3). This averaged values are shown as dots in Figures 10, 11 and B.13 and are the basis of the comparisons with Sentinel-2 measurements.

Vegetation indices derived from canopy reflectance are widely used in remote sensing, as they represent proxies for vegetation attributes (e.g., LAI, productivity). We calculated several vegetation indices (NDVI, EVI, MSI, in appendix: NDMI, kNDVI). NDVI is chlorophyll-sensitive. EVI (A. Huete et al. 2002) is responsive to canopy structural variations, including LAI, canopy type and plant physiognomy (X. Gao et al. 2000). We also analyzed kNDVI (Camps-Valls et al. 2021) as a modification of the NDVI. The NDMI is partly correlated with the water content of the canopy (Hardisky et al. 1983). Hunt Jr & Rock 1989 introduced the moisture stress index (MSI, Vogelmann & Rock 1986), which utilizes reflectance wavebands in the SWIR (1550–1750 nm) and NIRS (760–900 nm). Additionally to the vegetation indices, we analyzed the similarity index SAD (spectral angle distance, Kruse et al. 1993, see Appendix Figure B.12 and Table B.4).

In the mScope model, some code adjustments were made to account for the structure of the forest models and forests from the inventory. In forest models, it is possible that there are layers without leaves (vertical gaps). Adjustments were necessary to ensure that these layers had no influence on the reflectance spectrum. MScope calculates the probability of viewing a leaf in solar (P_S) and observer direction (P_O) by assuming a homogeneously distributed LAI in the forest.

$$P_S = e^{k \cdot xl \cdot LAI} \quad (3.1)$$

$$P_O = e^{K \cdot xl \cdot LAI} \quad (3.2)$$

with xl as negative cumulative layer thickness, k as extinction coefficient in direction

of the Sun, LAI as leaf area index of forest stand and K as extinction coefficient in the direction of the observer.

This leads to the situation that the probability is also influenced by layers with an LAI of 0. We have changed the calculation equivalently, allowing for different LAI values for the height layers.

$$P_S = e^{-k \cdot LAI(i)} \quad (3.3)$$

$$P_O = e^{-K \cdot LAI(i)} \quad (3.4)$$

with $LAI(i)$ as leaf area index in height layer i of forest stand, k as extinction coefficient in the direction of the Sun and with K as extinction coefficient in the direction of the observer.

The mScope code also includes a correction of P_S and P_O , which we also considered.

3.4 Results

First, to reduce the complexity of the analysis, we analyzed the reflectance of even-aged forests, where the RTM uses a low resolution (height layer size $\Delta h = 10$ m, Figure 10). In each layer a homogeneous leaf distribution is assumed. The even-aged forest stand number 17, which was dominated by one species, and stand number 5, which contained three species, were used as examples for this analysis. Reflectance was then calculated for simplified, detailed and spectra-averaged forest representations.

There were differences (up to 140%) in reflectance between the detailed (blue) and the simplified (orange) forest representation. The simplified representation consistently produced higher reflectance (especially in comparison to spectra-averaged representation). Both the modeling and satellite measurements show different reflectance spectra for the two forests. We found a higher similarity of reflectance for the detailed representation.

In the next part of the investigation, we increased the represented complexity of the forest by assuming a layer height of 0.5 m (Figure 11). Here, the forest model (here FORMIND) provided mScope with a higher resolution distribution of LAI and species-specific information over height. As in Figure 10, the results are again shown for the simplified, detailed and the spectra-averaged representation of the forests for both example sites.

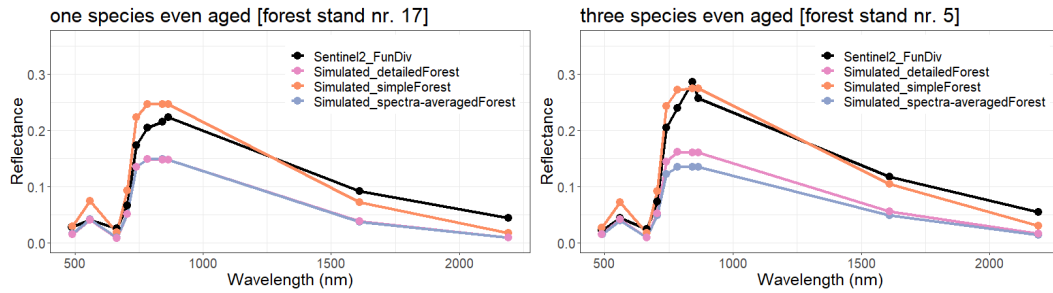


Figure 10. Reflectance spectra for detailed and simplified forest representation by using layers with a size of 10 m. Visualization of the calculated reflectance profiles for two forest stands (numbers 5 and 17) in Finland. Forest stand number 17 (left, mainly one species) and forest stand number 5 (right, three species). Both forest stands are even-aged (small standard deviation of tree heights: 4.6 m and 4.5 m). Each point represents the reflectance value averaged over the specific bands (corresponding to the bands of Sentinel-2). Sentinel measurements are shown in black and simulated reflectance is shown in orange/blue/pink from coupling a forest model (FORMIND) with mScope. We used 10 m height layers. The reflection of all other forest stands is shown in the Appendix (Figure B.6).

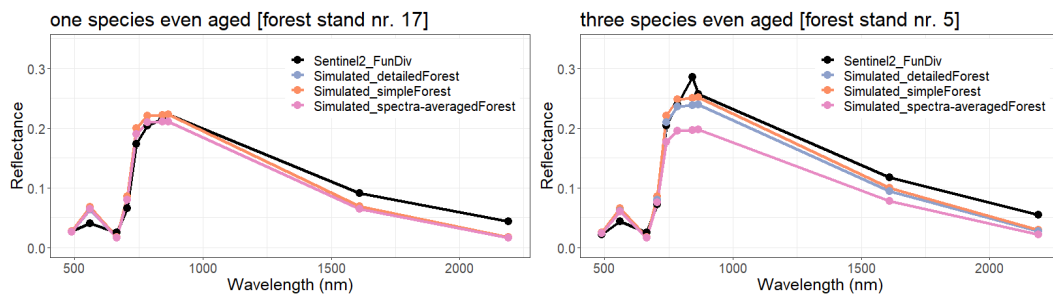


Figure 11. Reflectance spectra for detailed and simplified forest representation by using standard layers with a size of 0.5 m. Visualization of the calculated reflectance profiles for two forest stands (numbers 5 and 17) in Finland according to Figure 9 is shown. Each point represents the averaged reflectance value over the specific bands (corresponding to the bands of Sentinel-2). Sentinel measurements are shown in black and simulated reflectance is shown in orange/blue/pink from coupling a forest model (FORMIND) with mScope. We use here 0.5 m height layers. The reflection of all other forest stands is shown in the Appendix (Figure B.10) Additionally, the reflection for the complete spectra of all other forest stands is shown in the Appendix (Figure B.11). Additionally, we calculated the spectral angle distance for all comparisons (see Appendix Figure B.12).

All three versions produced comparable reflectance spectra (especially for forest stand number 17). The lowest reflectances were produced with the spectra-averaged forest representation (in particular for forest stand number 5 with underestimating the NIR values). For forest stand 17, the spectra-averaged forest representation produces the same reflectance values as the detailed forest representation version. As the forest

stand contains only one species, there is no averaging in the leaf parameters and spectra, and we obtained the same results for these versions. The results for all bands were in agreement with the Sentinel measurements.

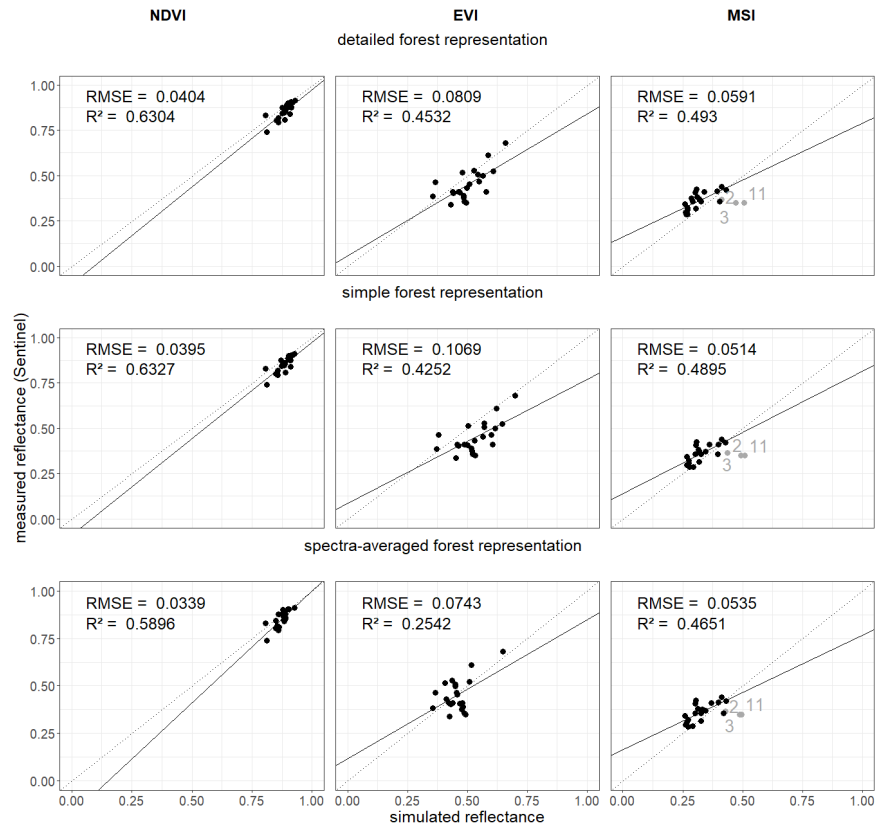


Figure 12. Comparison of vegetation indices. The vegetation indices (NDVI left, EVI middle, MSI right) are calculated from the reflectance values in the different wavebands once for the simulated reflectance spectra and for the satellite measurements. In each row, a different forest representation is used to calculate the results of the indices from the simulated spectra (1. detailed forest representation, 2. simple forest representation, 3. spectra-averaged forest representation; more information about the cases in Section 3.3.3). Each point represents a forest stand in Finland (gray points indicate birch forest stands that are not used to calculate the RMSE and R^2 —see Appendix Figures B.14–B.17). Results for the calculation of the NDMI and the kNDVI can be found in the Appendix (Figure B.19). We excluded five forest stands from our analysis due to inconsistencies in Sentinel-2 measurements (see Appendix Figures B.14–B.17).

The simulated reflectance spectra also enabled the calculation of vegetation indices (see Section 3.3.3). We analyzed NDVI, EVI and MSI (kNDVI and NDMI in Appendix Figure B.19) for each forest stand and for each forest representation (Figure 12). Each were then compared with indices calculated using the satellite observations. We obtained different results for all three forest representations when analyzing NDVI,

EVI and MSI. Lower R^2 and higher RMSE values are obtained for MSI. Measured Sentinel-2 values were close to each other. NDVI values from simulated reflectance spectra were within small ranges. We found an R^2 of 0.63 (detailed forest representation) when comparing simulated and measured NDVI values. For the EVI, there is a larger range of values. EVI led to a lower R^2 (about 0.45) and higher RMSE (0.08) compared to NDVI. The MSI of birch forest stands was overestimated (gray points). Detailed and simple forest representation show similar results for all three indices.

3.5 Discussion

In this work, we developed a new approach to study forest reflectance for radiation in the visible and near-infrared spectrum. For this, we coupled the individual-based forest model FORMIND with an adapted version of the radiative transfer model mScope. We then used the coupled models to reconstruct 28 forest stands in Finland and to calculate reflectance spectra for each. We analyzed three different concepts of forest representation: simple, detailed and spectra-averaged.

When we compared the simulated reflectance spectra with the Sentinel measurements, the best results were achieved for the detailed forest representation. However, the measured reflectance of forest stands with similar forest structure and species mixture shows large differences in five cases (Appendix Figures B.14–B.17). The analysis of these cases (outliers) suggests that factors other than LAI distribution and species composition are here responsible for the differences such as limitations in the atmospheric correction or overlapping of tree crowns in the neighborhood of the forest stands. In addition to this, the approach shows potential for improvement in the sensitivity of simulated reflectance for Sentinel-2 bands B01, B03 and B04. The quality of the simulated reflectance spectrum does not depend on certain species or forest structures (we did not find general relations). This study provides a baseline for further research. The coupling of individual-based forest models and multi-layer RTMs opens up the opportunity to analyze a vast range of forests with various structure and species mixtures and to gain a deeper understanding of the reflectance spectra of complex forests (e.g., influence of tree allometries, leaf parameters or role of understory).

An important aspect of our study is the representation of the forest in the RTM. The simple and detailed forest representations use an averaged leaf parameterization for each height layer (using the LAI of the occurring species as weighting factor). In the spectrum-averaged version, we simulated each occurring species as a monoculture forest and afterwards averaged the resulting reflectance spectra (using the LAI of the occurring species as weighting factor). Despite the non-linear nature of the RTM, the best results were obtained when the input leaf parameters were averaged (simple and

detailed concept). Less satisfactory results were obtained when the output reflectance (reflectance spectra for each species) was averaged (spectra-averaged concept). We obtained similar results for simple and detailed forest representation. The NDVI values were all within a smaller range. As only a few of the analyzed stands had an LAI below 2.5, we also observed a saturation of the NDVI values (A. R. Huete et al. 1997). Forest stands covering a broader spectrum of LAI values will allow for a more general comparison of satellite-based and modeled indices and should be conducted in future studies. For the EVI values, a lower correlation and a higher RMSE compared to the NDVI analysis was observed.

A challenge for the parameterization of radiative transfer models is the selection of suitable parameters (e. g. for leaf attributes, soil and leaf angle distribution). There are a large number of measurements available that include different leaf parameters. However, the leaf parameters of each species can vary depending on the site, the position of the leaf within the canopy, the day of the year of the measurement and environmental factors (Bussotti & Pollastrini 2015). Therefore, leaf parameterizations from sites with the most comparable environmental conditions should be used. A sensitivity analysis (Jacquemoud & Ustin 2019; Jacquemoud et al. 2009; Kattenborn et al. 2019, 2022) was used to analyze the influence of leaf parameters on the reflectance spectrum. In particular, higher sensitivity (Kothari et al. 2023) is observed for those parameters that influence the visible light spectrum (e.g., pigments). Using hyperspectral data, this approach can also be used to fit species parameters.

For soil reflectance, often a wet soil type is assumed (due to a lack of data) and, in this study, we followed this approach. Nevertheless, it is also possible to model the soil reflectance spectrum with an additional model (e.g., the BSM model by Verhoef et al. 2018).

In this study, we developed a forward modeling tool for connecting forest reflection with forest properties. There are further interesting analyses possible based on this approach. One example may be to analyze more complex forests, such as tropical forests. The information about reflectance can be used as an addition to, e.g., LiDAR measurements, to analyze forest structure and functions. It is useful to point out here that the forest model is not only able to investigate structural information but is also able to calculate characteristics of forest dynamics such as productivity. The combination of height-dependent information about forest structure with the information about light reflection spectra may give sufficient information about structure and species composition, resulting in the capability to derive, e.g., estimates of current carbon pools. In addition to the work by Rödiger et al. 2019, the presented approach makes it possible to improve the matching of satellite measurements (e.g., LiDAR profiles) to forest sim-

ulations considering spatially heterogeneous environmental and ecological conditions. As a result, it can improve the carbon estimates for large regions. Please note that the presented approach could be used to derive simulated LiDAR profiles (and thus may improve the LiDAR model used in the mentioned study).

Importantly, this approach can also be used to generate a large number of reflectance spectra for forests by simulating forests over time and tracking reflectance spectra. This may allow us to understand the dynamics of reflectance spectra during forest succession. Disturbed forests show similar characteristics as forests in the early and mid-successional phases. We can use this knowledge to characterize disturbed forests based on reflectance. This may help us to distinguish better between natural and disturbed forests.

However, forest simulations also include path dependencies. Not all types of forest may be covered in simulations, which might occur due to management or disturbances. To overcome this, the Forest Factory approach (Bohn & Huth 2017; Henniger, Huth, et al. 2023) generates a broad range of forest states covering various types of forest structures and species compositions. This approach can also be used to identify which forests or forest states provide the same reflectance spectrum, opening up the possibility of the inversion of reflectance spectra. On the one hand, we can relate a reflectance spectrum to a set of different forest structures. On the other hand, we could also attribute a reflectance spectrum to different leaf parameters (Pacheco-Labrador et al. 2019).

These types of studies could also be conducted for different climate scenarios, for different management strategies and regions/biomes (e.g., using the large set of available forest parameterizations for FORMIND R. Fischer et al. 2016; Henniger, Huth, et al. 2023). Lookup tables and artificial intelligence can help us analyze such large sets of forests and their reflectance spectra and, if desired, even offer the possibility to incorporate additional information about the forests using the forest model.

3.6 Conclusions

In this work, we have applied an adapted version of the radiative transfer model mScope to a complex vegetation structure modeled by the individual-based forest model FORMIND. We showed that the weighted averaging of leaf parameters could be a useful approach to simulate reflectance of forests with different species mixtures (simple/detailed representation). The investigated types of forest representation provide good simulated reflectance spectra (for optical and NIR-range) compared to satellite measurements. However, which type of forest representation provides the best results is influenced by forest structure. In respect to vegetation indices, the best results were

obtained assuming the simple or detailed forest representation. Good correlations were found between simulated and measured vegetation indices (especially NDVI). For future studies, we intend to take advantage of the detailed representation of the forest, and plan to study more heterogeneous forest stands, such as tropical forests. In combination with the forest model, many new perspectives emerge that provide the opportunity to better understand the relationship between forest reflectance and forest properties.

4 A new Approach to derive Productivity of Tropical Forests using Radar Remote Sensing Measurement

4.1 Abstract

Deriving forest net productivity (NPP) and carbon turnover time from remote sensing still remains challenging. This study presents a novel approach to estimate forest productivity by combining radar remote sensing measurements, machine learning and an individual based forest model. In this study we analyze the role of different spatial resolutions on predictions in the context of the Radar BIOMASS mission (by ESA). In our analysis, we use the forest gap model FORMIND in combination with a boosted regression tree to explore how spatial biomass distributions can be used to predict GPP, NPP, and carbon turnover time at different resolutions. We simulate different spatial biomass resolutions (4 ha, 1 ha, and 0.04 ha) in combination with different vertical resolutions (20, 10, and 2 meters). Additionally we analyzed the robustness of this approach and applied it to disturbed and mature forests. Disturbed forests have a strong influence on the predictions which leads to high correlations ($R^2 > 0.8$) at the spatial scale of 4 ha and 1 ha. Increased vertical resolution leads generally to better predictions for productivity (GPP, NPP). Increasing spatial resolution leads to better predictions for mature forests and lower correlations for disturbed forests. Our results emphasize the value of the forthcoming BIOMASS satellite mission and highlight the potential of deriving estimates for forest productivity from information on forest structure. If applied to more and larger areas, the approach might ultimately contribute to a better understanding of forest ecosystems.

4.2 Introduction

The carbon exchanges between the land surface and the atmosphere represent the largest fluxes within the global carbon cycle (Ciais et al. 2014). These fluxes are mediated by terrestrial ecosystems, where forests play a dominant role in the biosphere-atmosphere interface with evident impacts on climate via biophysical and biogeochemical feedbacks that affect water, carbon, energy fluxes (Bonan 2008). Therefore, understanding the link between the dynamics of forests (growth, competition, mortality and establishment of new trees) and the carbon cycle dynamics from short to long time scales are of fundamental interest and will improve our understanding of possible trajectories under future climate change scenarios (Friedlingstein et al. 2014).

To understand the fluxes related to forest dynamics and also to land use change, biomass is a central driver as it is a key component of the global carbon cycle and therefore iden-

tified as an Essential Climate Variable by the United Nations Framework Convention on Climate Change (UNFCCC 2004). Linking biomass to carbon fluxes is important to quantify ecosystem services and to develop adapted forest biomass management policies in the context of global warming. In particular, knowledge of biomass in disturbed forests is important because most carbon emissions from land use change are caused by deforestation.

Knowledge of the regional and global distribution of biomass is limited. Forest inventories are important for better understanding processes of forest dynamics and for calibrating remote sensing measurements. Promising knowledge of biomass distribution over large scales can be obtained from airborne and spaceborne measurements (Araza et al. 2023; Xiao et al. 2019). Latest satellite technology leads to increasing progress in quantifying biomass from space (e.g. Avitabile et al. 2016; Quegan et al. 2019; Saatchi et al. 2011). Technological developments are also leading to advances in the spatial resolution of satellite measurements. At the same time, such developments are associated with additional costs and possible additional error sources. These trade-offs raise the question which spatial resolution of satellite measurements are useful to investigate specific forests characteristics.

To better understand the potential of the expected remote sensing data from future satellite missions, we use the well-established forest model FORMIND (R. Fischer et al. 2016; Köhler & Huth 2010; Köhler & Huth 1998). This model was already used in various studies to better understand the link between remote sensing observations and forest dynamics (R. Fischer et al. 2019; Knapp et al. 2018; Rödig et al. 2019). It can simulate the development of carbon stocks, productivity and carbon fluxes from a single tree up to the whole forests.

With the help of the individual based Forest Model FORMIND we will investigate which horizontal and vertical resolutions of remote sensing biomass measurements are best suited to study the link between forest carbon stocks and forest productivity (GPP, NPP, carbon turnover time). We want to investigate these questions in the context of the P-band radar BIOMASS mission by the European Space Agency. The primary objective of the BIOMASS mission is to determine the distribution of forest aboveground biomass (AGB) worldwide (planned resolution 4 ha) and to reduce the major uncertainties in calculations of carbon stocks using radar space borne observation data.

In our analysis, we (i) introduce an approach where we use the forest gap model FORMIND to better understand the potential of the expected remote sensing data of the BIOMASS and subsequent satellite missions. We (ii) will investigate the relationship between horizontal and vertical structures of aboveground biomass and carbon dynamics (GPP, NPP and carbon turnover time) in forests at different spatial scales by using

a boosted regression tree. We (iii) want to investigate whether our research results are robust to different forest types.

4.3 Methods

We used the individual-based forest gap model FORMIND to simulate the vertical and horizontal biomass-distribution over time in an typical tropical forest. These spatial distributions are used to predict GPP, NPP and carbon turnover time (τ).

4.3.1 FORMIND

FORMIND is an individual-based forest gap model that is used to simulate the growth of forests. In high diversity forests, like in the tropics, species are classified into different plant functional types (R. Fischer et al. 2016, www.formind.org). In this study we have analyzed a tropical forest using the parameterization of Knapp et al. 2018. The parameterization has been developed for a tropical lowland rainforest (50 ha megaplot, 1000 m x 500 m) on Barro Colorado Island (BCI), Panama (9.15N, 79.85W), which has been continuously monitored for more than three decades. The inventory provides an important source of information for forest model parameterization (Knapp et al. 2018, Kazmierczak et al. 2014) and ground truthing for remote sensing studies (Lobo & Dalling 2014; Mascaro et al. 2011; Meyer et al. 2013) due to its remarkable spatial and temporal dimensions and the large number of studies associated with it. Thus, the parameterization which was developed for the forest gap model FORMIND has been extensively tested and other sites in south America provide similar results (Rödiger et al. 2017).

The model FORMIND simulates the following processes:

- *Establishment* - Seeds are distributed over the forest area. If light conditions are suitable, new trees can establish and compete for light and space.
- *Growth* - The growth of a tree is determined by its gross primary productivity (GPP), respiration and type-specific physiological parameters.
- *Mortality* - This process is described by a specific mortality rate. If one tree falls, neighboring trees can be damaged. Additional mortality occurs due to crowding in dense stands or due to low stem diameter increments (stress situations).

- *Competition* - One of the main driving factors of tree growth is light. FORMIND calculates the light condition in different height layers of the forest. A small tree in the shade of a large tree receive less light and therefore reduces its carbon production.

The model calculates the gross primary production (GPP), growth and maintenance respiration, and net primary production (NPP) for each tree. Based on tree mortality, carbon stocks and fluxes between atmosphere, forest stand and soil, the carbon balance of a forest stand can be derived.

Simulation setting: We simulated a total area of 100 ha of a tropical forest over a period of 320 years, consisting of 25 independent simulation runs with a resolution of 200 x 200 m each – which corresponds to the envisaged spatial resolution of the main products of the BIOMASS mission (i.e., aboveground biomass). The FORMIND model has a spatial resolution of 20 x 20 m, which enables us to analyze finer resolutions.

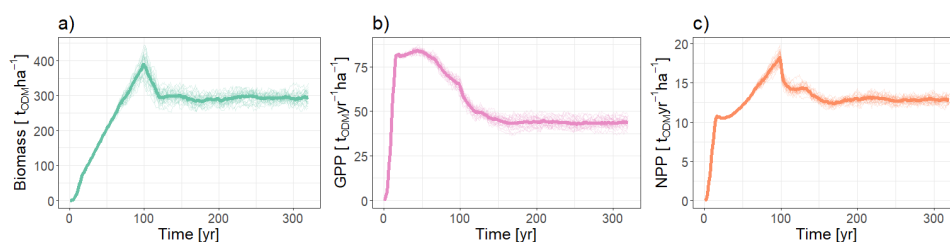


Figure 13. Simulated forest dynamics over time: a) aboveground biomass, b) GPP, and c) NPP. Each thin line represents the simulated forest dynamics for a 200 m x 200 m forest plot. The thick lines show the average over all 25 simulations.

The development of the simulated forests covers two phases: The successional phase in which forest grows from a bare ground until they reach the second, the equilibrium phase after 160 years (Fig. 13).

In the first phase, pioneer species in particular cause a rapid increase in biomass, GPP and NPP. At the peak, many of the even-aged large pioneer species die. Afterwards the forest takes over a diverse height structure of trees and a higher diversity in species, which is typical for mature forests.

4.3.2 Vertical biomass distribution

In our analysis, we divide the forest into different height layers with a layer size of Δh . To simplify the analysis, we assume here that biomass is equally distributed over

the height of a tree (Burt et al. 2021). Figure 14a shows an example of the biomass distribution in different height layers ($\Delta h = 20$ m) of a tree. For the tree A (with a height of 56 m) the same amount of biomass is allocated in the height layer from 0-20 m and 20-40 m (Fig. 14a). Since the tree is not taller than or equal to 60 m, there is less biomass in the 40-60 m height layer. With the help of the forest model we can analyze the vertical distribution of the biomass over the whole simulation area and investigate the development of biomass in different height layers during forest succession. Figure 14b shows the simulated biomass according to the different height layers with a resolution of 200 m x 200 m over time. The higher the height layer is located, the longer it takes for trees to grow into it and the longer it takes to reach the equilibrium phase.

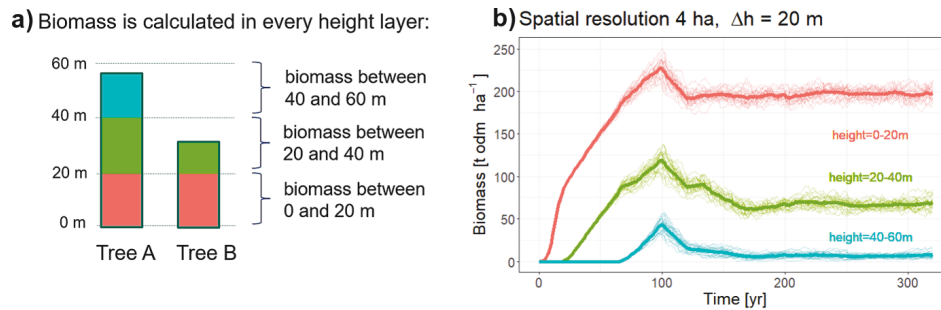


Figure 14. (a) Example for the assumed distribution of biomass over height (here height layers of 20m size) for two trees. The area of the rectangle symbolizes the share of biomass in the corresponding height layer. (b) Development of biomass for each 4 ha forest stand (thin lines) at a total simulated area of 100 ha for different height layers $\Delta h = 20$ m. The mean biomass for the different height layers are displayed by the corresponding thick lines (different colors corresponding to the different height layers).

4.3.3 Linking biomass distribution with carbon fluxes

The forest model FORMIND allows to link the derived biomass with corresponding carbon fluxes. GPP and NPP are derived directly from the model at a spatial resolution of 20 x 20m. In this study we also calculate the carbon turnover time (τ), which can be estimated as follows (Carvalhais et al. 2014):

$$\tau = \frac{\text{biomass}}{\text{GPP}} \quad (4.1)$$

The calculation of carbon turnover time is developed for forests in equilibrium state. For exploration we calculate τ also for forests which have not reached the equilibrium state.

We developed a framework to calculate GPP, NPP and τ in forests at different spatial resolutions. The resolution of biomass information varied both horizontally and vertically. In our analyses, we studied the tropical forest at scales of 200, 100 and 20 meters spatial resolution. On the vertical resolution, we have investigated one vertical layer ($\Delta h = 100$ m) and layers with a size of $\Delta h = 20$ m, $\Delta h = 10$ m and $\Delta h = 2$ m. The simulated forest is the same in all cases (100 ha), only the horizontal and vertical spatial resolution varies.

We applied boosted regression trees to quantify the predictability of forest productivity using the information of vertical and horizontal distributed biomass. Boosted regression trees are a machine learning algorithm using multiple decision (or regression) trees (Elith et al. 2008). Each model was trained in forward stage-wise procedures to predict one variable based on the vertical biomass distribution: GPP, NPP or τ using cross validation. One part of the forest data was used for training and the other part was used for the validation (for more information see Table A1).

The boosted regression tree uses iterative processes to minimize the squared error between predicted values and those of the data set. Hereby, part of the data was used for a fitting procedure and the rest was used for computing out-of-sample estimates of the loss function using the R package *dismo* 1.3-14 (Hijmans & Elith 2021). To get the best regression model, we varied the learning rates, bag fractions and interaction depths of the boosted regression tree algorithm and assumed a Laplace error structure (for more information on the boosted regression tree parameters and training settings see Appendix Tables C.1, C.2, C.3). The obtained best models were used for all further analyses. To determine the best regression model for disturbed and mature forests, we redo the whole training procedure using only data from the equilibrium phase (years 160-320) for mature forests and data before equilibrium phase (0-160 years) for disturbed forests. The best model was then used to predict the mature/disturbed forest data, which were not used for the training.

4.4 Results

4.4.1 Relationships between biomass and GPP, NPP and carbon turnover time τ

Here we analyze succession in tropical forests at the scale of 4 ha (total simulated area 100 ha). As a first step, we analyze the relationships between biomass and GPP, NPP and carbon turnover time (Fig. 15). Each point represents a forest with a given state (includes biomass, productivity or carbon turnover time, and the year of succession).

The relationships between biomass and the target variables are dynamic in the first 100 years. In the last 100 years, when the forest reaches its equilibrium state, the

forest states accumulate in a cloud of points (dark blue color). The fastest increase of target variable with increasing biomass is seen in the biomass-GPP relationship (Fig. 15c similar to Fig. 13). After the increase, the GPP remains stable but biomass still increase until GPP and biomass decrease and forests reach their equilibrium (point cloud). The biomass-NPP (Fig. 15b) and biomass- τ (Fig. 15c) relationships over time have a smaller increase and shorter stagnation of NPP as well as τ . In the biomass- τ relationship, τ continues to increase despite decreasing biomass in the late successional phase, in contrast to the other cases. The highest correlation between biomass and the target variables is found in the biomass-NPP relationship ($R^2 = 0.71$). No linear relationship can be found between GPP and biomass over time ($R^2 = 0.04$).

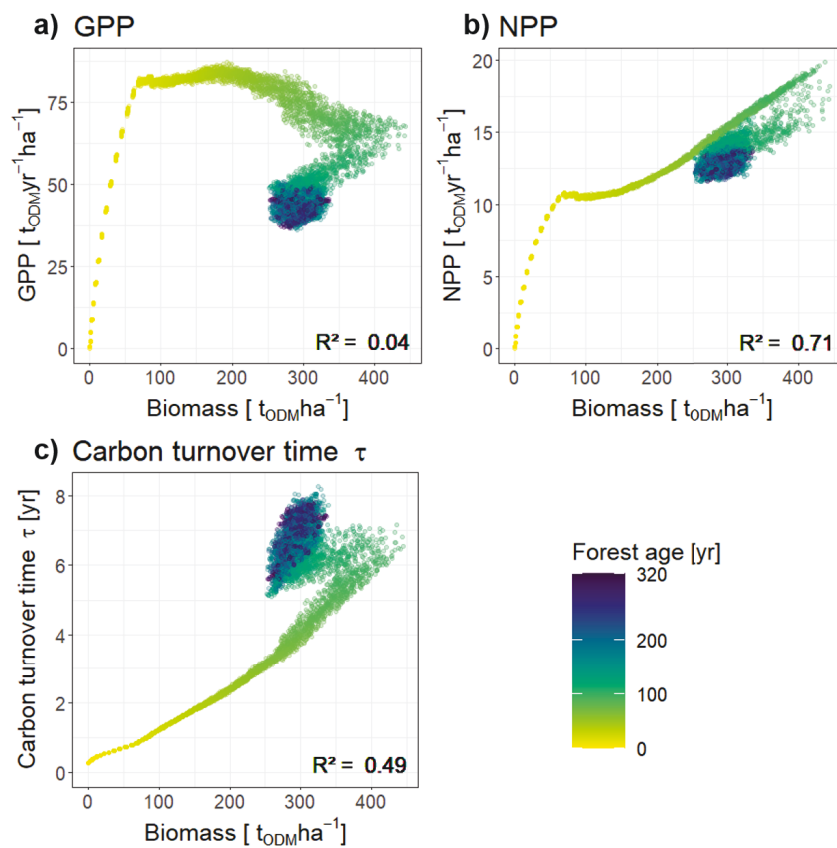


Figure 15. Relationship between biomass and (a) GPP, (b) NPP and (c) τ (carbon turnover time). The results are derived from forest simulations (tropical forest, see Section 4.3). Each point represents the relationship between biomass and the target variable for a 4-ha forest at a certain age (between 0 and 320 years indicated by color). Additionally, the R^2 is calculated (right bottom corner of each graph).

In a second step we examine how relationships change when we analyze biomass in different height layers (Fig. 16, here at the 40-60 m height layer). In all four cases we observe a strong increase GPP, NPP and carbon turnover time in the first years of of the simulation. Forest states in equilibrium state are found in all graphs (see also Fig.

15) after a dynamic phase of the relationships (0 - about 150 years). After about 70 years, when the first trees have reached reach the top layer, the relationships with NPP (Fig. 16b), τ (Fig. 16c) and total biomass (Fig. 16d) show a phase of linear increase. After this phase, NPP and total biomass decrease with decreasing biomass in the upper forest canopy (40 - 60 m). τ , on the other hand, remains stable despite decreasing biomass in the upper forest canopy. GPP (Fig. 16a) increases very rapidly and then decreases steadily until it reaches equilibrium. The highest correlation is found in the relationship between total biomass and biomass in the upper forest canopy ($R^2 = 0.63$). Between the NPP and the biomass in the upper layer we find still an R^2 of 0.39.

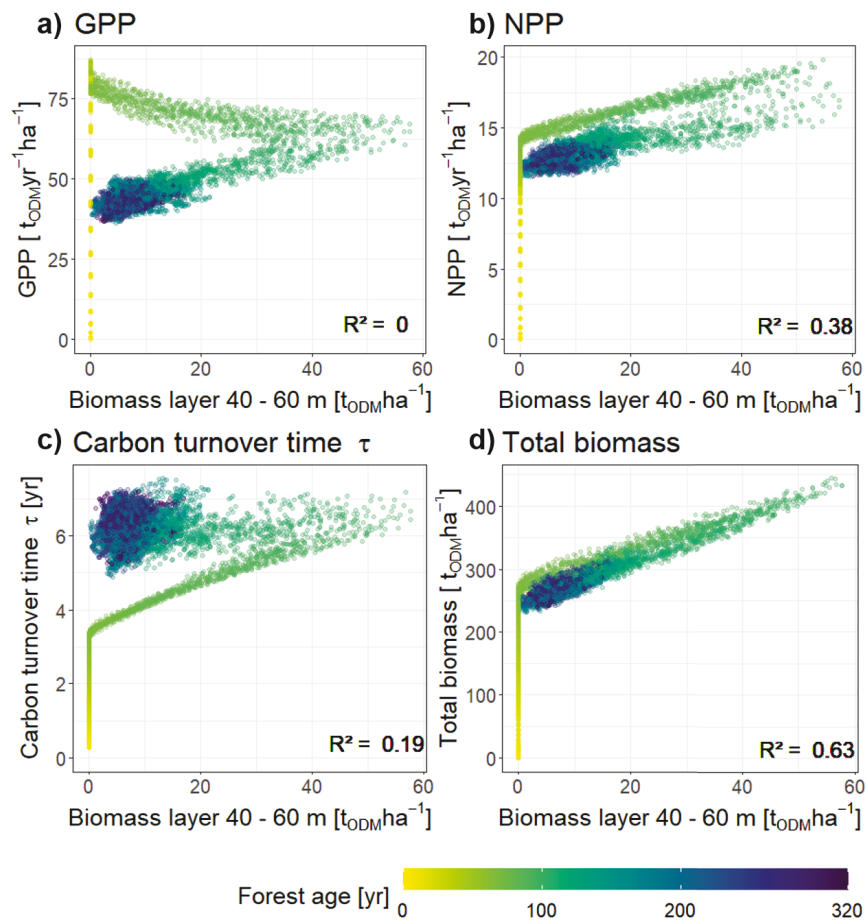


Figure 16. Relationship between the biomass in the height layer 40 – 60 m (with $\Delta h = 20$ m) and (a) GPP, (b) NPP and (c) τ (carbon turnover time) and (d) total biomass over time. The results are derived from forest simulations (tropical forest, see Section 4.3). Each point represents a 4-ha forest stand at a certain age (between 0 and 320 years indicated by color). Additionally, the R^2 value is calculated (right bottom corner of each graph).

4.4.2 Estimation of GPP, NPP and carbon turnover time τ with boosted regression trees using biomass information at different resolution

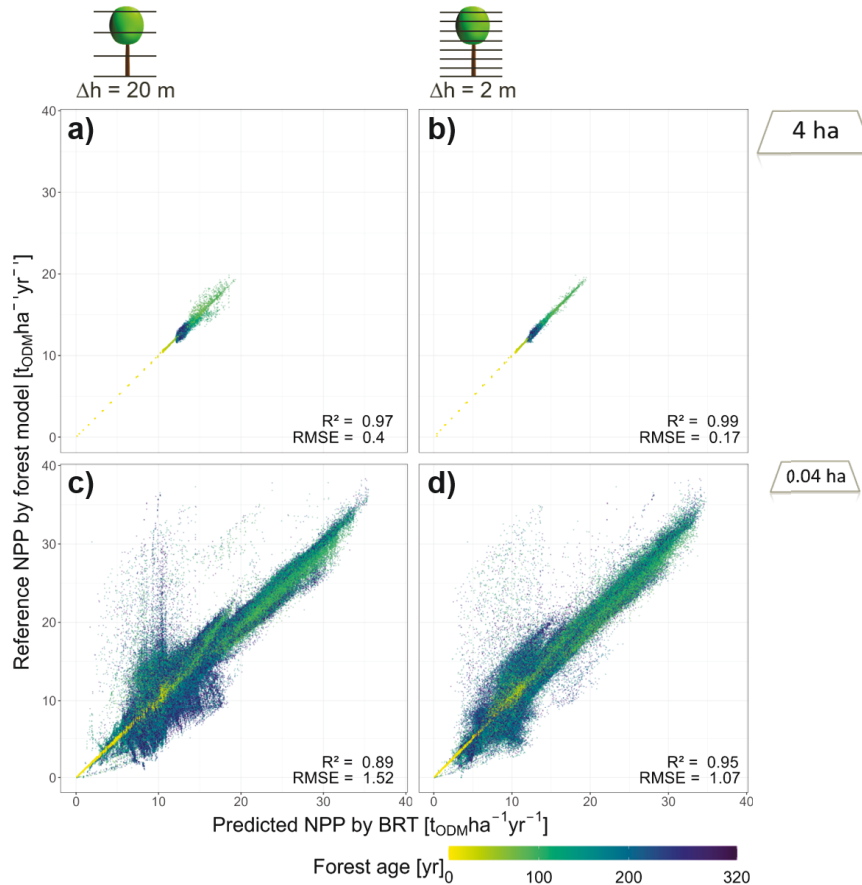


Figure 17. Comparison of predicted NPP derived from the boosted regression tree and reference NPP derived from a forest model. The predicted NPP is derived from the vertical biomass distributions in the simulated forests (in the height layers $\Delta h = 20$ m and $\Delta h = 2$ m). Each point represents a results for (a,b) a 4 ha forest and (c,d) a 0.04 ha forest (bottom row). The forest age is indicated by the color of points. Total simulated area was 100 ha (resulting in the different amount of points in top and bottom row). Results for GPP, carbon turnover times and other resolutions see Appendix C.4 – C.6.

After we explored the relationship of the biomass in one height layer with productivity and carbon turnover time we want to use the information of biomass of all height layers. We tested how a boosted regression tree can use this information and which influence has the spatial and vertical resolution on the estimations. Here we used a boosted regression tree (more information in Section 2.3.) to predict GPP, NPP and τ from the biomass distribution over height with different discretization (with Δh of 20 m, 10 m and 2 m). Additionally, we vary the spatial resolution of the analyzed forest stands (4 ha, 1 ha and 0.04 ha) to find out which cases (spatial area and Δh) provides good conditions for predicting productivity and carbon turnover time.

At first, we want to show an excerpt of our analysis and describe the results of the predicted NPP with the help of the boosted regression tree and biomass information with $\Delta h = 20$ m and 2 m and spatial resolution of 4 ha and 0.04 ha (Fig. 17). For all four cases the predictions of NPP for forests in early succession are near to the simulated NPP (yellow points).

The result for NPP after succession phase (forests after 160 years - blue points) are strongly influenced by the spatial resolution (from 4 ha resolution to 0.04 ha resolution). For 4 ha forests (Fig. 17a,b) we observe a small equilibrium cloud (already seen in Fig. 15 and 16) which is getting smaller with a higher number of height layers ($\Delta h = 2$). For 0.04 ha resolution (Fig. 17c,d) there are more forest plots over all (total simulated area 100 ha) but also more forest plots far from the 1:1 line. With higher vertical resolution (more height layers, $\Delta h = 2$, (Fig. 17d)) there are fewer outliers and estimations for forest stands are nearer at the 1:1 line (similar effect to 4-ha forests in Fig. 17b). We observe high correlation for all four cases ($R^2 > 0.89$). With higher spatial resolution the RMSE and the value range increase.

Comparing the estimated GPP by the boosted regression tree for 0.04 ha forests (Fig.

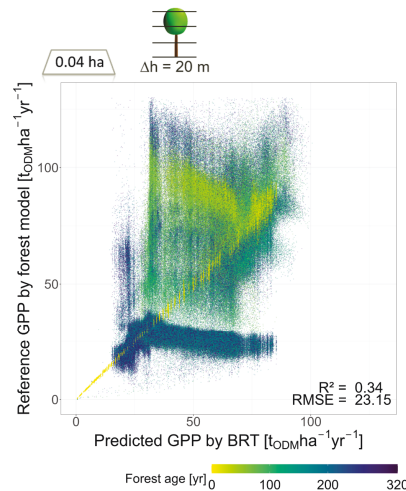


Figure 18. Comparison of predicted GPP derived from the boosted regression tree and reference GPP derived from a forest model. The predicted GPP is derived from the vertical biomass distributions in the simulated forests (with $\Delta h = 20$ m). Each point represents a results for a 0.04 ha forest. The forest age is indicated by the color of points. Total simulated area was 100 ha. Results for GPP, carbon turnover times and other resolutions see Appendix C.4 – C.6.

18) we find a lower correlation ($R^2 = 0.34$ for $\Delta h = 20$ m). We observe a high correlation for young forest stands (yellow points near 1:1 line like in Fig. 17), but for forests with an age between 50 and 100 years the boosted regression tree underestimates GPP. For 0.04 ha forests with an age over 100 years we observe under- and overestimation

of GPP.

Analyzing a broad range of spatial resolutions for the entire forest succession of 320 years (Fig. 19), we got a high predictive power of structural information, as the structure-derived values correlate good with productivity of the forest. Especially for forest stands of 4 ha and 1 ha the correlation is high for all investigated cases ($R^2 > 0.8$). Only GPP shows weaker correlations with for forest stands of 0.04 ha (details in Fig. 18a). The more information we have on the vertical biomass distribution in a forest, the better we can predict GPP, NPP and τ .

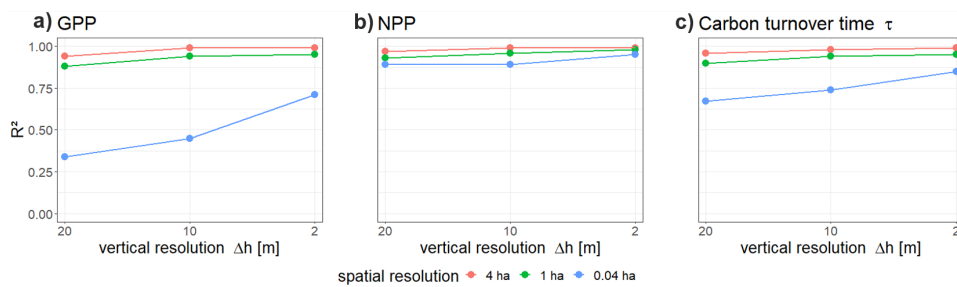


Figure 19. Comparison of correlation between the estimated and the reference (a) GPP, (b) NPP and (c) carbon turnover time (τ). The estimated values have been derived with a boosted regression tree using information of biomass distribution and the reference values were calculated by using a the FORMIND forest model. Each point represents the R^2 value of one comparison with a given spatial - (indicated by color) and vertical resolution (see Fig. 17, 18 and Appendix C.4 – C.6 on the bottom right).

As seen before we observed a high correlation for forest stands of 4 ha and 1 ha when training boosted regression tree with the biomass information of the whole forest succession (Fig. 17 and following in Fig. 20). In the next step, we want to analyze how the correlations change when the forest data set is splitted by assuming two categories of forests: mature forests in equilibrium with an age between 160 and 320 years and disturbed forests with an age between 0 and 160 years (Fig. 20). We assume here that the mature stage is reached after 160 years (see Fig. 13).

We obtain here that the high quality of the prediction with the boosted regression tree is mainly driven by the disturbed forests (Fig. 20a-c is similar to Fig. 19a-c). When analyzing mature forests of different sizes, we observe a different picture. The correlations for GPP, NPP and τ are overall lower (highest R^2 for NPP, lowest R^2 for GPP). In contrast to the analysis of the whole forest data set (Fig. 19) and disturbed forests (Fig. 20a-c), for mature forests (Fig. 20d-f) the weakest correlations are obtained for forest stands of 4 ha and 1 ha size. Mature forest stands with a higher vertical resolution ($\Delta h = 10$ m and 2 m) leading to better estimations of NPP, GPP and τ .

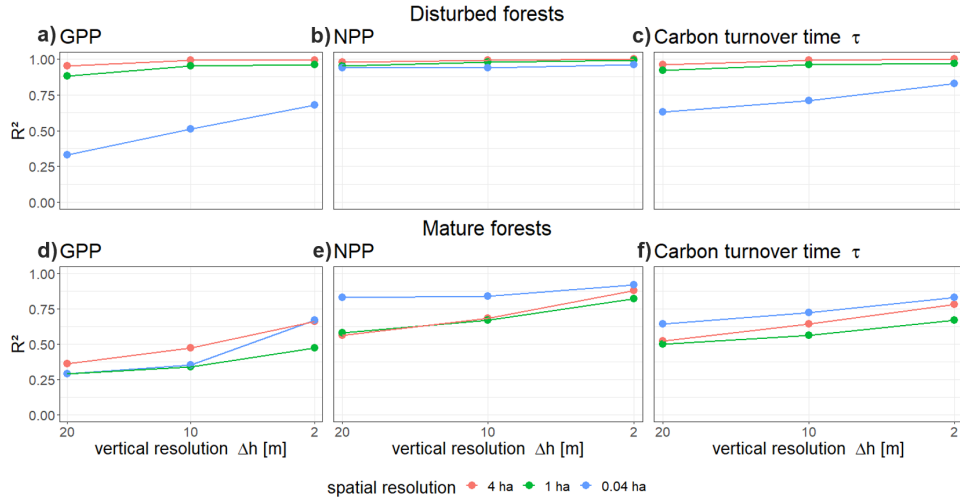


Figure 20. Comparison of correlation between the estimated and the reference GPP, NPP and carbon turnover time (τ) for different types of forests ((a-c) disturbed forests; (d-f) mature forests). The estimated values have been derived with a boosted regression tree using information of biomass distribution and the reference values have been derived using a forest model. Each point represents the R^2 value of one comparison with a given spatial - (indicated by color) and vertical resolution (see Appendix Fig. C.1 – C.3, on the bottom right).

4.5 Discussion

4.5.1 Summary

In our study we investigated a novel approach to estimate GPP, NPP and carbon turnover time for a tropical forest. Here, we used an individual forest model to simulate forest development of a typical tropical forest in combination with a boosted regression tree to analyze the relationships between biomass and productivity as well as carbon turnover time at different vertical and horizontal scales. This approach was used to explore which spatial resolutions (vertical and horizontal scale) are suitable to predict forest productivity for two different forests categories (mature and disturbed forests). When comparing simulated with predicted forest productivity we obtained surprisingly high correlations for GPP, NPP, carbon turnover time at the scale of 4 ha and 1 ha ($R^2 > 0.8$). In case of 0.04 ha scale, carbon turnover time and especially GPP show lower correlations (Fig. 19a,c). These results are mainly driven by disturbed forests (results shown in Fig. 19a-c are similar to the results shown in Fig. 20a-c), which is why we split the forest data. For the prediction of forest productivity of mature forests, we observe lower correlations. Among them, higher correlations were achieved using biomass information from spatial scales of 0.04 ha and $\Delta h = 2$ m ($R^2 > 0.67$).

4.5.2 Obtained trends

Increased vertical resolution (especially for $\Delta h = 2$ m) always leads to increasing correlations, for disturbed forests as well as mature forests. For disturbed forests (forests younger than 160 years) the better predictions results from one-to-one relations between biomass and productivity whereas these relations show ambiguities for mature forests (forests older than 160 years). For forests younger than 160 years (disturbed forests), one single biomass value relates to one productivity value, whereas the occurrence of one-to-one relations disappears for forests with above an age of 160 years (Fig. 15). In case of GPP this appearance of ambiguities already occurs for forest older than 100 years, which explains the lower correlation for GPP of disturbed forests (Fig. 20a).

The decreasing predictability of disturbed forest productivity with increasing spatial resolution (1 ha and 0.04 ha) seems counter intuitive at first glance. However, the higher the spatial resolution of biomass distribution is the more ambiguities occur in the relation to forest productivity and the more difficult it is to predict productivity. This becomes particularly clear for the estimation of GPP (for both category of forests). Nevertheless, higher spatial resolution measurements (at the scale of 0.04 ha) have the potential for improved prediction quality in connection with other sources of information (e.g. with measurements of additional forest attributes by remote sensing). For example height measurements derived from LiDAR (Köhler & Huth 2010) may have the potential of such synergy effects. Not only the quality of the prediction but also the categorization of forest types can be improved by additional information.

4.5.3 Limitations

Due to the strong influence of young forests (age < 160 years) on the boosted regression trees trained for the whole forest data, we decided to realize an additional analysis. So it turned out to be helpful to divide the training data for the boosted regression tree into different forest categories (using the age of a forest as an indicator). This partition of training data leads in our case to two different prediction possibilities (two BRT's, one for prediction of productivity of disturbed forest and one for mature forests). By splitting the forest data with the help of age we assume that complex disturbed ecosystems at the scale of 4 ha can be reconstructed from the set of forest states that occur during the first 160 years of succession. It is possible that other forest landscapes occur in which disturbance patterns have caused a different biomass distribution. Future studies should investigate whether these altered disturbance distributions have an im-

impact on prediction quality. We here propose to use thresholds of biomass values as an indicator for the categorizations of forest types which are provided by the BIOMASS mission anyway.

In addition to that, our approach of categorizing forests as disturbed or mature based on forest succession is only one way to categorize forests. Different classification may yield to different results. We analyzed the robustness of this approach by categorize the forest data and in a next step the analysis could be applied to other forest category selections and forest data compositions. Our study shows that predictions from satellite data might have weaknesses for specific forest categories that may not be directly apparent. In addition, the R packages available for boosted regression tree applications vary in their quality of results and computational requirements. This leads to additional challenges for the comparability of results (see Appendix Tables C.1, C.2, C.3).

4.5.4 Combining forest modeling and remote sensing

This study demonstrates how individual-based forest models can be used to explore the link between remote sensing measurements and ecosystem properties as well as forest processes. Forest models have a long tradition in analyzing the impact of climate change on forest structure, species composition and biogeochemical cycles (e.g. Bugmann & Fischlin 2002; Bugmann & Seidl 2022; Kienast 1991; Solomon 1986). As Maréchaux et al. 2021 pointed out, forest models are also able to conduct virtual experiments beyond empirical investigations (Fyllas et al. 2017; Morin et al. 2018; Schmitt et al. 2020), as well as to test assumptions about ecological processes (Maris et al. 2018; Mouquet et al. 2015). In addition they can be used to reveal potential of improvements and possible gaps in knowledge as well as to guide the design of further field experiments (Medlyn et al. 2016; Norby et al. 2016). Due to increasing confidence in forest models' ability to reflect the true behavior of the system (e.g., due to high agreement in model comparisons - Bugmann et al. 2019; Cramer et al. 2001; Fisher et al. 2018), they are increasingly being used to investigate important relationships between different forest characteristics (e.g., Bohn & Huth 2017; Bohn et al. 2018; Köhler & Huth 2010; Maréchaux & Chave 2017; Morin et al. 2020, 2011; Rödig et al. 2018; Sakschewski et al. 2016; Schmitt et al. 2020).

Studies which use forest models are more and more incorporating remote sensing measurements as the amount of remote sensing data is increasing significantly (Guanter et al. 2015). For example they derive information from remote sensing measurements about the heterogeneity of forest structure (R. Fischer et al. 2019) or estimate carbon dynamics of forests with the use of new allometric models parameterized with tree

crown data derived out of airborne measurements (Jucker et al. 2017). Forest models can be used to analyze correlations of height related biomass measurements to aboveground carbon stocks (Minh et al. 2013). Further, forest models can derive relationships for the interpretation of remote sensing measurements e.g. between forest height and carbon stock (Köhler & Huth 2010). In addition, forest models can be used to create virtual remote sensing data by combining them with forward modeling of radiative transfer or LiDAR to directly link measurements to forests (Henniger, Bohn, et al. 2023; Knapp et al. 2018).

Our study can be used (i) to explore general relationships between forest properties (here for biomass and NPP and possibly for other forest properties), (ii) to explore the potential of satellite measurements (here for the BIOMASS mission) as well as (iii) to combine remote sensing, field data and modeling knowledge (e.g. using inventories or airborne forest measurements in combination with remote sensing measurements on different spatial scales). For missions that are well advanced in planning, such as the BIOMASS satellite mission, the approach provides a first estimate of expected outcomes to derive estimates of forest productivity and carbon turnover time. Future missions may therefore benefit from similar approaches (with respect to improvements in vertical or spatial resolution) in the trade-off analysis of technical improvements.

4.5.5 Advantages of the presented approach

In this study, virtual measurements from future satellite missions are used to derive forest attributes (here GPP, NPP, carbon turnover time). Here, the possibility of further development of Level 3 satellite products is offered by the generation of virtual forests. The method provides good results (e.g. $R^2 > 0.8$ for NPP predictions see Fig. 19b) for the majority of investigated cases, in particular it provides also good results for predictions using high resolution measurements (especially for mature forests with $R^2 > 0.7$, Fig. 20e). This is promising and shows the potential of this novel method for estimating forest productivity from the spatial distribution of biomass within a forest. One highlight of our study is that we provide a method for predicting carbon turnover time, as other observation- and modeling based methods are characterized by large uncertainties (Fan et al. 2020). Additionally we use the advantage of individual based models in our approach, which allows us to investigate forest structure in detail and link it to e.g. forest productivity. In contrast, global vegetation models have a focus on large spatial scales, whereas individual-based models focus on smaller scales, as they consider processes at tree level and can thus also analyze structural dynamics of forests (Maréchaux et al. 2021). This allows the identification of structural properties

and functional characteristics of forest ecosystems at different spatial scales.

An important result of our study is to reveal the potential of tomographic biomass satellite measurements of high resolution in estimating the GPP, NPP and carbon turnover time. The developed approach enable the estimation of forest productivity based on Radar remote sensing measurements. Running et al. 2004 provide estimations of GPP and NPP (with resolution of 25 ha) by using NDVI and FPAR (Fraction of absorbed Photosynthetic Active Radiation) from MODIS satellite products in combination with additional models for the estimation of maintenance respiration (in case of NPP). The quality of these estimates varies for different tropical areas (Avitabile et al. 2016; Mitchard et al. 2013; Rödiger et al. 2019; Turner et al. 2006). In addition, BIOMASS measurements at long wavelengths in the P-band have the advantage of avoiding saturation effects.

4.5.6 Outlook

The developed framework, which uses machine learning techniques in combination with an individual-based forest model to derive forest relationships, can also be applied to other satellite missions (beyond radar). It is able to support satellite missions in exploring the potential of different spatial resolutions (horizontal and vertical) and the relationships between observation patterns and target variables of forests. Also the combination of different measurements (inventory data or other satellite measurements) can be analyzed with the presented mechanistic framework. In that way the here presented framework might be also interesting for remote sensing platforms like the ESA-NASA Joint Multi-Mission Algorithm and Analysis Platform (MAAP, Albinet et al. 2019). The platform is a collaborative project focused on improving the understanding of aboveground terrestrial carbon dynamics by sharing data, science algorithms and compute resources in order to foster and accelerate scientific research.

The used forest model FORMIND is also able to connect forest state and structure to top of canopy reflection (Henniger, Bohn, et al. 2023) or LiDAR measurements (Knapp et al. 2018). In this study we have analyzed a typical tropical forest and a possible next step is to applicate this approach to other tropical forests. Rödiger et al. 2018 applied FORMIND in combination with space-borne LiDAR measurements to the whole amazon and it would be possible to use this framework in combination with our approach. We expect similar trends with different values of GPP and NPP due to climate variations. Beyond that it is also possible to apply the approach to temperate and boreal forest (Bohn et al. 2018; Bruening et al. 2021; Henniger, Bohn, et al. 2023) where we expect even stronger variations of GPP and NPP driven by the shorter vegetation pe-

riods. Henniger et al. use the forest factory approach (Bohn & Huth 2017; Henniger, Huth, et al. 2023) to generate virtual forests instead of simulating forests development over time. This approach generates virtual forests with a forest generator that allows the analysis of vast virtual forest data (700,000 forest stands) for different eco regions, including forest conditions beyond those found in forest inventories. This approach is applied to a variety of other regions around the globe (Henniger, Huth, et al. 2023). Combining the forest factory approach with the here presented framework could help to generalize our results about what resolutions are appropriate for predicting different forest attributes (e.g., biomass and NEE).

4.6 Conclusion

This study presents a novel approach to analyze Radar remote sensing measurements to predict biomass and productivity at different spatial scales. It shows that structural details on forests facilitates a meaningful estimation of forest productivity. Further, the results demonstrate the influence of spatial resolution with differences between disturbed and mature forests. The predictions for mature forests profit from higher spatial resolutions whereas the prediction quality for disturbed forests decrease with higher horizontal resolutions and increase with higher vertical resolutions.

Overall, this study highlights the role of spatial resolution in analysis and emphasizes the need to consider both horizontal and vertical resolution when studying the relationship between biomass distribution in forests and productivity. It also shows that forest relationships do not apply equally to all forest categories. The results highlight the potential of the presented approach and, if applied to more and larger areas, can provide valuable insights, ultimately contributing to a better understanding of forest ecosystems and more informed decision making.

5 Conclusion

It remains a challenge to integrate field and remote sensing data into a coherent picture (Chave 2013; Estes et al. 2018; Levin 1992; Maréchaux et al. 2021) that contributes to the knowledge of forest ecosystem processes, especially in the context of climate change where disturbances are accelerating. Forest models can help to understand and predict behavior of forest ecosystems. Increasing computing power enables new application fields. In a collaboration of forest models, remote sensing, and new methodologies, there is potential to bring together knowledge and measurements at different spatial and temporal scales to understand what human behavior holds for the future.

In this thesis, an advanced forest generator was introduced that is applicable on a global scale. In addition, an individual-based forest model and a multi-layer radiative transfer model were coupled to simulate hyperspectral reflectance of forests. At least a new method for estimating forest productivity derived from RADAR remote sensing measurements was applied to typical tropical forests.

5.1 Main results, limitations and potentials

5.1.1 Virtual forests around the globe

To better understand terrestrial ecosystems, vegetation models simulate forests at different levels of complexity depending on continental or regional applications. Most models, especially forest models, analyze forests with a simulation starting from bare ground while exploring possible paths of forest development. In other ecosystems, there is a tradition of creating virtual environments to understand the behavior of ecosystems in general, to forecast, or to explore the diversity of such environments independent of path dependencies, as in forest succession. Examples of such powerful applications are the use of landscape generators which systematically and automatically generate realistic but simplified representations of land cover in agricultural landscapes and which can provide the input for models analyzing the impact of land use scenarios (Engel et al. 2012; Langhammer et al. 2019). In addition, weather generators are used in water engineering design and in agricultural, ecosystem, and hydrologic impact studies as a means to fill in missing data or to produce indefinitely long synthetic weather series from finite station records (Semenov & Barrow 1997; Wilks & Wilby 1999).

For forests, such a generator was developed by Bohn & Huth 2017, the so-called ‘forest factory approach’ which generate virtual forest stands that possibly could exist in Central Europe and was used for a multidimensional investigation of the relationships between structural properties, species diversity and productivity (Bohn & Huth 2017;

Bohn et al. 2018).

In this thesis (Chapter 2), we advanced the forest factory approach and presented the Forest Factory 2.0, a new open source software tool to simulate and analyze various forest states from different biomes on earth. This greatly accelerated the speed of forest stand generation, expanded the application of the forest factory approach to a global scale, and made the method and data available to the public. We demonstrated several benefits of the approach and provide insights into how this method can increase our knowledge on structure-function relationships of forests and overall forest functioning. Breaking with the tradition of investigating the development of individual forest stands over time, we used the Forest Factory 2.0 as a tool to gain knowledge about forests by analyzing the state space of forests, resulting from species pool and environmental factors.

The simple algorithm of the Forest Factory allows comparison of a large number of forest stands from different biomes (3 million forest stands per hour with a standard notebook) generated with the same process-driven architecture. This also provides a framework for a causal understanding of forest structure-function relationships (as we showed in Section 2.4.2). In this study, we presented a method to investigate the relationship between structure (maximum height, basal area, LAI, height heterogeneity) and productivity (biomass and AWP) as well as biodiversity (species evenness) of forests. With the Forest Factory 2.0, it is also possible to analyze other forest properties, such as stem-diameter increment or net ecosystem exchange. Additionally, it also enables us to explore the role of other more complex structural characteristics like stem size distribution or height-layer specific information. With this systematic approach we could investigate the causes of the differences and similarities of forest stands e.g. why forests with similar structure show different biomass or productivity values. Here, we generated forest datasets for seven ecoregions to illustrate the approach. It is also possible to use other parameterizations to generate forests for additional regions.

This approach is capable of creating not only forests that already exist, but also forests that could potentially exist. Using the Forest Factory 2.0 to generate forest states beyond the currently existing ones provides a fuller understanding of forests beyond the constraints of empirical data such as national forest inventories or remote sensing measurements. Some of these forest states may occur in future due to current changes in disturbance regimes or management, and for some forest states it may not even be clear which successional or disturbance pathways will lead to them.

It is also possible to combine the Forest Factory 2.0 approach with other forest models. The new approach of exploring forests in terms of states rather than simulations over time, along with the free coupling possibility of this approach, offers a promising path

to compare forest models and learn more about their capabilities and limitations. As Bugmann & Seidl 2022 pointed out: to date, we lack a comprehensive approach to quantify the (dis)similarity in models. In particular, the Forest Factory 2.0 opens up the possibility to use different forest models to generate different databases of forest stands, as shown in Chapter 2 with FORMIND, and to analyze them comparatively using the methods presented (especially in Figure 7, including comparison of different models instead of ecoregions).

In addition to the possibility of combining the Forest Factory 2.0 with other forest models, the Forest Factory 2.0 is also an additional test for parameterizations. Parameterizations of forest models are a collection of parameters from field studies and experiments, supplemented by parameters not known for real stands, which are then usually fitted by comparing the long-term simulations of forests (using the parameterization) with inventory data. By creating forest stands that are not part of long-term simulations of forest models or part of forest inventories, it is possible that some processes and their parameters result in characteristics that are not realistic. This may be indicative of problems that may also occur due the simulation of disturbances or changes in climatic conditions and that modelers should be aware of.

Another future application is the use of generated forest stands to initialize models for simulating forest development over a longer period of time. This allows, e.g. to analyze the future behavior of these forests under climate change, disturbances or management scenarios (natural extinction processes, establishment or invasion of new species). Again, the advantage is that we can initiate forest simulations with states beyond those that currently exist and gain information that we cannot obtain from inventory or remote sensing observations. For example, we can explore forest states that are more resilient to climate change and which are interesting for forest management. The presented way of analyzing forests in a digital universe of processes and mechanisms also offers new possibilities for data scientists. The freely available dataset of generated forest stands can be used to train artificial intelligence models (AI) that estimates additional forest/tree attributes from just a few attributes of forest stands. The resulting relationships could be used to gain a deeper understanding of forests on small scales using large-area remote sensing observations. In addition, all relationships shown in the graphs and the data product could be condensed into equations with symbolic regression AIs (Chen et al. 2019).

In contrast to the tradition of investigating the development of individual forest stands over time, we used the Forest Factory 2.0 as a tool to gain knowledge about forests by analyzing the state space of forests. We conducted a structural sensitivity analysis to analyze the relationships between structural properties and forest biomass as well as

productivity and species evenness. In this study we analyze the state space of forests in different biomes and demonstrate the potential of this approach for theoretical ecology. With the Forest Factory 2.0, researchers can generate virtual forests for their needs or use the open-source forest data to analyze a digital forest universe of forest states.

5.1.2 Creating virtual remote sensing measurements

To enlarge the knowledge of forest ecosystems there is the need of mapping relationships between a wide range of forest attributes. An important component in vegetation models is solar irradiance and the competition for light between plants. One simple way to calculate the light climate is based on Lambert–Beer’s law, which is often used by forest models and represents the competition of trees for light. Radiative transfer models (RTMs) calculate the light climate in the forests in a more detailed way.

Some Dynamic Global Vegetation Models necessarily have to include radiative transfer to calculate the energy balance between the vegetation and the atmosphere (e.g. CLM4.5 - Bonan et al. 2011, JULES - Mercado et al. 2007, ORCHIDEE-CAN - Naudts et al. 2015, ED2 - Medvigy et al. 2009). Most of them are coupled with simple RTMs to derive reflectance for a wavelength from 300 to 2500 nm. Often a two-stream approximation is used to represent radiative transfer which only use a few plant functional types and a low number of canopy layers.

Today, with new possibilities to observe forests not only from the ground, and with the availability of a new generation of RTMs (such as DART from Gastellu-Etchegorry et al. 2017 and mScope from Yang et al. 2017), there are new opportunities for individual-based models which include detailed information on the structure of vegetation. They endorse the fundamental premise that the structure of forests represents an important factor for ecosystem dynamics that is lost in more aggregated modeling approaches (Shugart et al. 2018).

The study in Chapter 3 introduced a new approach to study forest reflectance for radiation in the visible and near-infrared spectrum, and could help expand the number of potentially investigable forest attributes and their interrelationships. The approach includes the coupling of the individual-based forest model FORMIND with an adapted version of the radiative transfer model mScope. In this Chapter, the approach is introduced and applied by reconstructing 28 forest stands in Finland and calculating reflectance spectra for each forest stand. We analyzed three different concepts to transfer the detailed forest canopy information of FORMIND into the RTM: simple, detailed, and spectra-averaged. The simple concept of forest representation uses only reduced information about the forest and assumes a forest where leaves are evenly distributed and there is only one average species. The detailed concept uses more information and

represents a forest with multiple layers, with different amounts of leaves, and different species distributions. The spectra-averaged approach calculates the reflectance of as many forest stands as there are species (all forest stands are equal except for the tree species that occur) by using multiple layers with different amounts of leaves in the height layers (like in the detailed concept). The reflectance spectra are then condensed into one average spectrum.

When comparing the simulated reflectance spectra with the Sentinel measurements, the best results were achieved for the detailed forest representation. However, the measured reflectance of forests stands with similar forest structure and species mixture shows large differences in some cases. The analysis of these cases suggests that factors other than LAI distribution and species composition are here responsible such as limitations in the atmospheric correction or overlapping of tree crowns in the neighborhood of the investigated forest stands. In addition, the implementation of mScope in FORMIND could potentially be improved for simulated reflectance values, especially for Sentinel-2 bands B01, B03, and B04.

There are further interesting analyses possible based on this approach. One example is to analyze more complex forests, such as tropical forests, or a gradient from degraded to intact forests. The information about the forest reflectance provided by the approach can be used as an additional information to, e.g., LiDAR measurements, to analyze forest structure and functions. It is useful to point out here that the forest model is not only able to analyze structural information but is also able to calculate characteristics of forest dynamics such as productivity. The combination of height-dependent information about forest structure with the information about light reflection spectra may give sufficient information about structure and species composition, resulting in the capability to derive estimates of productivity.

Rödig et al. 2019 developed an approach to simulate carbon dynamics in the Amazon rainforest by integrating canopy height observations from space-based LiDAR to quantify spatial variations in forest condition and structure. In addition to the work by Rödig et al. 2019, the presented approach makes it possible to improve the matching of satellite measurements (e.g., LiDAR profiles) to forest simulations by including hyper- or multispectral reflectance measurements. This can improve the carbon estimates for large regions. Please note that the presented approach could be also used to derive simulated LiDAR profiles. The approach calculates the reflected radiation of leaves in different height layers with hyperspectral resolution (including the wavelength of LiDAR). Subsequently, it is possible to calculate the intensity of LiDAR beam reflectance in each height layer of the forest, which may allow the improvement of the LiDAR model (Knapp et al. 2018).

Importantly, this approach can also be used to generate a large number of reflectance spectra by simulating forests over time and tracking reflectance spectra. This can allow us to understand the dynamics of reflectance spectra during forest succession. Disturbed forests probably show similar characteristics as forests in the early- and mid-successional phases (West et al. 2012). We can use this knowledge to characterize disturbed forests based on reflected radiation. This may help us to distinguish better between natural and disturbed forests (Bolton et al. 2015).

This work provides a forward modeling approach for relating forest reflectance to forest characteristics. With this tool, it is possible to analyze a large set of forest stands with corresponding reflectances. This opens up the possibility to understand how reflectance is related to succession and different forest conditions.

5.1.3 Using virtual remote sensing measurements

Beyond making new type of remote sensing measurements usable for individual-based forest models, this thesis also gives an example on how to explore virtual remote sensing measurements. In Chapter 3 we explored the possibility to study forest reflectance using an individual-based forest model in addition to a radiative transfer model. This allows the calculation of an additional forest attribute (measurable by remote sensing) by using a forest model. Now we take it a step further. We investigate another forest attribute (which will be measured by a future satellite mission) and use this attribute to estimate more complex forest characteristics.

In the study of Chapter 4, virtual measurements from future satellite missions are used to derive important forest attributes (here GPP, NPP, carbon turnover time). An individual forest model (FORMIND) was used to simulate forest development of a typical tropical forest (on Barro Colorado Island, Panama) in combination with a boosted regression tree to analyze the relationships between spatial biomass distribution and productivity as well as carbon turnover time at different vertical and horizontal resolutions. This approach was used to explore which spatial resolutions of biomass (vertical and horizontal resolution) are suitable to predict forest productivity for two different forests categories (mature and disturbed forests). Here, the possibility of advancing satellite products (here biomass) to productivity estimates is offered by the generation of virtual forests, based on forest succession.

There are studies investigating the potential of RADAR-derived biomass measurements to estimate forest productivity (Bergen & Dobson 1999; Le Toan et al. 2004). In this study, we show how this approach could be applied to real RADAR remote sensing measurements (BIOMASS mission will be launched in 2024). Additionally, we investigate the quality of productivity estimates in relation to the design of the satellite

mission in terms of the resolution of biomass measurements.

The investigated method provides good results for the majority of investigated cases (e.g. $R^2 > 0.8$ for NPP predictions, Figure 19), in particular it provides also good results for predictions using high-resolution measurements (especially for mature forests with $R^2 > 0.7$, Figure 20). This is promising and shows the potential of the novel approach for estimating forest productivity from the spatial distribution of biomass within a forest. One innovation of the study lies in providing a method for predicting carbon turnover time, as other methods are characterized by large uncertainties (Fan et al. 2020). Additionally individual-based models allows to investigate forest structure in detail and link it to e.g. forest productivity in contrast to global vegetation models which provide a less detailed description of forest structure and operate on larger spatial scales (Maréchaux et al. 2021). This allows the identification of structural properties and functional characteristics of forest ecosystems at small spatial scales.

An important result of this study is the demonstration of the potential of high-resolution tomographic satellite measurements of biomass (at the spatial scale of 0.04 ha and vertical scale of $\Delta h = 2$ m) for estimating GPP, NPP and carbon turnover time. The developed approach enables the estimation of forest productivity based on RADAR remote sensing measurements. Running et al. 2004 provide estimations of GPP and NPP (with resolution of 25 ha) by using NDVI and FPAR (Fraction of absorbed Photosynthetic Active Radiation) from MODIS satellite products (multispectral measurements) in combination with additional models for the estimation of maintenance respiration (in case of NPP). The quality of these estimates varies for different areas (Avitabile et al. 2016; Mitchard et al. 2013; Rödig et al. 2019; Thurner et al. 2017). In addition, BIOMASS measurements using long wavelength P-band radiation have the advantage of avoiding saturation effects and are independent of day/night and cloud cover.

There are also some limiting factors in the design of the study. In this study we assumed that forest structure of disturbed forests is similar to the forest structure of early- and mid-successional forests. Based on this the used forest data of disturbed forests cover a wide and continuous range of biomass (bare ground to maximum biomass). Disturbances like fire and wind are influencing the age structure of forests and consequently the structure and biomass distribution of landscapes. It is possible that forest landscapes occur in which disturbance patterns have caused a different distribution of biomass. Future studies should investigate whether such altered disturbance distributions have an impact on prediction quality. Another idea of the study is to use thresholds of biomass values as an indicator for the categorizations of forest types which are provided by the BIOMASS mission anyway. For this study, this might imply that forests with a biomass between 250 and 350 $t_{odm}yr^{-1}ha^{-1}$ can be interpreted

as mature forests and disturbed forests have a biomass between 0 - 250 and 350 - 450 $t_{odm}yr^{-1}ha^{-1}$ (for better understanding see figure 13). Note that this concept seems to be suitable for a horizontal resolution of 4 ha, at other resolutions different thresholds for disturbed and mature forest might be more appropriate. Another interesting investigation would be to use other categories of forest types - the approach of classifying forests as disturbed or mature based on forest succession is only one way to categorize forests. Other classifications may yield different results.

The developed framework, which uses machine learning techniques in combination with an individual-based forest model to derive forest relationships, can also be applied also to other satellite missions (beyond RADAR). The presented approach is able to support satellite missions in exploring the potential of different spatial resolutions (horizontal and vertical) and the relationships between observation patterns and target variables of forests. Also, the combination of different measurements (inventory data or other satellite measurements) can be analyzed. In that way the here presented framework might be also interesting for remote sensing platforms like the ESA-NASA Joint Multi-Mission Algorithm and Analysis Platform (MAAP, Albinet et al. 2019). This platform is a collaborative project focused on improving the understanding of above-ground terrestrial carbon dynamics by sharing data, science algorithms and compute resources in order to foster and accelerate scientific research.

In this chapter we have analyzed a typical tropical forest. A possible next step is to apply this approach to other tropical forests. Rödiger et al. 2018 applied FORMIND in combination with space-borne LiDAR measurements to the whole Amazon. It would be possible to use this framework in combination with the approach of Rödiger et al. 2018. We would expect similar trends with different values of GPP and NPP due to variations in climate and soil conditions. Beyond that it is also possible to apply the approach to temperate and boreal forest (Bohn et al. 2018; Bruening et al. 2021; Henniger, Bohn, et al. 2023) where we expect even stronger variations of GPP and NPP driven by shorter vegetation periods.

The presented approach offers a number of innovations: the exploration of expected remote sensing data, the use of individual-based forest models for preliminary studies of satellite missions, the use of RADAR satellite measurements for the prediction of productivity and carbon turnover times and the exploration of the prediction quality for different forest types. The obtained results emphasize the value of the forthcoming BIOMASS satellite mission and highlight the potential of deriving estimates for forest productivity from information on forest structure. If applied to more and larger areas, the approach might ultimately contribute to a better understanding of forest ecosystems.

5.2 Outlook

Virtual forests around the globe

The Forest Factory 2.0 is already being used by five working groups (University of Maryland, Karlsruhe Institute for Technology, University of Göttingen/Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen, University of Lund and Helmholtz Centre of Environmental Research). It has been applied to several research topics, especially in combination with remote sensing.

Interaction with these groups has revealed potential for novel developments. The software tool (via RStudio) and the data of the generated virtual forests are available as open source. A graphical user interface (GUI) would make the Forest Factory 2.0 easier to use for groups not so familiar with programming.

Such a GUI could also help to make more use cases available or to make parameters easier to customize for users. For example, it could be interesting to generate forests with only selected tree species. Currently, forests with selected tree species can be filtered out of the data product or generated by changing the parameterization. In future, the initialization of the Forest Factory 2.0 may include an option to select the desired tree species. Given the case that users only want to study multi-layer forests, it would be interesting to allow also other height distributions that make these forests more likely (e.g., bimodal height distributions). Even if the desired type of forest is already included in the provided data, it saves the user time and lowers the barrier to use the Forest Factory 2.0 if subsequent filtering of the forest data is not required.

Another interesting research direction would be to allow different spatial resolutions for the Forest Factory if users want to create larger continuous forests without filtering and rearranging the ones already generated. An additional extension could be to allow different mechanisms for tree placement. It would be also possible to remove trees in the virtual forest stands to mimic interventions or management. Another alternative is to implement mechanisms which guarantee a denser packing of forest stands (for e.g. with thresholds for minimum density) which may widens the state space in Fig. 4 (Chapter 2.4.1). Nevertheless, the presented envelopes show that we can already cover a broad range of different forest types and structures with the current approach.

Combining virtual forests and virtual remote sensing measurements

The digital transformation is taking place in various domains, including ecology. Digital Twins are becoming part of the political sustainability agenda (e.g. in the Destination Earth program by the European Commission, Nativi & Craglia 2020) with the

vision to develop Digital Twins for climate, ocean and biodiversity. A Digital Twin of Earth ecosystems is an information system that provides users with a digital replication of the state and temporal evolution of parts of the Earth system (e.g. forest ecosystem), constrained by available observations and the laws of physics (Bauer et al. 2021). These parts, when put together, can result in a Digital Twin of the Earth.

The Forest Factory 2.0 is a tool for generating virtual forests. These forests could be interpreted as potential Digital Twins of forests. This means that they might be duplicates of real existing forests in nature and represent their forest states at a given point in time. Consequently, the generated forests either have a real, existing duplicate in nature or they do not.

If there is a duplicate, the chances of finding potential twins that have similar attributes are increasing with the growing amount of available forest data (especially from remote sensing, but also from field inventories). Finding these real duplicates allows to relate the known attributes, e.g. from remote sensing, to all other attributes that can be calculated with the coupled forest model. Furthermore, it is a starting point for the simulation of these forests over time and under different environmental conditions, making them to Digital Twins in the full sense (Bauer et al. 2021; de Koning et al. 2023).

If there is no counterpart, it could be that this forest state has not yet been reached (due to environmental conditions or lack of disturbance regimes) or because the forest cannot exist in nature, which the approach tries to avoid but cannot completely rule out (related to the description of parameterization testing, see Section 5.1.1). If the forest state has not been reached this gives information about unrealized forest states and enables an out of system analysis (outside of path dependencies). This means that it may be possible to overcome some of the difficulties in studying forest states that may be caused by climate change and invasive tree species, as traditionally forest models often are parameterized without these phenomena.

There are already studies which use virtual canopy surfaces, virtual forest point clouds and virtual trees (Frazer et al. 2011, 2005; Widlowski et al. 2015). These virtual trees or forest stands are not related to further forest attributes and are used e.g. to explore radiative transfer modelling. Using the forest factory approach would allow to have bigger insight into the functions and properties of the virtual forest. The generated forests would allow the calculation of additional properties (basal area, LAI, AWP, net ecosystem exchange) at tree - and forest stand level to explore the connection of remote sensing data and forest functions on small scales. This could improve the mechanistic understanding of forests in general and, more specifically, of forests that has limited access on the ground, such as those in the tropics.

The Forest Factory 2.0 generates a broad range of forest states covering various types of forest structures and species compositions (Chapter 2). Furthermore, the coupling of the Forest Factory 2.0 with other modules of forest models (here we use FORMIND) allows us to explore additional properties and characteristics of forest, for example to derive typical remote sensing information (e.g. RADAR in Chapter 4) and indices based on radiative transfer models (e.g. EVI in Chapter 3). Combining the forest factory approach with the presented framework in Chapter 4 could help to explore about what resolutions of future satellite measurements are appropriate for predicting different forest attributes on a larger scale (e.g., biomass and NEE).

An additional attribute which is measurable via remote sensing is Solar-induced chlorophyll fluorescence (SIF). SIF is a remotely sensed optical signal emitted during chemical photosynthesis. The past two decades have witnessed a strong increase in availability of SIF data at increasingly higher spatial and temporal resolutions, sparking applications in diverse research sectors (e.g., ecology, agriculture, hydrology, climate, and socioeconomics; Sun et al. 2023). SIF data is already used in several studies, which relate these data to forest attributes (Colombo et al. 2018; Hernández-Clemente et al. 2017; Lu et al. 2018; Pierrat et al. 2022). In Chapter 3 we used the radiative transfer model mScope (Yang et al. 2017) which also can calculate fluorescence induced by vegetation. A next interesting step would be to apply this part of mScope to forests e.g. by using FORMIND (like we did in Chapter 3 with reflectance) or by using the Forest Factory 2.0 (Chapter 2). This would enable to relate forest structure to SIF. This shows once again the potential to use the Forest Factory/FORMIND as a kind of interface to embed modules that can calculate worth exploring forest attributes, which enable the correlation between all calculated forest properties. The advantage of calculating SIF in addition to NDVI is that it provides independent, complementary information on seasonal vegetation transitions (X. Wang et al. 2020). It provides a means to monitor photosynthesis and productivity on a large scale in both deciduous and evergreen forests. This is in contrast to e.g. NDVI, which is limited in monitoring changes in photosynthesis in evergreen forests, and other indices that can only relate photosynthesis to the leaf level of the vegetation (Springer et al. 2017).

The Forest Factory 2.0 has been already used to create virtual forests in combination with virtual remote sensing measurements. For instance, a group in Maryland (Bruening et al. 2021) use the Forest Factory 2.0 to explore the relationship between LiDAR profiles and above-ground biomass in North America. In addition to that a working group at the KIT use the Forest Factory 2.0 to create synthetic LiDAR point clouds to train artificial intelligence to estimate forest biomass (Schäfer et al. 2023).

The philosophy to use forest models and the Forest Factory as a kind of interface to cal-

culate e.g. hyperspectral forest reflectance (Chapter 3) and LiDAR signals (described in Section 5.1.2) offers the possibility to link different remote sensing signals. The correlation between different vegetation indices and further between remote sensing signals derived from different remote sensing systems is an interesting and important research question. By using virtual forests, it is possible to create and correlate virtual remote sensing signals of different techniques (e.g. active and passive sensors) for a large number of forests. The FORMIND working group (Helmholtz Centre of Environmental Research) is developing a RADAR simulator which calculate coherence. Coherence is described as the complex correlation between the two incoming RADAR waves, which is an important component of active RADAR measurements. This enables the combination of RADAR, LiDAR and hyperspectral remote sensing measurements, which would be promising. It could provide information on the independence of different remote sensing techniques, and thus on the question of which remote sensing measurements should be combined to estimate specific forest attributes.

Currently in the SIMWALD project (Helmholtz Centre of Environmental Research) the Forest Factory 2.0 is used to create one million virtual forest stands and to calculate the corresponding hyperspectral reflectance spectra. This huge amount of generated data is used to train artificial intelligence (AI) and machine learning to link the created data (virtual forests and reflectance) to real existing forest states. The potential of this unique data set will be demonstrated in a show case where EnMAP measurements (German hyperspectral satellite mission launched in 2022) for selected forest areas in Germany will be linked to generated forests by AI. In addition to hyperspectral measurements, information on forest structure is also crucial and can be captured by RADAR measurements (like in Chapter 4). Therefore, the synergetic use of hyperspectral reflectance and RADAR data will be investigated. Especially after extreme weather events (e.g. droughts) the damage in forests could be identified quickly via remote sensing and necessary management interventions could be planned. The presented studies in this thesis provides the scientific basis for the SIMWALD project and allow an innovative fusion of satellite data with forest models. If successfully applied it should be possible to derive forest stand characteristics from hyperspectral measurements in high spatial and temporal resolution. This would be a new milestone in forest research, especially for the development of adapted forest management concepts and with regard to global climate change and its effects (Begon et al. 2016).

There are many regions for which the Forest Factory is already applicable. With the help of a these parameterizations, it is possible to create a large amount of possible forest states, which go beyond local inventories (e.g. from the Bundeswaldinventur). This allows the construction of regional forest inventories. In Chapter 2, the Forest

Factory 2.0 in combination with FORMIND is applied to seven ecoregions (out of 28 existing parameterizations, see formind.org). In combination with the parameterizations of other forest gap models the number of regions where the Forest Factory could possibly be applied is strongly increased. Historically, on-the-ground forest inventories have been used to document and derive information of forests for each individual tree. With today's remote sensing and computing capabilities, it is possible to conduct forest inventories by satellite. As described in this thesis (Section 1.3) satellites can only measure certain attributes of forests. However, the Forest Factory is a well-suited method to link remote sensing measurements to forest attributes. When the Forest Factory is applied to more ecoregions, there is the potential to generate trillions of forest stands, covering a wide variety of forest conditions for many different regions on Earth. This opens up the possibility of a global forest inventory conducted with the help of remote sensing measurements and the Forest Factory as an interface to link the measurements to the generated forests. Linking remote sensing measurements to generated forests (with their corresponding virtual remote sensing measurements) is a challenging task, especially in terms of computational power. Artificial intelligence such as neural networks are well known for solving assignment problems and combinatorial optimization, especially for vegetation and remote sensing (Kattenborn et al. 2021), and are able to perform such a task. A global forest inventory could be used as a starting point for simulations over time and would help resolve uncertainties in quantifying and understanding the global carbon cycle (Ruehr et al. 2023).

We are at an exciting point in time where the growing amount of remote sensing data and computational resources are opening up new possibilities for forest research. Ultimately, however, we are not yet at the point where we can simulate every single tree on Earth. DGVMs have their strength in simulating competition between different vegetation types at large scales, and individual-based forest models have advantages in representing competition and interaction between individual trees. A major challenge in vegetation modelling is to combine the strengths of both model families to provide a sufficiently detailed simulation of vegetation at large scales. Both types of models attempt to make progress in this direction. This thesis presents tools for bridging the gap between DGVMs and individual-based forest models to take advantage of the new opportunities provided by increasing remote sensing measurements, increasing computational capacity, and upcoming methods. These tools can help to incorporate forest structure and local biodiversity in large scale applications. We hope that this thesis will inspire scientists from different research fields to use the data generated and the methods presented to gain a deeper understanding of forest ecosystems and to help adapt policy instruments to preserve a livable planet for future generations.

References

- Albert, J. S., Carnaval, A. C., Flantua, S. G., Lohmann, L. G., Ribas, C. C., Riff, D., ... others (2023). Human impacts outpace natural processes in the amazon. *Science*, 379(6630), eabo5003.
- Albinet, C., Whitehurst, A. S., Jewell, L. A., Bugbee, K., Laur, H., Murphy, K. J., ... others (2019). A joint esa-nasa multi-mission algorithm and analysis platform (maap) for biomass, nisar, and gedi. *Surveys in Geophysics*, 40, 1017–1027.
- Ali, A. (2019). Forest stand structure and functioning: Current knowledge and future challenges. *Ecological Indicators*, 98, 665–677.
- Ali, A., & Mattsson, E. (2017). Disentangling the effects of species diversity, and intraspecific and interspecific tree size variation on aboveground biomass in dry zone homegarden agroforestry systems. *Science of the Total Environment*, 598, 38–48.
- Araza, A., Herold, M., De Bruin, S., Ciais, P., Gibbs, D. A., Harris, N., ... others (2023). Past decade above-ground biomass change comparisons from four multi-temporal global maps. *International Journal of Applied Earth Observation and Geoinformation*, 118, 103274.
- Argles, A. P., Moore, J. R., & Cox, P. M. (2022). Dynamic global vegetation models: Searching for the balance between demographic process representation and computational tractability. *PLOS Climate*, 1(9), e0000068.
- Argles, A. P., Moore, J. R., Huntingford, C., Wiltshire, A. J., Harper, A. B., Jones, C. D., & Cox, P. M. (2020). Robust ecosystem demography (red version 1.0): a parsimonious approach to modelling vegetation dynamics in earth system models. *Geoscientific Model Development*, 13(9), 4067–4089.
- Avitabile, V., Herold, M., Heuvelink, G. B., Lewis, S. L., Phillips, O. L., Asner, G. P., ... others (2016). An integrated pan-tropical biomass map using multiple reference datasets. *Global change biology*, 22(4), 1406–1420.
- Baeten, L., Verheyen, K., Wirth, C., Bruelheide, H., Bussotti, F., Finér, L., ... Scherer-Lorenzen, M. (2013). A novel comparative research platform designed to determine the functional significance of tree species diversity in european forests. *Perspectives in Plant Ecology, Evolution and Systematics*, 15(5), 281–291. doi: 10.1016/j.ppees.2013.07.002

- Bauer, P., Stevens, B., & Hazeleger, W. (2021). A digital twin of earth for the green transition. *Nature Climate Change*, *11*(2), 80–83.
- Beer, C., Lucht, W., Gerten, D., Thonicke, K., & Schmullius, C. (2007). Effects of soil freezing and thawing on vegetation carbon density in siberia: A modeling analysis with the lund-potsdam-jena dynamic global vegetation model (lpj-dgvm). *Global Biogeochemical Cycles*, *21*(1).
- Begon, M., Howarth, R. W., & Townsend, C. R. (2016). *Ökologie*. Springer-Verlag.
- Belward, A. S., & Skøien, J. O. (2015). Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS Journal of Photogrammetry and Remote Sensing*, *103*, 115–128.
- Bergen, K. M., & Dobson, M. C. (1999). Integration of remotely sensed radar imagery in modeling and mapping of forest biomass and net primary production. *Ecological Modelling*, *122*(3), 257–274.
- Blanco, J. A., Ameztegui, A., & Rodríguez, F. (2020). *Modelling forest ecosystems: A crossroad between scales, techniques and applications* (Vol. 425). Elsevier.
- Blanco, J. A., & Lo, Y.-H. (2023). Latest trends in modelling forest ecosystems: New approaches or just new methods? *Current Forestry Reports*, 1–11.
- Bohn, F. J., Frank, K., & Huth, A. (2014). Of climate and its resulting tree growth: Simulating the productivity of temperate forests. *Ecological Modelling*, *278*, 9–17.
- Bohn, F. J., & Huth, A. (2017). The importance of forest structure to biodiversity–productivity relationships. *Royal Society open science*, *4*(1), 160521.
- Bohn, F. J., May, F., & Huth, A. (2018). Species composition and forest structure explain the temperature sensitivity patterns of productivity in temperate forests. *Biogeosciences*, *15*(6), 1795–1813.
- Bolton, D. K., Coops, N. C., & Wulder, M. A. (2015). Characterizing residual structure and forest recovery following high-severity fire in the western boreal of canada using landsat time-series and airborne lidar data. *Remote Sensing of Environment*, *163*, 48–60.
- Bonan, G. (2008). Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *science*, *320*(5882), 1444–1449.

- Bonan, G., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., ... Swenson, S. C. (2011). Improving canopy processes in the community land model version 4 (clm4) using global flux fields empirically inferred from fluxnet data. *Journal of Geophysical Research: Biogeosciences*, 116(G2).
- Bonan, G., Levis, S., Sitch, S., Vertenstein, M., & Oleson, K. W. (2003). A dynamic global vegetation model for use with climate models: concepts and description of simulated vegetation dynamics. *Global Change Biology*, 9(11), 1543–1566.
- Bonan, G., Williams, M., Fisher, R., & Oleson, K. (2014). Modeling stomatal conductance in the earth system: linking leaf water-use efficiency and water transport along the soil–plant–atmosphere continuum. *Geoscientific Model Development*, 7(5), 2193–2222.
- Botkin, D. B., Janak, J., & Wallis, J. R. (1970). *A simulator for northeastern forest growth: a contribution of the hubbard brook ecosystem study and ibm research*. IBM Thomas J. Watson Research Center.
- 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.
- Botkin, D. B., Janak, J. F., & Wallis, J. R. (1973). Estimating the effects of carbon fertilization on forest composition by ecosystem simulation. *Carbon and the Biosphere*, 328–344.
- Brazhnik, K., Hanley, C., & Shugart, H. H. (2017). Simulating changes in fires and ecology of the 21st century eurasian boreal forests of siberia. *Forests*, 8(2), 49.
- Brazhnik, K., & Shugart, H. (2017). Model sensitivity to spatial resolution and explicit light representation for simulation of boreal forests in complex terrain. *Ecological Modelling*, 352, 90–107.
- Brovkin, V., Bendtsen, J., Claussen, M., Ganopolski, A., Kubatzki, C., Petoukhov, V., & Andreev, A. (2002). Carbon cycle, vegetation, and climate dynamics in the holocene: Experiments with the climber-2 model. *Global Biogeochemical Cycles*, 16(4), 86–1.
- Bruening, J. M., Fischer, R., Bohn, F. J., Armston, J., Armstrong, A. H., Knapp, N., ... Dubayah, R. (2021). Challenges to aboveground biomass prediction from waveform lidar. *Environmental Research Letters*, 16(12), 125013.
- Bugmann, H. (2001). A review of forest gap models. *Climatic Change*, 51(3-4), 259–305.

- Bugmann, H., & Fischlin, A. (2002). 12 comparing the behaviour of mountainous forest succession models in a changing climate. *Mountain environments in changing climates*.
- Bugmann, H., & Seidl, R. (2022). The evolution, complexity and diversity of models of long-term forest dynamics. *Journal of Ecology*, *110*(10), 2288–2307.
- Bugmann, H., Seidl, R., Hartig, F., Bohn, F., Brna, J., Cailleret, M., ... others (2019). Tree mortality submodels drive simulated long-term forest dynamics: assessing 15 models from the stand to global scale. *Ecosphere*, *10*(2), e02616.
- Bugmann, H., Yan, X., Sykes, M. T., Martin, P., Lindner, M., Desanker, P. V., & Cumming, S. G. (1996). A comparison of forest gap models: model structure and behaviour. *Climatic Change*, *34*, 289–313.
- Burt, A., Boni Vicari, M., Da Costa, A. C., Coughlin, I., Meir, P., Rowland, L., & Disney, M. (2021). New insights into large tropical tree mass and structure from direct harvest and terrestrial lidar. *Royal Society Open Science*, *8*(2), 201458.
- Burton, C., Betts, R., Cardoso, M., Feldpausch, T. R., Harper, A., Jones, C. D., ... Wiltshire, A. (2019). Representation of fire, land-use change and vegetation dynamics in the joint uk land environment simulator vn4. 9 (jules). *Geoscientific Model Development*, *12*(1), 179–193.
- Bussotti, F., & Pollastrini, M. (2015). Evaluation of leaf features in forest trees: Methods, techniques, obtainable information and limits. *Ecological indicators*, *52*, 219–230.
- Cabon, A., Kannenberg, S. A., Arain, A., Babst, F., Baldocchi, D., Belmecheri, S., ... others (2022). Cross-biome synthesis of source versus sink limits to tree growth. *Science*, *376*(6594), 758–761.
- Camps-Valls, G., Campos-Taberner, M., Álvaro Moreno-Martínez, Walther, S., Duveiller, G., Cescatti, A., ... Running, S. W. (2021). A unified vegetation index for quantifying the terrestrial biosphere. *Science Advances*, *7*(9), eabc7447. doi: 10.1126/sciadv.abc7447
- Cao, M., & Woodward, F. I. (1998). Dynamic responses of terrestrial ecosystem carbon cycling to global climate change. *Nature*, *393*(6682), 249–252.
- Carvalho, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., ... others (2014). Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature*, *514*(7521), 213–217.

- Chave, J. (2013). The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecology letters*, *16*, 4–16.
- Chen, Y., Angulo, M. T., & Liu, Y.-Y. (2019). Revealing complex ecological dynamics via symbolic regression. *BioEssays*, *41*(12), 1900069.
- Ciais, P., Sabine, C. L., Bala, G., Bopp, L., Brovkin, V. A., Canadell, J. G., ... Thornton, P. E. (2014). Carbon and other biogeochemical cycles. In *Climate change 2013 – the physical science basis: Working group i contribution to the fifth assessment report of the intergovernmental panel on climate change* (p. 465–570). Cambridge University Press. doi: 10.1017/CBO9781107415324.015
- Clements, F. (1916). Plant succession. an analysis of the development of vegetation. carnegie inst. pub. 242. 1936. nature and structure of the climax. *J. Ecol*, *24*(1), 253–84.
- Coffin, D. P., & Lauenroth, W. K. (1989). Disturbances and gap dynamics in a semiarid grassland: a landscape-level approach. *Landscape Ecology*, *3*, 19–27.
- Coffin, D. P., & Lauenroth, W. K. (1990). A gap dynamics simulation model of succession in a semiarid grassland. *Ecological Modelling*, *49*(3-4), 229–266.
- Colombo, R., Celesti, M., Bianchi, R., Campbell, P. K., Cogliati, S., Cook, B. D., ... others (2018). Variability of sun-induced chlorophyll fluorescence according to stand age-related processes in a managed loblolly pine forest. *Global change biology*, *24*(7), 2980–2996.
- Courchamp, F., Dunne, J. A., Le Maho, Y., May, R. M., Thebaud, C., & Hochberg, M. E. (2015). Fundamental ecology is fundamental. *Trends in ecology & evolution*, *30*(1), 9–16.
- Cramer, W., Bondeau, A., Woodward, F. I., Prentice, I. C., Betts, R. A., Brovkin, V., ... others (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.
- Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A., & Hansen, M. C. (2018). Classifying drivers of global forest loss. *Science*, *361*(6407), 1108–1111.
- Dale, V. H., Doyle, T. W., & Shugart, H. H. (1985). A comparison of tree growth models. *Ecological modelling*, *29*(1-4), 145–169.

- de Koning, K., Broekhuijsen, J., Kühn, I., Ovaskainen, O., Taubert, F., Endresen, D., . . . Grimm, V. (2023). Digital twins: dynamic model-data fusion for ecology. *Trends in Ecology & Evolution*.
- de Paula, M. D. (2017). *Forest fragmentation in space and time: New perspectives from remote sensing and forest modelling* (Unpublished doctoral dissertation). Helmholtz Centre for Environmental Research-UFZ.
- Deutschmann, T., Beirle, S., Frieß, U., Grzegorski, M., Kern, C., Kritten, L., . . . others (2011). The monte carlo atmospheric radiative transfer model mcartim: Introduction and validation of jacobians and 3d features. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112(6), 1119–1137.
- Dislich, C., Günter, S., Homeier, J., Schröder, B., & Huth, A. (2009). Simulating forest dynamics of a tropical montane forest in south ecuador. *Erdkunde*, 347–364.
- Dislich, C., & Huth, A. (2012). Modelling the impact of shallow landslides on forest structure in tropical montane forests. *Ecological Modelling*, 239, 40–53.
- Doyle, T. W. (1981). The role of disturbance in the gap dynamics of a montane rain forest: an application of a tropical forest succession model. In *Forest succession: Concepts and application* (pp. 56–73). Springer.
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of animal ecology*, 77(4), 802–813.
- Ellis, E. A., Romero Montero, J. A., & Hernández Gómez, I. U. (2017). Deforestation processes in the state of quintana roo, mexico: the role of land use and community forestry. *Tropical Conservation Science*, 10, 1940082917697259.
- Emmett, K. D., Renwick, K. M., & Poulter, B. (2021). Adapting a dynamic vegetation model for regional biomass, plant biogeography, and fire modeling in the greater yellowstone ecosystem: Evaluating lpj-guess-lmfirecf. *Ecological Modelling*, 440, 109417.
- Engel, J., Huth, A., & Frank, K. (2012). Bioenergy production and s kylark (a lauda arvensis) population abundance—a modelling approach for the analysis of land-use change impacts and conservation options. *Gcb Bioenergy*, 4(6), 713–727.
- Eriksson, H. M., Eklundh, L., Kuusk, A., & Nilson, T. (2006). Impact of understory vegetation on forest canopy reflectance and remotely sensed lai estimates. *Remote sensing of Environment*, 103(4), 408–418.

- Estes, L., Elsen, P. R., Treuer, T., Ahmed, L., Caylor, K., Chang, J., ... Ellis, E. C. (2018). The spatial and temporal domains of modern ecology. *Nature ecology & evolution*, 2(5), 819–826.
- Fan, N., Koirala, S., Reichstein, M., Thurner, M., Avitabile, V., Santoro, M., ... Carvalhais, N. (2020). Apparent ecosystem carbon turnover time: uncertainties and robust features. *Earth System Science Data*, 12(4), 2517–2536.
- FAO. (2020). Global forest resources assessments 2020: Main report. *The Food and Agricultural Organization of the United Nations: Rome, Italy*.
- FAO. (2022). The state of the world's forests 2022. forest pathways for green recovery and building inclusive, resilient and sustainable economies. *The Food and Agricultural Organization of the United Nations: Rome, Italy*.
- Feeley, K. J., Davies, S. J., Ashton, P. S., Bunyavejchewin, S., Nur Supardi, M., Kassim, A. R., ... Chave, J. (2007). The role of gap phase processes in the biomass dynamics of tropical forests. *Proceedings of the Royal Society B: Biological Sciences*, 274(1627), 2857–2864.
- Féret, J.-B., Gitelson, A., Noble, S., & Jacquemoud, S. (2017). Prospect-d: Towards modeling leaf optical properties through a complete lifecycle. *Remote Sensing of Environment*, 193, 204–215.
- Fischer, F. J., Maréchaux, I., & Chave, J. (2019). Improving plant allometry by fusing forest models and remote sensing. *New Phytologist*, 223(3), 1159–1165.
- Fischer, R. (2021). The long-term consequences of forest fires on the carbon fluxes of a tropical forest in africa. *Applied Sciences*, 11(10), 4696.
- Fischer, R., Armstrong, A., Shugart, H. H., & Huth, A. (2014). Simulating the impacts of reduced rainfall on carbon stocks and net ecosystem exchange in a tropical forest. *Environmental modelling & software*, 52, 200–206.
- Fischer, R., Bohn, F., Dantas de Paula, M., Dislich, C., Groeneveld, J., Gutiérrez, A. G., ... Huth, A. (2016). Lessons learned from applying a forest gap model to understand ecosystem and carbon dynamics of complex tropical forests. *Ecological Modelling*, 326, 124-133. doi: 10.1016/j.ecolmodel.2015.11.018
- Fischer, R., Ensslin, A., Rutten, G., Fischer, M., Schellenberger Costa, D., Kleyer, M., ... Huth, A. (2015). Simulating carbon stocks and fluxes of an african tropical montane forest with an individual-based forest model. *PLoS One*, 10(4), e0123300.

- Fischer, R., Knapp, N., Bohn, F., Shugart, H. H., & Huth, A. (2019). The relevance of forest structure for biomass and productivity in temperate forests: New perspectives for remote sensing. *Surveys in Geophysics*, *40*, 709–734.
- Fisher, R. A., Koven, C. D., Anderegg, W. R., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., . . . others (2018). Vegetation demographics in earth system models: A review of progress and priorities. *Global change biology*, *24*(1), 35–54.
- Francini, S., D’Amico, G., Vangi, E., Borghi, C., & Chirici, G. (2022). Integrating gedi and landsat: spaceborne lidar and four decades of optical imagery for the analysis of forest disturbances and biomass changes in italy. *Sensors*, *22*(5), 2015.
- Franklin, O., Harrison, S. P., Dewar, R., Farrior, C. E., Brännström, Å., Dieckmann, U., . . . others (2020). Organizing principles for vegetation dynamics. *Nature plants*, *6*(5), 444–453.
- Frazer, G. W., Magnussen, S., Wulder, M., & Niemann, K. (2011). Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of lidar-derived estimates of forest stand biomass. *Remote Sensing of Environment*, *115*(2), 636–649.
- Frazer, G. W., Wulder, M. A., & Niemann, K. O. (2005). Simulation and quantification of the fine-scale spatial pattern and heterogeneity of forest canopy structure: A lacunarity-based method designed for analysis of continuous canopy heights. *Forest ecology and management*, *214*(1-3), 65–90.
- 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.
- Friend, A. (1998). Parameterisation of a global daily weather generator for terrestrial ecosystem modelling. *Ecological Modelling*, *109*(2), 121–140.
- Friend, A., Stevens, A., Knox, R., & Cannell, M. (1997). A process-based, terrestrial biosphere model of ecosystem dynamics (hybrid v3. 0). *Ecological modelling*, *95*(2-3), 249–287.
- Fyllas, N. M., Bentley, L. P., Shenkin, A., Asner, G. P., Atkin, O. K., Díaz, S., . . . others (2017). Solar radiation and functional traits explain the decline of forest primary productivity along a tropical elevation gradient. *Ecology letters*, *20*(6), 730–740.

- Gao, X., Huete, A. R., Ni, W., & Miura, T. (2000). Optical–biophysical relationships of vegetation spectra without background contamination. *Remote sensing of environment*, 74(3), 609–620.
- Gao, Y., Skutsch, M., Paneque-Gálvez, J., & Ghilardi, A. (2020). Remote sensing of forest degradation: a review. *Environmental Research Letters*, 15(10), 103001.
- Gastellu-Etchegorry, J.-P., Lauret, N., Yin, T., Landier, L., Kallel, A., Malenovský, Z., ... Mitraka, Z. (2017). DART: Recent Advances in Remote Sensing Data Modeling With Atmosphere, Polarization, and Chlorophyll Fluorescence. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(6), 2640–2649. doi: 10.1109/JSTARS.2017.2685528
- Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., & Sitch, S. (2004). Terrestrial vegetation and water balance—hydrological evaluation of a dynamic global vegetation model. *Journal of hydrology*, 286(1-4), 249–270.
- Gibson, L., Lee, T. M., Koh, L. P., Brook, B. W., Gardner, T. A., Barlow, J., ... others (2011). Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature*, 478(7369), 378–381.
- Gleason, H. A. (1926). The individualistic concept of the plant association. *Bulletin of the Torrey botanical club*, 7–26.
- Grace, J., Mitchard, E., & Gloor, E. (2014). Perturbations in the carbon budget of the tropics. *Global Change Biology*, 20(10), 3238–3255.
- Guanter, L., Kaufmann, H., Segl, K., Foerster, S., Rogass, C., Chabrillat, S., ... others (2015). The enmap spaceborne imaging spectroscopy mission for earth observation. *Remote Sensing*, 7(7), 8830–8857.
- Guimarães, N., Pádua, L., Marques, P., Silva, N., Peres, E., & Sousa, J. J. (2020). Forestry remote sensing from unmanned aerial vehicles: A review focusing on the data, processing and potentialities. *Remote Sensing*, 12(6), 1046.
- Gutiérrez, A. G., & Huth, A. (2012). Successional stages of primary temperate rainforests of chiloé island, chile. *Perspectives in Plant Ecology, Evolution and Systematics*, 14(4), 243–256.
- Handbook, S. U., & Tools, E. (2015). Sentinel-2 user handbook. *ESA Standard Document Date*, 1, 1–64.

- Hardisky, M. H., Klemas, V., & Smart, R. (1983). The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of spartina alterniflora canopies. *Photogramm. Eng. Remote Sens*, *49*, 77–83.
- Harris, N. L., Gibbs, D. A., Baccini, A., Birdsey, R. A., De Bruin, S., Farina, M., ... others (2021). Global maps of twenty-first century forest carbon fluxes. *Nature Climate Change*, *11*(3), 234–240.
- Hartmann, H., Schuldt, B., Sanders, T. G., Macinnis-Ng, C., Boehmer, H. J., Allen, C. D., ... others (2018). Monitoring global tree mortality patterns and trends. report from the vw symposium ‘crossing scales and disciplines to identify global trends of tree mortality as indicators of forest health’. *New Phytologist*, *217*(3), 984–987.
- Heip, C. (1974). A new index measuring evenness. *Journal of the Marine Biological Association of the United Kingdom*, *54*(3), 555–557.
- Henniger, H., Bohn, F. J., Schmidt, K., & Huth, A. (2023). A new approach combining a multilayer radiative transfer model with an individual-based forest model: Application to boreal forests in finland. *Remote Sensing*, *15*(12), 3078.
- Henniger, H., Huth, A., Frank, K., & Bohn, F. J. (2023). Creating virtual forests around the globe and analysing their state space. *Ecological Modelling*, *483*, 110404.
- Hernández-Clemente, R., North, P. R., Hornero, A., & Zarco-Tejada, P. J. (2017). Assessing the effects of forest health on sun-induced chlorophyll fluorescence using the fluorflight 3-d radiative transfer model to account for forest structure. *Remote Sensing of Environment*, *193*, 165–179.
- Hernández-Gómez, I. U., Cerdán, C. R., Navarro-Martínez, A., Vázquez-Luna, D., Armenta-Montero, S., & Ellis, E. A. (2019). Assessment of the claslite forest monitoring system in detecting disturbance from selective logging in the selva maya, mexico. *Silva Fennica*, *53*(1).
- Hickler, T., Smith, B., Sykes, M. T., Davis, M. B., Sugita, S., & Walker, K. (2004). Using a generalized vegetation model to simulate vegetation dynamics in northeastern usa. *Ecology*, *85*(2), 519–530.
- Hickler, T., Vohland, K., Feehan, J., Miller, P. A., Smith, B., Costa, L., ... others (2012). Projecting the future distribution of european potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model. *Global Ecology and Biogeography*, *21*(1), 50–63.

- Hijmans, R. J., & Elith, J. (2021). Species distribution models. *Spatial Data Science (rsatial.org)*. Available online at <https://rsatial.org/sdm> [Google Scholar].
- Holtmann, A., Huth, A., Pohl, F., Rebmann, C., & Fischer, R. (2021). Carbon sequestration in mixed deciduous forests: The influence of tree size and species composition derived from model experiments. *Forests*, 12(6), 726.
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote sensing of environment*, 83(1-2), 195–213.
- Huete, A. R., Liu, H., Batchily, K., & van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of tm images for eos-modis. *Remote Sensing of Environment*, 59(3), 440-451. doi: 10.1016/S0034-4257(96)00112-5
- Humphries, H. C., Coffin, D. P., & Lauenroth, W. K. (1996). An individual-based model of alpine plant distributions. *Ecological modelling*, 84(1-3), 99–126.
- Hunt Jr, E. R., & Rock, B. N. (1989). Detection of changes in leaf water content using near-and middle-infrared reflectances. *Remote sensing of environment*, 30(1), 43–54.
- Hurt, G. C., Moorcroft, P. R., And, S. W. P., & Levin, S. A. (1998). Terrestrial models and global change: challenges for the future. *Global Change Biology*, 4(5), 581–590.
- Huth, A., & Ditzer, T. (2001). Long-term impacts of logging in a tropical rain forest—a simulation study. *Forest Ecology and management*, 142(1-3), 33–51.
- Huth, A., Drechsler, M., & Köhler, P. (2004). Multicriteria evaluation of simulated logging scenarios in a tropical rain forest. *Journal of environmental management*, 71(4), 321–333.
- Jacquemoud, S., & Ustin, S. (2019). *Leaf optical properties*. Cambridge University Press.
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., ... Ustin, S. L. (2009). Prospect+ sail models: A review of use for vegetation characterization. *Remote sensing of environment*, 113, S56–S66.
- Joos, F., Gerber, S., Prentice, I., Otto-Bliesner, B. L., & Valdes, P. J. (2004). Transient simulations of holocene atmospheric carbon dioxide and terrestrial carbon since the last glacial maximum. *Global Biogeochemical Cycles*, 18(2).

REFERENCES

- Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., ... others (2017). Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Global change biology*, 23(1), 177–190.
- Kammesheidt, L., Köhler, P., & Huth, A. (2001). Sustainable timber harvesting in venezuela: a modelling approach. *Journal of applied ecology*, 38(4), 756–770.
- Kattenborn, T., Fassnacht, F. E., & Schmidlein, S. (2019). Differentiating plant functional types using reflectance: which traits make the difference? *Remote Sensing in Ecology and Conservation*, 5(1), 5–19.
- Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on convolutional neural networks (cnn) in vegetation remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 173, 24–49.
- Kattenborn, T., Richter, R., Guimarães-Steinicke, C., Feilhauer, H., & Wirth, C. (2022). Anglecarn: Predicting the temporal variation of leaf angle distributions from image series with deep learning. *Methods in Ecology and Evolution*, 13(11), 2531–2545.
- Kazmierczak, M., Wiegand, T., & Huth, A. (2014). A neutral vs. non-neutral parametrizations of a physiological forest gap model. *Ecological modelling*, 288, 94–102.
- Keane, R. E., Ryan, K. C., & Running, S. W. (1996). Simulating effects of fire on northern rocky mountain landscapes with the ecological process model fire-bgc. *Tree Physiology*, 16(3), 319–331.
- Keenan, R. J., Reams, G. A., Achard, F., de Freitas, J. V., Grainger, A., & Lindquist, E. (2015). Dynamics of global forest area: Results from the fao global forest resources assessment 2015. *Forest Ecology and Management*, 352, 9–20.
- Kicklighter, D. W., Bruno, M., DZönges, S., Esser, G., Heimann, M., Helfrich, J., ... others (1999). A first-order analysis of the potential role of co₂ fertilization to affect the global carbon budget: a comparison of four terrestrial biosphere models. *Tellus B: Chemical and Physical Meteorology*, 51(2), 343–366.
- Kienast, F. (1991). Simulated effects of increasing atmospheric co₂ and changing climate on the successional characteristics of alpine forest ecosystems. *Landscape Ecology*, 5, 225–238.

- Knapp, N., Fischer, R., & Huth, A. (2018). Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states. *Remote Sensing of Environment*, 205, 199–209.
- Koca, D., Smith, B., & Sykes, M. T. (2006). Modelling regional climate change effects on potential natural ecosystems in sweden. *Climatic change*, 78(2-4), 381–406.
- Köhler, P., Chave, J., Riéra, B., & Huth, A. (2003). Simulating the long-term response of tropical wet forests to fragmentation. *Ecosystems*, 6, 0114–0128.
- Köhler, P., & Huth, A. (2004). Simulating growth dynamics in a south-east asian rainforest threatened by recruitment shortage and tree harvesting. *Climatic Change*, 67(1), 95–117.
- Köhler, P., & Huth, A. (2007). Impacts of recruitment limitation and canopy disturbance on tropical tree species richness. *Ecological Modelling*, 203(3-4), 511–517.
- Köhler, P., & Huth, A. (2010). Towards ground-truthing of spaceborne estimates of above-ground life biomass and leaf area index in tropical rain forests. *Biogeosciences*, 7(8), 2531–2543.
- Kokhanovsky, A. A., Kuusk, A., Lang, M., & Kuusk, J. (2013). Database of optical and structural data for the validation of forest radiative transfer models. *Light Scattering Reviews 7: Radiative Transfer and Optical Properties of Atmosphere and Underlying Surface*, 109–148.
- Kothari, S., Beauchamp-Rioux, R., Blanchard, F., Crofts, A. L., Girard, A., Guilbeault-Mayers, X., ... others (2023). Predicting leaf traits across functional groups using reflectance spectroscopy. *New Phytologist*, 238(2), 549–566.
- Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., ... Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, 19(1).
- Kruse, F. A., Lefkoff, A., Boardman, J., Heidebrecht, K., Shapiro, A., Barloon, P., & Goetz, A. (1993). The spectral image processing system (sips)—interactive visualization and analysis of imaging spectrometer data. *Remote sensing of environment*, 44(2-3), 145–163.
- Kumagai, T., Katul, G. G., Porporato, A., Saitoh, T. M., Ohashi, M., Ichie, T., & Suzuki, M. (2004). Carbon and water cycling in a bornean tropical rainforest under current and future climate scenarios. *Advances in Water Resources*, 27(12), 1135–1150.

- Kuusik, A. (2018). Canopy radiative transfer modeling. In S. Liang (Ed.), *Comprehensive remote sensing* (pp. 9–22). Oxford: Elsevier. doi: 10.1016/B978-0-12-409548-9.10534-2
- Köhler, P., & Huth, A. (1998). The effects of tree species grouping in tropical rain-forest modelling: Simulations with the individual-based model formind. *Ecological Modelling*, *109*(3), 301–321. doi: 10.1016/S0304-3800(98)00066-0
- Langhammer, M., Thober, J., Lange, M., Frank, K., & Grimm, V. (2019). Agricultural landscape generators for simulation models: A review of existing solutions and an outline of future directions. *Ecological Modelling*, *393*, 135–151.
- Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., ... others (2011). Parameterization improvements and functional and structural advances in version 4 of the community land model. *Journal of Advances in Modeling Earth Systems*, *3*(1).
- Lechner, A. M., Foody, G. M., & Boyd, D. S. (2020). Applications in remote sensing to forest ecology and management. *One Earth*, *2*(5), 405–412.
- Le Toan, T., Quegan, S., Woodward, I., Lomas, M., Delbart, N., & Picard, G. (2004). Relating radar remote sensing of biomass to modelling of forest carbon budgets. *Climatic change*, *67*(2-3), 379–402.
- Levin, S. A. (1992). The problem of pattern and scale in ecology: the robert h. macarthur award lecture. *Ecology*, *73*(6), 1943–1967.
- Lindenmayer, D. B., & Likens, G. E. (2009). Adaptive monitoring: a new paradigm for long-term research and monitoring. *Trends in ecology & evolution*, *24*(9), 482–486.
- Lindenmayer, D. B., Likens, G. E., Haywood, A., & Miezi, L. (2011). Adaptive monitoring in the real world: proof of concept. *Trends in ecology & evolution*, *26*(12), 641–646.
- Liu, J., & Ashton, P. S. (1995). Individual-based simulation models for forest succession and management. *Forest ecology and management*, *73*(1-3), 157–175.
- Lobo, E., & Dalling, J. W. (2014). Spatial scale and sampling resolution affect measures of gap disturbance in a lowland tropical forest: implications for understanding forest regeneration and carbon storage. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1778), 20133218.

- Loreau, M. (2010). Linking biodiversity and ecosystems: towards a unifying ecological theory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1537), 49–60.
- Lu, X., Liu, Z., An, S., Miralles, D. G., Maes, W., Liu, Y., & Tang, J. (2018). Potential of solar-induced chlorophyll fluorescence to estimate transpiration in a temperate forest. *Agricultural and Forest Meteorology*, 252, 75–87.
- Lucht, W., Prentice, I. C., Myneni, R. B., Sitch, S., Friedlingstein, P., Cramer, W., ... Smith, B. (2002). Climatic control of the high-latitude vegetation greening trend and pinatubo effect. *Science*, 296(5573), 1687–1689.
- Lulla, K., Nellis, M. D., Rundquist, B., Srivastava, P. K., & Szabo, S. (2021). *Mission to earth: Landsat 9 will continue to view the world* (Vol. 36) (No. 20). Taylor & Francis.
- Ma, X., Mahecha, M. D., Migliavacca, M., van der Plas, F., Benavides, R., Ratcliffe, S., ... Wirth, C. (2019). Inferring plant functional diversity from space: the potential of Sentinel-2. *Remote Sensing of Environment*, 233, 111368. doi: 10.1016/j.rse.2019.111368
- Ma, X., Migliavacca, M., Wirth, C., Bohn, F. J., Huth, A., Richter, R., & Mahecha, M. D. (2020). Monitoring plant functional diversity using the reflectance and echo from space. *Remote Sensing*, 12(8), 1248.
- Malhi, Y. (2010). The carbon balance of tropical forest regions, 1990–2005. *Current Opinion in Environmental Sustainability*, 2(4), 237–244. doi: 10.1016/j.cosust.2010.08.002
- Maris, V., Huneman, P., Coreau, A., Kéfi, S., Pradel, R., & Devictor, V. (2018). Prediction in ecology: promises, obstacles and clarifications. *Oikos*, 127(2), 171–183.
- Maréchaux, I., & Chave, J. (2017). An individual-based forest model to jointly simulate carbon and tree diversity in amazonia: description and applications. *Ecological Monographs*, 87(4), 632–664.
- Maréchaux, I., Langerwisch, F., Huth, A., Bugmann, H., Morin, X., Reyer, C. P., ... Bohn, F. J. (2021). Tackling unresolved questions in forest ecology: The past and future role of simulation models. *Ecology and Evolution*, 11(9), 3746–3770. doi: 10.1002/ece3.7391

- Mascaro, J., Asner, G. P., Muller-Landau, H. C., van Breugel, M., Hall, J., & Dahlin, K. (2011). Controls over aboveground forest carbon density on Barro Colorado Island, Panama. *Biogeosciences*, 8(6), 1615–1629.
- McDowell, N. G., Allen, C. D., Anderson-Teixeira, K., Aukema, B. H., Bond-Lamberty, B., Chini, L., ... others (2020). Pervasive shifts in forest dynamics in a changing world. *Science*, 368(6494), eaaz9463.
- Medlyn, B. E., De Kauwe, M. G., Zaehle, S., Walker, A. P., Duursma, R. A., Luus, K., ... others (2016). Using models to guide field experiments: A priori predictions for the CO₂ response of a nutrient- and water-limited native eucalypt woodland. *Global Change Biology*, 22(8), 2834–2851.
- Medvigy, D., Clark, K., Skowronski, N., & Schäfer, K. (2012). Simulated impacts of insect defoliation on forest carbon dynamics. *Environmental Research Letters*, 7(4), 045703.
- Medvigy, D., Wofsy, S., Munger, J., Hollinger, D., & Moorcroft, P. (2009). Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem demography model version 2. *Journal of Geophysical Research: Biogeosciences*, 114(G1).
- Melton, J., & Arora, V. (2016). Competition between plant functional types in the Canadian terrestrial ecosystem model (CTEM) v. 2.0. *Geoscientific Model Development*, 9(1), 323–361.
- Menaut, J.-C., Gignoux, J., Prado, C., & Clobert, J. (1990). Tree community dynamics in a humid savanna of the Côte d'Ivoire: modelling the effects of fire and competition with grass and neighbours. *Journal of Biogeography*, 17, 471–481.
- Mercado, L. M., Huntingford, C., Gash, J. H., Cox, P. M., & Jogireddy, V. (2007). Improving the representation of radiation interception and photosynthesis for climate model applications. *Tellus B: Chemical and Physical Meteorology*, 59(3), 553–565.
- Meyer, V., Saatchi, S., Chave, J., Dalling, J., Bohlman, S., Fricker, G., ... Hubbell, S. (2013). Detecting tropical forest biomass dynamics from repeated airborne lidar measurements. *Biogeosciences*, 10(8), 5421–5438.
- 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.

- Minh, D. H. T., Le Toan, T., Rocca, F., Tebaldini, S., d'Alessandro, M. M., & Villard, L. (2013). Relating p-band synthetic aperture radar tomography to tropical forest biomass. *IEEE Transactions on Geoscience and Remote Sensing*, 52(2), 967–979.
- Missions, L. (2016). Using the usgs landsat8 product. *US Department of the Interior-US Geological Survey–NASA*.
- 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, 1–13.
- Montero, D., Aybar, C., Mahecha, M. D., Martinuzzi, F., Söchting, M., & Wieneke, S. (2023). A standardized catalogue of spectral indices to advance the use of remote sensing in earth system research. *Scientific Data*, 10(1), 197.
- 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.
- Morales, P., Hickler, T., Rowell, D. P., Smith, B., & Sykes, M. T. (2007). Changes in european ecosystem productivity and carbon balance driven by regional climate model output. *Global Change Biology*, 13(1), 108–122.
- Morin, X., Damestoy, T., Toigo, M., Castagneyrol, B., Jactel, H., de Coligny, F., & Meredieu, C. (2020). Using forest gap models and experimental data to explore long-term effects of tree diversity on the productivity of mixed planted forests. *Annals of Forest Science*, 77(2), 1–19.
- Morin, X., Fahse, L., Jactel, H., Scherer-Lorenzen, M., García-Valdés, R., & Bugmann, H. (2018). Long-term response of forest productivity to climate change is mostly driven by change in tree species composition. *Scientific reports*, 8(1), 5627.
- Morin, X., Fahse, L., Scherer-Lorenzen, M., & Bugmann, H. (2011). Tree species richness promotes productivity in temperate forests through strong complementarity between species. *Ecology letters*, 14(12), 1211–1219.
- Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., ... others (2015). Predictive ecology in a changing world. *Journal of applied ecology*, 52(5), 1293–1310.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403(6772), 853–858.

- Nativi, S., & Craglia, M. (2020). Destination earth. *Use Cases Analysis. European Commission Publication. Joint Research Centre* <https://op.europa.eu/en/publication-detail/-/publication/767e8063-2499-11eb-9d7e-01aa75ed71a1/language-en>.
- Naudts, K., Ryder, J., McGrath, M. J., Otto, J., Chen, Y., Valade, A., ... others (2015). A vertically discretised canopy description for orchidee (svn r2290) and the modifications to the energy, water and carbon fluxes. *Geoscientific Model Development*, 8(7), 2035–2065.
- Neumann, M., Ferro-Famil, L., & Reigber, A. (2009). Estimation of forest structure, ground, and canopy layer characteristics from multibaseline polarimetric interferometric sar data. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3), 1086–1104.
- New, M., Hulme, M., & Jones, P. (2000). Representing twentieth-century space–time climate variability. part ii: Development of 1901–96 monthly grids of terrestrial surface climate. *Journal of climate*, 13(13), 2217–2238.
- Newnham, R., & Smith, J. (1964). Development and testing of stand models for douglas fir and lodgepole pine. *The Forestry Chronicle*, 40(4), 494–504.
- Norby, R. J., De Kauwe, M. G., Domingues, T. F., Duursma, R. A., Ellsworth, D. S., Goll, D. S., ... others (2016). Model–data synthesis for the next generation of forest free-air co₂ enrichment (face) experiments. *New Phytologist*, 209(1), 17–28.
- Pacheco-Labrador, J., Perez-Priego, O., El-Madany, T. S., Julitta, T., Rossini, M., Guan, J., ... others (2019). Multiple-constraint inversion of scope. evaluating the potential of gpp and sif for the retrieval of plant functional traits. *Remote Sensing of Environment*, 234, 111362.
- Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., ... Hayes, D. (2011). A Large and Persistent Carbon Sink in the World's Forests. *Science*, 333(6045), 988–993. doi: 10.1126/science.1201609
- Pan, Y., 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.
- Paquette, A., & Messier, C. (2011). The effect of biodiversity on tree productivity: from temperate to boreal forests. *Global Ecology and Biogeography*, 20(1), 170–180.

- Paulick, S., Dislich, C., Homeier, J., Fischer, R., & Huth, A. (2017). The carbon fluxes in different successional stages: modelling the dynamics of tropical montane forests in south ecuador. *Forest Ecosystems*, *4*, 1–11.
- Peet, R. K. (1975). Relative diversity indices. *Ecology*, *56*(2), 496–498.
- Peng, C. (2000). From static biogeographical model to dynamic global vegetation model: a global perspective on modelling vegetation dynamics. *Ecological modelling*, *135*(1), 33–54.
- Pettorelli, N., Laurance, W. F., O'Brien, T. G., Wegmann, M., Nagendra, H., & Turner, W. (2014). Satellite remote sensing for applied ecologists: opportunities and challenges. *Journal of Applied Ecology*, *51*(4), 839–848.
- Pfeiffer, M., Spessa, A., & Kaplan, J. O. (2013). A model for global biomass burning in preindustrial time: Lpj-lmfire (v1. 0). *Geoscientific Model Development*, *6*(3), 643–685.
- Pierrat, Z., Magney, T., Parazoo, N. C., Grossmann, K., Bowling, D. R., Seibt, U., ... others (2022). Diurnal and seasonal dynamics of solar-induced chlorophyll fluorescence, vegetation indices, and gross primary productivity in the boreal forest. *Journal of Geophysical Research: Biogeosciences*, *127*(2), e2021JG006588.
- Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., ... Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *science*, *344*(6187), 1246752.
- Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., ... others (2021). Mapping global forest canopy height through integration of gedi and landsat data. *Remote Sensing of Environment*, *253*, 112165.
- Prentice, I. C., Bondeau, A., Cramer, W., Harrison, S. P., Hickler, T., Lucht, W., ... Sykes, M. T. (2007). Dynamic global vegetation modeling: quantifying terrestrial ecosystem responses to large-scale environmental change. *Terrestrial ecosystems in a changing world*, 175–192.
- Prentice, I. C., & Cowling, S. A. (2013). Dynamic global vegetation models. In *Encyclopedia of biodiversity* (pp. 670–689). Elsevier.
- Prentice, I. C., Cramer, W., Harrison, S. P., Leemans, R., Monserud, R. A., & Solomon, A. M. (1992). Special paper: a global biome model based on plant physiology and dominance, soil properties and climate. *Journal of biogeography*, 117–134.

- Prentice, I. C., Kaplan, J. O., Sykes, M. T., & Thonicke, K. (2000). Lpj-a coupled model of vegetation dynamics and the terrestrial carbon cycle.
- Pugh, T. A., Arneth, A., Kautz, M., Poulter, B., & Smith, B. (2019). Important role of forest disturbances in the global biomass turnover and carbon sinks. *Nature geoscience*, *12*(9), 730–735.
- Qin, Y., Xiao, X., Wigneron, J.-P., Ciais, P., Brandt, M., Fan, L., ... others (2021). Carbon loss from forest degradation exceeds that from deforestation in the brazilian amazon. *Nature Climate Change*, *11*(5), 442–448.
- Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J.-F., Minh, D. H. T., ... others (2019). The european space agency biomass mission: Measuring forest above-ground biomass from space. *Remote Sensing of Environment*, *227*, 44–60.
- Quillet, A., Peng, C., & Garneau, M. (2010). Toward dynamic global vegetation models for simulating vegetation–climate interactions and feedbacks: recent developments, limitations, and future challenges. *Environmental Reviews*, *18*(NA), 333–353.
- Radočaj, D., Obhodaš, J., Jurišić, M., & Gašparović, M. (2020). Global open data remote sensing satellite missions for land monitoring and conservation: A review. *Land*, *9*(11), 402.
- Reyer, C. P. O., Silveyra Gonzalez, R., Dolos, K., Hartig, F., Hauf, Y., Noack, M., ... Frieler, K. (2020). The PROFOUND Database for evaluating vegetation models and simulating climate impacts on European forests. *Earth System Science Data*, *12*(2), 1295–1320. doi: 10.5194/essd-12-1295-2020
- Robert, A. (2003). Simulation of the effect of topography and tree falls on stand dynamics and stand structure of tropical forests. *Ecological Modelling*, *167*(3), 287–303.
- Rödig, E., Cuntz, M., Heinke, J., Rammig, A., & Huth, A. (2017). Spatial heterogeneity of biomass and forest structure of the amazon rain forest: Linking remote sensing, forest modelling and field inventory. *Global ecology and biogeography*, *26*(11), 1292–1302.
- Rödig, E., Cuntz, M., Rammig, A., Fischer, R., Taubert, F., & Huth, A. (2018). The importance of forest structure for carbon fluxes of the amazon rainforest. *Environmental Research Letters*, *13*(5), 054013.

- Rödig, E., Knapp, N., Fischer, R., Bohn, F. J., Dubayah, R., Tang, H., & Huth, A. (2019). From small-scale forest structure to amazon-wide carbon estimates. *Nature communications*, *10*(1), 5088.
- Roebroek, C. T., Duveiller, G., Seneviratne, S. I., Davin, E. L., & Cescatti, A. (2023). Releasing global forests from human management: How much more carbon could be stored? *Science*, *380*(6646), 749–753.
- Ruehr, S., Keenan, T. F., Williams, C., Zhou, Y., Lu, X., Bastos, A., ... Terrer, C. (2023). Evidence and attribution of the enhanced land carbon sink. *Nature Reviews Earth & Environment*, 1–17.
- Rüger, N., Gutiérrez, A. G., Kissling, W. D., Armesto, J. J., & Huth, A. (2007). Ecological impacts of different harvesting scenarios for temperate evergreen rain forest in southern chile—a simulation experiment. *Forest Ecology and Management*, *252*(1-3), 52–66.
- Rüger, N., Williams-Linera, G., Kissling, W. D., & Huth, A. (2008). Long-term impacts of fuelwood extraction on a tropical montane cloud forest. *Ecosystems*, *11*, 868–881.
- Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., & Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *Bioscience*, *54*(6), 547–560.
- Saatchi, S., Marlier, M., Chazdon, R. L., Clark, D. B., & Russell, A. E. (2011). Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass. *Remote Sensing of Environment*, *115*(11), 2836–2849.
- Sakschewski, B., Von Bloh, W., Boit, A., Poorter, L., Peña-Claros, M., Heinke, J., ... Thonicke, K. (2016). Resilience of amazon forests emerges from plant trait diversity. *Nature climate change*, *6*(11), 1032–1036.
- Sakschewski, B., von Bloh, W., Boit, A., Rammig, A., Kattge, J., Poorter, L., ... Thonicke, K. (2015). Leaf and stem economics spectra drive diversity of functional plant traits in a dynamic global vegetation model. *Global Change Biology*, *21*(7), 2711–2725.
- Sato, H., Itoh, A., & Kohyama, T. (2007). Seib-dgvm: A new dynamic global vegetation model using a spatially explicit individual-based approach. *Ecological Modelling*, *200*(3-4), 279–307.

- Schäfer, J., Winiwarter, L., Weiser, H., Novotný, J., Höfle, B., Schmidlein, S., ... Fassnacht, F. E. (2023). Assessing the potential of synthetic and ex situ airborne laser scanning and ground plot data to train forest biomass models. *Forestry: An International Journal of Forest Research*, cpad061.
- Schaphoff, S., Lucht, W., Gerten, D., Sitch, S., Cramer, W., & Prentice, I. C. (2006). Terrestrial biosphere carbon storage under alternative climate projections. *Climatic change*, 74, 97–122.
- Schmitt, S., Maréchaux, I., Chave, J., Fischer, F. J., Pioniot, C., Traissac, S., & Hérault, B. (2020). Functional diversity improves tropical forest resilience: Insights from a long-term virtual experiment. *Journal of Ecology*, 108(3), 831–843.
- Sellers, P. J. (1985). Canopy reflectance, photosynthesis and transpiration. *International journal of remote sensing*, 6(8), 1335–1372.
- Semenov, M. A., & Barrow, E. M. (1997). Use of a stochastic weather generator in the development of climate change scenarios. *Climatic change*, 35(4), 397–414.
- Sethi, S. S., Kovac, M., Wiesemüller, F., Miriyev, A., & Boutry, C. M. (2022). Biodegradable sensors are ready to transform autonomous ecological monitoring. *Nature Ecology & Evolution*, 6(9), 1245–1247.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379–423.
- Shugart, H. H. (1998). *Terrestrial ecosystems in changing environments*. Cambridge University Press.
- Shugart, H. H., Asner, G. P., Fischer, R., Huth, A., Knapp, N., Le Toan, T., & Shuman, J. K. (2015). Computer and remote-sensing infrastructure to enhance large-scale testing of individual-based forest models. *Frontiers in Ecology and the Environment*, 13(9), 503–511.
- Shugart, H. H., Foster, A., Wang, B., Druckenbrod, D., Ma, J., Lerda, M., ... Yan, X. (2020). Gap models across micro-to mega-scales of time and space: examples of tansley's ecosystem concept. *Forest Ecosystems*, 7, 1–18.
- Shugart, H. H., & Noble, I. (1981). A computer model of succession and fire response of the high-altitude eucalyptus forest of the brindabella range, australian capital territory. *Australian Journal of Ecology*, 6(2), 149–164.

- Shugart, H. H., et al. (1984). *A theory of forest dynamics. the ecological implications of forest succession models*. Springer-Verlag.
- Shugart, H. H., & Smith, T. (1996). A review of forest patch models and their application to global change research. *Climatic Change*, 34(2), 131–153.
- Shugart, H. H., Wang, B., Fischer, R., Ma, J., Fang, J., Yan, X., ... 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.
- 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. *Journal of Environmental Management*, 5, 161-179.
- Shugart, H. H., & West, D. C. (1980). Forest succession models. *BioScience*, 30(5), 308–313.
- Shuman, J. K., Foster, A. C., Shugart, H. H., Hoffman-Hall, A., Krylov, A., Loboda, T., ... Sochilova, E. (2017). Fire disturbance and climate change: implications for russian forests. *Environmental Research Letters*, 12(3), 035003.
- Siccama, T., Botkin, D., Bormann, F., & Likens, G. (1969). Computer simulation of a northern hardwood forest. *Bull. Ecol. Soc. Amer*, 50, 93.
- Sims, D. A., Rahman, A. F., Cordova, V. D., El-Masri, B. Z., Baldocchi, D. D., Flanagan, L. B., ... others (2006). On the use of modis evi to assess gross primary productivity of north american ecosystems. *Journal of Geophysical Research: Biogeosciences*, 111(G4).
- Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., ... others (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.
- Smith, B. (2001). Lpj-guess-an ecosystem modelling framework. *Department of Physical Geography and Ecosystems Analysis, INES, Sölvegatan*, 12, 22362.
- Smith, B., Knorr, W., Widlowski, J.-L., Pinty, B., & Gobron, N. (2008). Combining remote sensing data with process modelling to monitor boreal conifer forest carbon balances. *Forest Ecology and Management*, 255(12), 3985–3994.
- Solomon, A. M. (1986). Transient response of forests to co 2-induced climate change: simulation modeling experiments in eastern north america. *Oecologia*, 68, 567–579.

- Song, X.-P., Hansen, M. C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E. F., & Townshend, J. R. (2018). Global land change from 1982 to 2016. *Nature*, *560*(7720), 639–643.
- Spahni, R., Wania, R., Neef, L., Van Weele, M., Pison, I., Bousquet, P., ... others (2011). Constraining global methane emissions and uptake by ecosystems. *Biogeosciences*, *8*(6), 1643–1665.
- Springer, K. R., Wang, R., & Gamon, J. A. (2017). Parallel seasonal patterns of photosynthesis, fluorescence, and reflectance indices in boreal trees. *Remote Sensing*, *9*(7), 691.
- Stocker, T. (2014). *Climate change 2013: the physical science basis: Working group I contribution to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge university press.
- Stovall, A. E., Shugart, H., & Yang, X. (2019). Tree height explains mortality risk during an intense drought. *Nature Communications*, *10*(1), 4385.
- Su, Y., Guo, Q., Xue, B., Hu, T., Alvarez, O., Tao, S., & Fang, J. (2016). Spatial distribution of forest aboveground biomass in china: Estimation through combination of spaceborne lidar, optical imagery, and forest inventory data. *Remote Sensing of Environment*, *173*, 187–199.
- Sun, Y., Gu, L., Wen, J., van Der Tol, C., Porcar-Castell, A., Joiner, J., ... others (2023). From remotely sensed solar-induced chlorophyll fluorescence to ecosystem structure, function, and service: Part I—harnessing theory. *Global Change Biology*.
- Tang, G., Beckage, B., & Smith, B. (2012). The potential transient dynamics of forests in new england under historical and projected future climate change. *Climatic change*, *114*, 357–377.
- Taubert, F., Fischer, R., Groeneveld, J., Lehmann, S., Müller, M. S., Rödig, E., ... Huth, A. (2018). Global patterns of tropical forest fragmentation. *Nature*, *554*(7693), 519–522.
- Taubert, F., Frank, K., & Huth, A. (2012). A review of grassland models in the biofuel context. *Ecological Modelling*, *245*, 84–93.
- Thonicke, K., Venevsky, S., Sitch, S., & Cramer, W. (2001). The role of fire disturbance for global vegetation dynamics: coupling fire into a dynamic global vegetation model. *Global Ecology and Biogeography*, *10*(6), 661–677.

- Thurner, M., Beer, C., Ciais, P., Friend, A. D., Ito, A., Kleidon, A., ... others (2017). Evaluation of climate-related carbon turnover processes in global vegetation models for boreal and temperate forests. *Global Change Biology*, 23(8), 3076–3091.
- Tredennick, A. T., Hooker, G., Ellner, S. P., & Adler, P. B. (2021). A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, 102(6), e03336.
- Trudgill, S. (2007). Tansley, ag 1935: The use and abuse of vegetational concepts and terms. *ecology* 16, 284—307. *Progress in Physical Geography*, 31(5), 517–522.
- Trugman, A., Fenton, N., Bergeron, Y., Xu, X., Welp, L., & Medvigy, D. (2016). Climate, soil organic layer, and nitrogen jointly drive forest development after fire in the north american boreal zone. *Journal of Advances in Modeling Earth Systems*, 8(3), 1180–1209.
- Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Zhao, M., ... others (2006). Evaluation of modis npp and gpp products across multiple biomes. *Remote sensing of environment*, 102(3-4), 282–292.
- UNFCCC. (2004). Implementation plan for the global observing system for climate in support of the united nations framework convention on climate change. *IGO submissions*, 27.
- Urban, D. L., & Shugart, H. H. (1992). Individual-based models of. *Plant Succession: Theory and prediction*, 11, 249.
- van der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., & Su, Z. (2009). An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance. *Biogeosciences*, 6(12), 3109–3129. doi: 10.5194/bg-6-3109-2009
- van der Sande, M. T., Peña-Claros, M., Ascarrunz, N., Arets, E. J., Licona, J. C., Toledo, M., & Poorter, L. (2017). Abiotic and biotic drivers of biomass change in a neotropical forest. *Journal of Ecology*, 105(5), 1223–1234.
- Van der Sande, M. T., Poorter, L., Balvanera, P., Kooistra, L., Thonicke, K., Boit, A., ... others (2017). The integration of empirical, remote sensing and modelling approaches enhances insight in the role of biodiversity in climate change mitigation by tropical forests. *Current opinion in environmental sustainability*, 26, 69–76.

- Verhoef, W. (1984). Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sensing of Environment*, *16*(2), 125–141. doi: 10.1016/0034-4257(84)90057-9
- Verhoef, W., Jia, L., Xiao, Q., & Su, Z. (2007). Unified optical-thermal four-stream radiative transfer theory for homogeneous vegetation canopies. *IEEE Transactions on geoscience and remote sensing*, *45*(6), 1808–1822.
- Verhoef, W., Van Der Tol, C., & Middleton, E. M. (2018). Hyperspectral radiative transfer modeling to explore the combined retrieval of biophysical parameters and canopy fluorescence from flex–sentinel-3 tandem mission multi-sensor data. *Remote sensing of environment*, *204*, 942–963.
- Vermote, E., Kotchenova, S., & Ray, J. (2011). Modis surface reflectance user’s guide. *MODIS Land Surface Reflectance Science Computing Facility, version, 1*, 1–40.
- Vilfan, N., van der Tol, C., Muller, O., Rascher, U., & Verhoef, W. (2016). Fluspect-b: A model for leaf fluorescence, reflectance and transmittance spectra. *Remote Sensing of Environment*, *186*, 596–615. doi: 10.1016/j.rse.2016.09.017
- Vogelmann, J., & Rock, B. (1986). Assessing forest decline in coniferous forests of vermont using ns-001 thematic mapper simulator data. *International Journal of Remote Sensing*, *7*(10), 1303–1321.
- Wang, Q., Adiku, S., Tenhunen, J., & Granier, A. (2005). On the relationship of ndvi with leaf area index in a deciduous forest site. *Remote sensing of environment*, *94*(2), 244–255.
- Wang, X., Dannenberg, M. P., Yan, D., Jones, M. O., Kimball, J. S., Moore, D. J., ... Smith, W. K. (2020). Globally consistent patterns of asynchrony in vegetation phenology derived from optical, microwave, and fluorescence satellite data. *Journal of Geophysical Research: Biogeosciences*, *125*(7), e2020JG005732.
- Wania, R., Ross, I., & Prentice, I. (2010). Implementation and evaluation of a new methane model within a dynamic global vegetation model: Lpj-whyme v1. 3. *Geoscientific Model Development Discussions*, *3*(1), 1–59.
- Watt, A. S. (1923). On the ecology of british beechwoods with special reference to their regeneration. *Journal of Ecology*, *11*(1), 1–48.
- West, D. C., Shugart, H. H., & Botkin, D. (2012). *Forest succession: concepts and application*. Springer Science & Business Media.

- Whittaker, R. H. (1953). A consideration of climax theory: the climax as a population and pattern. *Ecological monographs*, 23(1), 41–78.
- Whittaker, R. H., et al. (1970). Communities and ecosystems. *Communities and ecosystems*.
- Widlowski, J.-L., Mio, C., Disney, M., Adams, J., Andredakis, I., Atzberger, C., ... others (2015). The fourth phase of the radiative transfer model intercomparison (rami) exercise: Actual canopy scenarios and conformity testing. *Remote Sensing of Environment*, 169, 418–437.
- Wilks, D. S., & Wilby, R. L. (1999). The weather generation game: a review of stochastic weather models. *Progress in physical geography*, 23(3), 329–357.
- Xiao, J., Chevallier, F., Gomez, C., Guanter, L., Hicke, J. A., Huete, A. R., ... others (2019). Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. *Remote Sensing of Environment*, 233, 111383.
- Yang, P., Verhoef, W., & van der Tol, C. (2017). The mSCOPE model: A simple adaptation to the SCOPE model to describe reflectance, fluorescence and photosynthesis of vertically heterogeneous canopies. *Remote Control of Environment*, 201, 1–11. doi: 10.1016/j.rse.2017.08.029
- Yuan, Z., Wang, S., Ali, A., Gazol, A., Ruiz-Benito, P., Wang, X., ... Loreau, M. (2018). Aboveground carbon storage is driven by functional trait composition and stand structural attributes rather than biodiversity in temperate mixed forests recovering from disturbances. *Annals of Forest Science*, 75(3), 1–13.
- Zeng, Y., Hao, D., Huete, A., Dechant, B., Berry, J., Chen, J. M., ... others (2022). Optical vegetation indices for monitoring terrestrial ecosystems globally. *Nature Reviews Earth & Environment*, 3(7), 477–493.
- Zhang, K., de Almeida Castanho, A. D., Galbraith, D. R., Moghim, S., Levine, N. M., Bras, R. L., ... others (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.
- Zhang, Y., & Chen, H. Y. (2015). Individual size inequality links forest diversity and above-ground biomass. *Journal of Ecology*, 103(5), 1245–1252.
- Zhao, Q., Yu, L., Du, Z., Peng, D., Hao, P., Zhang, Y., & Gong, P. (2022). An overview of the applications of earth observation satellite data: impacts and future trends. *Remote Sensing*, 14(8), 1863.

REFERENCES

- Zolkos, S. G., Goetz, S. J., & Dubayah, R. (2013). A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote sensing of environment*, 128, 289–298.

A Appendix Chapter 2

Additional Information on the Method Section

Pre-defined forest stand attributes

For each forest stand we select a minimum and maximum tree height h_{min} and h_{max} (a), a species pool (b) and a total crown volume of trees (c). The Forest Factory 2.0 receives as an initial input a minimum and maximum height of the trees H_{min} and H_{max} (this applies to all trees in all forest stands). From the range of H_{min} and H_{max} a forest stand specific h_{min} and h_{max} (a) is chosen randomly for each forest stand. We assume an equally distributed probability distribution in the mentioned range. The parameterization of each ecoregion defines the total species pool for Forest Factory 2.0. For each forest stand, a forest stand-specific species pool (b) is chosen. For this purpose, a number between 1 and the number of species in the total species pool is chosen uniformly distributed. It determines how many species the forest stand-specific species pool should contain. Each species has an equal probability of being included in the forest stand-specific species pool until the next to last species (selected number of species pool) is selected. To ensure that it is possible to plant trees within the selected height range between h_{min} and h_{max} (which is different for each forest stand) we check whether at least one of the species selected so far has a maximum attainable height greater than or equal to h_{max} . If this is not the case, the last species is selected so that a tree with h_{max} could be placed. Forest Factory 2.0 receives as input also a maximum total crown volume. This total crown volume can be seen as a kind of crown density (proportion of crown volume to forest stand volume). None of the generated forest stands have a total crown volume above this input value. For each forest stand, between 0 and the maximum total crown volume, a forest stand specific maximum total crown volume (c) is randomly chosen, assuming an equal probability distribution.

Normalized Shannon Index

The Shannon Index H (Shannon 1948) and the species evenness EH (Shannon equitability Index by Heip 1974, Peet 1975) is calculated by:

$$H = \sum_{i=1}^S p_i \cdot \ln(p_i)$$

$$E_H = \frac{H}{\ln(S)}$$

with p_i ... Proportion of trees of species i in the total number of trees

with S ... Set of all species in the initial species pool derived by the parameterization of

the ecoregion (we treat pfts as species here)

Information about data product

Table A.1

Description of the Data product of 700,000 forests stands from 7 different ecoregions. Each forest patch has an area of 20 m x 20 m. The total number of trees is accumulated over all forest stands from one ecoregion. The species mix is described through the number of pft's. In addition, the mean and maximum values of basal area and biomass are shown.

Region	Number of forest stands	Total number of trees	Number of pfts	Mean basal area einheit	max basal area einheit	mean biomass einheit	max biomass einheit
Amazon	100,000	1379,989	3	26.5	99.3	272	1430
Panama	100,000	611,013	4	15.6	57.6	166.3	752
Germany	100,000	1258,893	8	26.5	99.8	163.7	1462
US	100,000	816,839	9	41.4	99.9	383.7	1623
Ecuador	100,000	3851,850	7	20.7	37.5	116.9	228
Malaysia	100,000	1623,988	4	19.9	67.2	269.9	1067
Tanzania	100,000	1222,930	6	32.8	99.7	300.6	1,091

Reduced climate information

The climate information for Hainich climate (Germany) includes daily temperature values, radiation values and precipitation values. In the case of reduced climate information, other values are used for productivity calculation: mean yearly light intensity above canopy during day-length, length of daily photosynthetic active period, i.e. day-length, relative length of wet and dry season. More information about these variables can be found in respective studies and parameterizations (see main text Table 3).

Realization of envelopes in R

For the analysis of the state space of forests with the help of envelopes we use the function `geom_mark_hull` of the R package `R/mark_hull.R`. It uses the package `concaveman` (github.com/mapbox/concaveman) which allows to adjust concavity of the resulting hull. We choose the following parameters: `con. Cap = 0` and `concavity = 2`.

Additional results for biomass, productivity and species evenness of forest stands

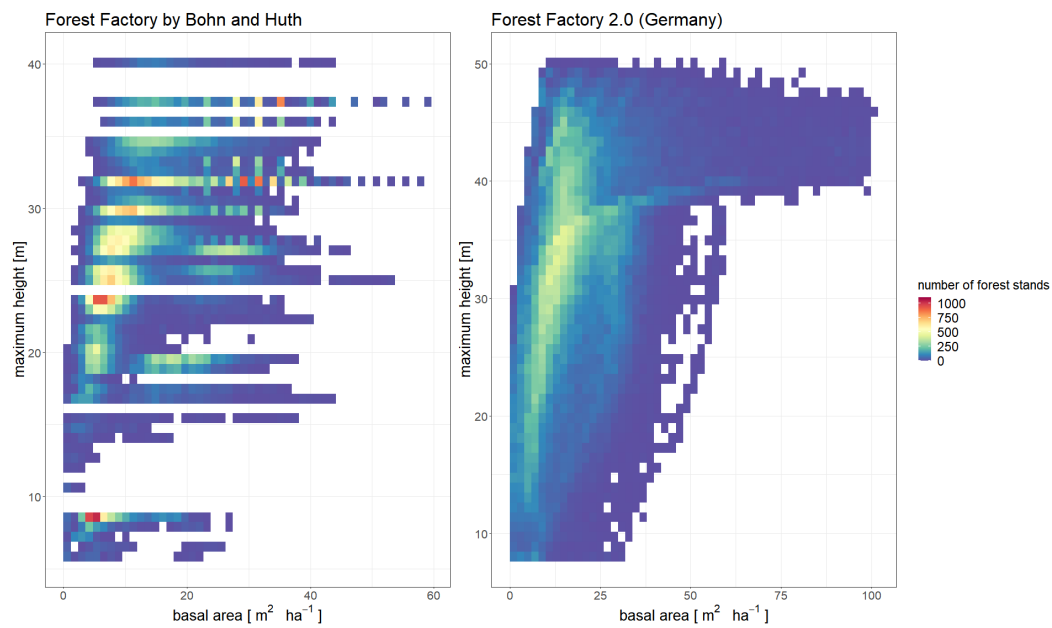


Figure A.1. Comparison of Forest Factory by Bohn & Huth 2017 and Forest Factory 2.0. Distribution of forest properties for the generated forests of the Forest Factory by Bohn & Huth 2017 (left) and Forest Factory 2.0 (right). We used here as structural properties basal area and maximum height per forest stand. The color of each cell in the graph represents the number of forest stands with the respective value of the properties.

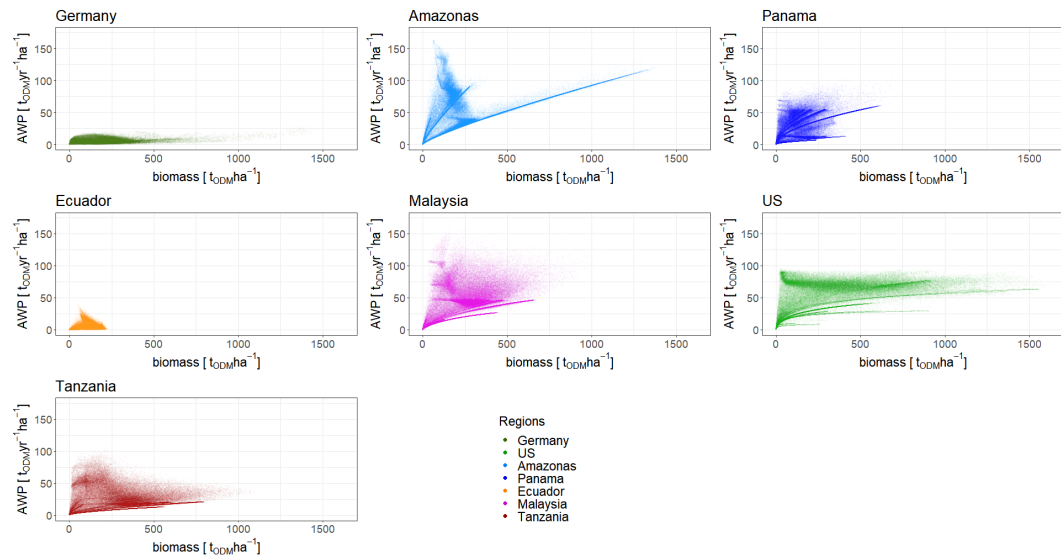


Figure A.2. Relationship between biomass and above-ground productivity for analyzed ecoregions. Each point represents a forest stand in the respective ecoregion (100,000 per ecoregion).

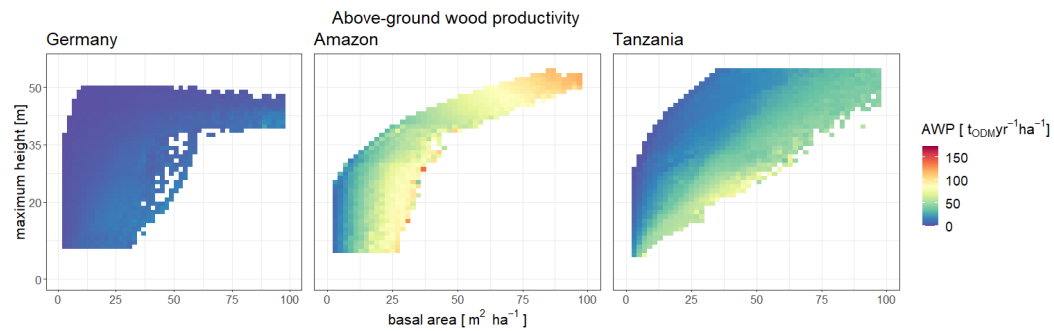


Figure A.3. Relationship between structural properties and above-ground wood productivity. We analyzed here structural properties (basal area and maximum height per forest stand) and above-ground productivity for three selected ecoregions (Germany, Amazon, Tanzania). The color of each cell in the graph represents the mean value of the AWP of all forest stands within one cell. In difference to Fig. 5 (main text), we use for AWP always the same color legend (for all three ecoregions).

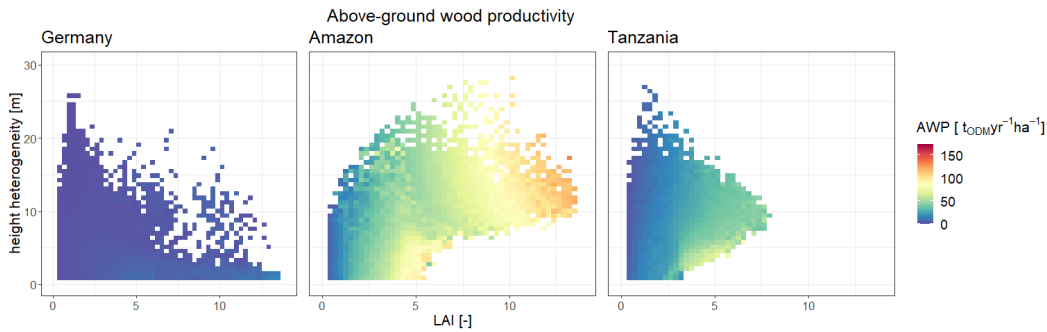


Figure A.4. Relationship between structural properties and above-ground wood productivity. We analyzed here structural properties (height heterogeneity and LAI per forest stand) and above-ground productivity for three selected ecoregions (Germany, Amazon, Tanzania). The color of each cell in the graph represents the mean value of the AWP of all forest stands within one cell. In difference to Fig. 6 (main text), we use for AWP always the same color legend (for all three ecoregions).

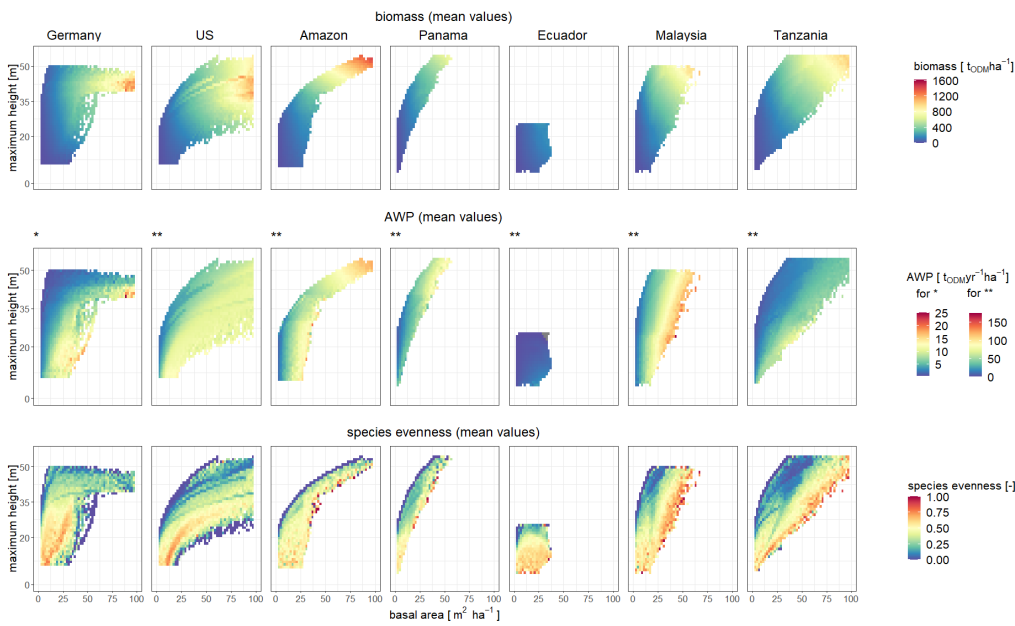


Figure A.5. Relationship between structural properties and functional characteristics of forest stands. Relationship between structural properties (basal area and maximum height per forest stand) and biomass (first row), above-ground productivity (second row) and species evenness (third row) for all analyzed ecoregions. The color of each cell in the graph represents the mean value of the investigated property of all forest stands within one cell.

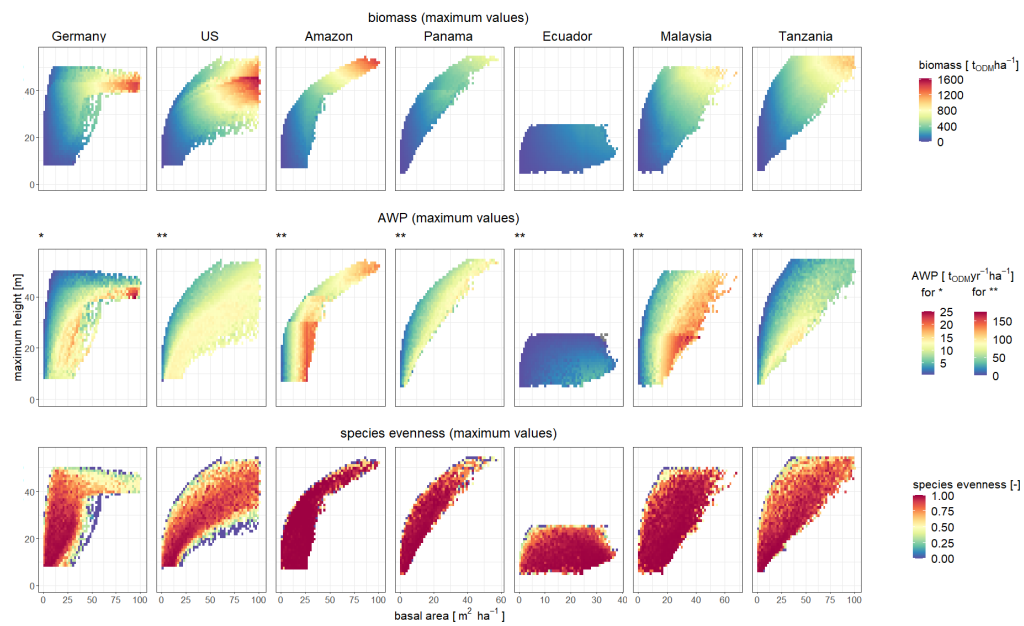


Figure A.6. Maximum values of functional characteristics of forest stands. Relationship between structural properties (basal area and maximum height per forest stand) and biomass (a-g), above-ground productivity (h-n) and species evenness (o-u) for all analyzed ecoregions. The color of each cell in the graph represents the maximum value of the investigated property of all forest stands within one cell.

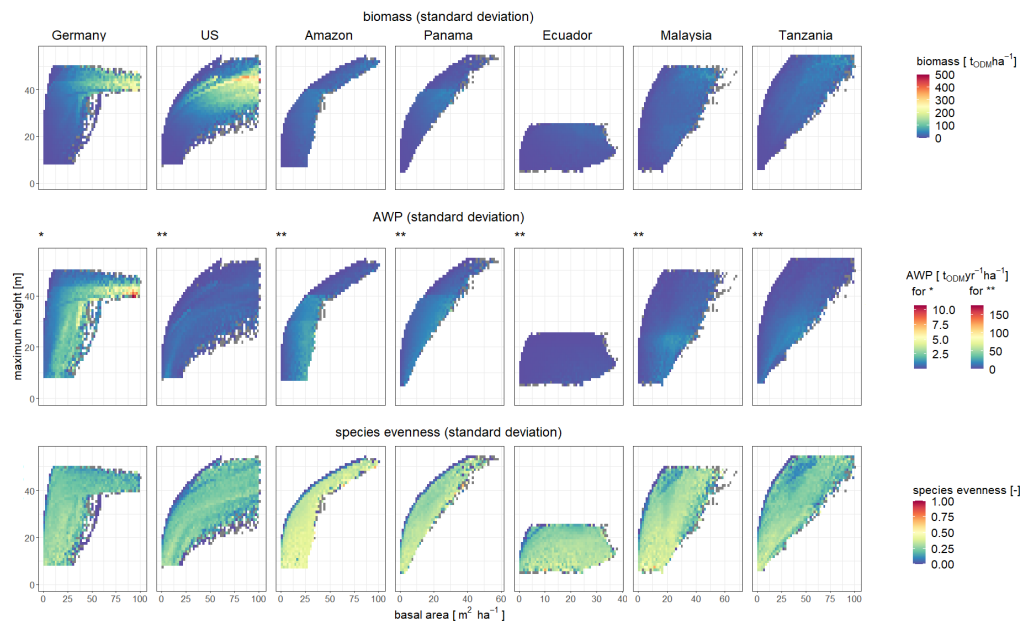


Figure A.7. Standard deviation of functional characteristics of forest stands. Relationship between structural properties (basal area and maximum height per forest stand) and biomass (a-g), above-ground productivity (h-n) and species evenness (o-u) for all analyzed ecoregions. The color of each cell in the graph represents the standard deviation of the investigated property of all forest stands within one cell. Cells consisting of only one forest stand has no standard deviation (gray color).

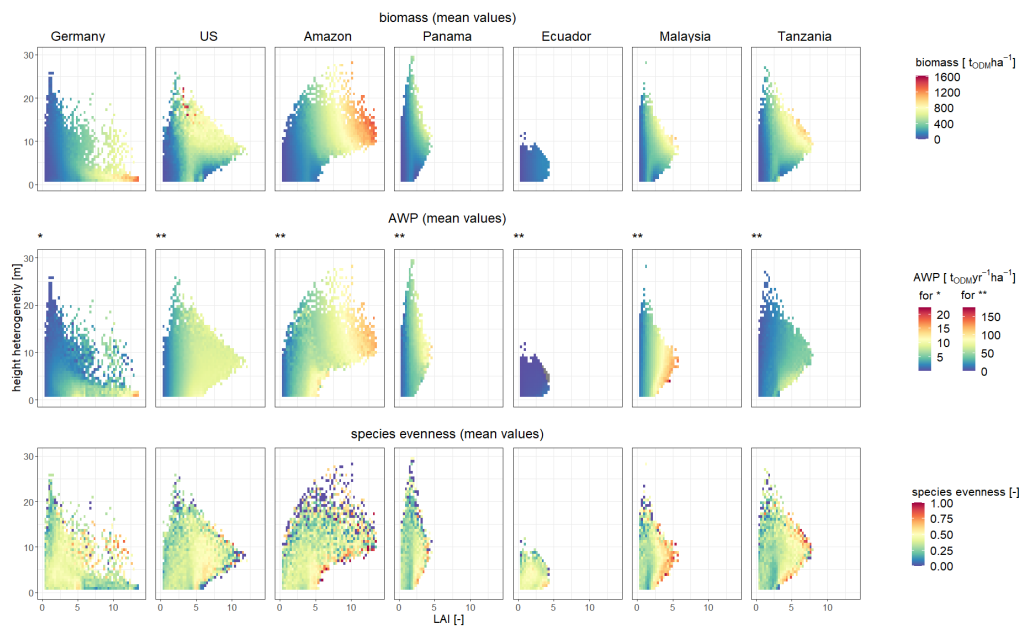


Figure A.8. Relationship between structural properties and functional characteristics of forest stands. Relationship between structural properties (LAI and height heterogeneity per forest stand) and biomass (a-g), above-ground productivity (h-n) and species evenness (o-u) for all analyzed ecoregions. The color of each cell in the graph represents the mean value of the investigated property of all forest stands within one cell.

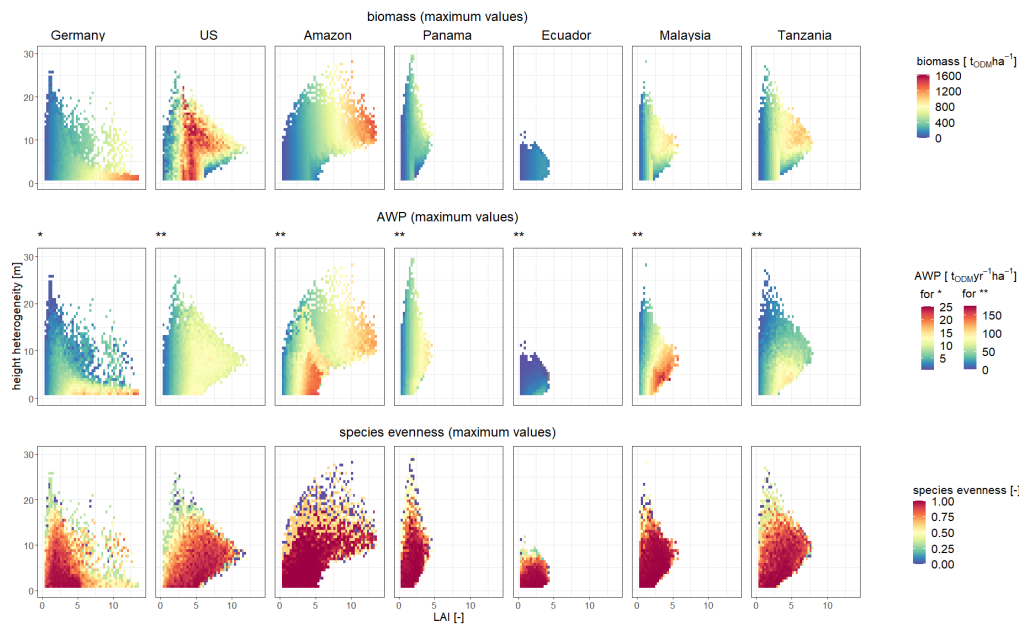


Figure A.9. Maximum values of functional characteristics of forest stands. Relationship between structural properties (LAI and height heterogeneity per forest stand) and biomass (a-g), above-ground productivity (h-n) and species evenness (o-u) for all analyzed ecoregions. The color of each cell in the graph represents the maximum value of the investigated property of all forest stands within one cell.

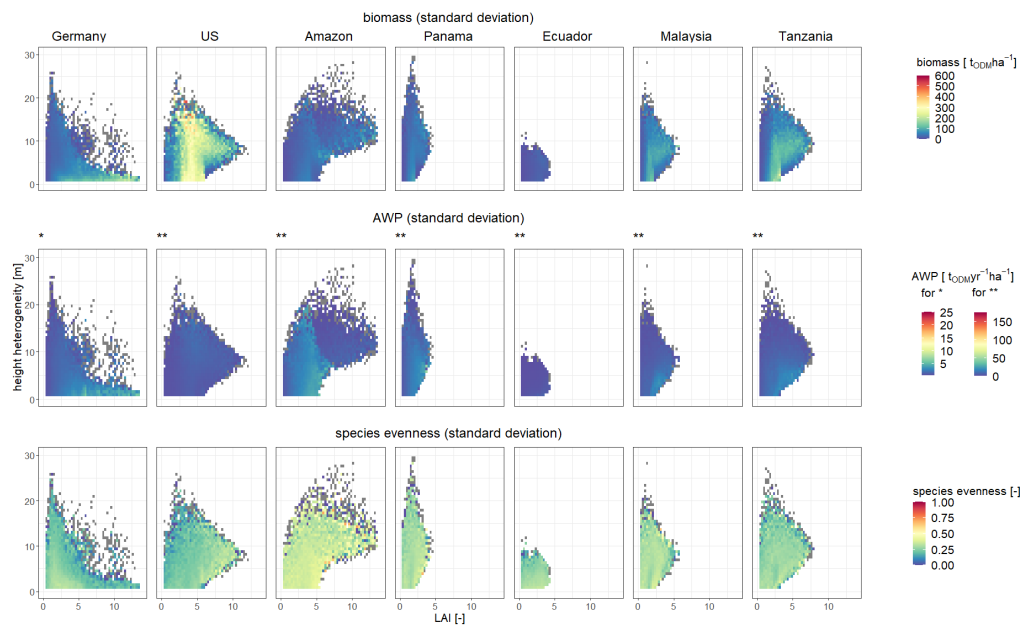


Figure A.10. Standard deviation of functional characteristics of forest stands. Relationship between structural properties (LAI and height heterogeneity per forest stand) and biomass (a-g), above-ground productivity (h-n) and species evenness (o-u) for all analyzed ecoregions. The color of each cell in the graph represents the standard deviation of the investigated property of all forest stands within one cell. Cells consisting of only one forest stand has no standard deviation (gray color).

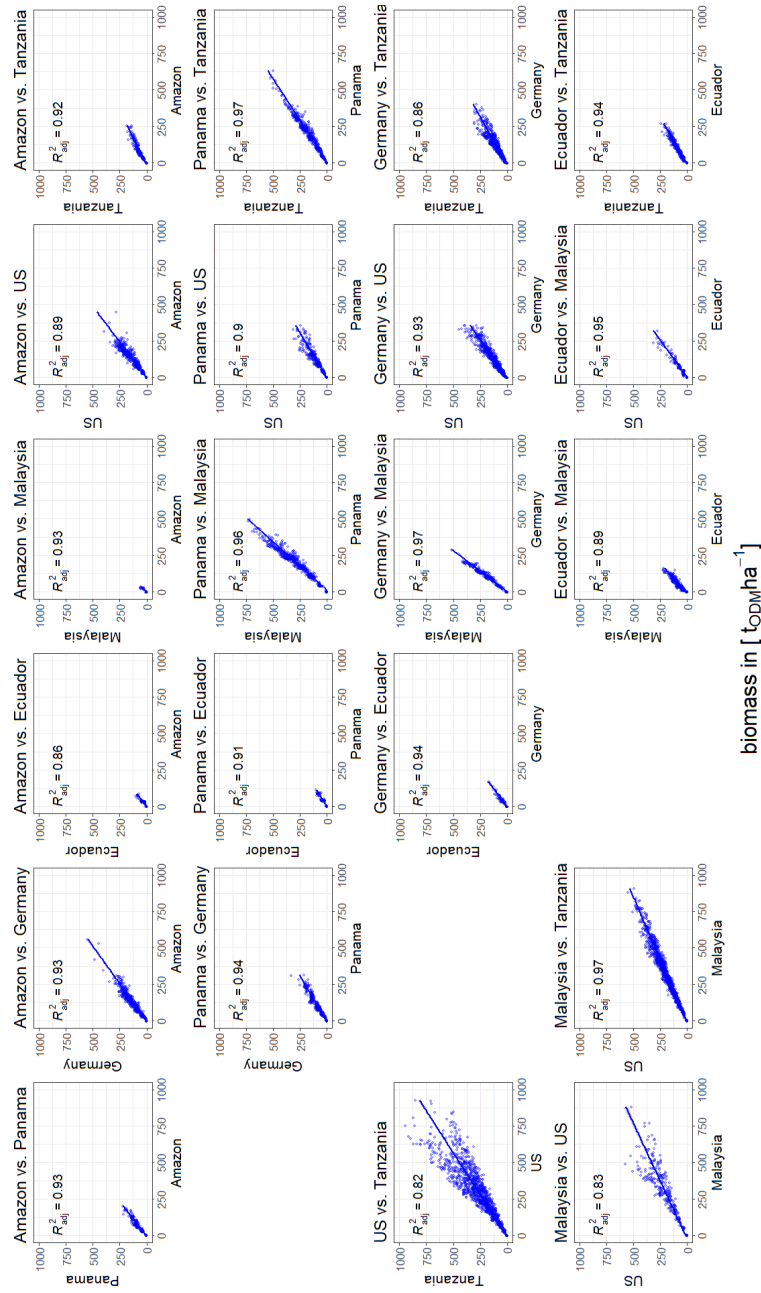


Figure A.11. Comparison of biomass from different regions derived from forests with a similar state space (by a 2% quantile of the four structural properties LAI, basal area, tree height heterogeneity and maximum tree height). Each graph compares forest stands out of two ecoregions. We show the pairwise comparisons for all ecoregions. Each point represents the mean values of biomass $t_{odmyr}^{-1}ha^{-1}$ for both regions.

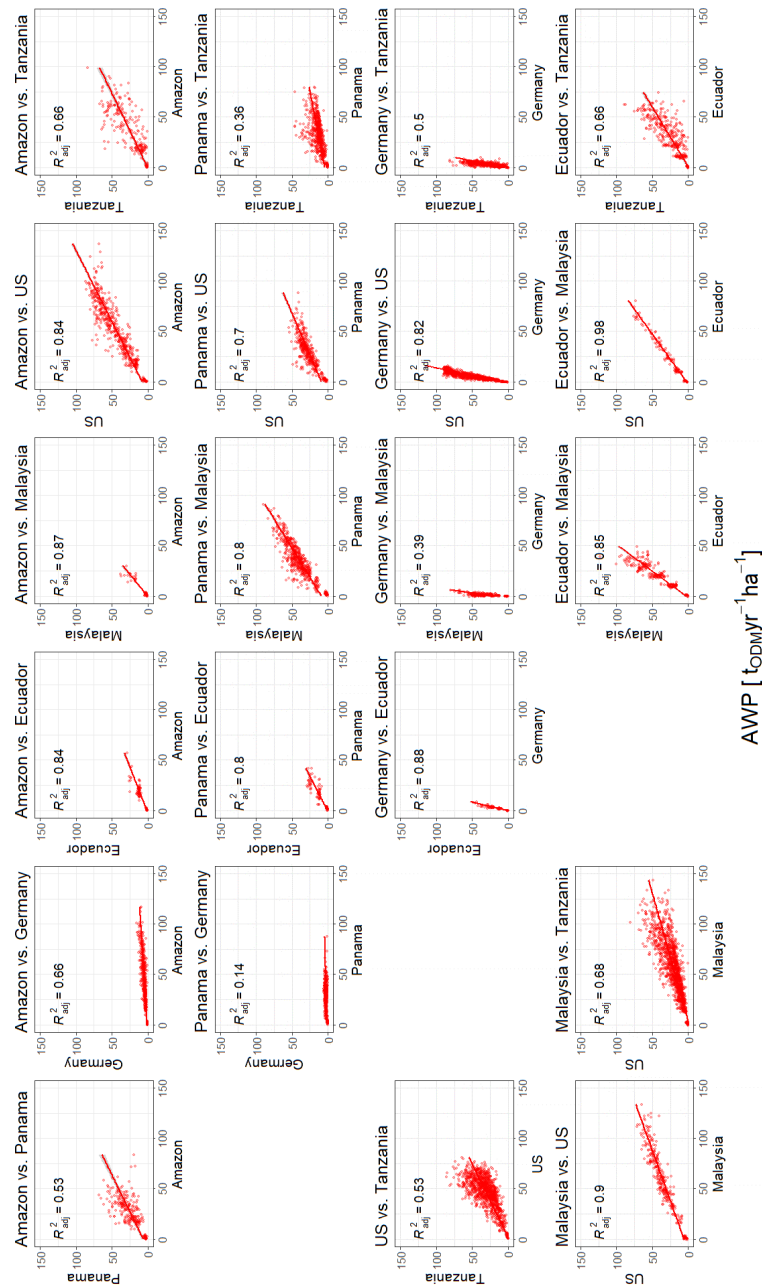


Figure A.12. Comparison of above-ground productivity derived from forests with a similar state space (by a 2% quantile of the four structural properties LAI, basal area, tree height heterogeneity and maximum tree height). Each graph compares forest stands out of two ecoregions. We show the pairwise comparisons for three illustrative ecoregions. Each point represents the mean values of AWP $t_{odmyr}^{-1}ha^{-1}$ for both regions.

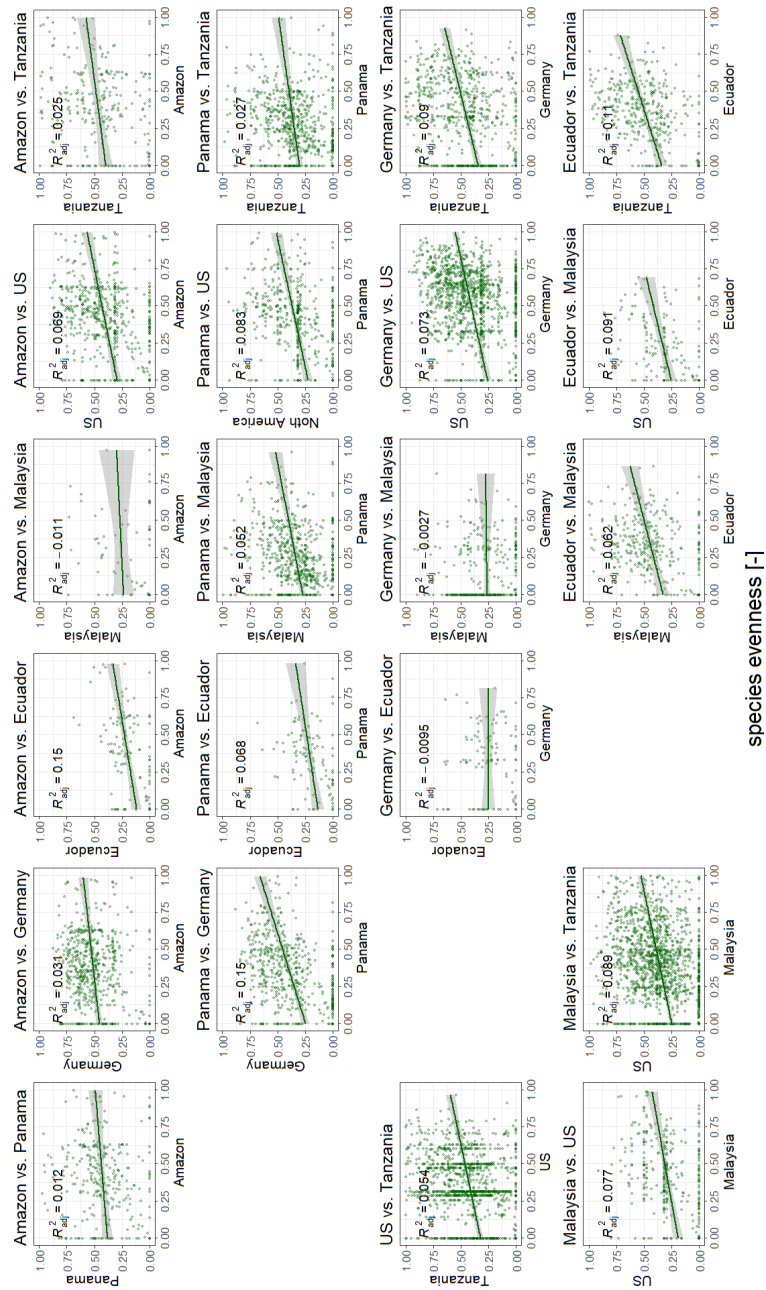


Figure A.13. Comparison of species evenness derived from forests with a similar state space (by a 2% quantile of the four structural properties LAI, basal area, tree height heterogeneity and maximum tree height). Each graph compares forest stands out of two ecoregions. We show the pairwise comparisons for three illustrative ecoregions. Each point represents the mean values of species evenness [-] for both regions.

B Appendix Chapter 3

Additional Information on the Method Section



Figure B.1. Study map of the 28 forest stands in the region North Karelia (blue area), Finland. The forest stands are distributed over an area of 150 km \times 150 km.

Table B.1

Leaf parameters

Leaf Parameter	Picea Abies	Pinus Silvestrys	Betula (Pendula & Pubescens)
Cab [$\mu\text{g cm}^{-2}$]	21.94	23.92	36.71
Cdm [g cm^{-2}]	0.024	0.025	0.006
Cw [g cm^{-2}]	0.03	0.03	0.0117
Cs [–]	0.01	0.01	0.01
Car [$\mu\text{g cm}^{-2}$]	4.40	4.50	8.62
N [–]	1.25	1.24	1.77

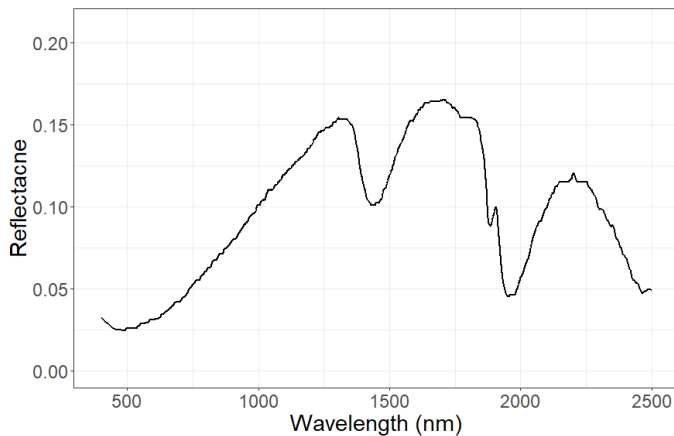


Figure B.2. Soil reflection. Shown is assumed the reflection of wet soil. We assume for all forest stands the same soil reflectance.

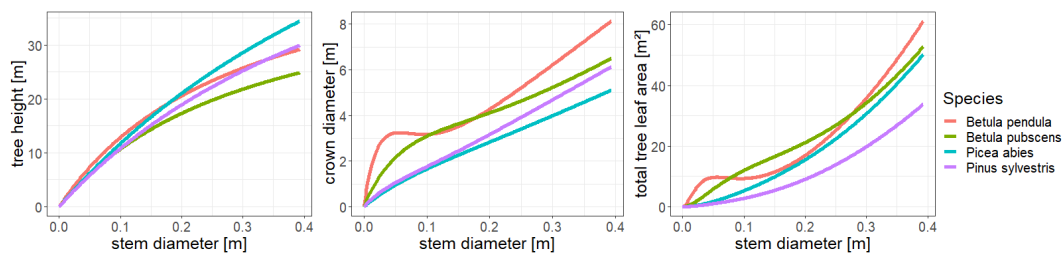


Figure B.3. Analysis of allometries in FORMIND. Shown are the relationships between stem diameter and tree height, crown diameter as well as the total leaf area of a tree for the tree species simulated in FORMIND.

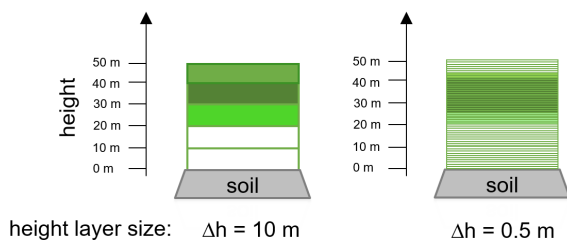


Figure B.4. Concept of height layers. Shown is the same forest with a different resolution of height layers ($\Delta h = 10$ m and $\Delta h = 0.5$ m). Each layer includes different species mixtures (indicated by color). We use 5 height layers in the case $\Delta h = 10$ m and 100 height layers in the case $\Delta h = 0.5$ m. All layers which contain leaves contribute to the resulting reflectance spectrum. Empty height layers (no leaves are in the layer) do not influence the reflectance spectrum (details see Section 3.3.4).

Table B.2

Attributes of the forest stands from Finland used for this study (28 plots, 30 m × 30 m).

Plot Number	Basal Area	Maximum Height	Height Heterogeneity	Species Richness	Species Evenness	Bio-mass	LAI
[–]	[m ² ha ⁻¹]	[m]	[m]	[–]	[–]	[t _{odm} ha ⁻¹]	[–]
1	28.78	30.09	5.31	3	0.49	148.30	3.93
2	19.63	28.63	3.47	2	0.35	106.83	3.23
3	16.31	26.38	2.88	2	0.26	90.74	2.69
4	22.25	29.21	4.54	2	0.41	118.74	3.46
5	19.87	29.71	4.64	5	0.74	109.47	2.26
6	32.84	30.80	4.04	2	0.04	170.06	4.84
7	17.60	23.84	3.36	4	0.16	94.34	3.13
8	17.33	24.59	2.37	3	0.17	95.09	2.94
9	25.66	26.10	3.69	5	0.25	133.24	2.39
10	27.03	30.01	4.59	2	0.05	139.91	3.98
11	17.02	22.70	2.92	3	0.27	85.02	3.15
12	27.35	23.32	3.39	3	0.65	114.86	4.07
13	20.38	21.67	3.28	3	0.59	88.96	2.64
14	18.37	23.34	2.77	1	0.00	86.37	1.67
15	21.37	26.22	3.38	3	0.44	105.38	2.37
16	27.37	26.44	4.08	3	0.10	139.98	2.52
17	24.54	28.80	4.52	2	0.38	125.61	2.58
18	30.99	28.45	4.24	2	0.07	152.92	4.71
19	30.12	28.74	3.61	3	0.47	162.74	3.56
20	30.42	30.30	3.92	2	0.35	164.55	4.52
21	17.60	23.34	2.45	2	0.06	84.49	1.60
22	27.16	25.85	3.96	2	0.43	120.19	3.54
23	29.31	27.40	4.08	2	0.16	139.35	4.44
24	25.45	21.70	2.63	3	0.38	89.55	5.04
25	30.72	27.96	4.61	3	0.62	145.02	4.58
26	26.66	34.41	5.57	4	0.69	138.13	3.67
27	22.03	22.77	2.95	3	0.55	95.38	2.64
28	22.10	25.73	3.21	3	0.50	114.78	2.97

Basal area is defined by the cross-sectional area of trees at breast height. Maximum height describes the highest tree height in the forest stand. Height heterogeneity describes the standard deviation of tree height. Richness describes the number of species in the forest. Evenness is defined by the normalized Shannon index. Biomass describes the sum of all tree biomass.

Table B.3

Spectral configuration of the 10 Sentinel-2A bands used in this study Ma et al. 2019.

Spectral Band	Center Wavelength [nm]	Band Name	Band Width [nm]	Spatial Resolution [m]
B02	490	blue	65	10
B03	560	green	35	10
B04	665	red	30	10
B05	705	red-edge 1	15	20
B06	740	red-edge 2	15	20
B07	783	red-edge 3	20	20
B08	842	NIR 1	115	10
B08a	865	NIR 2	20	20
B11	1610	SWIR 1	90	20
B12	2190	SWIR 2	180	20

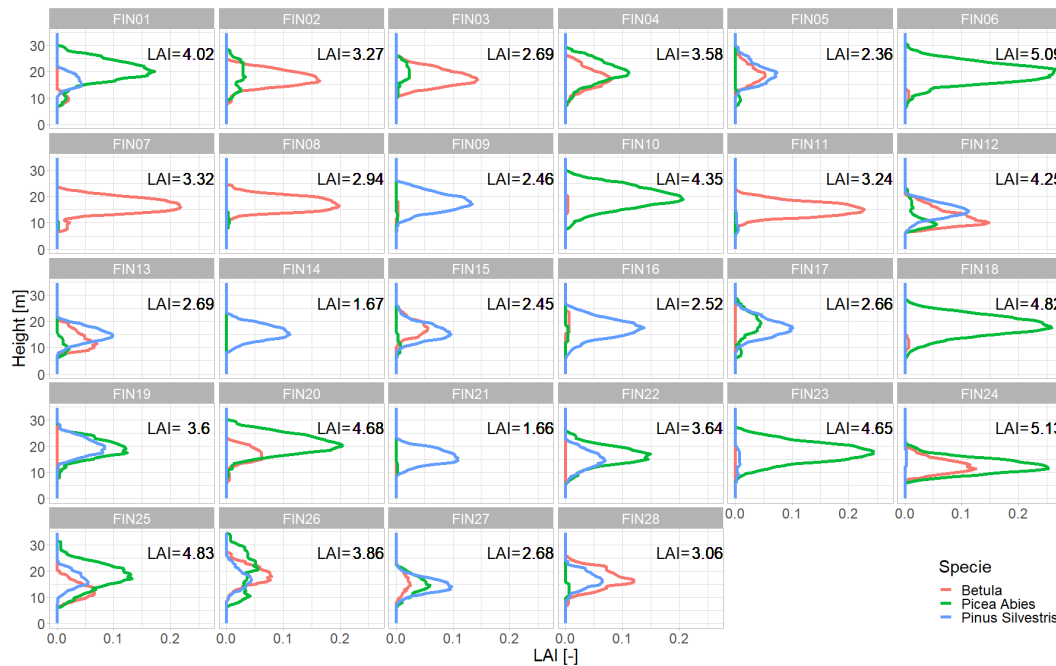


Figure B.5. LAI Profiles for each forest stand. Shown are the LAI values (x-axis) in each height layer (y-axis) per species (red—Betula, green—Picea Abies, blue—Pinus Sylvestris). The colored lines show the LAI for a particular species (sum of all trees of the species in the plot). Therefore, the sum of all the lines gives the LAI profile of all trees in the plot.

Additional Information on the Result Section

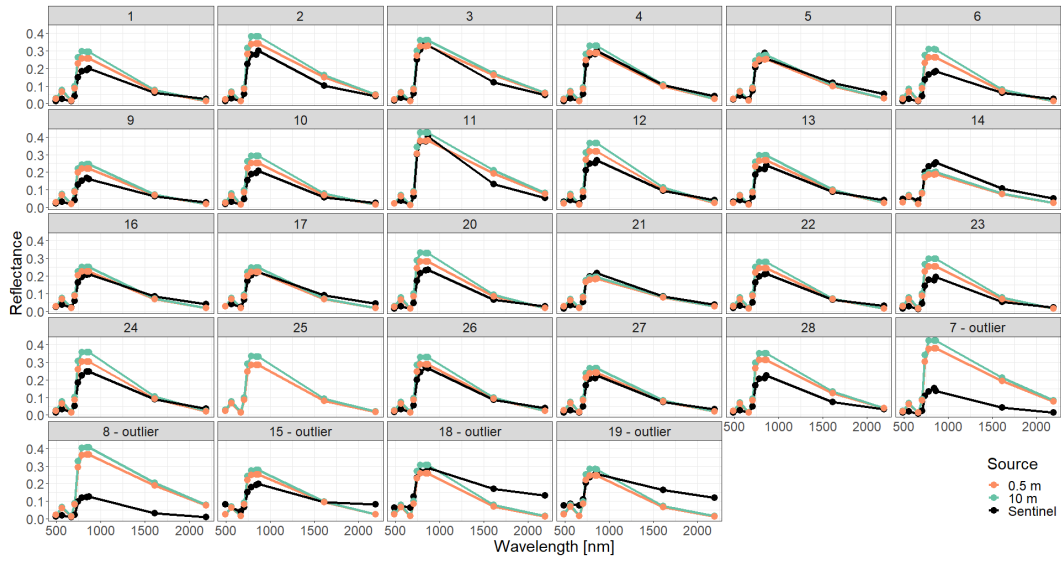


Figure B.6. Comparison of simulated reflectance spectra with Sentinel measurements assuming a simple forest representation using different descriptions of the vertical forest structure (0.5 m or 10 m height layers). The Sentinel 2 spectrum of plot 25 is not provided by Ma et al. 2019.

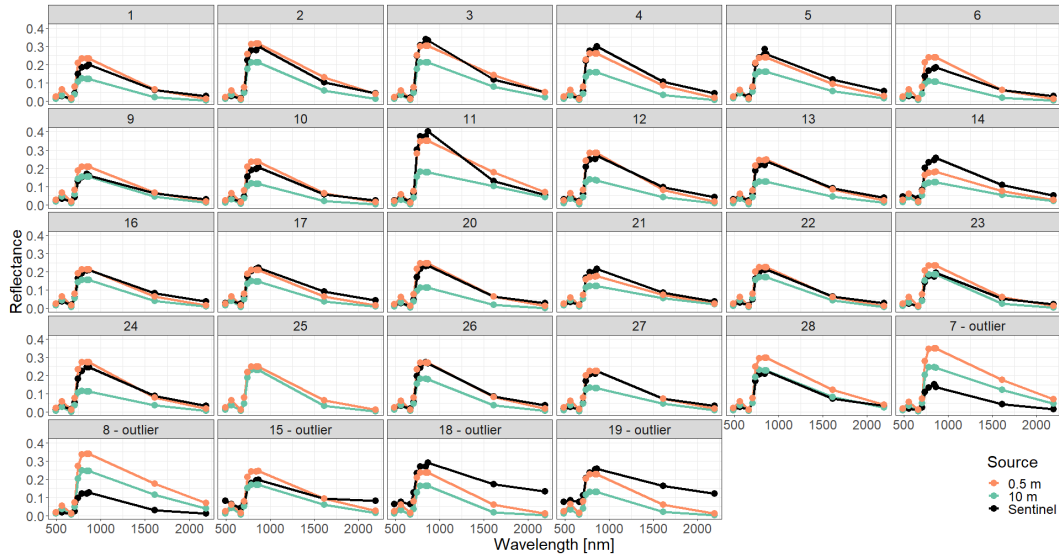


Figure B.7. Comparison of simulated reflectance spectra with Sentinel measurements assuming a detailed forest representation using different descriptions of the vertical forest structure (0.5 m or 10 m height layers).

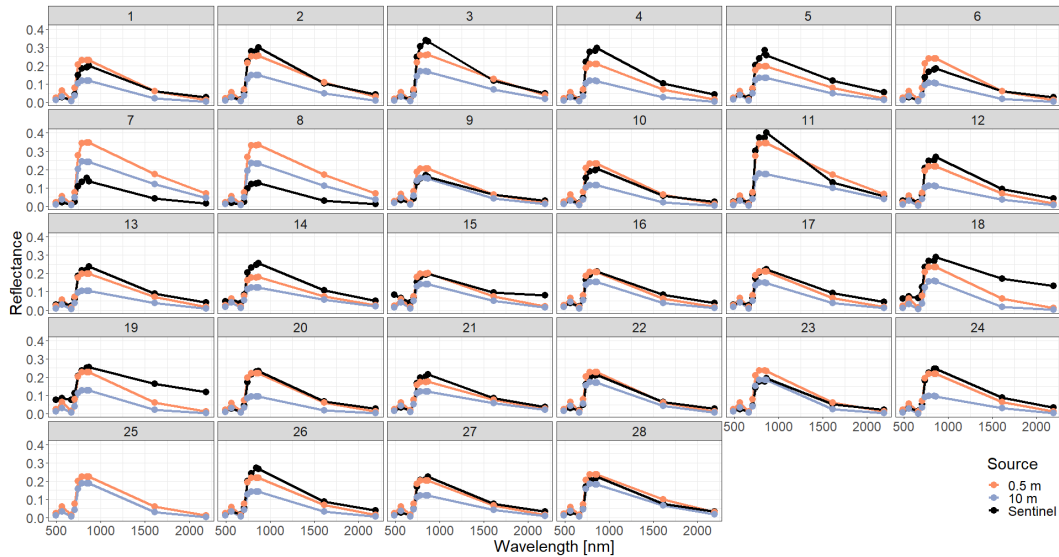


Figure B.8. Comparison of simulated reflectance spectra with Sentinel measurements assuming a spectra-averaged forest representation using different descriptions of the vertical forest structure (0.5 m or 10 m height layers).

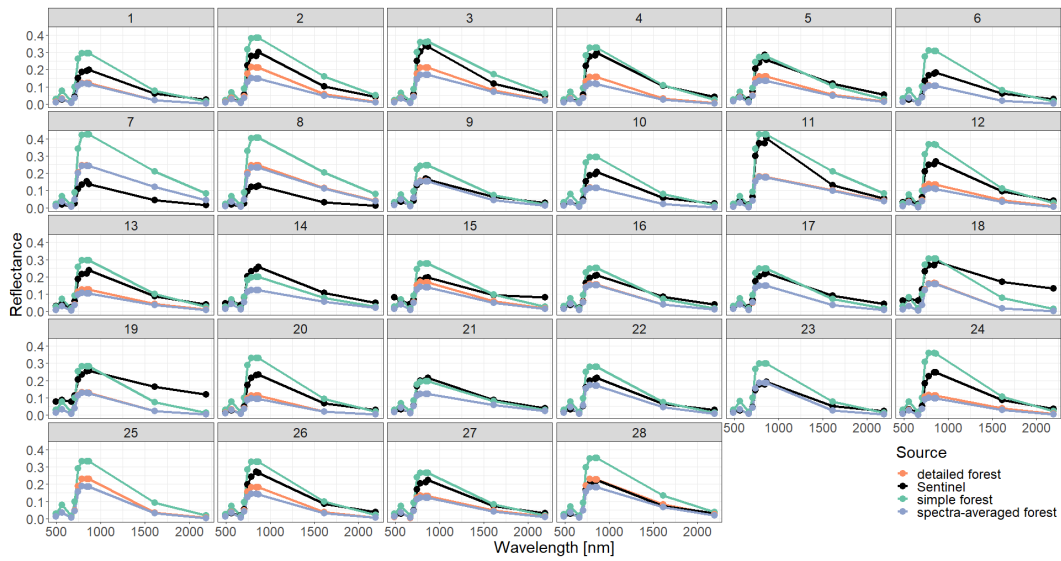


Figure B.9. Comparison of simulated reflectance spectra with Sentinel measurements assuming different types of forest representations (simple, detailed and spectra-averaged) and using 10 m height layers for the description of vertical forest structure.

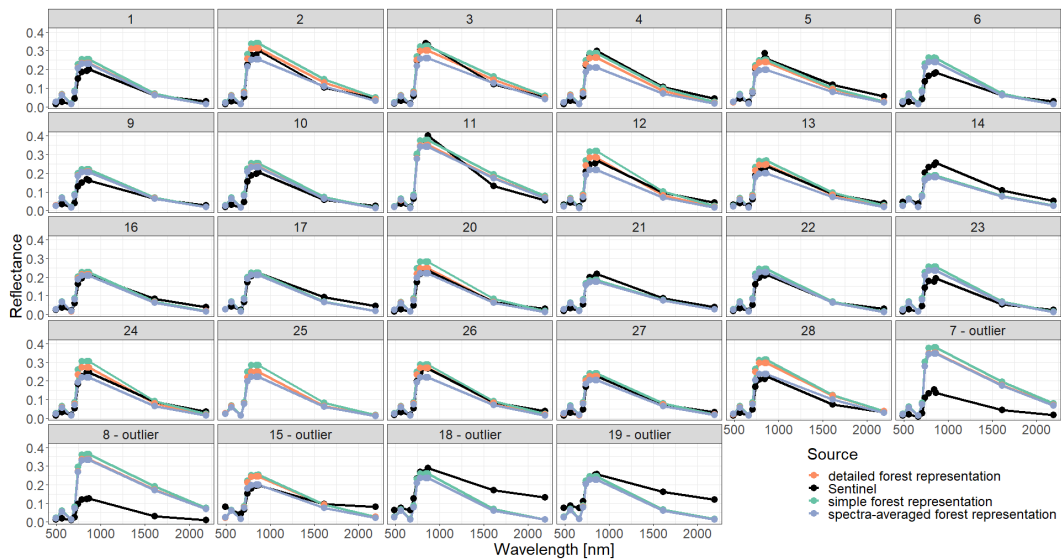


Figure B.10. Comparison of simulated reflectance spectra with Sentinel measurements assuming different types of forest representations (simple, detailed, spectra-averaged) and using 0.5 m height layers for the description of vertical forest structure.

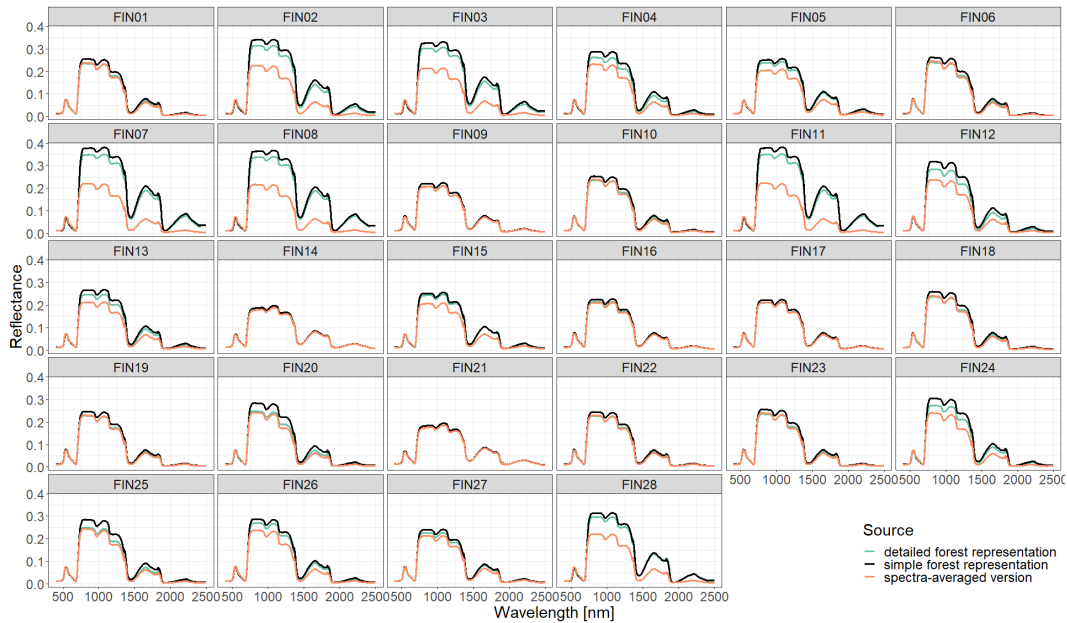


Figure B.11. Comparison of simulated reflectance spectra assuming different types of forest representations (simple, detailed, spectra averaged) and using 0.5 m height layers for the description of vertical forest structure. The resolution of the simulated reflectance wavelengths is 1 nm.

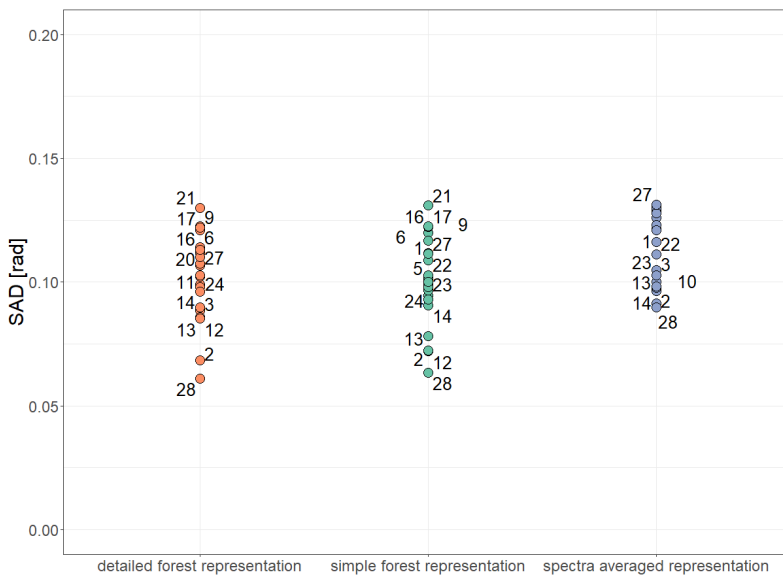


Figure B.12. Comparison of simulated reflectance spectra and measured reflectance spectra using a distance index (Spectral Angle Distance, 0 rad: identical, $\frac{\pi}{2}$ rad: different). Results are shown for 28 forest plots (dots) and different forest representations. Comparison has been conducted for 10 wavebands of Sentinel-2.

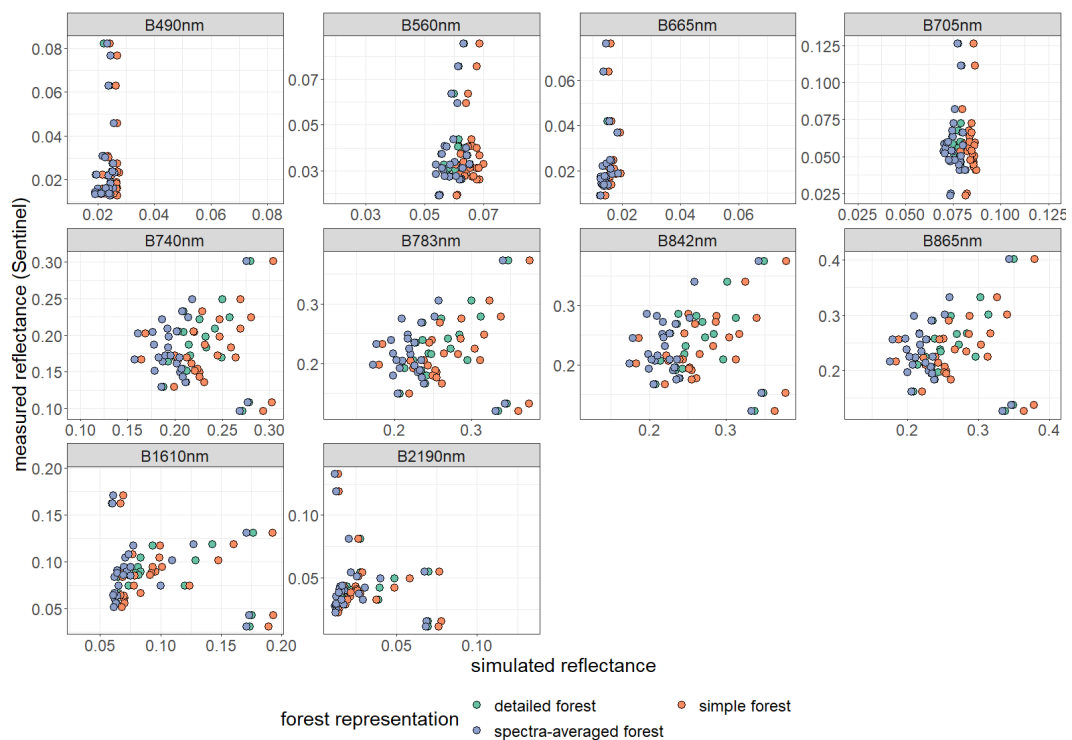


Figure B.13. Comparison of simulated and measured reflectance of 28 forest stands (dots) and different forest representations (indicated by colors). Please note that the scales used for the illustration of measured and simulated reflectance differs for each band. Reflectance has been averaged for 10 wavebands (described by centered wavelengths; for more information on wavebands see Table B.3).

Analysis of Selected Forest Stands (Outliers)

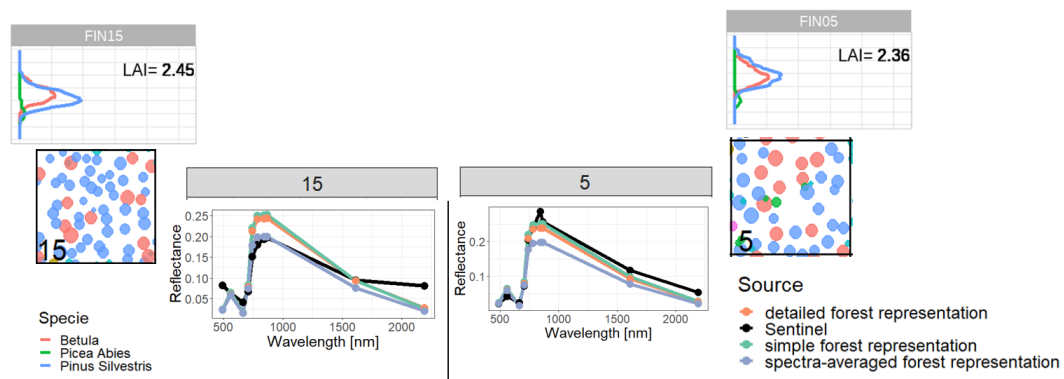


Figure B.14. Comparison of reflectance spectra and additional information of forest stand 15 (classified as outlier) with forest stand 5. We compare forest properties of an outlier (left side) with forest properties of a forest with similar attributes, which is not an outlier (right side). Therefore, we compare the LAI profile (outer sides top), the reflectance spectra (in the middle) and the species composition (outer sides bottom). Despite the similar LAI distribution and species composition we obtained different Sentinel-measurements for reflectance, but similar simulated reflectance.

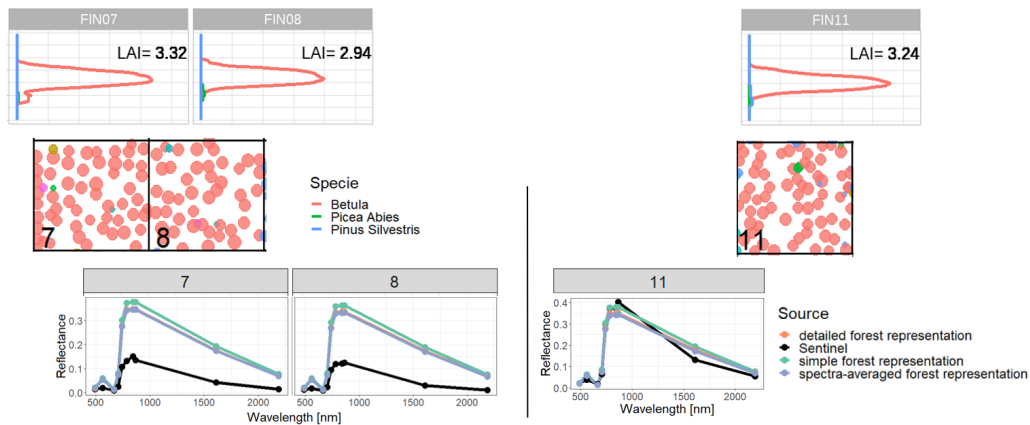


Figure B.15. Comparison of reflectance spectra and additional information of forest stands 7 and 8 (classified as outliers) with forest stand 11. We compare forest properties of two outliers (left side) with forest properties of a forest with similar attributes, which is not an outlier (right side). Therefore, we compare the LAI profile (outer sides top), the reflectance spectra (in the middle) and the species composition (outer sides bottom). Despite the similar LAI distribution and species composition, we obtained different Sentinel measurements for reflectance, but similar simulated reflectance.

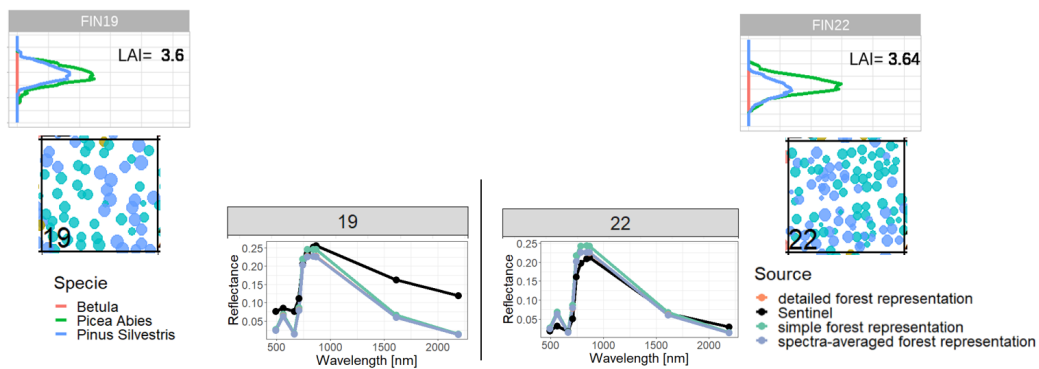


Figure B.16. Comparison of reflectance spectra and additional information of forest stand 19 (classified as outlier) with forest stand 22. We compare forest properties of an outlier (left side) with forest properties of a forest with similar attributes, which is not an outlier (right side). Therefore, we compare the LAI profile (outer sides top), the reflectance spectra (in the middle) and the species composition (outer sides bottom). Despite the similar LAI distribution and species composition, we obtained different Sentinel measurements for reflectance, but similar simulated reflectance.

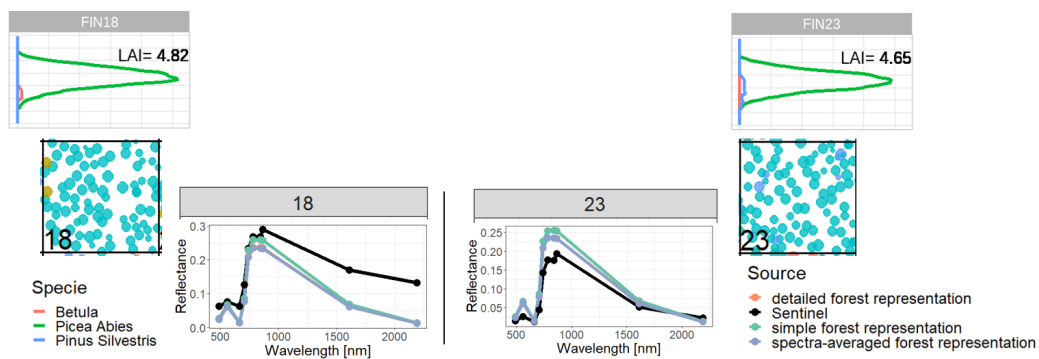


Figure B.17. Comparison of reflectance spectra and additional information of forest stand 18 (classified as outlier) with forest stand 23. We compare forest properties of an outlier (left side) with forest properties of a forest with similar attributes, which is not an outlier (right side). Therefore, we compare the LAI profile (outer sides top), the reflectance spectra (in the middle) and the species composition (outer sides bottom). Despite the similar LAI distribution and species composition, we obtained different Sentinel measurements for reflectance, but similar simulated reflectance.

Analysis of LAI and Additional Indices

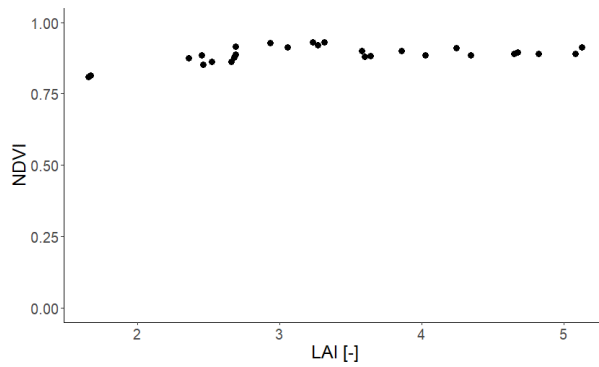


Figure B.18. Relationship between LAI (x-axis, field data) and NDVI (y-axis, Sentinel-2 measurements) of 28 Finland forest stands.

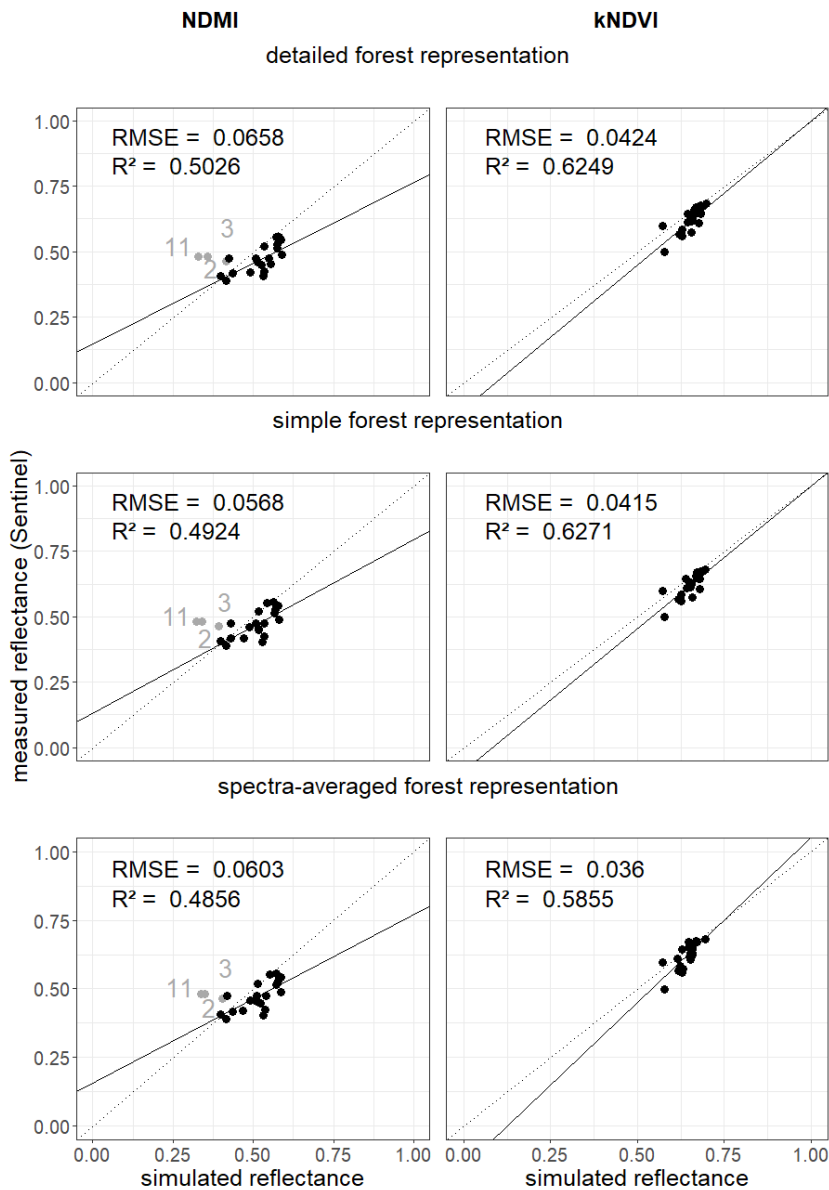


Figure B.19. Comparison of vegetation indices for 28 forest stands in Finland. The vegetation indices (NDMI left, kNDVI right) for the measured reflectance spectra in the different wavebands for the simulated reflectance spectra (x-axis) and for the satellite measurements (y-axis). In each row, a different forest representation is assumed (1. detailed forest representation, 2. simple forest representation, 3. spectra-averaged forest representation; more information about the cases in Section 3.3.3). Each point represents a forest stand in Finland (gray points indicate outliers that are not used to calculate the RMSE and R^2 —see Appendix Figures B.14–B.17).

Table B.4

Analysis of vegetation indices for different forest representations. The mean spectral angle distance is calculated as average of the SAD of 28 forest stands for each forest representation (height layer size = 0.5 m). More details about the vegetation indices and spectral angle distance can be found in Section 3.3.4, Figures 12 and B.19.

	Simple Forest	Detailed Forest	Spectra Averaged Forest
NDVI			
R^2	0.63	0.63	0.59
bias R^2	-0.086	-0.097	-0.177
RMSE	0.04	0.04	0.033
MAE	0.033	0.034	0.027
EVI			
R^2	0.43	0.45	0.25
bias R^2	0.086	0.059	0.116
RMSE	0.107	0.081	0.074
MAE	0.092	0.069	0.062
MSI			
R^2	0.49	0.49	0.47
bias R^2	0.141	0.162	0.164
RMSE	0.051	0.059	0.054
MAE	0.041	0.05	0.043
NDMI			
R^2	0.49	0.5	0.49
bias R^2	0.133	0.15	0.156
RMSE	0.057	0.066	0.060
MAE	0.046	0.056	0.049
kNDVI			
R^2	0.63	0.62	0.59
bias R^2	-0.089	-0.098	-0.15
RMSE	0.041	0.042	0.036
MAE	0.034	0.035	0.029
mean SAD	0.101	0.103	0.113

C Appendix Chapter 4

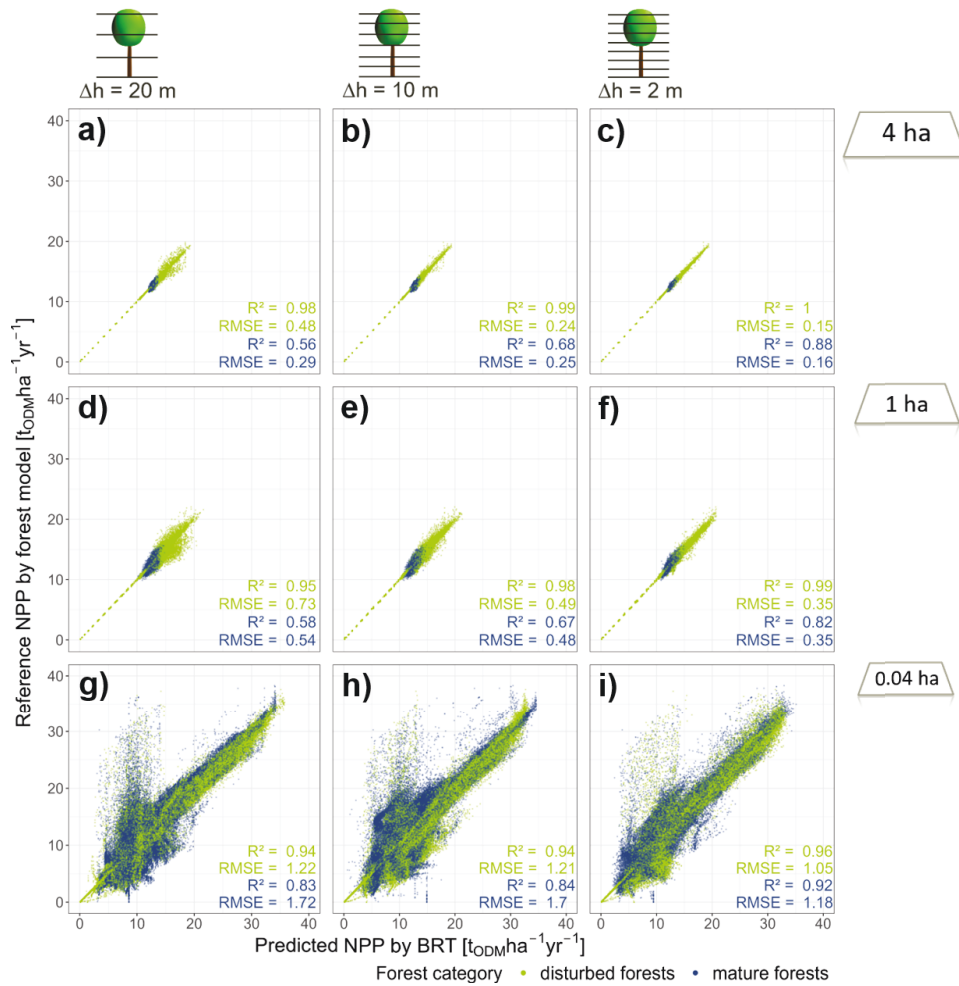


Figure C.1. Comparison of predicted NPP derived from the boosted regression tree and reference NPP derived from simulations using the forest model FORMIND. The comparison is done for disturbed forests (green) and natural forests (blue). The predicted NPP from biomass measurements is derived from the simulated vertical and horizontal biomass distributions (descriptions above and on the right side). Each point represents a result for a forest. Total simulated area was 100 ha (resulting in the different amount of points in the rows).

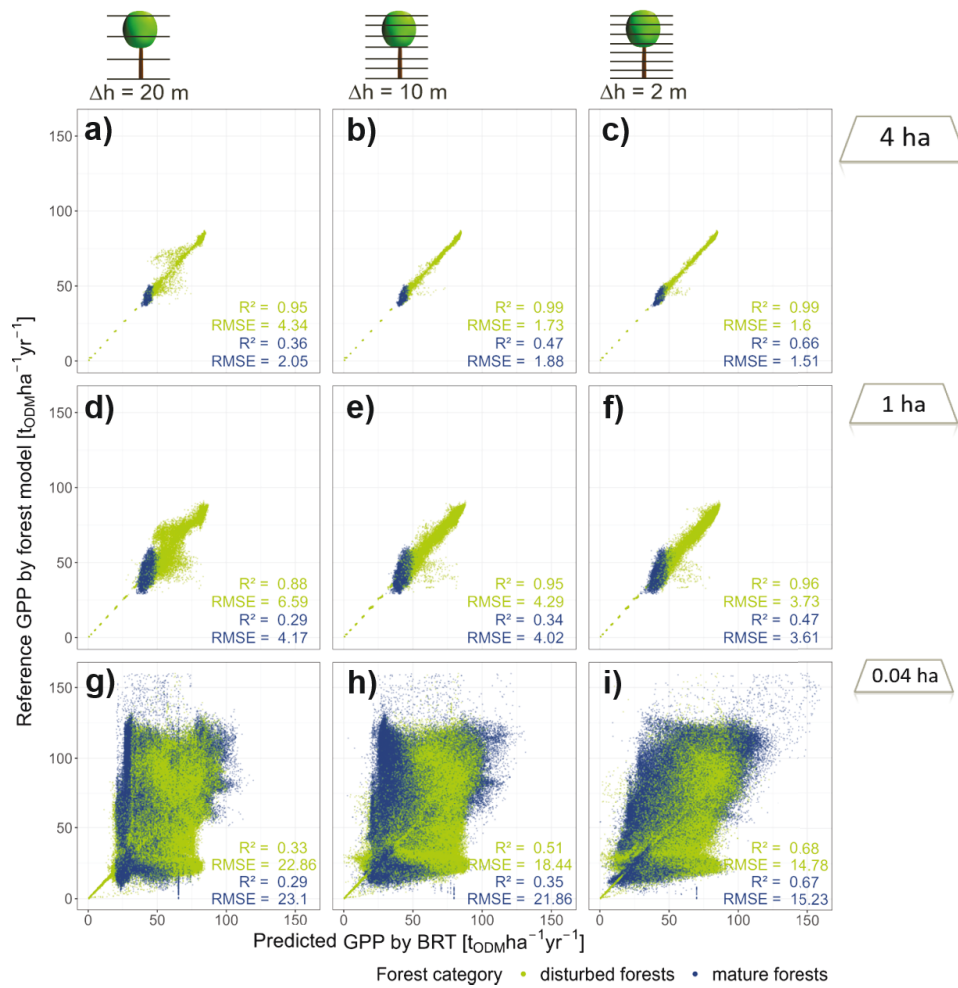


Figure C.2. Comparison of predicted GPP derived from the boosted regression tree and reference GPP derived from simulations using the forest model FORMIND. The comparison is done for disturbed forests (green) and natural forests (blue). The predicted GPP from biomass measurements is derived from the simulated vertical and horizontal biomass distributions (descriptions above and on the right side). Each point represents a result for a forest. Total simulated area was 100 ha (resulting in the different amount of points in the rows).

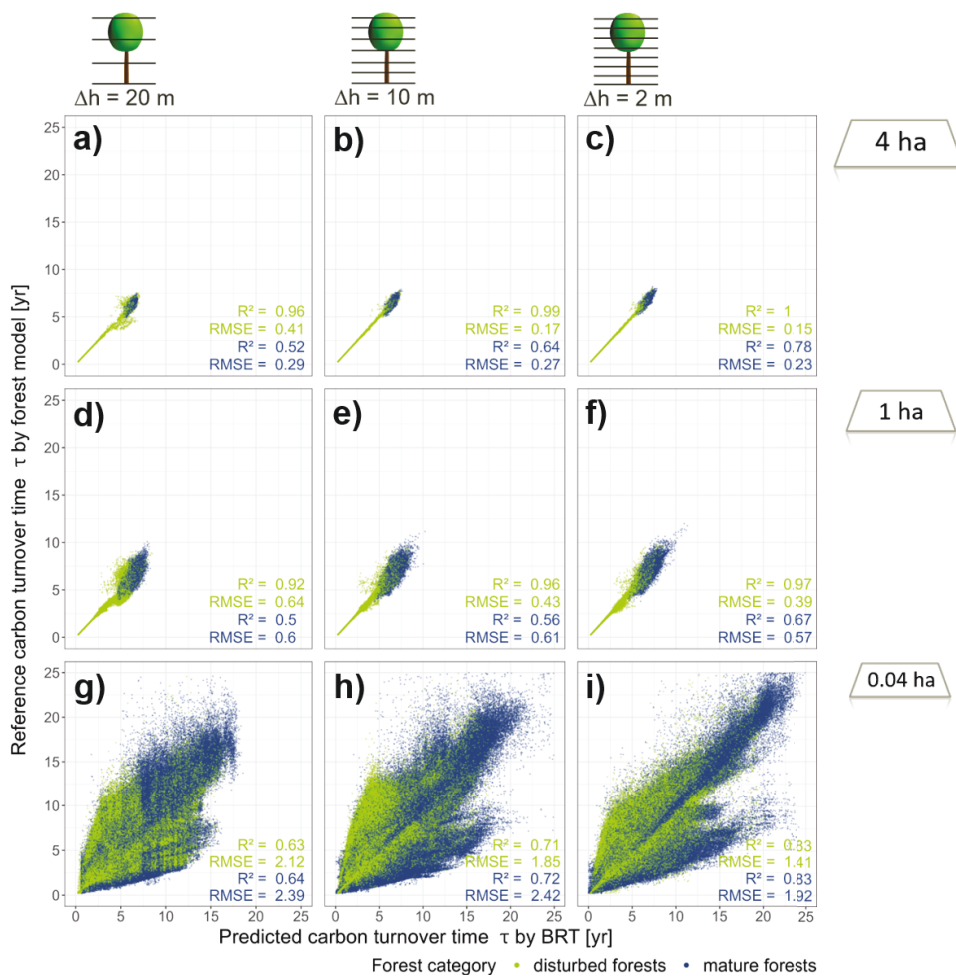


Figure C.3. Comparison of predicted carbon turnover time (τ) derived from the boosted regression tree and reference τ derived from simulations using the forest model FOR-MIND. The comparison is done for disturbed forests (green) and natural forests (blue). The predicted τ from biomass measurements is derived from the simulated vertical and horizontal biomass distributions (descriptions above and on the right side). Each point represents a result for a forest. Total simulated area was 100 ha (resulting in the different amount of points in the rows).

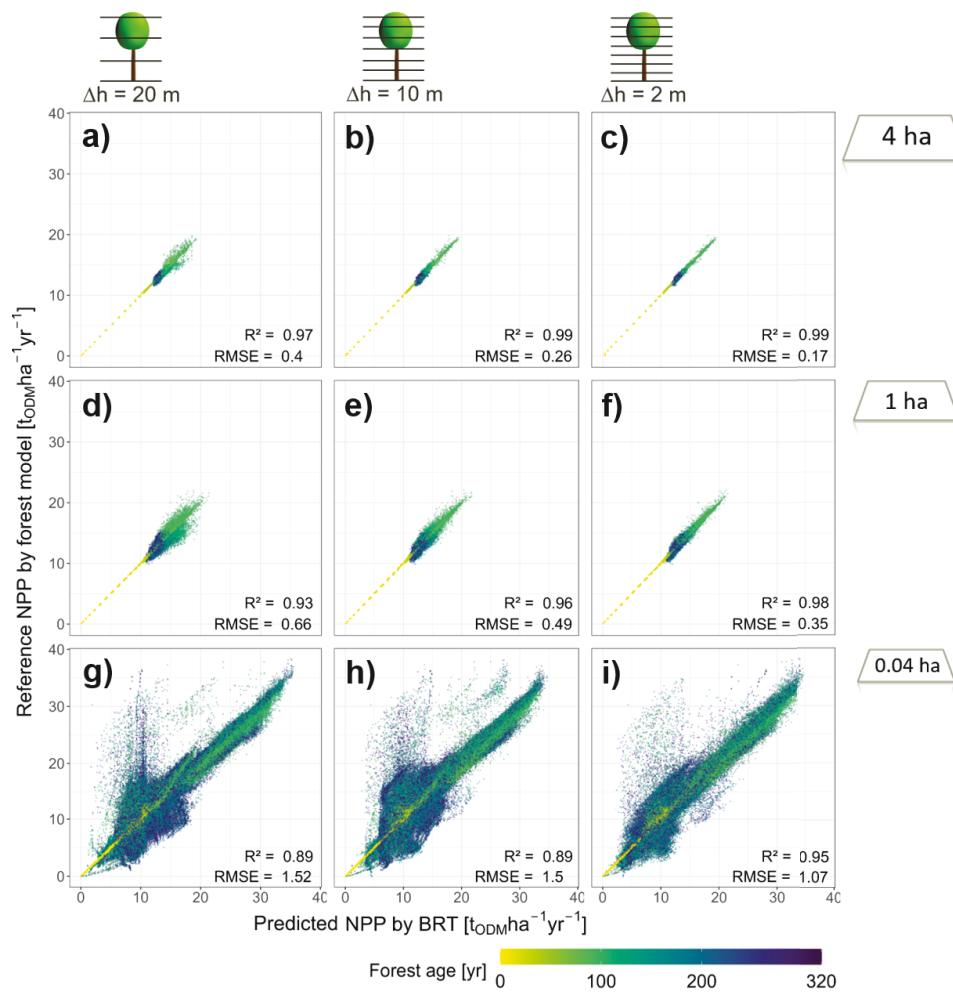


Figure C.4. Comparison of predicted NPP derived from the boosted regression tree and reference NPP derived from simulations using the forest model FORMIND. The comparison is done for the whole forest data. The predicted NPP from biomass measurements is derived from the simulated vertical and horizontal biomass distributions (descriptions above and on the right side). Each point represents a results for a forest with a certain age (color of points). Total simulated area was 100 ha (resulting in the different amount of points in the rows).

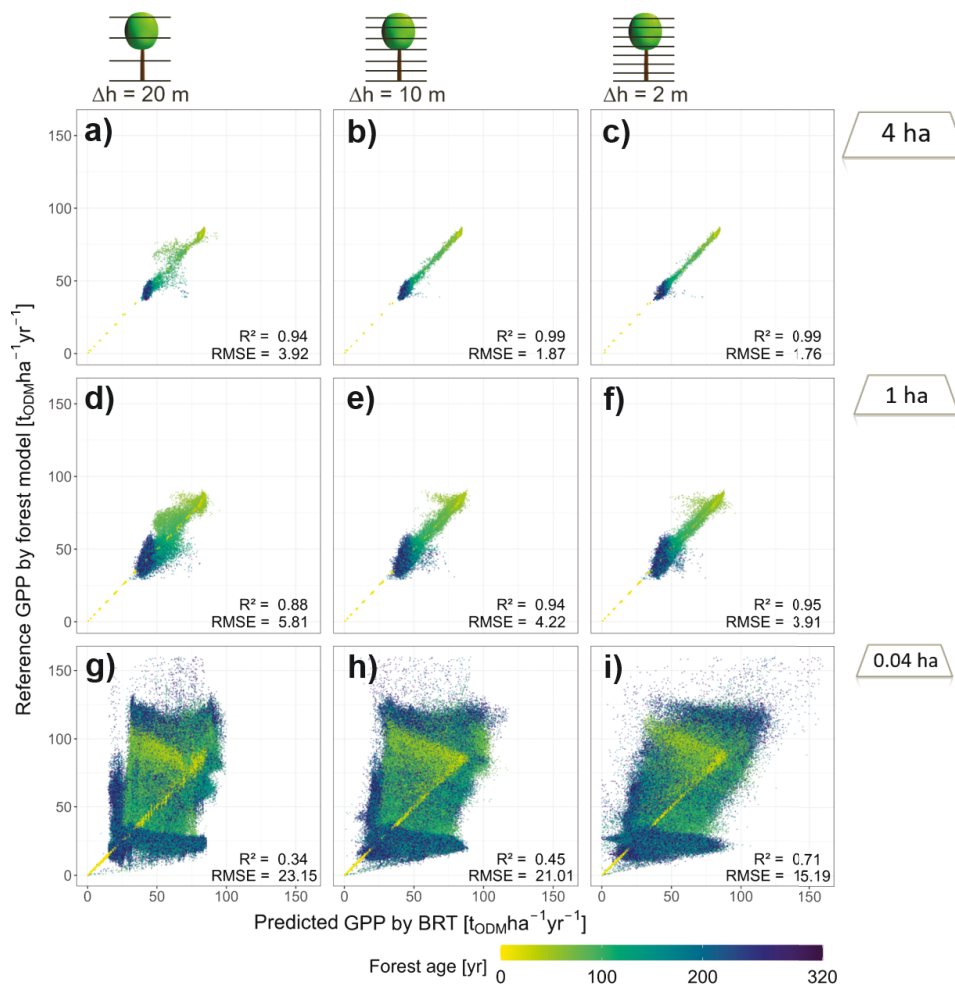


Figure C.5. Comparison of predicted GPP derived from the boosted regression tree and reference GPP derived from simulations using the forest model FORMIND. The comparison is done for the whole forest data. The predicted GPP from biomass measurements is derived from the simulated vertical and horizontal biomass distributions (descriptions above and on the right side). Each point represents a result for a forest with a certain age (color of points). Total simulated area was 100 ha (resulting in the different amount of points in the rows).

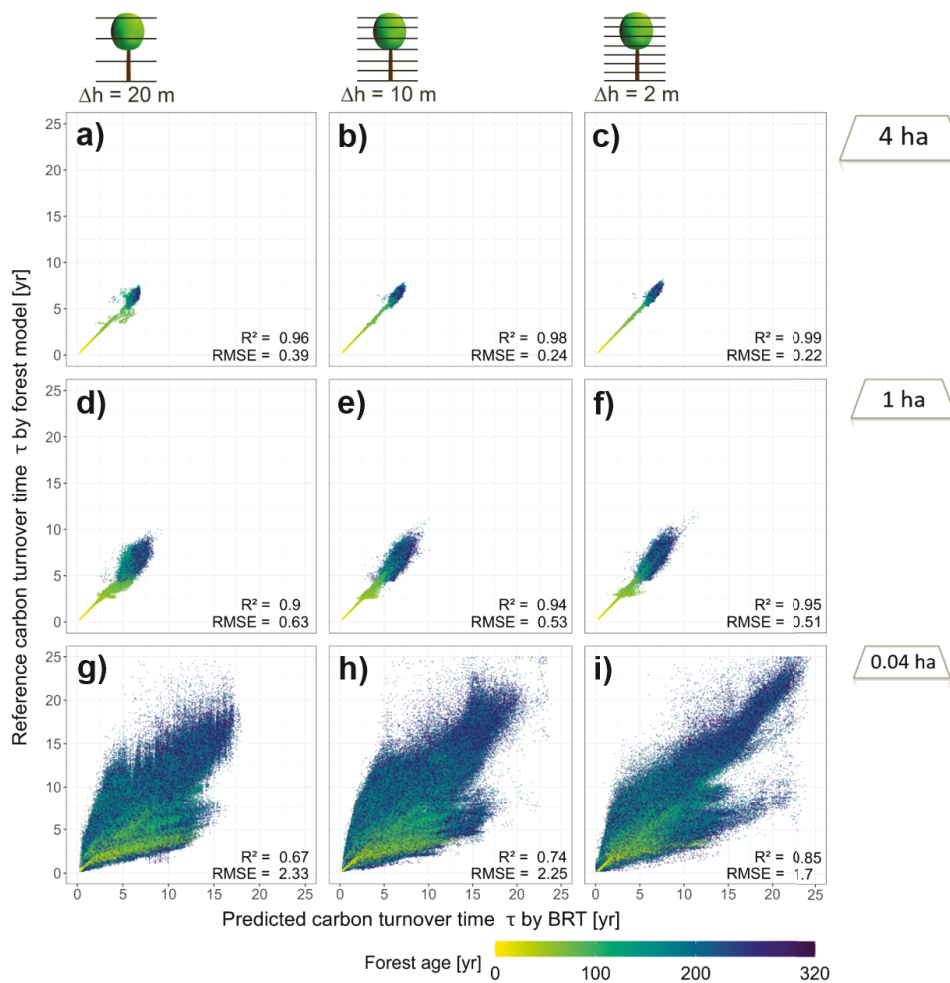


Figure C.6. Comparison of predicted carbon turnover time (τ) derived from the boosted regression tree and reference τ derived from simulations using the forest model FORMIND. The comparison is done for the whole forest data. The predicted τ from biomass measurements is derived from the simulated vertical and horizontal biomass distributions (descriptions above and on the right side). Each point represents a result for a forest with a certain age (color of points). Total simulated area was 100 ha (resulting in the different amount of points in the rows).

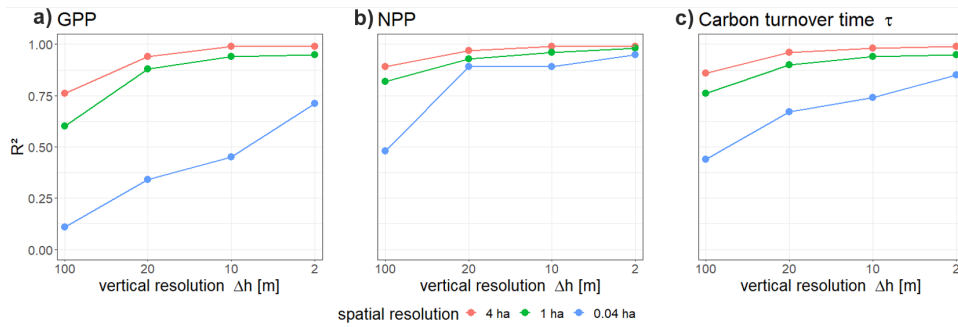


Figure C.7. Comparison of correlation between the estimated and the reference GPP, NPP and carbon turnover time (τ). The estimation was done with a boosted regression tress using information of biomass distribution and the calculation was done with the FORMIND forest model. Each point represents the R^2 value of one comparison with a given spatial - (indicated by color) and vertical resolution. Here we include also an analysis with no height layer ($\Delta h = 100$ m, compared to 19 in the main text).

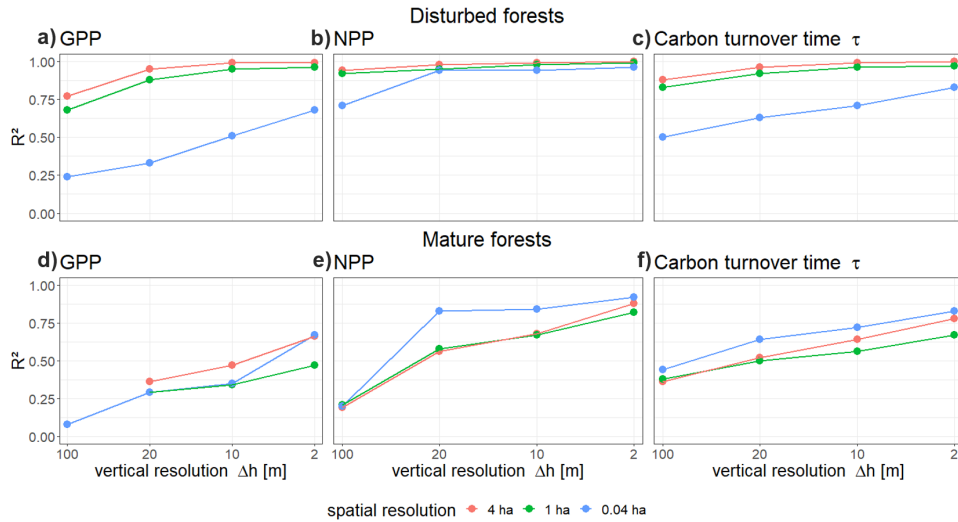


Figure C.8. Comparison of correlation between the estimated and the reference GPP, NPP and carbon turnover time (τ) for different types of forests (disturbed forests - top, mature forests - bottom). The estimation was done with a boosted regression tress using information of biomass distribution and the calculation was done with the FORMIND forest model. Each point represents the R^2 value of one comparison with a given spatial - (indicated by color) and vertical resolution. Here we include also an analysis with no height layer ($\Delta h = 100$ m, compared to 20 in the main text).

Table C.1

Parameters of the derived Boosted regression trees (BRTs) for the whole forest data (0 - 320 years of simulation). Depending on different spatial resolutions (vertical and horizontal) of the biomass data we derived different boosted regression trees. Range of settings describe all possible combinations of settings (LR = learning rate, BF = bag fraction, TC = tree complexity). Not mentioned settings use the standard values of the R package *dismo* 1.3-14. The amount of training data is given in relative and absolute values. Additionally, the parameters of the final BRTs for GPP, NPP and carbon turnover times τ are shown (normal typesetting, italic, underlined). We also included the analysis with no height layers (vertical resolution $\Delta h = 100$ m Figure C.7 and C.8).

Data used for BRT (resolution)	Range of setting [(LR), (BF),(TC)]	amount of training data relative (total)	Used setting (GPP,NPP, τ) (LR,BF,TC)
4 ha, $\Delta h = 100$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(3975)	(0.01, 0.66, 3), (0.01, 0.5, 3), (0.01, 0.66, 3)
4 ha, $\Delta h = 20$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(3975)	(0.05, 0.5, 3) (0.05, 0.66, 3), (0.05, 0.5, 3)
4 ha, $\Delta h = 10$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5)]	50%(3975)	(0.05, 0.66, 5) (0.05, 0.66, 5), (0.05, 0.5, 5)
4 ha, $\Delta h = 2$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5, 7)]	50%(3975)	(0.05, 0.5, 7) (0.05, 0.5, 7), (0.05, 0.66, 7)
1 ha, $\Delta h = 100$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(15900)	(0.05, 0.66, 3) (0.01, 0.66, 3), (0.05, 0.66, 3)
1 ha, $\Delta h = 20$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(15900)	(0.05, 0.5, 3) (0.05, 0.66, 3), (0.05, 0.5, 3)
1 ha, $\Delta h = 10$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5)]	50%(15900)	(0.05, 0.66, 5) (0.05, 0.5, 5), (0.05, 0.5, 5)
1 ha, $\Delta h = 2$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5, 7)]	50%(15900)	(0.05, 0.5, 7) (0.05, 0.5, 7), (0.05, 0.5, 7)
0.04 ha, $\Delta h = 100$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	2%(15900)	(0.05, 0.5, 3) (0.05, 0.5, 3), (0.05, 0.5, 3)
0.04 ha, $\Delta h = 20$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	2%(15900)	(0.05, 0.5, 3) (0.05, 0.5, 3), (0.05, 0.5, 3)
0.04 ha, $\Delta h = 10$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5)]	2%(15900)	(0.05, 0.5, 5) (0.05, 0.5, 5), (0.05, 0.5, 5)
0.04 ha, $\Delta h = 2$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5, 7)]	2%(15900)	(0.05, 0.5, 7) (0.05, 0.5, 7), (0.05, 0.5, 7)

Table C.2

Parameters of the derived Boosted regression trees (BRTs) for disturbed forests (0 - 160 years of simulation). Depending on different spatial resolutions (vertical and horizontal) of the biomass data we derived different boosted regression trees. Range of settings describe all possible combinations of settings (LR = learning rate, BF = bag fraction, TC = tree complexity). Not mentioned settings use the standard values of the R package *dismo* 1.3-14. The amount of training data is given in relative and absolute values. Additionally, the parameters of the final BRTs for GPP, NPP and carbon turnover times τ are shown (normal typesetting, italic, underlined). We also included the analysis with no height layers (vertical resolution $\Delta h = 100$ m Figure C.7 and C.8).

Data used for BRT (resolution)	Range of setting [(LR), (BF),(TC)]	amount of training data relative (total)	Used setting (GPP,NPP, τ) (LR,BF,TC)
4 ha, $\Delta h = 100$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(1987)	<u>(0.05, 0.5, 3),</u> <u>(0.01, 0.66, 3),</u> <u>(0.05, 0.66, 3)</u>
4 ha, $\Delta h = 20$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(1987)	<u>(0.05, 0.5, 3),</u> <u>(0.05, 0.66, 3),</u> <u>(0.05, 0.66, 3)</u>
4 ha, $\Delta h = 10$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5)]	50%(1987)	<u>(0.05, 0.5, 5),</u> <u>(0.05, 0.66, 5),</u> <u>(0.05, 0.66, 5)</u>
4 ha, $\Delta h = 2$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5, 7)]	50%(1987)	<u>(0.05, 0.5, 7)</u> <u>(0.05, 0.5, 7),</u> <u>(0.05, 0.66, 7)</u>
1 ha, $\Delta h = 100$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(7950)	<u>(0.05, 0.66, 3),</u> <u>(0.01, 0.66, 3),</u> <u>(0.01, 0.5, 3)</u>
1 ha, $\Delta h = 20$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(7950)	<u>(0.05, 0.5, 3),</u> <u>(0.05, 0.66, 3),</u> <u>(0.05, 0.5, 3)</u>
1 ha, $\Delta h = 10$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5)]	50%(7950)	<u>(0.05, 0.66, 5),</u> <u>(0.05, 0.5, 5),</u> <u>(0.05, 0.66, 5)</u>
1 ha, $\Delta h = 2$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5, 7)]	50%(7950)	<u>(0.05, 0.66, 7),</u> <u>(0.05, 0.5, 7),</u> <u>(0.05, 0.5, 7)</u>
0.04 ha, $\Delta h = 100$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	2%(7950)	<u>(0.01, 0.66, 3),</u> <u>(0.05, 0.5, 3),</u> <u>(0.01, 0.5, 3)</u>
0.04 ha, $\Delta h = 20$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	2%(7950)	<u>(0.05, 0.5, 3),</u> <u>(0.05, 0.5, 3),</u> <u>(0.05, 0.5, 3)</u>
0.04 ha, $\Delta h = 10$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5)]	2%(7950)	<u>(0.05, 0.5, 5),</u> <u>(0.05, 0.5, 5),</u> <u>(0.05, 0.5, 5)</u>
0.04 ha, $\Delta h = 2$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5, 7)]	2%(7950)	<u>(0.05, 0.5, 7),</u> <u>(0.05, 0.5, 7),</u> <u>(0.05, 0.5, 7)</u>

Table C.3

Parameters of the derived Boosted regression trees (BRTs) for mature forests (160 - 320 years of simulation). Depending on different spatial resolutions (vertical and horizontal) of the biomass data we derived different boosted regression trees. Range of settings describe all possible combinations of settings (LR = learning rate, BF = bag fraction, TC = tree complexity). Not mentioned settings use the standard values of the R package dismo 1.3-14. The amount of training data is given in relative and absolute values. Additionally, the parameters of the final BRTs for GPP, NPP and carbon turnover times τ are shown (normal typesetting, italic, underlined). We also included the analysis with no height layers (vertical resolution $\Delta h = 100$ m Figure C.7 and C.8). Please note, the R package did not find suitable parameters for the prediction of GPP in two cases (even with smaller step size and smaller learning rate).

Data used for BRT (resolution)	Range of setting [(LR), (BF),(TC)]	amount of training data relative (total)	Used setting (GPP,NPP,τ) (LR,BF,TC)
4 ha, $\Delta h = 100$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(1987)	–, <i>(0.01, 0.5, 3),</i> <u>(0.01, 0.66, 3)</u>
4 ha, $\Delta h = 20$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(1987)	<u>(0.05, 0.66, 3),</u> <i>(0.01, 0.66, 3),</i> <u>(0.05, 0.66, 3)</u>
4 ha, $\Delta h = 10$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5)]	50%(1987)	<u>(0.05, 0.66, 5),</u> <i>(0.05, 0.66, 5),</i> <u>(0.05, 0.5, 5)</u>
4 ha, $\Delta h = 2$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5, 7)]	50%(1987)	<u>(0.05, 0.66, 7),</u> <i>(0.05, 0.66, 7),</i> <u>(0.05, 0.5, 7)</u>
1 ha, $\Delta h = 100$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(7950)	–, <i>(0.01, 0.66, 3),</i> <u>(0.01, 0.5, 3)</u>
1 ha, $\Delta h = 20$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	50%(7950)	<u>(0.05, 0.66, 3),</u> <i>(0.05, 0.66, 3),</i> <u>(0.05, 0.66, 3)</u>
1 ha, $\Delta h = 10$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5)]	50%(7950)	<u>(0.05, 0.66, 5),</u> <i>(0.05, 0.66, 5),</i> <u>(0.05, 0.66, 5)</u>
1 ha, $\Delta h = 2$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5, 7)]	50%(7950)	<u>(0.05, 0.5, 7)</u> <i>(0.05, 0.66, 7),</i> <u>(0.05, 0.5, 7)</u>
0.04 ha, $\Delta h = 100$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	2%(7950)	<u>(0.01, 0.5, 3),</u> <i>(0.01, 0.5, 3),</i> <u>(0.01, 0.66, 3)</u>
0.04 ha, $\Delta h = 20$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (3)]	2%(7950)	<u>(0.05, 0.66, 3),</u> <i>(0.05, 0.66, 3),</i> <u>(0.05, 0.5, 3)</u>
0.04 ha, $\Delta h = 10$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5)]	2%(7950)	<u>(0.05, 0.5, 5),</u> <i>(0.05, 0.66, 5),</i> <u>(0.05, 0.66, 5)</u>
0.04 ha, $\Delta h = 2$ m	[(0.05, 0.01, 0.005), (0.3, 0.5, 0.66), (5, 7)]	2%(7950)	<u>(0.05, 0.66, 7),</u> <i>(0.05, 0.66, 7),</i> <u>(0.05, 0.66, 7)</u>

Erklärung über die Eigenständigkeit der erbrachten wissenschaftlichen Leistung

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich / unentgeltlich geholfen.

Kapitel 1

- Kommentare & Sprachliche Korrekturen: Andreas Huth, Friedrich Bohn

Kapitel 2

- Anteil an Idee, Konzept und Interpretation: Friedrich Bohn, Andreas Huth
- Kommentare & Sprachliche Korrekturen: Andreas Huth, Karin Frank

Kapitel 3

- Anteil an Idee, Konzept und Interpretation: Andreas Huth, Friedrich Bohn
- Kommentare & Sprachliche Korrekturen: Andreas Huth, Friedrich Bohn

Kapitel 4

- Anteil an Idee, Konzept und Interpretation: Friedrich Bohn, Andreas Huth
- Kommentare & Sprachliche Korrekturen: Andreas Huth, Friedrich Bohn

Kapitel 5

- Kommentare & Sprachliche Korrekturen: Andreas Huth, Friedrich Bohn

Weitere Personen waren an der inhaltlichen materiellen Erstellung der vorliegenden Arbeit nicht beteiligt.

Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Hans Henniger

Leipzig, September 5, 2023

Lebenslauf

Bildungsweg

- 2010 – 2014 Friedrich Schiller Universität Jena "Wirtschaftsmathematik",
Bachelor of Science mit Gesamtnote 2.8
- 2014 – 2017 Universität Osnabrück "Umweltsysteme und Ressourcenmanagement",
Master of Science mit Gesamtnote 1,6
- 2020 – 2022 Machtkritische Mediationsausbildung
Anerkennung als zertifizierter Mediator
- 2020 – heute Universität Leipzig Politikwissenschaft
Bachelor of Arts mit Zwischennote 1,6
- 2023 – heute Universität Osnabrück Promotionstudium

Veröffentlichungen

- 05/2023 Henniger, H., Frank, F., Huth, A., and Bohn, F. J.
“Creating virtual forests around the globe and analysing their state space”
Ecological Modelling, 483, 110404
- 06/2023 Henniger, H., Bohn, F. J., and Huth, A.
“A New Approach Combining a multi-layer Radiative Transfer Model with an
Individual-Based Forest Model: Application to Boreal Forests in Finland”
Remote Sensing, 15(12), 3078
- 11/2023 Henniger, H., Huth, A., and Bohn, F. J.
“A new Approach to derive Productivity of Tropical Forests using
RADAR Remote Sensing Measurements”
Royal Society Open Science
- 12/2023 Schäfer, J., Winiwarter, L., Weiser, H., Novotný, J., Höfle, B.,
Schmidtlein, S., Henniger, H., Stereńczak, K. and Fassnacht, F. E.
“Assessing the potential of synthetic and ex situ airborne laser
scanning and ground plot data to train forest biomass models.”
Forestry: An International Journal of Forest Research

Erfahrungen

- 2017 – heute Wissenschaftlicher Mitarbeiter am Helmholtz Zentrum für
 Umweltforschung (UFZ)
- 2017 – heute Vorstandssprecher, Projektleiter, Mitglied bei Gesellschaftsdanken e.V.
- 2018 – 2019 Energiewende Botschafter bei 100 prozent erneuerbar stiftung
- 2019 Analyst bei CO-Firm und Pricewaterhousecoopers
- 2018 Wissenschaftlicher Mitarbeiter am Institut für Meteorologie und
 Klimaforschung (IMK-IFU)

September 5, 2023