

ISSN 1860-0387

PHD DISSERTATION 3 | 2019

PHD DISSERTATION 3 | 2019  
Helmholtz Centre for Environmental Research - UFZ  
Department of Ecological Modelling

Nikolai Knapp | Remote Sensing of Forests: Analyzing Biomass Stocks, Changes and ...

Nikolai Knapp

## Remote Sensing of Forests: Analyzing Biomass Stocks, Changes and Variability with Empirical Data and Simulations

Helmholtz Centre for  
Environmental Research - UFZ  
Permoserstraße 15  
04318 Leipzig | Germany  
[www.ufz.de](http://www.ufz.de)

NOT FOR SALE.

3 | 2019

 HELMHOLTZ  
CENTRE FOR  
ENVIRONMENTAL  
RESEARCH - UFZ



**Remote Sensing of Forests:  
Analyzing Biomass Stocks, Changes and  
Variability with Empirical Data and  
Simulations**

**Dissertation**

for the degree of Doctor of Natural Sciences (Dr. rer. nat.)

University of Osnabrück  
School of Mathematics/Computer Science

submitted by  
Nikolai T. Knapp  
from Miltenberg

November 2018

**Supervisors:**

Prof. Dr. Andreas Huth

Dr. Rico Fischer

Nikolai T. Knapp: *Remote Sensing of Forests: Analyzing Biomass Stocks, Changes and Variability with Empirical Data and Simulations*, © November 2018

## ABSTRACT

---

Forests are an important component in the earth system. They cover nearly one third of the land surface, store about as much carbon as the entire atmosphere and host more than half of the planet's biodiversity. Forests provide ecosystem services such as climate regulation and water cycling and they supply resources. However, forests are increasingly at risk worldwide, due to anthropogenic deforestation, degradation and climate change. Concepts for counteracting this development require abilities to monitor forests and predict possible future developments. Given the vast size of forest cover along with the variety of forest types, field measurements and experiments alone cannot provide the solution for this task. Remote sensing and forest modeling enable a broader and deeper understanding of the processes that shape our planet's forests.

Remote sensing from airborne and spaceborne platforms can provide detailed measurements of forest attributes ranging from landscape to global scale. The challenge is to interpret the measurements in an appropriate way and derive biophysical properties. This requires a good understanding of the interaction between radiation and the vegetation. Forest models are tools that synthesize our knowledge about processes, such as tree growth, competition, disturbances and mortality. They allow simulation experiments which go beyond the spatial and temporal scales of field experiments.

In this thesis, several major challenges in forest ecology and remote sensing were addressed. The main variable of interest was forest biomass, as it is the most important variable for forest carbon mapping and for understanding the role of vegetation in the global carbon cycle. For the purpose of biomass estimation, remote sensing derived canopy height and structure measurements were combined with field data, forest simulations and remote sensing simulations. The goals were: 1) to integrate remote sensing measurements into a forest model; 2) to understand the effects of spatial scale and disturbances on biomass estimation using a variety of remote sensing metrics; 3) to develop approaches for quantifying biomass changes over time with remote sensing and 4) to overcome differences among forest types by considering several structural aspects in the biomass estimation function.

In the first study, a light detection and ranging (lidar) simulator was developed and integrated in the forest model FORMIND. The model was parameterized for the tropical rainforest on Barro Colorado Island (BCI, Panama). The output of the lidar simulator was validated against real airborne lidar data from BCI. Undisturbed and disturbed forests were simulated with FORMIND to identify the most well suited lidar metric for biomass estimation. The objective hereby was to achieve a low normalized root mean squared error (nRMSE) over the entire range of forest structures caused by disturbances and succession. Results identified the mean top-of-canopy height (TCH) as the best lidar-derived predictor. The accuracy strongly depended on spatial scale and relative errors  $< 10\%$  could be achieved if the spatial resolution of the produced biomass map was  $\geq 100$  m and the spatial resolution of the remote sensing input was  $\leq 10$  m. These results could provide guidance for biomass mapping efforts.

In the second study, forest simulations were used to explore approaches for estimating changes in forest biomass over time based on observed changes in canopy height. In an ideal situation, remote sensing provides measurements of canopy height

above ground which allows the estimation of biomass stocks and changes. However, this requires sensors which are able to detect canopy surface and terrain elevation, and some sensors can only detect the surface (e.g., X-band radar). In such cases, biomass change has to be estimated from height change using a direct relationship. Unfortunately, such a relationship is not constant for forests in different successional stages, which can lead to considerable biases in the estimates of biomass change. A solution to this problem was found, where missing information of canopy height was compensated by integrating metrics of canopy texture. Applying this improved approach enables estimations of biomass losses and gains after disturbances at 1-ha resolution. In mature forests with very small changes in height and biomass all tested approaches have limited capabilities, as was revealed by an application using TanDEM-X derived canopy height from BCI.

In the third study, a general biomass estimation function, which links remote sensing-derived structure metrics to forest biomass, was developed. General in this context means that it can be applied in different forest types and different biomes. For this purpose a set of predictor metrics was explored, with each predictor representing one of the following structural aspects: mean canopy height, maximal possible canopy height, maximal possible stand density, vertical canopy structure and wood density. The derived general equation resulted in almost equally accurate biomass estimates across the five considered sites (nRMSE = 12.4%,  $R^2 = 0.74$ ) as site-specific equations (nRMSE = 11.7%,  $R^2 = 0.77$ ). The contributions of the predictors provide a better understanding of the variability in the height-to-biomass relationship observed across forest types.

The thesis has laid foundations for a close link between remote sensing, forest modeling and forest inventories. Several ongoing projects carry this further, by 1) disentangling and quantifying the uncertainty in biomass remote sensing, 2) trying to predict forest productivity based on structure and 3) detecting single trees from lidar to be used as forest model input. These methods can in the future lead to an integrated forest monitoring and information system, which assimilates remote sensing measurements and produces predictions about forest development. Such tools are urgently needed to reduce the risks forests are facing worldwide.

## ZUSAMMENFASSUNG

---

Wälder sind ein wichtiger Bestandteil des Systems Erde. Sie bedecken fast ein Drittel der Landoberfläche, speichern etwa so viel Kohlenstoff wie die gesamte Atmosphäre und beherbergen mehr als die Hälfte der biologischen Vielfalt des Planeten. Wälder bieten Ökosystemdienstleistungen wie die Regulierung von Klima und Wasserkreisläufen und liefern Ressourcen. Allerdings sind Wälder weltweit zunehmend gefährdet durch anthropogene Abholzung, Degradierung und den Klimawandel. Konzepte zur Bekämpfung dieser Entwicklung erfordern die Fähigkeit, Wälder zu monitoren und mögliche zukünftige Entwicklungen vorherzusagen. Angesichts der riesigen Waldflächen und der Vielfalt der Waldtypen können Feldmessungen und Experimente allein keine Lösungen für diese Aufgaben bieten. Fernerkundung und Waldmodellierung ermöglichen ein breiteres und tieferes Verständnis der Prozesse, die die Wälder unseres Planeten prägen.

Die Fernerkundung von Flugzeugen oder Satelliten aus kann detaillierte Messungen von Waldeigenschaften liefern um diese für ganze Landschaften bis hin zum globalen Maßstab zu kartieren. Die Herausforderung besteht darin, die Messungen richtig zu interpretieren und biophysikalische Eigenschaften abzuleiten. Dies erfordert ein gutes Verständnis der Wechselwirkung zwischen Strahlung und Vegetation. Waldmodelle sind Werkzeuge, die unser Wissen über Prozesse wie Baumwachstum, Konkurrenz, Störungen und Mortalität bündeln. Sie ermöglichen Simulationsexperimente, die über die räumlichen und zeitlichen Skalen von Feldexperimenten hinausgehen.

In dieser Arbeit wurden mehrere große Herausforderungen in der Waldökologie und Fernerkundung behandelt. Die wichtigste Variable von Interesse war die Waldbiomasse, da sie die zentrale Variable für die Kohlenstoffkartierung und für das Verständnis der Rolle der Vegetation im globalen Kohlenstoffkreislauf darstellt. Zum Zwecke der Biomasseabschätzung wurden aus der Fernerkundung abgeleitete Waldhöhen- und Strukturmessungen mit Felddaten, Waldsimulationen und Fernerkundungssimulationen kombiniert. Ziele waren: 1) die Integration von Fernerkundungsmessungen in ein Waldmodell; 2) die Auswirkungen von räumlicher Skala und Störungsereignissen auf die Biomasseschätzung für eine Vielzahl von Fernerkundungsmetriken zu verstehen; 3) die Entwicklung von Ansätzen zur Quantifizierung von Biomasseänderungen über die Zeit mittels Fernerkundung und 4) die Überwindung von Unterschieden zwischen Waldtypen durch Berücksichtigung mehrerer struktureller Aspekte in der Biomasseschätzfunktion.

In der ersten Studie wurde ein Simulator für das Fernerkundungsverfahren 'Light Detection And Ranging' (Lidar) entwickelt und in das Waldmodell FORMIND integriert. Das Modell wurde für den tropischen Regenwald auf Barro Colorado Island (BCI, Panama) parametrisiert. Die Ergebnisse des Lidarsimulators wurden mit realen Lidardaten von BCI verglichen. Ungestörte und gestörte Wälder wurden mit FORMIND simuliert, um die am besten geeignete Lidarmetrik für die Biomasseschätzung zu identifizieren. Ziel war es dabei einen niedrigen Schätzfehler (normalized Root Mean Squared Error = nRMSE) über den gesamten Bereich der durch Störungen und Sukzessionen entstandenen Waldstrukturen zu erreichen. Die Ergebnisse identifizierten die mittlere Top-of-Canopy-Höhe (TCH) als den besten aus Lidar abgeleiteten Prädiktor. Die Genauigkeit hing stark von der räumlichen Skala ab und relative Feh-

ler  $< 10\%$  konnten erreicht werden, wenn die räumliche Auflösung der erzeugten Biomassekarte  $\geq 100$  m und die räumliche Auflösung der Fernerkundungsdaten  $\leq 10$  m betrug. Diese Ergebnisse könnten als Orientierungshilfe für die Kartierung von Biomasse dienen.

In der zweiten Studie wurden Waldsimulationen verwendet, um Ansätze zur Abschätzung von Veränderungen der Waldbiomasse im Zeitverlauf basierend auf beobachteten Veränderungen der Waldhöhe zu untersuchen. Im Idealfall liefert die Fernerkundung Messungen der Baumkronenhöhe über dem Boden, die eine Abschätzung der Biomassebestände und -veränderungen ermöglichen. Dies erfordert jedoch Sensoren, die in der Lage sind, die Waldoberflächenhöhe und die Geländehöhe zu erfassen, und einige Sensoren können nur die Oberfläche erfassen (z.B. X-Band-Radar). In solchen Fällen muss die Biomassenänderung aus der Höhenänderung in direkter Beziehung abgeschätzt werden. Leider ist ein solcher Zusammenhang bei Wäldern in verschiedenen Sukzessionsphasen nicht konstant, was zu erheblichen Verfälschungen bei den Schätzungen der Biomassenänderungen führen kann. Eine Lösung für dieses Problem wurde gefunden, indem fehlende Information über die Waldhöhe durch die Integration von Metriken über die Waldoberflächentextur kompensiert wurde. Die Anwendung dieses verbesserten Ansatzes ermöglicht die Abschätzung von Biomasseverlusten und -gewinnen nach Störungsereignissen mit einer Auflösung von 1 ha. In alten Wäldern mit sehr geringen Höhen- und Biomassenänderungen stoßen alle getesteten Ansätze an ihre Grenzen, wie eine Anwendung mit aus TanDEM-X abgeleiteten Waldhöhen auf BCI gezeigt hat.

In der dritten Studie wurde eine allgemeine Biomasseschätzfunktion entwickelt, welche aus der Fernerkundung abgeleitete Strukturmetriken mit der Waldbiomasse verknüpft. Allgemein bedeutet in diesem Zusammenhang, dass die Funktion in verschiedenen Waldtypen und Biomen eingesetzt werden kann. Zu diesem Zweck wurde eine Reihe von Prädiktormetriken untersucht, wobei jeder Prädiktor einen der folgenden strukturellen Aspekte darstellt: mittlere Waldhöhe, maximal mögliche Waldhöhe, maximal mögliche Bestandsdichte, vertikale Waldstruktur und Holzdichte. Die abgeleitete allgemeine Gleichung führte zu fast ebenso genauen Biomasseschätzungen über die fünf betrachteten Standorte ( $nRMSE = 12,4\%$ ,  $R^2 = 0,74$ ) wie standortspezifische Gleichungen ( $nRMSE = 11,7\%$ ,  $R^2 = 0,77$ ). Die Beiträge der Prädiktoren liefern ein besseres Verständnis für die Unterschiede in der Höhen-Biomasse-Beziehung, welche sich für verschiedene Waldtypen beobachten lassen.

Die Arbeit hat die Grundlage für eine enge Verbindung zwischen Fernerkundung, Waldmodellierung und Waldinventuren geschaffen. Mehrere laufende Projekte führen das fort, indem sie 1) die Unsicherheit bei der Biomassefernerkundung auflösen und quantifizieren, 2) versuchen, die Waldproduktivität auf der Grundlage der Struktur vorherzusagen und 3) einzelne Bäume aus Lidardaten erkennen, die als Input für das Waldmodell verwendet werden. Diese Methoden können in Zukunft zu einem integrierten Waldmonitoring- und Informationssystem führen, das Fernerkundungsmessungen assimiliert und Vorhersagen über die Waldentwicklung liefert. Solche Instrumente sind dringend erforderlich, um die Risiken zu reduzieren, denen die Wälder weltweit ausgesetzt sind.

## PUBLICATIONS

---

### PUBLISHED CHAPTERS OF THIS THESIS

Knapp, N, Fischer, R, and Huth, A (2018a). "Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states." In: *Remote Sensing of Environment* 205, pp. 199–209. ISSN: 00344257. DOI: 10.1016/j.rse.2017.11.018. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425717305679> ;

Knapp, N, Huth, A, Kugler, F, Papathanassiou, K, Condit, R, Hubbell, SP, and Fischer, R (2018b). "Model-assisted estimation of tropical forest biomass change: A comparison of approaches." In: *Remote Sensing* 10.5, pp. 1–23. ISSN: 20724292. DOI: 10.3390/rs10050731. URL: <https://www.mdpi.com/2072-4292/10/5/731> ;

### CONTRIBUTIONS TO RELATED PUBLICATIONS

Shugart, HH, Asner, GP, Fischer, R, Huth, A, Knapp, N, Le Toan, T, and Shuman, JK (2015). "Computer and remote-sensing infrastructure to enhance large-scale testing of individual-based forest models." In: *Frontiers in Ecology and the Environment* 13.9, pp. 503–511. ISSN: 1540-9295. DOI: 10.1890/140327. URL: <http://www.esajournals.org/doi/10.1890/140327> ;

Fischer, R, Bohn, F, Dantas de Paula, M, Dislich, C, Groeneveld, J, Gutiérrez, AG, Kazmierczak, M, Knapp, N, Lehmann, S, Paulick, S, Pütz, S, Rödig, E, Taubert, F, Köhler, P, and Huth, A (2016). "Lessons learned from applying a forest gap model to understand ecosystem and carbon dynamics of complex tropical forests." In: *Ecological Modelling* 326, pp. 124–133. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2015.11.018. URL: <http://www.sciencedirect.com/science/article/pii/S0304380015005505> ;

Getzin, S, Fischer, R, Knapp, N, and Huth, A (2017). "Using airborne LiDAR to assess spatial heterogeneity in forest structure on Mount Kilimanjaro." In: *Landscape Ecology* 32.9, pp. 1881–1894. ISSN: 15729761. DOI: 10.1007/s10980-017-0550-7. URL: <http://link.springer.com/10.1007/s10980-017-0550-7> ;

Fischer, R, Knapp, N, Bohn, F, and Huth, A (2019). "Remote Sensing Measurements of Forest Structure Types for Ecosystem Service Mapping." In: *Atlas of Ecosystem Services: Drivers, Risks, and Societal Responses*. Ed. by M Schröter, A Bonn, S Klotz, R Seppelt, and C Baessler. 1st ed. Springer International Publishing, p. 430. ISBN: 978-3-319-96229-0. DOI: 10.1007/978-3-319-96229-0 ;



## CONTENTS

---

1	INTRODUCTION	1
1.1	The Role of Forests in the Earth System	2
1.1.1	Background	2
1.1.2	Threats to Forests	4
1.2	Forest Measurement in the Field	6
1.2.1	Forest Inventory	6
1.2.2	Forest Biomass	7
1.3	Forest Remote Sensing	9
1.3.1	Remote Sensing Technology	9
1.3.2	Passive Sensors	9
1.3.3	Active Sensors: Synthetic Aperture Radar	10
1.3.4	Active Sensors: Lidar	12
1.3.5	Remote Sensing Products about Forests	13
1.3.6	Remote Sensing of Forest Structure	15
1.3.7	Remote Sensing of Forest Biomass	18
1.4	Forest Modeling	19
1.4.1	Forest Models	19
1.4.2	Linking Forest Models and Remote Sensing	21
1.4.3	The Model FORMIND	21
1.5	Objectives of this Thesis	23
2	LINKING LIDAR AND FOREST MODELING TO ASSESS BIOMASS ESTIMATION ACROSS SCALES AND DISTURBANCE STATES	25
2.1	Abstract	26
2.2	Introduction	26
2.3	Materials and Methods	28
2.3.1	Study Area	28
2.3.2	Lidar Data	29
2.3.3	Lidar Model Description	29
2.3.4	Forest Model Description	31
2.3.5	Simulation Experiment	31
2.3.6	Lidar-based Biomass Prediction	32
2.4	Results	34
2.4.1	Forest and Lidar Simulation Results	34
2.4.2	Biomass Prediction from Top-of-Canopy Height	35
2.4.3	Biomass Prediction based on Various Lidar Metrics	36
2.5	Discussion	39
2.5.1	Lidar Simulations	39
2.5.2	Biomass Prediction from Lidar Height	40
2.5.3	The Role of Structure Metrics	41
2.5.4	Prediction Errors	41
2.5.5	Linking Remote Sensing with Dynamic Forest Models	42
2.6	Conclusion	43
2.7	Acknowledgments	43

<b>3</b>	<b>MODEL-ASSISTED ESTIMATION OF TROPICAL FOREST BIOMASS CHANGE: A COMPARISON OF APPROACHES</b>	<b>45</b>
3.1	Abstract . . . . .	46
3.2	Introduction . . . . .	46
3.3	Materials and Methods . . . . .	49
3.3.1	Study Area . . . . .	49
3.3.2	Forest Model Description . . . . .	49
3.3.3	Simulations . . . . .	50
3.3.4	Biomass, Height and Change Calculations . . . . .	50
3.3.5	Texture Calculations . . . . .	51
3.3.6	Biomass Change Estimation . . . . .	52
3.3.7	Evaluation Experiments by Bootstrapping . . . . .	54
3.3.8	Application on TanDEM-X Data . . . . .	54
3.4	Results . . . . .	55
3.4.1	Simulation Results . . . . .	55
3.4.2	Theoretical Considerations about the $\Delta H$ -to- $\Delta AGB$ Relationship . . . . .	58
3.4.3	Results for the 50-ha Plot . . . . .	58
3.5	Discussion . . . . .	60
3.5.1	Performance of the Approaches . . . . .	61
3.5.2	Comparison with Other Studies . . . . .	62
3.5.3	Outcome of TanDEM-X Application . . . . .	63
3.5.4	Perspectives . . . . .	64
3.6	Conclusion . . . . .	65
3.7	Acknowledgments . . . . .	65
<b>4</b>	<b>FOREST STRUCTURE METRICS TO GENERALIZE BIOMASS ESTIMATION FROM LIDAR ACROSS CONTINENTS</b>	<b>67</b>
4.1	Abstract . . . . .	68
4.2	Introduction . . . . .	68
4.3	Materials and Methods . . . . .	70
4.3.1	Study Sites . . . . .	70
4.3.2	Inventory Data . . . . .	70
4.3.3	Lidar Data . . . . .	71
4.3.4	Forest Structure Metrics . . . . .	72
4.3.5	Multivariate Regression Analysis . . . . .	74
4.3.6	Site-specific Reference Regression Models . . . . .	74
4.4	Results . . . . .	75
4.4.1	Forest Structure at Different Sites . . . . .	75
4.4.2	Site-specific Basal Area Estimation . . . . .	77
4.4.3	Generalized Basal Area Estimation . . . . .	77
4.4.4	Site-specific Aboveground Biomass Estimation . . . . .	78
4.4.5	Generalized Aboveground Biomass Estimation . . . . .	78
4.4.6	Comparison of Results . . . . .	83
4.5	Discussion . . . . .	84
4.5.1	The Role of Mean Canopy Height . . . . .	84
4.5.2	The Role of Stand Density . . . . .	84
4.5.3	The Role of Maximum Height . . . . .	85
4.5.4	The Role of Vertical Heterogeneity . . . . .	85
4.5.5	The Role of Wood Density . . . . .	86

4.5.6	Generalization and Outlook . . . . .	86
4.6	Conclusion . . . . .	87
4.7	Acknowledgments . . . . .	87
5	SYNTHESIS AND OUTLOOK . . . . .	89
5.1	Main Results and Limitations . . . . .	90
5.1.1	Integrating Remote Sensing Simulations in a Forest Model . . . . .	90
5.1.2	Effects of Scale and Disturbance States on Biomass Estimation . . . . .	91
5.1.3	Detecting Biomass Changes over Time with Remote Sensing . . . . .	92
5.1.4	Generalization of Structural Relationships Across Forest Types . . . . .	92
5.2	Ideas for Future Research . . . . .	93
5.2.1	Disentangling Uncertainty Using Simulations . . . . .	93
5.2.2	Remote Sensing of Forest Productivity . . . . .	96
5.2.3	Model Initialization at Individual Tree Level . . . . .	97
5.3	Vision . . . . .	99
A	APPENDIX OF CHAPTER 2 . . . . .	103
A.1	Supplementary Information . . . . .	103
A.1.1	Lidar Model Parameterization . . . . .	103
A.1.2	FORMIND Model Parameterization . . . . .	104
A.1.3	Comparison of Simulation-Derived Lidar-to-Biomass Relations with Reference Data . . . . .	105
A.1.4	Sensitivity of Lidar Simulations to Alternative Tree Representations	107
A.2	Supplementary Graphics . . . . .	109
A.2.1	Graphics Concerning FORMIND Output Patterns . . . . .	109
A.2.2	Graphics Concerning Biomass Prediction from Top-of-Canopy Height . . . . .	110
A.2.3	Graphics Concerning Biomass Prediction from Various Lidar Metrics . . . . .	118
A.2.4	Graphics Concerning the Sensitivity of Lidar Simulations to Alternative Tree Representations . . . . .	120
B	APPENDIX OF CHAPTER 3 . . . . .	123
C	APPENDIX OF CHAPTER 4 . . . . .	125
	LIST OF FIGURES . . . . .	131
	LIST OF TABLES . . . . .	134
	ACRONYMS . . . . .	135
	DANKSAGUNG . . . . .	139
	ERKLÄRUNG . . . . .	141



INTRODUCTION

---

1.1 THE ROLE OF FORESTS IN THE EARTH SYSTEM

1.1.1 Background

Forests are often called the green lungs of our planet earth. Through the process of photosynthesis they assimilate carbon dioxide and produce oxygen. This fact makes forests indispensable for human existence. This thesis is concerned with improving our understanding of the role forests play in the earth system by investigating how we can measure key attributes of forests, in particular biomass, in a spatially continuous way for quantifying the ecosystem services provided by forests worldwide.

With an **extent** of around 41 million square kilometers forests are a dominant land cover type, covering 28% of the global landmass and 8% of the earth surface (Hansen et al., 2013). The exact numbers depend on how forests are defined. A frequently used **definition** is the one by the United Nations Food and Agriculture Organization (FAO) which defines them as land areas of at least 0.5 ha size with a minimum tree crown cover of 10% and a minimum tree height at maturity of 5 m (Chazdon et al., 2016). Forests provide numerous **ecosystem functions and services**, such as carbon sequestration, air filtering, water storage and cycling, cloud formation via evapotranspiration, soil protection, wildlife habitat, resource supply as well as recreational and cultural services (Ninan and Inoue, 2013). More than half of the earth's plant and animal species are found in tropical forests (Thomas and Baltzer, 2002). Many **people's livelihoods** directly depend on forests with up to 1.2 billion people being smallholder farmers using agroforestry practices, up to 350 million belonging to forest-dwelling indigenous peoples and up to 140 million people working in forest-based enterprises (Chao, 2012). Within the global **economy**, the forest sector generates annual revenue of ~600 billion USD, which corresponds to 0.8% of the world's gross domestic product (FAO, 2015).

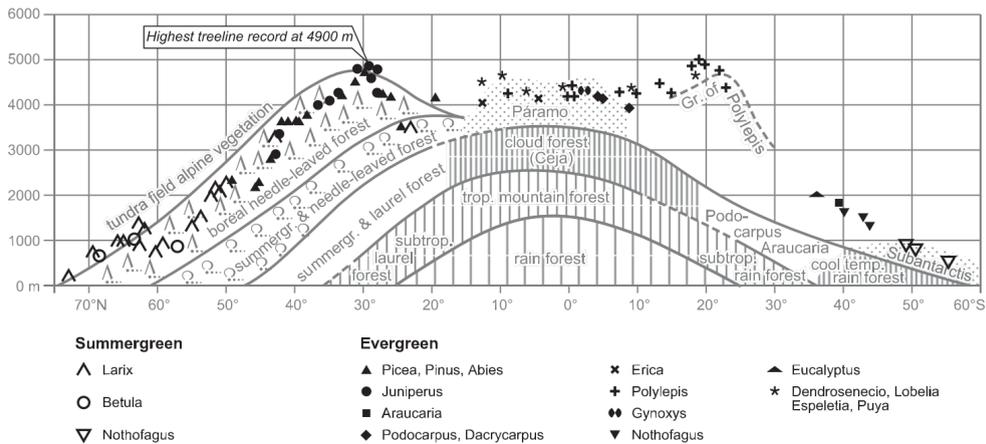


Figure 1.1: Latitudinal and altitudinal ranges of the different forest types with some dominant tree genera (Miehe et al., 2007).

The **geographic distribution** of vegetation types or biomes is determined by temperature and precipitation (Whittaker, 1970). In different climatic zones different forest types have evolved. A total of 60,065 **tree species** have been scientifically described, representing around 20% of all known plant species (Beech et al., 2017). Of those tree species 96% reside in the tropical, 3.7% in the temperate and 0.3% in the boreal climate zone (Fine and Ree, 2006). In particular tropical forests harbor a large number

of plant and animal species, partly with high degrees of endemism, which makes them hotspots of **biodiversity** (Myers et al., 2000). At the margins of their distribution, forests can persist under harsh climatic conditions and occur at all latitudes between 72° N (*Larix gmelinii*) and 55° S (*Nothofagus antarctica*) and in altitudes as high as 4,900 m in the Himalayas (*Juniperus tibetica*) and 4,810 m in the Andes (*Polylepis tarapacana*; Figure 1.1; Hoch and Körner, 2005; Miehe et al., 2007).

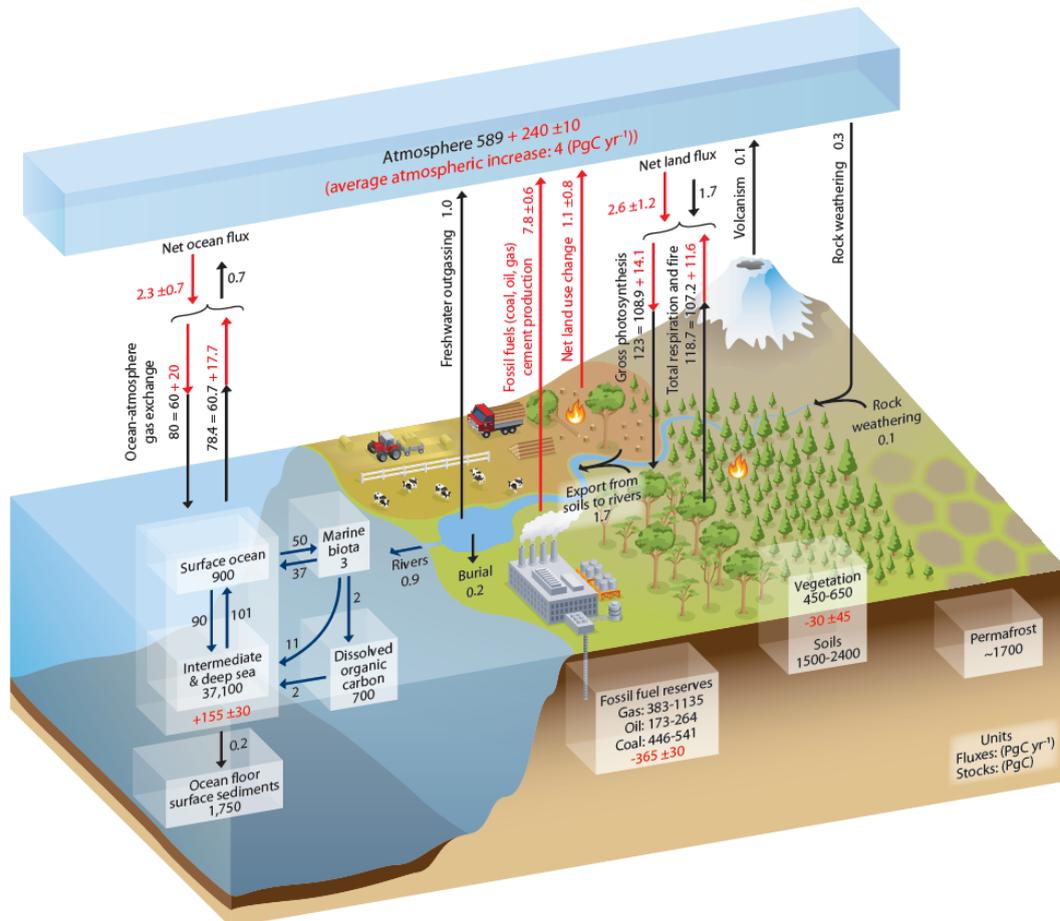


Figure 1.2: Stocks and fluxes in the global carbon cycle in Pg (equal to Gt). Red numbers highlight the anthropogenic contributions (IPCC, 2013).

Forests are crucial in **climate regulation**. In low latitudes they contribute to cooling mainly through atmospheric CO<sub>2</sub> uptake and evapotranspiration. In high latitudes they have a net warming effect due to their low albedo compared to unforested land, in particular under snow covered conditions (Bonan, 2008). Although it is clear that forests are a major compartment in the **global carbon cycle**, their carbon content is quite uncertain and has so far been estimated indirectly as the residual of the other compartments (Houghton et al., 2009; Houghton et al., 2018). The fifth assessment report of the Intergovernmental Panel on Climate Change (IPCC) states the range of estimates for the global vegetation carbon pool as 450 to 650 Gt. For comparison, the global soil carbon pool is estimated to be between 1,500 and 2,400 Gt plus an additional permafrost carbon pool of around 1,700 Gt. The atmospheric carbon pool was at 589 Gt in the preindustrial age and now holds an additional 240 Gt from anthropogenic emissions, leading to a total of 829 Gt. Currently, annual emissions

from fossil fuel burning are at  $7.8 \text{ Gt yr}^{-1}$  compared to  $1.1 \text{ Gt yr}^{-1}$  caused by land use change (predominantly deforestation). Parts of these emissions are taken up by the oceans ( $2.3 \text{ Gt yr}^{-1}$ ) and land vegetation ( $2.6 \text{ Gt yr}^{-1}$ ), such that the net atmospheric carbon increase is at about  $4 \text{ Gt yr}^{-1}$  (Figure 1.2; IPCC, 2013; Le Quéré et al., 2016). The mechanisms behind the terrestrial **net carbon sink** are only partly understood and a matter of current research and debate. Factors contributing to the sink are long term forest recovery from past disturbances and harvesting as well as afforestation and reforestation programs, but also climate change,  $\text{CO}_2$  fertilization and nitrogen deposition may play a role (Pan et al., 2011; Houghton et al., 2018).

### 1.1.2 *Threats to Forests*

Humans have used forests as a timber resource supply for centuries. Overexploitation resulting in regional **deforestation** has led to collapses of various civilizations in human history (Diamond, 2005). Timber shortage has, however, also incentivized the development of forest management practices in several parts of the world, which led to the concept of **sustainability** in 18th century Germany (Carlowitz, 1713) and ultimately to modern-day **forestry**. Nevertheless, forest loss is a pervasive problem of the 21st century with global implications. Since the middle of the 20th century tropical regions have undergone massive deforestation with annual deforestation rates of several percent in certain countries (Whitmore, 1990). According to satellite-based analyses, the tropics have lost 0.5% of forest cover annually throughout the 1990s (Achard et al., 2002) and 2000s (Hansen et al., 2013) with partly decreasing and partly increasing rates at regional level. Additionally, each year 0.2% of the area was affected by degradation (e.g., due to selective logging). On the other hand, regrowing forests on abandoned land amounted to 0.08% per year (Achard et al., 2002).

**Drivers** behind changes in forest cover are manifold. The ultimate causes of deforestation are population pressure, poor governance and trade liberalization (Laurance, 1999). They lead to conversion of forest areas into agricultural or urban land and free them to logging for timber extraction. Also natural disasters such as wildfires, wind throws or insect outbreaks contribute to forest loss. Not all losses of forest cover result in permanent deforestation. A recent study identified four main drivers of forest cover change, each of which accounted for ~25% of the global total (Curtis et al., 2018). The dominant drivers differ across geographical regions (Figure 1.3): **Commodity driven deforestation**, i.e., the conversion into agricultural land to produce products for the global market, is responsible for much of tropical forest loss in South America (mainly cattle ranging, soybean and sugarcane plantations) and Southeast Asia (mainly oil palm plantations). It is the most severe form of forest loss, as it causes a permanent change of land use, while all other major drivers cause temporary losses which are later followed by forest recovery. Much of Africa and Central America are subject to **shifting agriculture**, which is acting at small spatial scales (~1 ha). In the temperate forest regions across all continents **forest management** is the major source of forest cover change. In the boreal forest regions of North America and Siberia **wildfires** are the dominant driver.

Besides deforestation, forest degradation is a severe problem. **Degradation** refers to declines in ecosystem service provision (e.g., biodiversity, carbon stocks), while the forest area stays constant (Sasaki and Putz, 2009). Degradation can be caused by **selective logging** activities and by forest **fragmentation** resulting from expansions

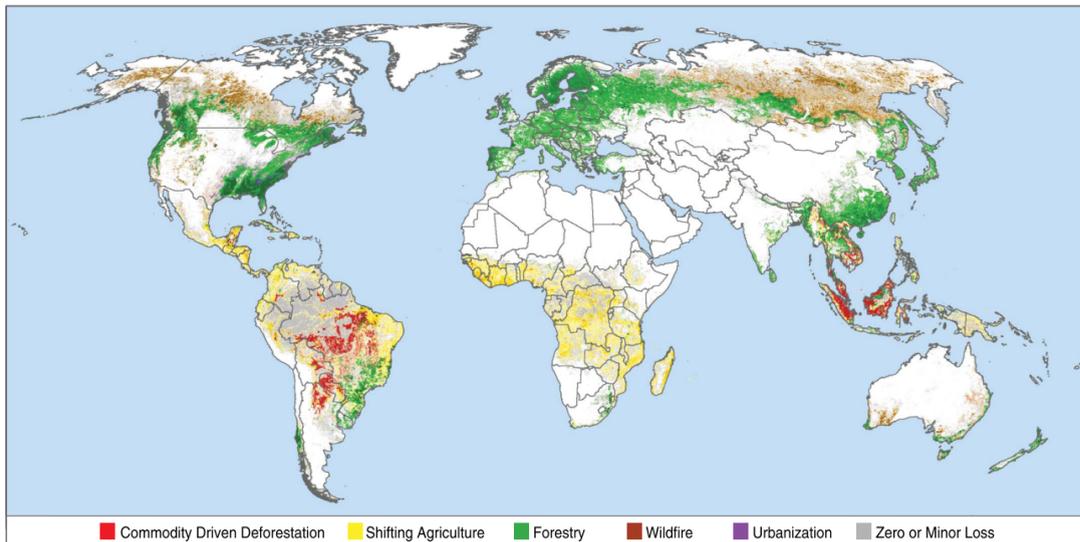


Figure 1.3: Global map of the main drivers of forest cover loss (Curtis et al., 2018).

of roads, crop lands and settlements. In the Brazilian Amazon carbon emissions due to selective logging have been estimated to be at 60% of the ones from deforestation (Asner, 2009). In tropical forests, tree mortality rates are elevated near the forest edges compared to the forest interior, which may cause fragmentation-induced carbon losses at about 31% of the losses due to deforestation (Brinck et al., 2017).

Apart from threats arising directly from human activities, **climate change** is posing an additional threat to forest ecosystems. As forests are a key component in the global climate system, their responses to climate change may trigger complex **feedbacks** which are difficult to predict. For the Amazon region, climate simulations using general circulation models have predicted an average temperature increase of 3.3 °C for the 21st century and an intensification of dry seasons (Malhi et al., 2008). This will probably increase the risk of **droughts** and fires and decrease cloud formation. The resilience of forests to altered climate regimes is uncertain. In the worst case, altered climate could lead to large scale forest dieback (Rammig et al., 2010). This could turn terrestrial carbon sinks into sources, which would be an irreversible tipping point for climate change (Lenton et al., 2008). In other regions of the world forests might profit from altered climate regimes.

Different measures for reducing deforestation and thereby also mitigating additional greenhouse gas emissions have been put into effect. Since 1990, the proportion of global land area with a protection status has almost doubled and is currently at 15%. For forests the proportion of **protected areas** is even 20% (Juffe-Bignoli et al., 2014). **Certification** schemes such as the Forest Stewardship Council (FSC), the Programme for the Endorsement of Forest Certification Schemes (PEFC) or Rainforest Alliance certify forest products produced according to sustainable forest management standards (Auld et al., 2008). With increasing awareness on the consumers' side these labels aim to promote the transition towards a more sustainable forest use. Since 2000, the area of certified forests has grown by a factor 30 such that today more than 10% of the global forest area is either FSC or PEFC certified. Still, certification is mostly restricted to boreal and temperate regions, accounting for 90% of certified forest area (MacDicken et al., 2015). To fight tropical deforestation, the United Nations Framework Convention on Climate Change (UNFCCC) has created a framework

for the Reduction of Emissions from Deforestation and forest Degradation and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries (**REDD+**). It is designed to provide financial incentives for countries that can report avoided emissions due to avoided deforestation compared to a predefined baseline scenario (Gibbs et al., 2007). To ensure the proper implementation of these protection mechanisms regular monitoring, reporting and verification (MRV) are required, which can be achieved, e.g., by combining field inventory plot networks with remote sensing (De Sy et al., 2012).

## 1.2 FOREST MEASUREMENT IN THE FIELD

### 1.2.1 *Forest Inventory*

**Forest inventories** are used to characterize forests and quantify different attributes. Many countries conduct systematic forest inventories at the national level to estimate their forest resources. Additionally, scientists have established research plots targeted towards different objectives. There exist more than 3 million vegetation plots worldwide (Global Index of Vegetation-Plot Databases<sup>1</sup>; Dengler et al., 2011) and single studies based on nearly half a million forest plots have become possible (Crowther et al., 2015). Depending on the specific purposes the designs of forest inventories can vary tremendously.

The most common tree attribute to be measured is the stem **diameter at breast height** (DBH). The reasons for conducting measurements at breast height, which is defined as 1.3 m above ground, are, firstly, practical considerations and, secondly, the fact, that tree stems tend to develop irregularities (e.g., buttresses) near ground for static support. The cross sectional area of the stem at breast height is called **basal area** of the tree. Summing the basal areas of all individual trees in a forest stand leads to stand basal area, i.e., the cross sectional area of all trees in the stand at 1.3 m height. Stand basal area is a simple aggregated statistic to characterize stand density with a single number. More information on the stand's size and age structure is contained in its **stem diameter distribution** (Taubert et al., 2013).

Apart from DBH, inventories may also record the species of the tree, the spatial position of the stem foot within the plot, the tree top height (commonly estimated with laser range finders) and crown attributes, such as crown width and crown base height. Additionally, plants of other life forms, such as palms, bushes, bamboo, shrubs, ferns, lianas and epiphytes, may be included. Inventories can be classified according to their **sampling design** into angle-count sampling (Bitterlich, 1952), nested sampling and complete enumeration. In angle-count sampling, trees are recorded by an observer turning 360° around a center point and including only trees that exceed a certain apparent size (viewing angle) from the viewing point. This results in small trees being recorded only in close proximity to the center point, while large trees are also recorded at larger distances from the center point. As there are usually many more small than large trees, this method reduces the sampling effort. Alternatively, a nested plot design can be used, where the smaller tree size classes are only sampled in subplots of the entire plot area. Complete enumeration on the other hand requires the recording of all trees (above a fixed minimum size threshold of, e.g., 1, 5 or 10 cm) in the entire plot area. Commonly, plot areas have either circular or rectangular shape.

<sup>1</sup> [www.givd.info](http://www.givd.info); accessed on August 29th 2018

National forest inventories have the objective to quantify standing timber resources (and productivity through repeated inventories) representatively to allow accurate upscaling for entire countries. Thus, their design usually involves many small plots (~500 m<sup>2</sup>, often angle-count sampling) distributed on a regular grid across the country (e.g., the 47,000 plots of the German national forest inventory; Bundeswaldinventur; Thünen-Institut, 2012). Such plots are less well suited for studying ecological processes in forests. Hence, forest ecologists have established long term observation plots covering large areas of many hectares (megaplots of 25 to 50 ha), where complete enumeration censuses take place at regular time intervals. To facilitate site comparisons and synthesis **collaborative networks** have formed, which bundle the data from individual sites and also provide guidelines and protocols for the inventories. Examples for such networks are the ForestGEO network coordinated by the Smithsonian Institution (Anderson-Teixeira et al., 2015) and the ForestPlots network (Lopez-Gonzalez et al., 2011).

### 1.2.2 Forest Biomass

The products derived from inventories go beyond the variables that can directly be measured in field plots. Forest **aboveground biomass** and closely related variables, such as harvestable timber volume and aboveground carbon density, are of interest in commercial and scientific applications. The only way to directly measure aboveground biomass of a forest is destructive sampling, i.e., all trees need to be cut and weighted. The common non-destructive approach, however, is to calculate aboveground biomass using established **allometric equations** (Chave et al., 2014). These equations have been derived from harvested trees and describe the relationships between DBH and tree height, crown size, stem volume and biomass. In recent years, **terrestrial laser scanning** has gained popularity as a method to collect volumetric information of trees and whole inventory plots in a non-destructive manner (Calders et al., 2015).

To convert volume into biomass knowledge about **wood density** (wood specific gravity) is required. As wood densities of trees are species-specific, large databases have been compiled (Chave et al., 2009). Wood densities can vary over more than an order of magnitude – from 0.1 t m<sup>-3</sup> for Balsa wood (*Ochroma pyramidale*) to 1.39 t m<sup>-3</sup> for Ébano wood (*Caesalpinia sclerocarpa*) – with a median of 0.6 t m<sup>-3</sup> across all species. Chemically, wood is mainly composed of three biopolymers: cellulose (~45%), hemicellulose (~25%) and lignin (~25%). The remainder consists of volatile organic compounds (~5%) such as fatty acids, sugars, resin acids, waxes and terpenes. The carbon content of wood lies in a range of 44 to 51% (Elias and Potvin, 2003; Thomas and Malczewski, 2007) and 0.48 or 0.5 are the most common factors used to convert forest aboveground biomass into aboveground **carbon density**.

The term forest biomass commonly refers to the aboveground woody dry biomass of living trees. **Belowground biomass** (roots) is estimated in a non-destructive way by its relation to aboveground biomass. The non-woody part of aboveground biomass (foliage) is rather small and rarely quantified explicitly. **Dead wood** (coarse woody debris) can reach large proportions of living biomass. It is related to the latter via turnover rates (mortality, decomposition), which are site-specific and climate-dependent. Analyses of a large number of tropical forest plots in the ForC database (Anderson-Teixeira et al., 2016) show that on average the total living forest biomass consists of ~81% aboveground woody biomass, ~16% root biomass and ~3% foliage biomass. The



## 1.3 FOREST REMOTE SENSING

### 1.3.1 *Remote Sensing Technology*

Remote sensing refers to the acquisition of information about objects without making physical contact. This is achieved by recording reflected radiation in the electromagnetic spectrum. The origins of remote sensing go back to military observations, but also civilian applications have a long tradition in particular the use of aerial imagery in forestry. With the development of earth observation satellites the field of remote sensing has grown rapidly over the past decades. Remote sensing products are nowadays used in basically all scientific disciplines with a geospatial context (Campbell and Wynne, 2011).

Remote sensing techniques can be differentiated by sensor platforms and sensor types. Both have influence on the spatial scale of the collected data. **Spatial scale** here refers to a) spatial resolution (grain or pixel size), i.e., the area which is represented by a single digital number in the resulting image, and b) spatial extent (image size or swath width), i.e., the total area covered by a single image. There are three main **platform types**: 1) airborne, 2) spaceborne and 3) unmanned aerial vehicles (UAV). **Airborne** systems, i.e., sensors carried by airplanes or helicopters, have the longest tradition. They are used for mapping at regional scales with resolutions usually ranging from cm to m scale and extends ranging from 100-m to km scale. **Spaceborne** systems may be carried by satellites or mounted on the international space station (ISS) and are built for acquisitions at global scale. Their resolutions may reach sub-m scale but are more commonly in the 10-m to 50-km scale and extents of single images can reach hundreds of km. For spaceborne systems also the temporal resolution is an important parameter. Temporal resolution refers to the revisiting interval, i.e., how frequently a particular point in space is observed by the sensor. There is always a tradeoff between spatial and temporal resolution. **UAV** may refer to multicopters or fixed-wing drones, which are used for mapping at local scales. UAV-based sensors can collect very detailed data at cm resolution for limited extents of a few hectares. With regard to **costs**, the total costs for a single campaign increase from UAV through airborne campaigns to spaceborne missions. However, regarding cost efficiency the costs per area sampled are usually the lowest for spaceborne data (0 - 13 € km<sup>-2</sup>), with many products of publicly funded missions being available for free, and increase through airborne (30 - 200 € km<sup>-2</sup>) to UAV data (720 € km<sup>-2</sup>; Ørka and Hauglin, 2016).

### 1.3.2 *Passive Sensors*

Sensor types can be classified into passive and active sensors. Passive sensors record radiation originating from the sun and being reflected from objects on the ground. **Optical passive sensor** systems typically collect data in the visible (VIS,  $\lambda = 400 - 700$  nm) and infrared (IR,  $\lambda = 700$  nm - 1000  $\mu$ m) regions of the electromagnetic spectrum (Figure 1.5 and Table 1.1). There also exist **passive microwave sensors** ( $\lambda = 1$  mm - 1 m). Optical passive sensors are further split into multispectral and hyperspectral sensors (Pettorelli et al., 2014). **Multispectral** sensors collect data in a small number (3 to 15) of bands, which only cover parts of the full reflectance spectrum (e.g., satellites Landsat, Roy et al., 2014, and Sentinel-2, Drusch et al., 2012). **Hyperspectral** sensors collect data in a large number (commonly > 200) of narrow

## Types of Remote Sensing Sensors

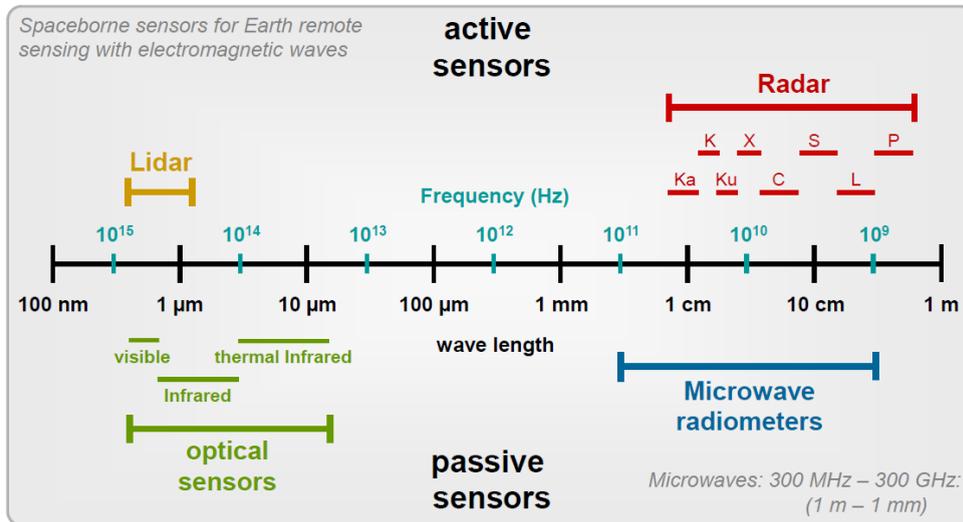


Figure 1.5: Wavelengths and frequencies of different remote sensing sensor types (Moreira, 2015).

bands, which cover the reflectance spectrum continuously (e.g., EnMAP mission, Guanter et al., 2015). As the amount of energy reflected from an area is limited, there is a tradeoff between spatial and spectral resolution. Many sensors additionally have a panchromatic channel that records reflectance across a broad bandwidth and therefore provides higher spatial resolution than the individual band channels. Multispectral satellites provide data in a wide range of scales. The Moderate-resolution Imaging Spectroradiometer (MODIS) provides data at coarse spatial (250 - 1000 m) but high temporal (2 days) resolution (Justice et al., 2002). Medium resolution satellites such as Landsat-8 (Roy et al., 2014) and Sentinel-2 (Drusch et al., 2012) acquire data at 10 - 30 m pixel resolution but intervals between cloud-free scenes of the same point on earth are in the weeks to months range (depending on the regional climate). High resolution satellites (e.g., IKONOS, Quickbird, RapidEye, SPOT, WorldView) belong to commercial enterprises. They can provide data at spatial resolutions of 0.5 - 5 m on demand. Passive optical sensors are the most common sensors on airborne and UAV platforms, including ordinary digital cameras.

### 1.3.3 Active Sensors: Synthetic Aperture Radar

Active sensors actively emit signals and record the reflection of their own emissions. In the case of radio detection and ranging (radar), the signal is in the **radio** wavelengths (0.8 - 100 cm). In the case of light detection and ranging (lidar) the signal is in the VIS or NIR range. The common technology to acquire remote sensing imagery with radar sensors is called **synthetic aperture radar** (SAR). To achieve high resolution images with pixels in the m range by using radio waves in the cm range at observation distances in the km range, the required antenna length along track, called aperture, would have to be very large (hundreds of meters). SAR solves this problem by combining several acquisitions with a small physical antenna (aperture in the m range), which are taken along track, to create a synthetic aperture in the 100-m range.

Table 1.1: Wavelength ranges of the different regions and bands in the electromagnetic spectrum.

Region	Band	Wavelengths
Ultraviolet (UV)	UV	10-400 nm
Visible (VIS)	Violet	400-450 nm
	Blue	450-495 nm
	Green	495-570 nm
	Yellow	570-590 nm
	Orange	590-620 nm
	Red	620-700 nm
Infrared (IR)	Near (NIR)	700 nm-1.5 $\mu\text{m}$
	Shortwave (SWIR)	1.5-3 $\mu\text{m}$
	Midwave (MWIR)	3-8 $\mu\text{m}$
	Longwave/Thermal (LWIR/TIR)	8-15 $\mu\text{m}$
Radio/Microwaves	Far (FIR)	15-1000 $\mu\text{m}$
	Ka	0.8-1.1 cm
	K	1.1-1.7 cm
	Ku	1.7-2.4 cm
	X	2.4-3.8 cm
	C	3.8-7.5 cm
	S	7.5-15 cm
	L	15-30 cm
P	30-100 cm	

SAR measurements can be decomposed into information on a) backscatter intensity and b) the phase of the wave (Moreira et al., 2013).

**Backscatter intensity** refers to the relative amount of energy emitted by the system, which is scattered back in the sensor direction after interacting with the objects on the ground (scatterers). Phase refers to the precise position of the received signal within one wave cycle compared to the emitted signal. SAR **polarimetry** operates with filters, such that only horizontally (H) or vertically (V) polarized waves are emitted or received, respectively. Polarimetric channels are obtained by combining different filter orientations for emission and reception (HH, VV, HV, VH). Based on their structure, different ground surfaces show different backscatter behavior in the different polarimetric channels, which allows a classification of land cover types (Lee et al., 2001). The phase information is used in the field of SAR **interferometry** (InSAR) to construct interferograms. Interferograms visualize phase differences between two acquisitions, either from a pair of satellites in so called bistatic configuration (i.e., located at a certain distance from each other, called spatial baseline) or the same satellite at two points in time (temporal baseline; Bamler & Hartl, 1998). This enables the derivation of terrain information at cm precision, which is used to produce digital elevation models (DEM) and to quantify tectonic deformations caused by earth quakes, e.g., using the TanDEM-X satellites (Krieger et al., 2007).

SAR systems can operate at different wavelengths and different letters of the alphabet are used for specific **bands** within the spectrum (Table 1.1). Waves are scattered mainly by objects of sizes similar to the wavelength. Hence, in forests, backscatter at X-band (e.g., 3 cm) is mostly caused by foliage and small twigs, at L-band (e.g., 27 cm) by branches and at P-band (e.g., 70 cm) by tree trunks and the ground (Treuhaft et al., 2004). A particular problem with interferometric applications in the context of vegetation is that leaves and branches are constantly moving due to weather influences. Hence, the phase difference between successive acquisitions at two points in time is disturbed (temporal decorrelation). Therefore, a bistatic configuration is of advantage, in which two receiver antennas separated by a spatial baseline capture the same signal simultaneously (e.g., TanDEM-X). Such a constellation enables interferometric derivation of vegetation height and by combining polarimetric and interferometric data (**PolInSAR**) the derivation of vertical vegetation structure (Cloude and Papathanassiou, 1998; Kugler et al., 2014). If multiple acquisitions with different baselines are combined for 3D reconstruction, this is called SAR **tomography** (Reigber and Moreira, 2000).

#### 1.3.4 Active Sensors: Lidar

Lidar remote sensing follows a different principle and enables a more direct retrieval of 3D information, compared to SAR. A lidar sensor emits a short **laser pulse** and measures the time between emission and reception of the reflected signal. Half of this two-way travel time multiplied by the speed of light ( $\sim 3 \cdot 10^8 \text{ m s}^{-1}$ ) results in the distance between sensor and reflecting object. Together with the exact GNSS position and orientation of the sensor at time of emission, the 3D position of the reflecting surface can be calculated. Especially in vegetated areas, not all energy of the pulse is reflected from a single point. The area where the laser hits the land surface (footprint) may contain several reflecting objects (leaves, branches, tree trunks, ground) at different distances, which all contribute to the reflected signal (Figure 1.6; Lefsky et al., 2002a). So-called **full-waveform** lidar systems record the returned energy as a function of time. **Discrete return** lidar systems only record peaks in the returned energy, and the number of recorded returns per pulse varies among different sensor systems (e.g., only first return, only last return or multiple returns). Lidar systems are further differentiated into **small-footprint** systems (diameters of 0.1 - 1 m) and **large-footprint** systems (diameters of 5 - 100 m). Typical products of large-footprint full-waveform lidar systems are vertical **canopy profiles**. Typical products of small-footprint discrete return lidar systems are **3D point clouds** (Wulder et al., 2012).

Important large-footprint lidar systems in vegetation remote sensing are the Scanning Lidar Imager of Canopies by Echo Recovery (SLICER; airborne with 10-m footprint; Blair et al., 1994; Lefsky et al., 1999), the Land, Vegetation and Ice Sensor (LVIS; airborne with 25-m footprint; Blair et al., 1999; Dubayah et al., 2010) and the Geoscience Laser Altimeter System (GLAS with 65-m footprint) aboard the Ice, Cloud, and land Elevation Satellite (ICESat; Schutz et al., 2005; Los et al., 2012). The Global Ecosystem Dynamics Investigation (GEDI) mission is planned to launch in late 2018 and will consist of a 25-m footprint lidar mounted on the international space station (Stavros et al., 2017). Well known small-footprint lidar systems are the Airborne Taxonomic Mapping System (AToMS) aboard the Carnegie Airborne Observatory 2 (CAO-2; Asner et al., 2012c), and NASA Goddard's Lidar, Hyperspectral and Thermal

airborne imager (G-LiHT; Cook et al., 2013), both of which allow simultaneous lidar and hyperspectral data acquisitions. Apart from these experimental, scientific systems, there are many commercially operated airborne lidar systems worldwide, which are, e.g., commissioned by state authorities for land surveying. Emerging fields in lidar remote sensing are UAV-borne systems, as sensors are becoming more lightweight (Guo et al., 2017), and multispectral systems, which bear lasers in several spectral bands (Wichmann et al., 2015).

### 1.3.5 Remote Sensing Products about Forests

The most widely known products derived from remote sensing are **land cover maps**. Such maps are usually created by applying classification algorithms on the spectral information at pixel level. However, there are a multitude of different options within this general framework, leading to many different products claiming to map the same land cover type. In the case of forest, as mentioned earlier, the precise definition of forest matters and thus products providing **relative canopy cover** as a continuous number rather than a discrete class are preferred. Such products allow for the application of tree cover thresholds which match the different forest definitions, respectively. Global datasets of this kind are the MODIS **vegetation continuous field (VCF)** (Hansen et al., 2002) and more fine scale Landsat-based derivatives (Hansen et al., 2013; Sexton et al., 2013). They are regularly updated and used in various countries to track deforestation over time (Hansen et al., 2016). They also serve as a foundation for analyzing patterns of forest fragmentation (Taubert et al., 2018). Recently, the first SAR-based (TanDEM-X) global forest cover map has been published (Martone et al., 2018).

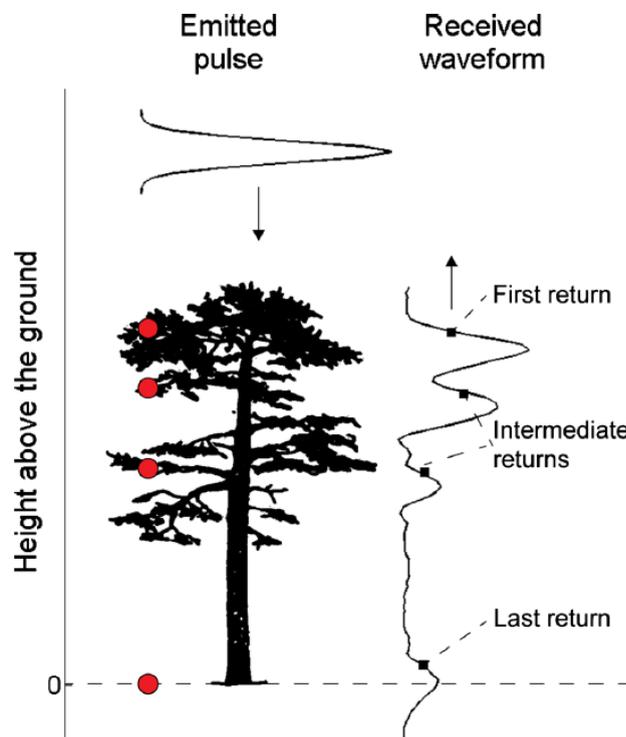


Figure 1.6: Lidar signal generation for one laser beam that hits a tree on multiple branches and the ground (Lindberg and Holmgren, 2017).

**Vegetation indices** are another common type of product. The idea behind vegetation indices is to combine the spectral information of the different bands in a way such that vegetated areas can be easily discriminated from non-vegetated areas. The most widely used vegetation index is the **normalized difference vegetation index** (NDVI; Rouse et al., 1974), which is defined as the difference between NIR and red divided by the sum of NIR and red band. It makes use of the fact that, plants, in contrast to soil and rock, show high reflectance in the NIR and high absorptance in the red band. Many other vegetation indices using different mathematical combinations and different bands have been suggested in the literature (Bannari et al., 1995). They serve to monitor the health status of vegetation and detect, e.g., stress caused by droughts or infestations. They are also often used to estimate important biophysical parameters such as **leaf area index** (LAI), i.e., the one-sided leaf area per unit ground area [ $\text{m}^2 \text{m}^{-2}$ ], and the **fraction of absorbed photosynthetic active radiation** (FAPAR; Myneni et al., 1997).

The **water content** in vegetation can be estimated based on passive microwave radiation emitted by soils. This emission has to pass the vegetation which is covering the soil and based on the transmission the so called **vegetation optical depth** (VOD) can be derived (Liu et al., 2011). VOD can serve in drought stress monitoring (Anderegg et al., 2018).

A further step is the delineation of **biophysical** and **biochemical plant traits**. Here, based on the spectral information the chemical composition of leaves is estimated. For this purpose commonly spectral libraries are used, which contain many reference spectra and associated chemical leaf composition values obtained in laboratory measurements. With such an approach concentrations of, e.g., nitrogen, phosphorus, chlorophyll a and b and carotenoids can be derived from hyperspectral remote sensing data (Asner and Martin, 2009). This information can further be used to monitor stress and forest health, classify forest types, quantify functional diversity or even identify tree species (Jetz et al., 2016; Lausch et al., 2016).

**Tree species identification** using remote sensing is a difficult task. Best results have so far been achieved in relatively species poor forests of the boreal or temperate zones (Immitzer et al., 2012). While different functional groups, e.g., conifers and broadleaves, show different characteristics in their reflectance spectra, many species within those groups are hard to distinguish based on the limited available spectral information. In deciduous forests the phenology across the year (leaf flushing, flowering, autumn coloring, senescence) can provide additional features to distinguish species (Fassnacht et al., 2016). In hyperdiverse tropical forests with small phenological changes current technology cannot provide taxonomic information at species level. However, the diversity of spectral traits and functional types can serve as an indicator for taxonomic diversity (Jetz et al., 2016; Asner et al., 2017).

**Individual tree crown delineation** (ITCD) is an area that has gained much attention in recent years. While most remote sensing products are area-based products, meaning they represent aggregated values for a certain mapping unit, ITCD allows the derivation of object-based, i.e., tree-based attributes (Coomes et al., 2017). Mostly 3D data from small footprint lidar systems is used to identify single trees (Lindberg and Holmgren, 2017), but algorithms also exist for passive optical imagery (Jing et al., 2012). Obtained products include **tree numbers** (stand density), **tree heights**, **crown diameters**, **crown shapes** and more. The incorporation of ITCD information can provide tree-based versions of many of the discussed and commonly area-based remote sensing products (e.g. species identification, biomass).

With the goal of monitoring global change, the Global Climate Observation System (GCOS) has defined a list of **essential climate variables** (ECV; Hollmann et al., 2013) and correspondingly the Group on Earth Observations Biodiversity Observation Network (GEOBON) has defined a list of **essential biodiversity variables** (EBV; Pereira et al., 2013). These variables have been suggested for continuous monitoring in space and time, which requires remote sensing approaches (Pettorelli et al., 2016). Several of the variables in both lists are directly related to vegetation, such as leaf area index, fraction of absorbed photosynthetic active radiation, forest structure and forest biomass. In the following, the possibilities of how to estimate the latter two with remote sensing will be presented.

### 1.3.6 Remote Sensing of Forest Structure

**Canopy height** can be obtained in various ways. Lidar is certainly the most straightforward technology for canopy height measurements, providing direct information on the position of canopy elements and the ground. However, also optical and SAR imagery may allow the extraction of forest height. Several fine resolution optical images showing a scene from different viewing angles can be combined to derive 3D information via photogrammetry. SAR imagery collected in a bistatic configuration can be processed to derive 3D information via interferometry or PolInSAR (Figure 1.7). Depending on the data source this 3D information may be in the form of discrete point clouds or continuous profiles of backscatter intensity across vertical canopy layers. One problem with such data is that the signal is increasingly attenuated when traveling through the canopy and this attenuation depends on a number of factors, such as signal wavelength, viewing geometry, leaf angle distribution and moisture. Hence, comparing remote sensing products of vertical canopy structure derived from different sensor systems or just different acquisition settings can be challenging.

The most important and also most robust remote sensing products to analyze canopy height are gridded **digital elevation models** (DEM). There are several types of DEMs relevant for vegetation mapping (Morsdorf et al., 2008): The elevation of the upper vegetation surface is accessible in **digital surface models** (DSM) and the underlying ground elevation in **digital terrain models** (DTM; Aryal et al., 2017). Pixel values in DSMs and DTMs represent elevation above a geographic coordinate reference ellipsoid. If DTM elevation is subtracted from DSM elevation a terrain-normalized digital surface model (nDSM) is obtained, where pixel values represent height above ground. If buildings and other man-made constructions are masked out and only vegetation pixels are retained the nDSM is commonly referred to as a **canopy height model** (CHM). An important condition for any attempt to estimate vegetation height is that accurate information about terrain elevation is available. Thus, either the remote sensing data itself has to contain a strong enough ground signal, as commonly the case for lidar and long-wave SAR, or a DTM from another data source (lidar) has to be combined with the DSMs derived from, e.g., photogrammetry or short-wave SAR (Kugler et al., 2014). Obtaining precise terrain elevations is particularly problematic in areas with steep slopes, where, e.g., the ground signal of large-footprint lidar is smeared over the whole elevation gradient in the footprint (Chen, 2010) and in very densely vegetated areas where the signal can hardly reach the ground and interpolation over large areas with missing ground information is necessary.

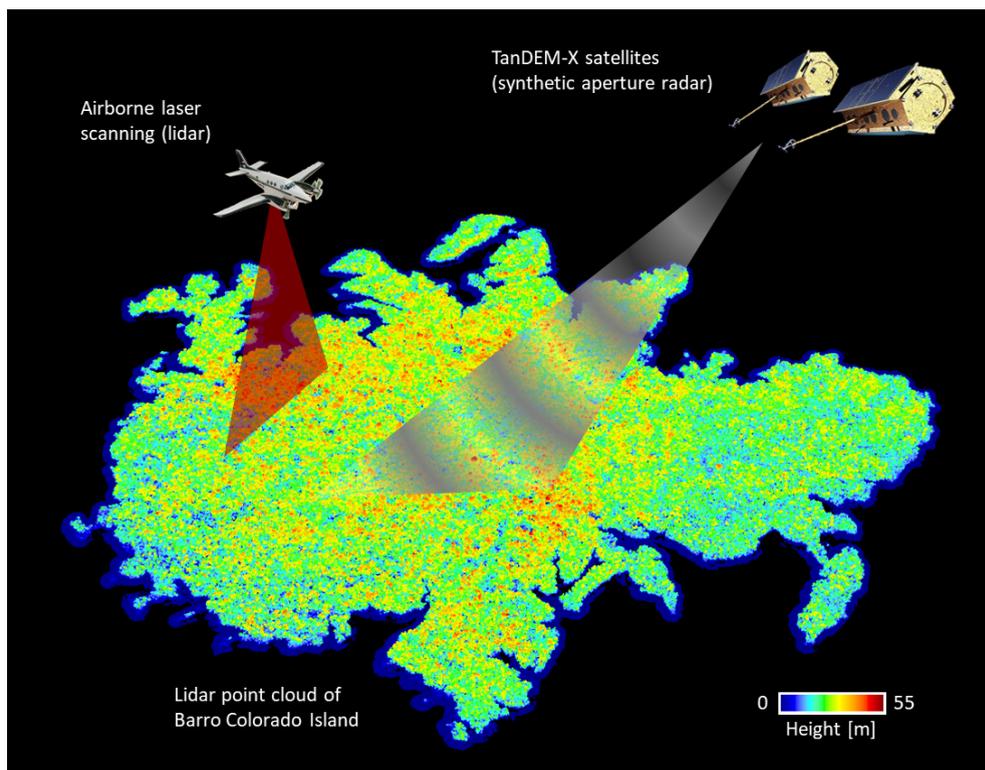


Figure 1.7: Active remote sensing techniques (lidar and SAR) for canopy height measurements. The point cloud of Barro Colorado Island was normalized with the underlying digital terrain model so that it represents vegetation height over ground.

The simplest way to quantify forest height is the **maximum tree height**. It can be extracted from remote sensing and inventory data. However, it only depends on the height of the one single largest tree in the stand and does not tell much about the height structure of the rest of the stand. Hence, average canopy height is more informative than maximum canopy height, albeit less easy to define. A simple way to calculate average canopy height from measured tree heights in inventories is the arithmetic mean. However, in a naturally regenerating stand there are usually many more small than large trees, but large trees have larger crown projection areas and contribute more to the stand's canopy surface. Foresters have long been accounting for this fact, by using a size-weighted mean tree height to describe the canopy height of a stand. This mean height of all trees weighted by each tree's basal area is commonly referred to as **Lorey's height** in forestry (Lorey, 1878).

In the remote sensing case it is implicit that large trees with large crowns contribute more to the distribution of the signal than small trees with small crowns. Hence, the mean height of a vertical remote sensing signal is related to field-based Lorey's height. There are different terms for the mean canopy height in the literature, depending on how exactly the remote sensing data was processed. Vertical canopy profiles may either represent vertical energy distribution or simply the count of points of a discrete point cloud per height bin. In either case a weighted mean height can be calculated using the profile values (energy or point count) as weighting variable. This is usually referred to as **mean canopy height** or mean canopy profile height (MCH; Equation 2.2 on page 32; Lefsky et al., 2002b). If the data is not yet aggregated to a vertical profile the calculation of MCH is even simpler: For a terrain-normalized

3D point cloud MCH is simply the mean height of all points (Figure 1.8). Instead of MCH, sometimes also the quadratic mean (QMCH; Equation 2.3 on page 32) or the median (e.g., HOME, for height of median energy) are used (Drake et al., 2002; Lefsky et al., 2002b). As mentioned earlier, the vertical signal distribution may be strongly influenced by acquisition parameters. This problem can be avoided by using only the rather robust information on canopy surface height, i.e., basing calculations on the CHM distribution instead of the full vertical profile. The mean height of all pixels in a CHM is often called **mean top-of-canopy height** (TCH; Figure 1.8; Equation 2.4 on page 32; Asner and Mascaro, 2014). Correspondingly, also the quadratic mean (QTCH) or median height of the CHM can be derived. Often a set of several **height quantiles** of the vertical (full profile or CHM) distribution is calculated, which are frequently named by the acronym RH for relative height followed by the number of the specific percentile (e.g., RH10, RH20, etc., with RH50 being the median height and RH100 being the maximum signal height corresponding to the maximum tree height in the considered area; Dubayah et al., 2010). Less common is the alternative, where instead of height percentiles the signal density percentages in a priori defined height strata are quantified (Hudak et al., 2012).

The most common application of a canopy density metric is the one where only two height bins are considered: the lower one being defined as canopy gaps and the upper one as canopy cover. In this way **gap fraction** and **fractional canopy cover** can be quantified. Their relative proportions vary from sparse to dense stands but also depending on the height threshold that marks the border between gap and canopy (Meyer et al., 2018). The **vertical heterogeneity** of the canopy can be quantified with metrics that describe the shape of the vertical signal distribution. Common metrics are standard deviation, coefficient of variation, Gini index, Shannon index, skewness and kurtosis (Bouvier et al., 2015). These metrics measure whether the tree crowns in a stand are rather concentrated in one or distributed over several layers and whether the canopy is denser at the high or low end. Lidar profiles can also be used to reconstruct **vertical foliage profiles** (VFP, i.e., leaf area density estimates in each height layer) by inverting the light extinction in the canopy (Equation 4.4 on page 72; MacArthur and Horn, 1969; Harding et al., 2001; Tang et al., 2012).

### 1.3.7 Remote Sensing of Forest Biomass

Forest biomass cannot be derived directly from remote sensing. At the center of biomass estimation algorithms there is always a statistical relationship between field-estimated biomass in **ground-truth plots** and remote sensing metrics derived over these plots. Data from all the mentioned sensor types (passive optical, radar and lidar) have been employed for biomass estimation (Fassnacht et al., 2014; Lu et al., 2014). Clear relationships between optical reflectance or vegetation indices and biomass can only be observed in areas of sparse vegetation. Such relationships show **saturation** over closed canopies, which is commonly the case at biomass stocks around 100 to 150 t ha<sup>-1</sup> (Lu et al., 2012). With SAR-data there are basically two ways of biomass estimation (Santoro and Cartus, 2018): One approach is to link SAR backscatter intensities to biomass. Such relationships only work well with long wave SAR (L-band) and saturate in a biomass range at around 200 t ha<sup>-1</sup>. The other approach is using InSAR or PolInSAR to derive canopy height and relate this height to biomass.

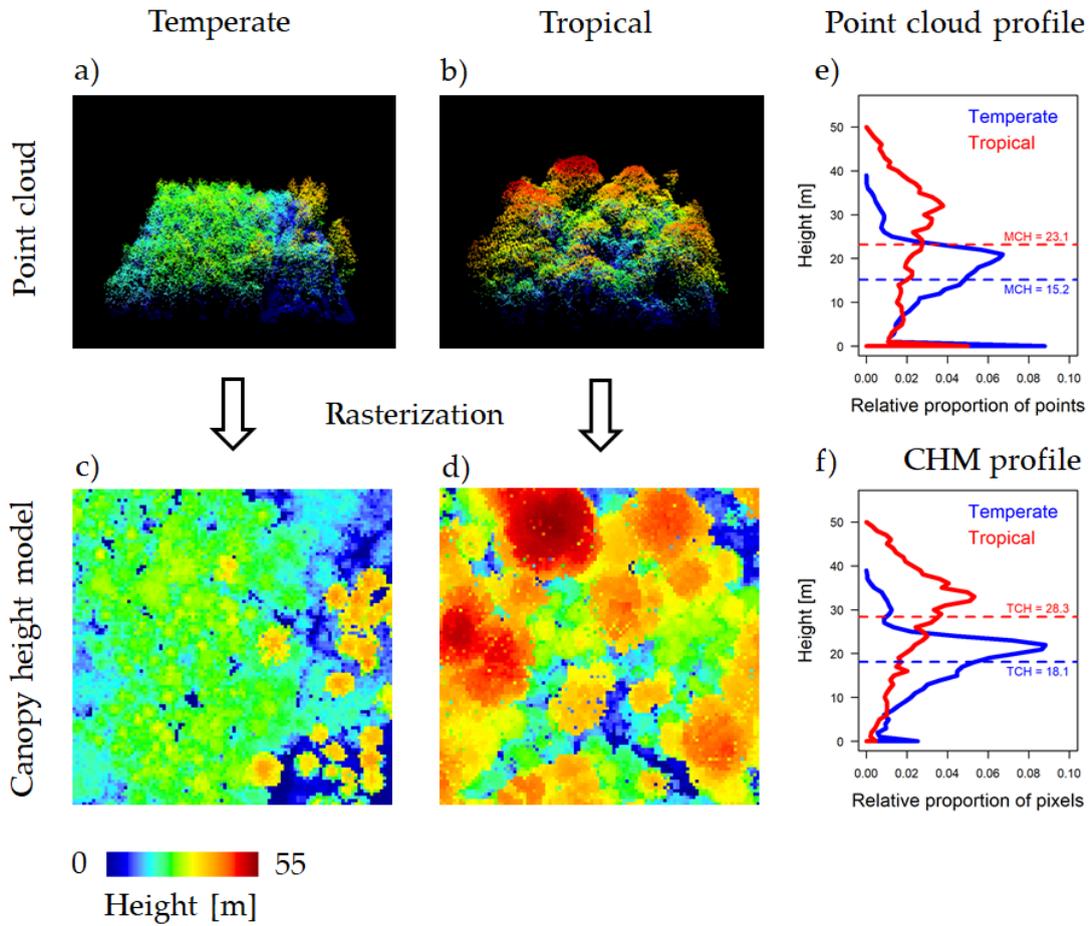


Figure 1.8: Lidar point clouds (1 ha, terrain normalized) of a temperate (a) and a tropical (b) forest and different aggregation steps. A 2D raster representation of the highest returns in a certain pixel resolution (e.g., 1 m  $\times$  1 m) is called canopy height model (CHM, c+d). The count of points in each height bin (e.g., 1 m) is called point cloud profile (e) and the mean height of all points is called mean canopy height (MCH). The count of CHM pixels in each height bin (e.g., 1 m) is called CHM profile (f) and the mean height of all pixels is called mean top-of-canopy height (TCH).

Using canopy height (from lidar, SAR or photogrammetry) is the most frequently applied approach for biomass estimation today, as it does not suffer from the saturation problems observed with other methods (Zolkos et al., 2013). Various combinations of the different **canopy height metrics** have been used in different studies (Chen, 2013). It is still a subject of ongoing research which metrics work best at which spatial scales and how the relationships differ under different conditions (e.g., disturbance regimes, forest types). Those questions are a main subject of this thesis. The estimation approach for biomass from forest height and structure metrics can be transferred to many other inventory variables such as basal area, quadratic mean stem diameter, timber volume or tree density (Næsset, 2002). With repeated acquisitions this methodology can be used to quantify **changes over time** (Cao et al., 2016).

In recent years there have been first attempts to map forest biomass at global scale (Saatchi et al., 2011, Baccini et al., 2012; Avitabile et al., 2016). The largest problem with **wall-to-wall biomass maps** is that there is no wall-to-wall canopy height data available. Spaceborne lidar sensors (ICESat GLAS, GEDI) can only provide point samples of canopy height with large unsampled areas in between. Spaceborne SAR

systems are either not bistatic, making InSAR canopy height retrieval impossible, or have too short wavelengths (TanDEM-X) for a separation of canopy and terrain height. Hence, all global biomass maps so far have used a three-stage approach consisting of the following steps: 1) Field-based biomass is linked to canopy height measurements from ICESat GLAS at footprint locations that coincide with ground-truth plots. 2) Predictions of biomass based on this first relationship are done for all footprint locations of ICESat GLAS. These predicted biomass values are then linked to surrogate variables which are available wall-to-wall (e.g., passive optical reflectance, SAR backscatter intensity, climate and topography layers). 3) Wall-to-wall predictions of biomass are made based on the second relationship. However, this second relationship suffers from the saturation problems described earlier. Hence, uncertainties in the derived biomass maps are large and at regional scale the existing maps differ strongly from each other (Mitchard et al., 2013). A similar three-stage approach is behind the high resolution biomass maps published at national and subnational level (Columbia, Hawaii, Panama, Peru; Asner et al., 2011; Asner et al., 2012a; Asner et al., 2014), where airborne lidar data was only collected over transects across the countries and wall-to-wall predictions are based on surrogate variables.

The scientific community is working on the development of methods for large scale wall-to-wall canopy height measurements for biomass mapping and monitoring. This involves data **fusion**, i.e., the synergistic combination of datasets, and the development of new target-tailored **future missions**. A promising example for fusion is the combination of TanDEM-X data with the upcoming spaceborne lidar GEDI (Qi and Dubayah, 2016), which will be sampling much more densely than the previous ICESat GLAS. A powerful new mission concept is Tandem-L, a bistatic long-wave SAR satellite system, which will provide wall-to-wall measurements of vertical forest structure and terrain height at high spatial and temporal resolution (Moreira et al., 2015).

## 1.4 FOREST MODELING

### 1.4.1 *Forest Models*

Models are simplified representations of reality. By reducing the complexity of nature to the essential processes of interest, ecological models help us to **synthesize** and test **ecological theory** for a better understanding of the environment. They can be used to conduct **simulation experiments** on how ecosystems may behave under different **scenarios** and allow **predictions** for future trajectories. In particular, forest ecology and management often rely on process-based forest models. Field experiments in forestry can take decades and centuries, due to the long time scales at which forests grow. Hence, in forest science, models are often the only way for conducting experiments and analyzing results within the lifetime of a single researcher (Pretzsch, 2009).

Since the beginning of forest science (Carlowitz, 1713), foresters have tried to predict timber growth into the future. Relationships derived from long term observations have led to **yield tables**, which were the first simple forest models (Moser Jr, 1980). However, yield tables are purely empirical and can only be applied to stands of relatively simple compositions (e.g., even-aged plantations with a single or a few species), under a fixed set of growing condition. They are unsuited for forests with a mixed-age and diverse species composition and cannot account for environmental

changes. To be able to predict the development of forests with more complex structures and under variable conditions a variety of forest models have been developed in the second half of the 20th century. The increase in computational capacity has facilitated this development over the recent decades (Shugart et al., 2018).

Forest models can be broadly categorized into three main types (Huth, 1999): 1) **Whole stand models** are focused on aggregated state variables, such as total timber volume or biomass. 2) **Size class models** are able to simulate the development of the stem size distribution over time (Kohyama, 1992). Mathematically, both model types can be formulated as sets of partial differential equations. 3) Single tree models are **individual based models** (IBM), which simulate recruitment, competition, growth and mortality at the individual tree level (Shugart, 1984; Bugmann, 2001). Apart from this classification based on the level of detail, models can also be broadly categorized along a gradient from purely empirical to rather process-oriented. **Empirical** models rely more on statistical relationships derived from field measurements, while **process-oriented** models try to incorporate underlying physiological processes. The difference can be illustrated by considering the simulation of annual stem diameter increment: An empirical approach would be to model diameter increment as a direct function of current stem diameter and a competition-based reduction factor. A process-oriented approach would be to derive diameter increment from a photosynthesis- and respiration-driven biomass balance (Huth, 1999).

The strength of single tree models lies in the **structural realism**, i.e., their output (and input if simulations are initialized with an existing stand) has the same structure and level of detail as forest inventory data. This design allows the explicit modeling of processes that act at the individual level (Grimm and Berger, 2016). IBMs, however, have certain caveats: Their **complexity** often requires a large number of parameters which makes it difficult to constrain the models. Furthermore, simulations can require long processing times, depending on the number of considered processes, the number of species and the simulation time step (Shugart et al., 2018).

A strategy to reduce processing time is to avoid modeling the pairwise interaction of trees explicitly. Instead competition among trees is modeled in an aggregated way by assuming that a tree affects all its neighboring trees within a certain area (patch) in the same way. Thus, within such a patch the trees do not have explicit positions in space. Their crowns can be imagined as discs or cylinders in a certain height that extend over the entire patch. Competition and thus succession is modeled within each patch independently of the neighboring patches. The first model following this idea was JABOWA (Botkin et al., 1972) and many modifications of the original idea have led to a whole model family which have been called **gap models** (Shugart Jr and West, 1980). The name refers to the size of the patches, which corresponds to the size of a typical tree fall gap. Since then, more than one hundred different gap models have been developed, all of which differ slightly with regard to considered processes and fields of application. At least 12 of them have been used continuously over the past decades with each having more than one hundred scientific citations (Shugart et al., 2018).

Forest gap model concepts are currently being integrated in **dynamic global vegetation models** (DGVM) for analyses of atmosphere vegetation feedbacks in earth system models (Fisher et al., 2018). DGVMs have traditionally used very simple representations of vegetation, based on few plant functional types (PFT) and single layer canopies (Sitch et al., 2003). By incorporating individual based submodules vegetation demography and the resulting structure can be represented more accurately (Moor-

croft et al., 2001). Modern computer clusters meanwhile even allow direct applications of forest gap models at continental scale using **regionalized** drivers (Rödiger et al., 2017b). Constraining such large scale applications of models often relies on remote sensing products. But this is just one of multiple synergies that arise from linking remote sensing with forest models (Shugart et al., 2015).

#### 1.4.2 *Linking Forest Models and Remote Sensing*

Forest models are increasingly used in combination with remote sensing data for various purposes. There are at least four different motivations behind linking remote sensing and ecological models with each other (Plummer, 2000): 1) model **parameterization** and **initialization** (provision of functional relationships and starting conditions), 2) model **validation** (comparison against empirical data), 3) model **calibration** (iterative improvement of the model based on comparisons between output and empirical data, also referred to as assimilation, constraining, tuning or updating) and 4) remote sensing **interpretation** and **understanding** (inversion of remotely sensed signal into biophysical variables). Usually, a combination of several of these aspects plays a role. The link can either be achieved by empirical **parametric regression**, by non-parametric statistical approaches (**machine learning**) or by **physically based inversions** using radiative transfer models or look-up tables (Verrelst et al., 2018). **Radiative transfer models** are mathematical descriptions of the scattering of radiation considering canopy and ground surface parameters. They allow the forward simulation of synthetic remote sensing products based on ecosystem model output, which can then be compared to real-world remote sensing products (Jacquemoud et al., 2009; Kuusk, 2018). The ecosystem model output serves as synthetic ground-truth data (e.g., synthetic inventory tables in case of a forest model).

Earliest applications that combined remote sensing products with ecosystem models already tried to extract "invisible" variables from remote sensing by correlating NDVI to process model-derived estimates of photosynthesis and net primary productivity (NPP; Running and Nemani, 1988). In later applications, structural forest parameters have served to initialize forest simulations across landscapes, either using SAR-derived biomass (Ranson et al., 2001) or lidar-derived canopy height (Hurt et al., 2004) as the linking variables. Falkowski et al. (2010) used machine learning to impute sample point inventory data with wall-to-wall lidar data to parameterize and initialize a forest model at individual tree-level and landscape scale. In the context of better understanding and interpreting remote sensing variables, synthetic ground-truth and synthetic remote sensing data have been used to explore possibilities for estimating biomass and other stand variables. Studies addressed the effects of sampling, measurement scale and geolocation errors (Hurt et al., 2010; Köhler and Huth, 2010; Frazer et al., 2011).

#### 1.4.3 *The Model FORMIND*

FORMIND is an individual-based and process-oriented forest gap model designed to simulate the dynamics of species-rich forests (Köhler and Huth, 1998; Fischer et al., 2016). The model simulates the processes of establishment, growth, competition and mortality of trees on a regular grid of patches with the dimensions of a typical treefall gap (20 m × 20 m). Within each patch the trees do not have explicit spatial positions, following the gap model concept. By combining many patches, large forest areas up

to hundreds of hectares can be simulated. Tree species with similar ecological traits are aggregated into **plant functional types (PFT)** to facilitate parameterization for diverse forests and reduce computation time. The PFTs may, e.g., represent different successional types (from pioneers to climax species) or size classes (from understory to emergent species).

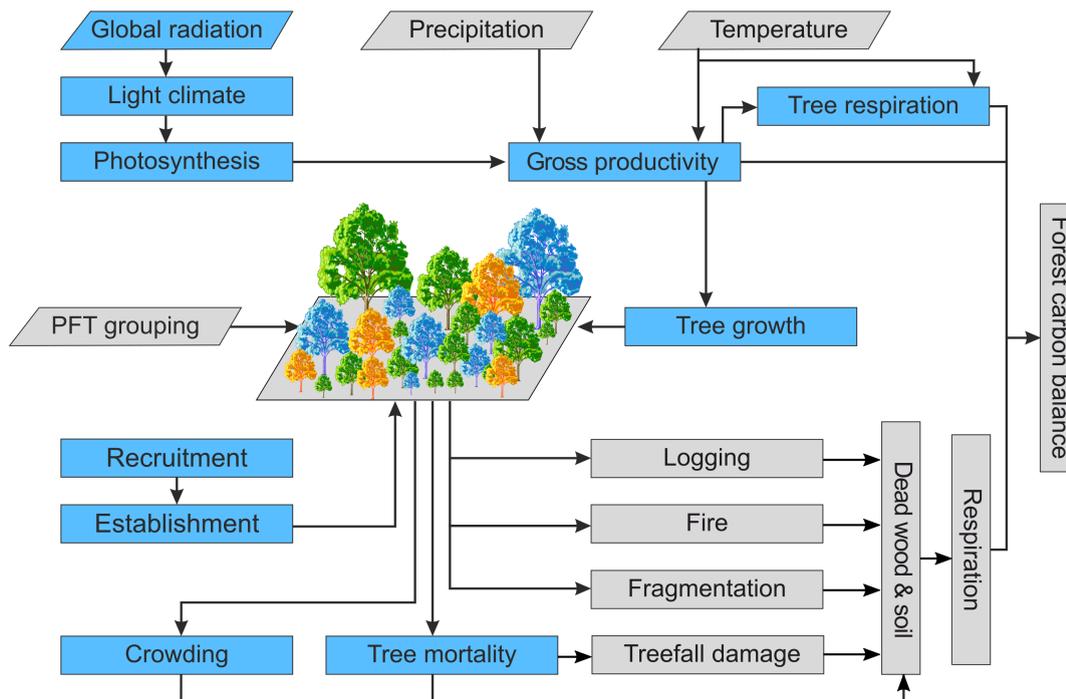


Figure 1.9: Flowchart of the main processes simulated in the individual-based forest model FORMIND. The core processes are represented by blue boxes. Optional processes and inputs are represented by grey boxes (e.g., climate dependence, disturbance regimes, full ecosystem carbon balance). The grouping of species into plant functional types (PFT) is indicated by the colored trees.

In each simulated time step, the following main processes take place (Figure 1.9):

- 1) **Establishment:** Seeds are distributed over the forest area. If light conditions are suitable, new saplings can establish and compete for light and space in the patch.
- 2) **Competition:** The main driving factor of the model is light. Radiation intensity within each patch decreases from the top to the ground according to a light extinction function. The light extinction depends on the combined vertical leaf area profile of all trees in the patch. The productivity of each tree is determined by the available light in its height layer.
- 3) **Growth** of each tree is dependent on its gross primary productivity (GPP), respiration and PFT-specific physiological and allometric parameters.
- 4) **Mortality:** Trees die stochastically according to a PFT-specific mortality rate. If a tree falls, it can damage neighboring trees in adjacent patches. Optionally, tree-density-, size- or growth-rate-dependent mortality can be activated.

Besides these core processes, FORMIND has the following optional features:

- 5) **Environmental limitations:** Forest dynamics can be influenced by climatic variables, like precipitation and temperature.
- 6) **Disturbances** such as fires, landslides, logging and fragmentation can be simulated.
- 7) **Carbon cycle:** Gross primary production, respiration and net primary production are calculated for each individual tree. Based

on this the carbon balance for a whole forest can be derived, including soil respiration and net ecosystem productivity.

FORMIND has been used for more than 20 years to study different aspects related to forest management, disturbance regimes and climate change. Currently, 26 different parameterizations are being maintained (November 2018), including published, site-specific versions for Brazil (Groeneveld et al., 2009; Dantas de Paula et al., 2015), Ecuador (Dislich et al., 2009; Paulick et al., 2017), Ethiopia (Hiltner et al., 2016), French Guiana (Hiltner et al., 2018), Germany (Bohn et al., 2014; Rödiger et al., 2017a), Madagascar (Fischer et al., 2014; Armstrong et al., 2018), Malaysia (Köhler and Huth, 1998; Huth et al., 2004), Panama (Kazmierczak et al., 2014; Knapp et al., 2018a), Tanzania (Fischer et al., 2015; Fischer et al., 2018), Venezuela (Kammesheidt et al., 2001), as well as two large-scale versions for the tropical rainforests of the Amazon basin (Rödiger et al., 2017b) and temperate forests of Central Europe (Bohn et al., 2017), and a grassland version (GRASSMIND), which transfers the gap model concept to herbaceous plant communities (Taubert et al., 2012). The detailed model description was published with Fischer et al. (2016) and can also be found on [www.formind.org](http://www.formind.org).

## 1.5 OBJECTIVES OF THIS THESIS

The main objective of this thesis was to improve our understanding of remote sensing of forest structure for the estimation of biomass. To this end, analyses of empirical data (remote sensing and field inventories) were combined with process-oriented modeling. The work addresses the three aspects of 1) biomass stocks, 2) biomass change and 3) variability among forest types, each of which is the focus of one of the three following chapters.

In Chapter 2 about remote sensing of **biomass stocks**, a lidar simulation model was developed and integrated into the forest model FORMIND. The model was parameterized for a tropical rainforest in Panama using field inventory and airborne lidar data from the Barro Colorado Island (BCI) megaplot site. The objectives of the study were: 1) Establishing a lidar simulation model that is able to produce synthetic lidar-like data for dynamic forest model output. 2) Testing a wide variety of lidar metrics for their ability to predict biomass of a tropical rainforest at various spatial scales. 3) Investigating the influence of disturbances on the lidar-to-biomass relationships. This chapter was published in the journal *Remote Sensing of Environment* (Knapp et al., 2018a).

In Chapter 3 about remote sensing of **biomass change**, different approaches (direct and indirect) for estimating biomass change over time from canopy height change were compared. Again, BCI served as a test site and canopy height models derived from FORMIND simulations and from the TanDEM-X satellites were analyzed to answer the following questions: 1) How can aboveground biomass change be estimated from canopy height change, if the goal is to cover the full range of tropical forest succession including disturbed forests? 2) Under which conditions does each of the different approaches work best? 3) Can canopy texture information improve estimates of biomass change in the absence of canopy height information? This chapter was published in the journal *Remote Sensing* (Knapp et al., 2018b).

In Chapter 4 about **variability among forest types** in the relationships between canopy height metrics and biomass, data from five megaplot sites representing different continents and biomes was brought together. For each site, a set of metrics

## INTRODUCTION

was computed from lidar data, with each metric characterizing a different aspect of forest structure. The goals of the study were: 1) Finding a generic approach for basal area and aboveground biomass estimation that can be applied across all sites without causing prediction bias at any individual site. 2) Investigating the contributions of the different structural attributes in the estimation functions. This chapter is ready for submission to the journal *Remote Sensing of Environment*.

LINKING LIDAR AND FOREST MODELING TO ASSESS BIOMASS  
ESTIMATION ACROSS SCALES AND DISTURBANCE STATES

---

## 2.1 ABSTRACT

Light detection and ranging (lidar) is currently the state-of-the-art remote sensing technology for measuring the 3D structures of forests. Studies have shown that various lidar-derived metrics can be used to predict forest attributes, such as aboveground biomass. However, finding out which metric works best at which scale and under which conditions requires extensive field inventories as ground-truth data. The goal of our study was to overcome the limitations of inventory data by complementing field-derived data with virtual forest stands from a dynamic forest model. The simulated stands were used to compare 29 different lidar metrics for their utility as predictors of tropical forest biomass at different spatial scales. We used the process-based forest model FORMIND, developed a lidar simulation model, based on the Beer-Lambert law of light extinction, and applied it to a tropical forest in Panama. Simulation scenarios comprised undisturbed primary forests and stands exposed to logging and fire disturbance regimes, resulting in mosaics of different successional stages, totaling 3.7 million trees on 4,200 ha. The simulated forest was sampled with the lidar model. Several lidar metrics, in particular height metrics, showed good correlations with forest biomass, even for disturbed forest. Estimation errors (nRMSE) increased with decreasing spatial scale from < 10% (200-m scale) to > 30% (20-m scale) for the best metrics. At the often used 1-ha scale, the top-of-canopy height obtained from canopy height models with fine to relatively coarse pixel resolutions (1 to 10 m) yielded the most accurate biomass predictions, with nRMSE < 6% for undisturbed and nRMSE < 9% for disturbed forests. This study represents the first time dynamic modeling of a tropical forest has been combined with lidar remote sensing to systematically investigate lidar-to-biomass relationships for varying lidar metrics, scales and disturbance states. In the future, this approach can be used to explore the potential of remote sensing of other forest attributes, e.g., carbon dynamics, and other remote sensing systems, e.g., spaceborne lidar and radar.

## 2.2 INTRODUCTION

Due to their important role in the global carbon cycle and ongoing deforestation and degradation, tropical forests are of particular interest to biomass remote sensing. Tropical forest carbon accounting and monitoring of deforestation are important tasks in the context of REDD+ and global climate modeling. In recent years, remote sensing has led to considerable improvements in this field (Gibbs et al., 2007; De Sy et al., 2012; Pan et al., 2013). Airborne small-footprint lidar (light detection and ranging) is currently the state-of-the-art technology for measuring the 3D structure of forests (Lefsky et al., 2002a; Wulder et al., 2012; Mascaro et al., 2014). Various lidar metrics correlate well with different forest attributes. In particular, lidar-derived height metrics have commonly been used to predict forest aboveground biomass (AGB) and carbon density (ACD) (Drake et al., 2002; Asner et al., 2009; Dubayah et al., 2010; Jubanski et al., 2013; Asner and Mascaro, 2014). The major challenges in biomass estimation based on lidar data are that 1) the calibration of the prediction functions relies on field data that must be collected manually in inventory plots; and 2) there are many different metrics available using different spatial scales, and the task is to find the combination that provides accurate AGB predictions.

In inventory plots, tree diameters at breast height (DBH) are typically measured, from which AGB is calculated via known allometric equations (e.g., Chave et al., 2005; Chave et al., 2014; Chen, 2015). Lidar data are acquired for the same inventory plots to build regression models between lidar-based structure metrics and ground-based AGB. A wide range of metrics can be calculated from lidar data. To date, no standard approach for AGB estimation from lidar has been established and different studies have applied different metrics (Chen, 2013; Lu et al., 2014). Several publications have compared metrics among each other for different forest types (e.g., Lefsky et al., 1999; Lefsky et al., 2002b; Dubayah et al., 2010; Jubanski et al., 2013). However, there has not been a comparison of a wide range of metrics on a single tropical forest dataset. Lidar metrics can generally be divided into metrics which are based on the full 3D point cloud of lidar returns and metrics which are based on canopy height models (CHM), i.e., the rasterized canopy surfaces which are derived from the uppermost returns of the point clouds (Chen, 2013). The full 3D point cloud contains more information about the vertical canopy structure than the corresponding CHM. On the other hand, the vertical distribution of lidar returns also depends on technical properties of the specific sensor, making point-cloud-based metrics less robust and comparable between different studies than CHM-based metrics (Næsset, 2009; Asner and Mascaro, 2014). Many commonly used metrics can be calculated based on both types of data. Those metrics include mean heights (Lefsky et al., 2002b; Asner and Mascaro, 2014), relative height quantiles (the heights below which a certain percentage of returns or pixels falls; Patenaude et al., 2004; Dubayah et al., 2010; Meyer et al., 2013), and metrics of heterogeneity such as the standard deviation of heights or the Shannon diversity index of the height profiles (Stark et al., 2012). Other metrics, such as the ratio of above ground returns to total returns or fractional canopy cover above a certain height, that can be derived either from point clouds or CHMs describe relative vegetation cover.

An important aspect of AGB prediction from remote sensing is spatial resolution. Resolution means, first, spatial resolution of the remote sensing data from which different metrics are calculated and, second, the spatial resolution of the output map, i.e., the grain size of the units for which the metrics are calculated to produce an AGB prediction. The resolution of the data is determined by the sensor's technical specifications and the capacities to store and process data. The resolution of the mapping units is influenced by the desired estimation accuracy and the desired spatial detail of the mapped product. Köhler and Huth (2010), Mascaro et al. (2011b) and Chen et al. (2016) showed how errors in AGB estimations from mean lidar heights decreased with increasing grain sizes and that a grain of approximately 1 ha is required to achieve errors of < 10%.

Fitting any of the described lidar metrics to measured AGB relies on field inventory data. Forest inventory plots are limited in number, size and structural variety. The collection of inventory data is costly and laborious and most studies in the past made use of tens to a few hundred plots (Fassnacht et al., 2014). Those plots are often located in old growth forests. Hence, available data sets might not cover the full structural complexity of forests over their entire successional range (noteworthy exceptions are, e.g., Dubayah et al., 2010; Poorter et al., 2016). For lidar-to-AGB-calibration, a broad range of different forest succession states that cover the range of all possible AGB stocks and associated forest structures is preferable. To overcome this limitation, we propose a new approach in which we complement in situ measurements with simulated forest stands (Figure 2.1). We used an individual-based forest model (FORMIND, Fischer et al., 2016) to simulate a large virtual inventory dataset, covering

the full range of succession stages by including forest disturbances in the simulations. The model was parameterized to represent the well-studied lowland tropical rainforest of Barro Colorado Island, Panama (Condit et al., 2001; Kazmierczak et al., 2014). We developed a lidar model to sample lidar data of simulated forest stands.

The research goals of this study were 1) to establish a lidar simulation model that is able to produce synthetic lidar-like data for dynamic forest model output; 2) to test a wide variety of lidar metrics for their ability to predict AGB of a tropical rainforest at various spatial scales; and 3) to investigate the influence of disturbances on the lidar-to-biomass relationships.

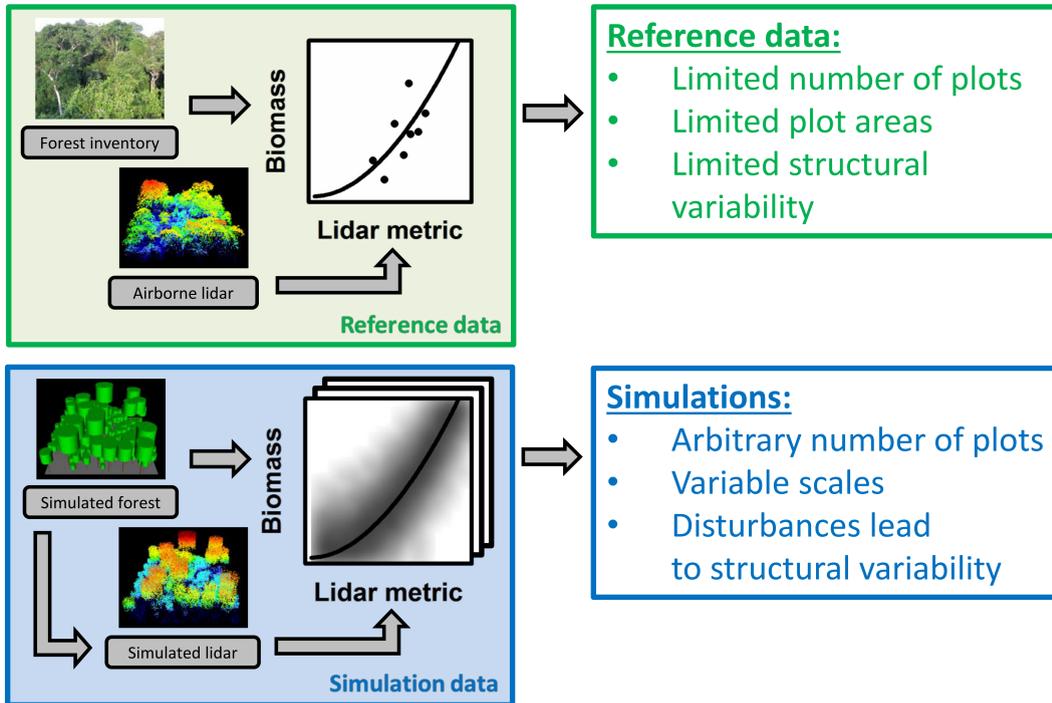


Figure 2.1: Workflow of the study. Reference data from field inventories and an airborne lidar campaign were used to parameterize and calibrate a forest model and a lidar model. With the models, large quantities of simulated inventory and simulated lidar data were generated, allowing for a systematic analysis of lidar-to-biomass relationships under different disturbance regimes and for various spatial scales.

## 2.3 MATERIALS AND METHODS

### 2.3.1 Study Area

The study focused on the tropical forest on Barro Colorado Island (BCI), Panama (9.15° N, 79.85° W). BCI is a 15 km<sup>2</sup> island located in Lake Gatun, an artificial water body created by the construction of the Panama Canal (Condit et al., 2001). It is covered with semi-deciduous tropical lowland rainforest. The minimum forest age is estimated to range from 300 to 1,500 years (Bohlman and O'Brien, 2006; Meyer et al., 2013; Lobo and Dalling, 2014). The climate is characterized by average daily maximum and minimum temperatures of 30.8 and 23.4 °C and an annual precipitation sum of approximately 2,600 mm, with a dry season from January to April (Condit

et al., 2001). A 50-ha rainforest observation plot is located on the central plateau of the island, with terrain altitudes varying between 120 and 160 m above sea level (Lobo and Dalling, 2014). Since the establishment of the plot in the early 1980s, each tree in the 1000 m × 500 m area with a DBH ≥ 1 cm has been measured during censuses in five year intervals (Condit, 1998; Hubbell et al., 1999; Hubbell et al., 2005). Estimates of the mean canopy height are 24.6 ± 8.2 m, and those of the mean AGB are 281 ± 20 t ha<sup>-1</sup> (Chave et al., 2003).

### 2.3.2 Lidar Data

An airborne discrete point cloud lidar dataset was collected on BCI in August 2009 with a multi-pulse scanning laser altimeter (Optech ALTM Gemini system; BLOM Sistemas Geoespaciales SLU, Madrid, Spain, Lobo and Dalling, 2014). The terrain elevation was subtracted from the point cloud to obtain the relative height above ground. Point densities ranged from 0 to 60 m<sup>-2</sup> with a median of 10 m<sup>-2</sup> and a 5th-percentile of 4 m<sup>-2</sup>. To avoid locally varying point densities, caused by flight swath overlaps, the point clouds were thinned by random subsampling of 4 returns in each square meter. A 1-m resolution canopy height model (CHM) was derived from the highest returns in each square meter. Data processing was performed using LAStools (Isenburg, 2011) and R (R Development Core Team, 2014).

### 2.3.3 Lidar Model Description

The purpose of the lidar model is the simulation of a lidar scan of a given forest stand. More specifically, it generates point clouds of discrete returns as usually produced by small-footprint lidar systems. As input, a tree list has to be provided. The list can either be real forest inventory data or data generated by a forest model (Figure 2.2a). The basic elements of the model are trees, lidar pulses and lidar returns. Trees are characterized by their position (X- and Y-coordinate), height, crown length, crown radius, crown shape and leaf area index (LAI). The model operates in a 3D space represented by an array of cuboid voxels. Each vertical column of voxels represents one modeled lidar pulse. Lidar returns are points in 3D space, characterized by their X-, Y- and Z-coordinates.

From the tree list, a voxel representation of the entire forest is created. Thus, voxels that could potentially produce a lidar return, because they belong to a tree crown or the ground, are distinguished from empty space voxels. The voxel forest is then scanned with a virtual lidar. The simulation follows a probabilistic approach. Instead of explicitly simulating the branches and foliage and their interaction with laser beams within the tree crowns, the model assumes that the tree crown space is a homogeneous, turbid medium filled with a certain leaf area density (LAD). The probability of having a lidar return from a certain point decreases as the distance the laser beam has to travel through the medium before reaching the point increases. This relationship is analogous to the Beer-Lambert light-extinction law (Campbell and Norman, 2012). Thus, the probability for a lidar return  $P$  for each tree and ground voxel (Figure 2.2c) can be calculated as a function of cumulative leaf area index LAI above the voxel (Figure 2.2b).

$$P(LAI) = P_0 \cdot e^{-k \cdot LAI} \quad (2.1)$$

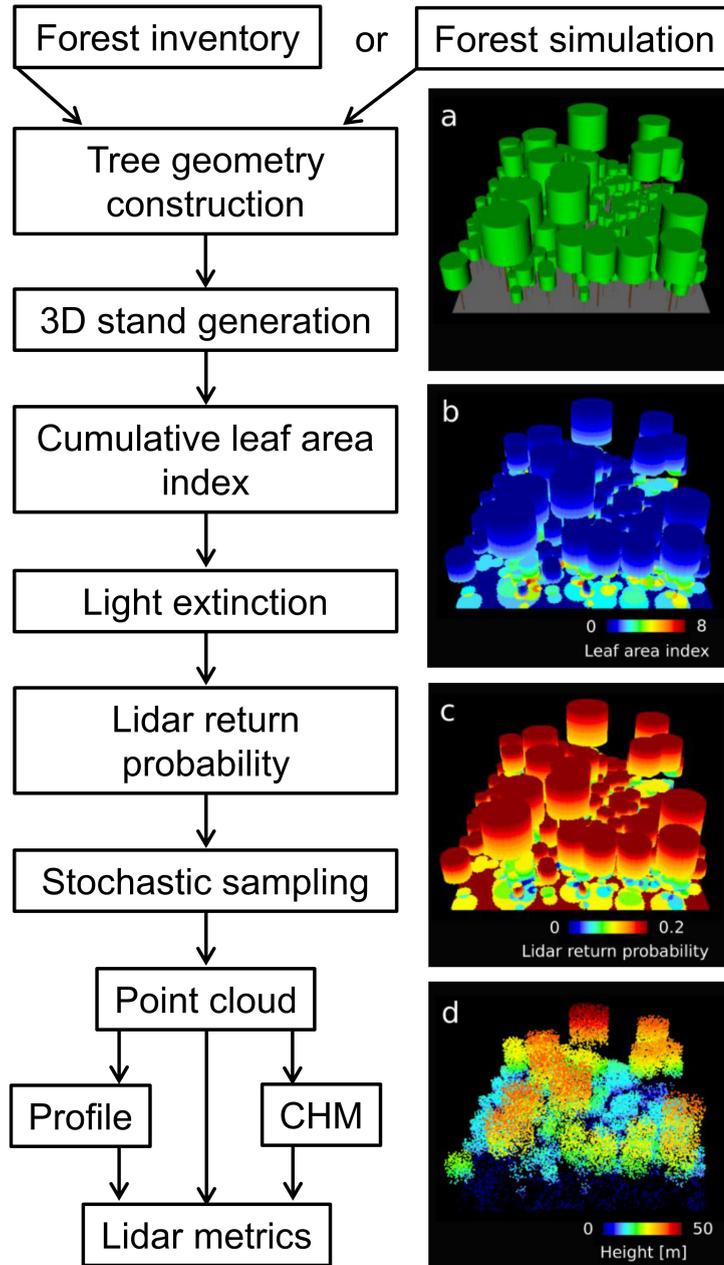


Figure 2.2: Principle of the lidar model. Inputs to the workflow can either be forest model output or field inventory data. The pictures on the right side show intermediate products: a) Visualization of a forest stand; b) voxel representation with colors indicating the cumulative leaf area index; c) voxel representation with colors indicating the probability of containing a lidar return; d) simulated lidar point cloud with colors indicating height above ground.

$P_0$  in Equation 2.1 represents the probability of obtaining a return from the very upper voxel, where the laser beam hits a tree or the ground for the first time. The parameter  $k$  is the exponential extinction coefficient, which determines how fast the return probability decreases after entering the crown space. The decision regarding whether each voxel will contain a return is taken stochastically, based on the calculated return probability. Ultimately, this leads to a discrete point cloud (Figure 2.2d). The

voxel resolution was set to  $0.5 \text{ m} \times 0.5 \text{ m}$  along the horizontal direction and  $1 \text{ m}$  along the vertical direction. The parameters  $P_0$  and  $k$  were calibrated such that simulated point cloud profiles derived for subareas of the 50-ha inventory data set matched the airborne lidar profiles of those subareas (details see Appendix A.1.1). The resulting value for  $k = 0.2$  can be confirmed by literature (Campbell and Norman, 2012; Jones, 2013). For  $P_0$  we found 0.2 to be a good value, leading to simulated point densities that were similar to the airborne reference point cloud.  $P_0$  being smaller than 1 can be interpreted by the heterogeneity of leaves, branches and empty space within the tree crown. This means that a laser beam entering the idealized cylindrical tree crown does not necessarily trigger a return in the first voxel.

#### 2.3.4 Forest Model Description

FORMIND belongs to the group of forest gap models (Botkin et al., 1972; Shugart, 1984; Bugmann, 2001). As such, the model simulates the processes of establishment, growth, competition and mortality of trees on spatial patches with the dimensions of a typical treefall gap ( $20 \text{ m} \times 20 \text{ m}$ ). By combining many patches, large forest areas of hundreds of hectares can be simulated. FORMIND is an individual-based model (IBM) in which the individuals represent trees that belong to different plant functional types (PFTs). One PFT may contain several species with similar ecological traits. FORMIND has been applied to many tropical forest sites and has proven capable of accurately reproducing patterns observed in these complex ecosystems (Fischer et al., 2016). The individual-based model architecture allows for the inclusion of disturbances such as logging or forest fires in a structurally realistic way. A detailed description of FORMIND including the modules for logging and fire disturbance can be found in Fischer et al. (2016). The supplements contain descriptions of the parameterization of the lidar model and the forest model (Table A.1). Before using the forest model output for remote sensing analyses, the structural validity of the simulated old growth stands was confirmed by visually comparing biomass stocks (Figure A.1) and stem size distributions (Figure A.2) of all PFTs to the values obtained from the inventory data.

#### 2.3.5 Simulation Experiment

Using FORMIND, we simulated the development of a  $16 \text{ ha}$  ( $400 \text{ m} \times 400 \text{ m}$ ) area of the BCI forest over several thousand years and stored the results at 20-yr intervals. The simulations were repeated with different disturbance regimes. The first run comprised 2,000 yr without any external disturbance, simulating only natural gap dynamics. In the second run, forest fires were introduced as a source of spatially heterogeneous disturbance to clear parts of the area regularly and enable natural succession and regrowth. Fire occurrence was drawn from a Poisson distribution such that the mean interval between two fire events was 25 yr. Fire size at each fire event was drawn from an exponential distribution, such that on average 50% of the total area was affected. More information on the fire module used is provided in Fischer (2013) and Fischer et al. (2016). The third scenario included selective logging. At a logging cycle of 99 yr, all trees with  $\text{DBH} > 30 \text{ cm}$  were felled and removed. More information on the logging module used is provided in Huth et al. (2004). For all three runs, the first 200 yr were discarded as spin-up. For each of the remaining simulation years, a virtual

lidar campaign using the lidar model was conducted. The disturbance frequencies and intensities were not intended to represent realistic disturbances scenarios in the study region. The intention was to sample many stands at each stage along the full successional range, using the disturbance modules to regularly set the forest back to an early stage. The selective logging acts on the whole area, while the fires move in a spatially explicit way through the simulated area, causing mosaics of unaffected forest next to cleared areas where succession starts over. Such patchy landscapes are typical for many forest regions, although the reasons for the structures may be as diverse as clear cuts, wind blowdowns, fires or natural areas without vegetation, e.g., grasslands or water bodies. Thus, these simulations produce landscapes that can be used as general examples of heterogeneous landscapes.

### 2.3.6 Lidar-based Biomass Prediction

We analyzed forest plots measuring 20, 33, 50, 100 or 200 m (side length). At each spatial scale, a range of 29 different lidar metrics (Table 2.1) were tested for their suitability as single predictors of AGB. Metrics were either derived from point clouds (PC) or canopy height models (CHM). CHMs were constructed from point clouds by rasterizing the highest lidar returns in each pixel of a given pixel size. Point-cloud-based metrics comprised the mean canopy profile height (MCH), which is the mean height of all lidar returns, and the quadratic mean canopy profile height (QMCH), where high returns receive a larger weighting than low returns. For a given point cloud profile  $p_{PC}$  that consists of lidar return counts at height bins  $h_i$ , MCH and QMCH can be calculated from Equations 2.2 and 2.3, respectively.

$$MCH = \frac{\sum_{i=1}^{i_{max}} (p_{PC,i} \cdot h_i)}{\sum_{i=1}^{i_{max}} p_{PC,i}} \quad (2.2)$$

$$QMCH = \sqrt{\frac{\sum_{i=1}^{i_{max}} (p_{PC,i} \cdot h_i^2)}{\sum_{i=1}^{i_{max}} p_{PC,i}}} \quad (2.3)$$

where  $p_{PC,i}$  is the lidar return count in height bin  $h_i$ . A metric similar to MCH can be derived from the vertical CHM profile instead of the point cloud profile. This metric corresponds to the mean of all pixel values of the CHM, and is commonly referred to as the mean top-of-canopy height (TCH, Equation 2.4).

$$TCH = \frac{\sum_{i=1}^{i_{max}} (p_{CHM,i} \cdot h_i)}{\sum_{i=1}^{i_{max}} p_{CHM,i}} \quad (2.4)$$

A CHM can be derived from a point cloud at variable pixel resolutions, by taking the height of the highest return that falls into each pixel. Thus, TCH always depends

on the pixel size used. We calculated TCH from CHMs with pixel side lengths of 1, 5, 10, 20, 33, 50 and 100 m. Note that, once the pixel size equals the plot size for which AGB is calculated, TCH is equal to the maximal height in the plot, which is also referred to as  $H_{\max}$  or RH100 in the literature. Another method for measuring forest height from lidar data is by using relative height quantiles of either the point cloud or the CHM. These quantiles represent the heights below which a certain percentage of the returns or CHM pixels fall. We calculated RH25, RH50 and RH75 for the point clouds and 1-m resolution CHMs.

Other metrics, however, capture the vertical heterogeneity of the forest. Those metrics include the standard deviation (SD) of heights (point-cloud- or CHM-based), the coefficient of variation (CV, Equations 2.5 and 2.6), the skewness of the vertical point cloud profile (Equation 2.7, where  $N$  is the total number of points and  $h_i$  is the height of each point  $i$ ), the Shannon Index (Equation 2.8, where  $i_{\max}$  is the number of height layers and  $p_i$  is the count of points in the layer  $i$ ) as a measure of entropy of the profile and the P:H ratio (Equation 2.9, where  $i_{\max}$  is the number of height layers,  $p_i$  is the count of points in the layer  $i$  and  $h_i$  is height of layer  $i$ ), which describes the height of the densest part of the point cloud (peak in the profile) relative to the maximal height (Marvin et al., 2014).

$$CV_{PC} = \frac{SD_{PC}}{MCH} \quad (2.5)$$

$$CV_{CHM} = \frac{SD_{CHM}}{TCH} \quad (2.6)$$

$$Skewness = \frac{1}{N} \cdot \sum_{i=1}^N \left( \frac{h_i - MCH}{SD_{PC}} \right)^3 \quad (2.7)$$

$$Shannon\ Index = - \sum_{i=1}^N p_i \cdot \ln(p_i) \quad (2.8)$$

$$P : H\ ratio = \frac{h(\max_{i \in [1, i_{\max}]}(p_i))}{\max_{i \in [1, i_{\max}]}(h_i)} \quad (2.9)$$

Furthermore, we calculated vegetation density metrics. Based on the point clouds, the count of aboveground returns divided by either the count of ground returns  $N_{AGR}/N_{GR}$  or the count of total returns  $N_{AGR}/N_{TR}$  was calculated. Based on the CHMs, the fractional canopy cover (FCC) was derived by defining different height thresholds below which a CHM-pixel was considered a canopy gap. We calculated FCC0, FCC10 and FCC20 using the forest floor, 10 m and 20 m as height thresholds, respectively.

Each lidar metric LM was fit to the dependent variable AGB using a power law model (Equation 2.10) and maximum likelihood estimation in R.

$$AGB = a \cdot LM^b \quad (2.10)$$

Table 2.1: List of the lidar metrics and the underlying data (PC = point cloud, CHM = canopy height model). CHM usually refers to 1-m resolution rasters, except for TCH where various resolutions were tested.

Lidar metric	Description	Data
MCH	Mean canopy profile height	PC
QMCH	Quadratic mean canopy profile height	PC
TCH	Mean top-of-canopy height (at variable CHM pixel resolutions), e.g., TCH <sub>5</sub> is based on 5-m pixels	CHM
RH	Relative height quantile, e.g., RH50 is the 50-percentile of heights	PC or CHM
SD	Standard deviation of heights	PC or CHM
CV	Coefficient of variation of heights (normalized SD)	PC or CHM
Skewness	Skewness of the vertical profile	PC
Shannon Index	Entropy of the vertical profile	PC
P:H ratio	Relative height of the peak in the vertical profile	PC
N <sub>AGR</sub> /N <sub>GR</sub>	Ratio of aboveground returns to ground returns	PC
N <sub>AGR</sub> /N <sub>TR</sub>	Ratio of aboveground returns to total returns	PC
FCC	Fractional canopy cover, e.g., FCC10 is the relative share of pixels higher than 10 m	CHM

If possible, such relationships were derived for plots with side lengths of 20, 33, 50, 100 and 200 m. Relationships could not be derived in cases where pixel size exceeded plot size or where the maximum likelihood estimation did not provide a parameter  $b$  different from zero. The AGB-prediction accuracy for the different power law functions was quantified as the normalized root mean square error (nRMSE) [%]. The measure was calculated as the RMSE of  $n$  AGB predictions against  $n$  observations, normalized by the mean observed AGB (Equation 2.11).

$$nRMSE = \sqrt{\frac{\sum_{i=1}^n (predAGB_i - obsAGB_i)^2}{n}} \cdot \frac{1}{mean(obsAGB)} \quad (2.11)$$

The power law parameters and additional statistics (mean, RMSE, bias,  $R^2$ , slope and intercept of linear fits between predictions and observations) for all metrics, scales and datasets (672 models) can be found in the digital supplementary materials of Knapp et al. (2018a) in Table A.2.

## 2.4 RESULTS

### 2.4.1 Forest and Lidar Simulation Results

The forest simulations could reproduce AGB succession over time for the four PFTs. An overshoot of total AGB around a forest age of 100 yr was observed (Figure A.1). The duration of the primary succession and the biomass overshoot are consistent with observations by Mascaro et al. (2012). Furthermore, the stem size distributions for

all four PFTs matched well between the model and reference data (Figure A.2). The AGB distributions of reference data and undisturbed and disturbed FORMIND runs can be found in Figure 2.3, and for the undisturbed case, the simulated distributions are in good agreement with previously reported distributions based on field data (Chave et al., 2003). At all scales the range of AGB in undisturbed simulations was smaller than the observed range of AGB in the field reference data. In the disturbance scenarios, the range of AGB values increased. At the small 20 m × 20 m scale, the real forest contained extremely high local AGB values (max. 2,022 t ha<sup>-1</sup>) caused by single large trees. Such extreme values were not reached in the simulations.

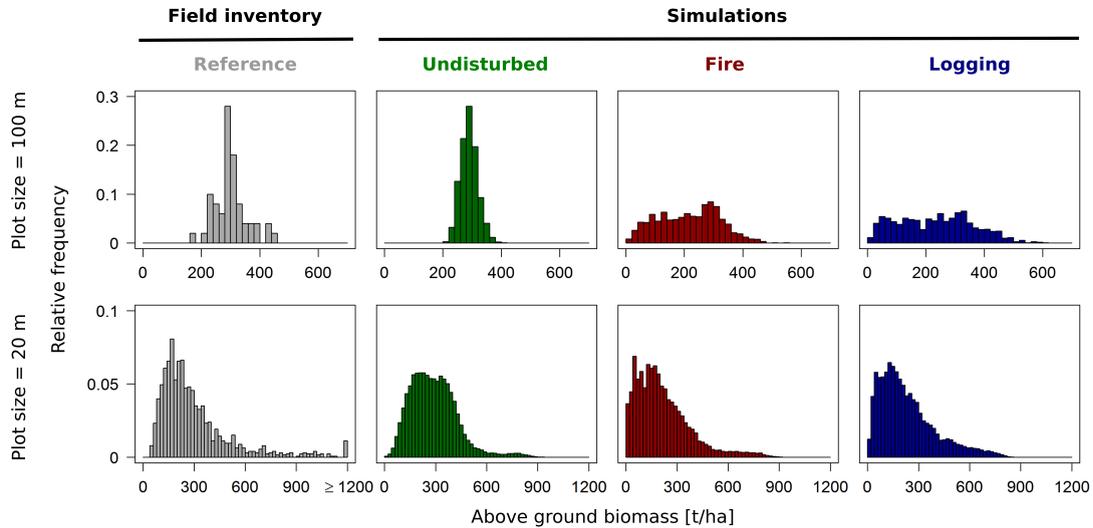


Figure 2.3: Relative test frequency distributions of aboveground biomass (AGB). Columns represent the BCI field data (50 ha) and output of FORMIND simulations from different disturbance scenarios (1,400 ha each). Rows represent different spatial resolutions. Notice the different axis scaling in each row.

Using the lidar simulation approach, synthetic lidar data were generated for the simulated forest stands. Lidar simulation outputs, such as the vertical point cloud profile (Figure 2.4) and CHMs, closely resembled their airborne equivalents. In Appendix A.1.4 we present how alternative assumptions about the tree geometry affect the simulated lidar profiles and metrics (Figures A.15 to A.18).

#### 2.4.2 Biomass Prediction from Top-of-Canopy Height

Based on the simulated stands, we analyzed 4,200 ha of forest (3.7 million trees with  $DBH \geq 3$  cm) with respect to the relationships between forest height (TCH) and biomass (AGB). We generated undisturbed (1,400 ha), fire-disturbed (1,400 ha) and logging-disturbed (1,400 ha) stands. Figure 2.5 shows the relationships observed for different plot sizes (20 to 100 m) assuming a fine resolution (pixel size = 1 m). The disturbed stands (fire and logging were pooled) cover a wider range of TCH and AGB values than the undisturbed stands. The fitted relationships for undisturbed and disturbed forest stands are similar. The scattering around the regression lines decreases with increasing plot size. If we decrease the pixel resolution from 1 to 10 m (Figure 2.6), we observe a change in the TCH-to-AGB relationship. Curves become flatter because averaging over lidar point height maxima in 10 m × 10 m pixels

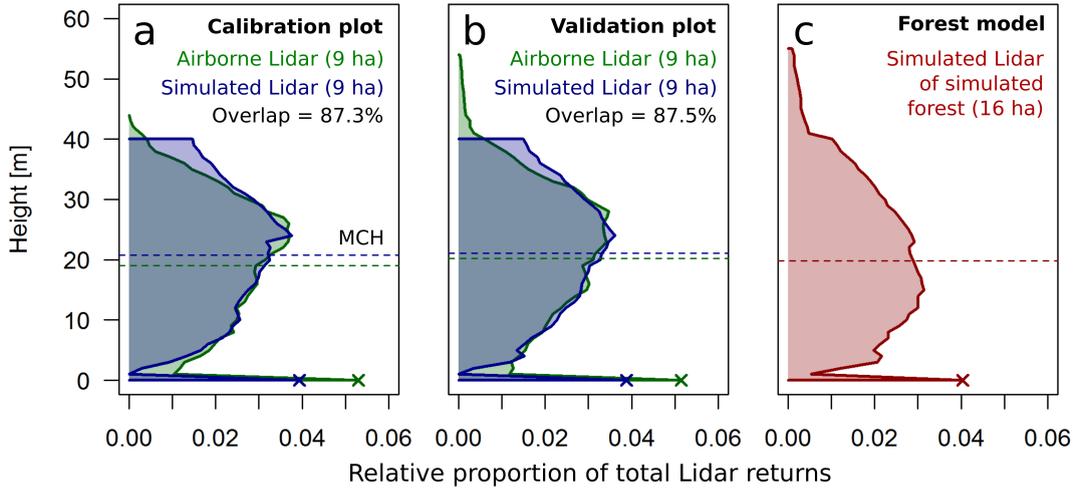


Figure 2.4: Vertical lidar profiles of a) the 9 ha in the southwestern corner of the BCI megaplot, airborne and simulated based on inventory data; b) the same for the 9 ha in the northeastern corner of the BCI megaplot; and c) the simulated lidar profile of 16 ha simulated forest in FORMIND in the old growth stage (age 500 yr). Dashed lines mark the mean canopy profile height (MCH), and 'x' symbols mark the ground return peaks.

leads to higher TCH-values than averaging over the lidar point height maxima in all  $1\text{ m} \times 1\text{ m}$  pixels. Thus, the coarser the pixel resolution is, the higher the TCH value for a given stand becomes. For the 1-m and the 10-m pixel resolution, we observe similar relations for disturbed and undisturbed forests, respectively. More extensive analyses and graphics that consider the BCI reference data and treat the different disturbance regimes separately can be found in Appendix A.2.2 (Figure A.4 and following).

The general trends were that the nRMSE of the TCH-based AGB predictions increased with decreasing plot size and with increasing pixel size (Figure 2.7). The prediction accuracy at each scale was better for the undisturbed forest dataset than for the disturbed forest dataset, indicated by generally lower nRMSE for each plot size and pixel size combination for the undisturbed forest as compared to the disturbed forest (Figure 2.7). For the disturbed dataset and large plot sizes (100 and 200 m), we observed slightly better prediction accuracies at medium pixel resolutions (5 and 10 m) than at fine pixel resolutions (1 and 2 m). The analysis shows that to achieve, a plot-level biomass estimation error  $< 10\%$ , plot sizes of  $\geq 100\text{ m}$  are required. At such plot sizes, any pixel size would be sufficient to predict AGB for undisturbed forests with the desired accuracy, but for disturbed forests, the errors exceed 10% and increase strongly at pixel sizes  $\geq 20\text{ m}$ .

### 2.4.3 Biomass Prediction based on Various Lidar Metrics

In addition to TCH, we analyzed 21 other metrics concerning their capability to predict biomass using power law equations. For this analysis, we no longer distinguished between the different disturbance regimes and pooled all forest stands. Figure 2.8 shows nRMSE values for all lidar metrics, for which it was possible to fit a power law model, at the plot scales of 100 and 20 m. From left to right, the metrics are sorted by increasing nRMSE at the 100-m plot size. The figure shows that the best ten

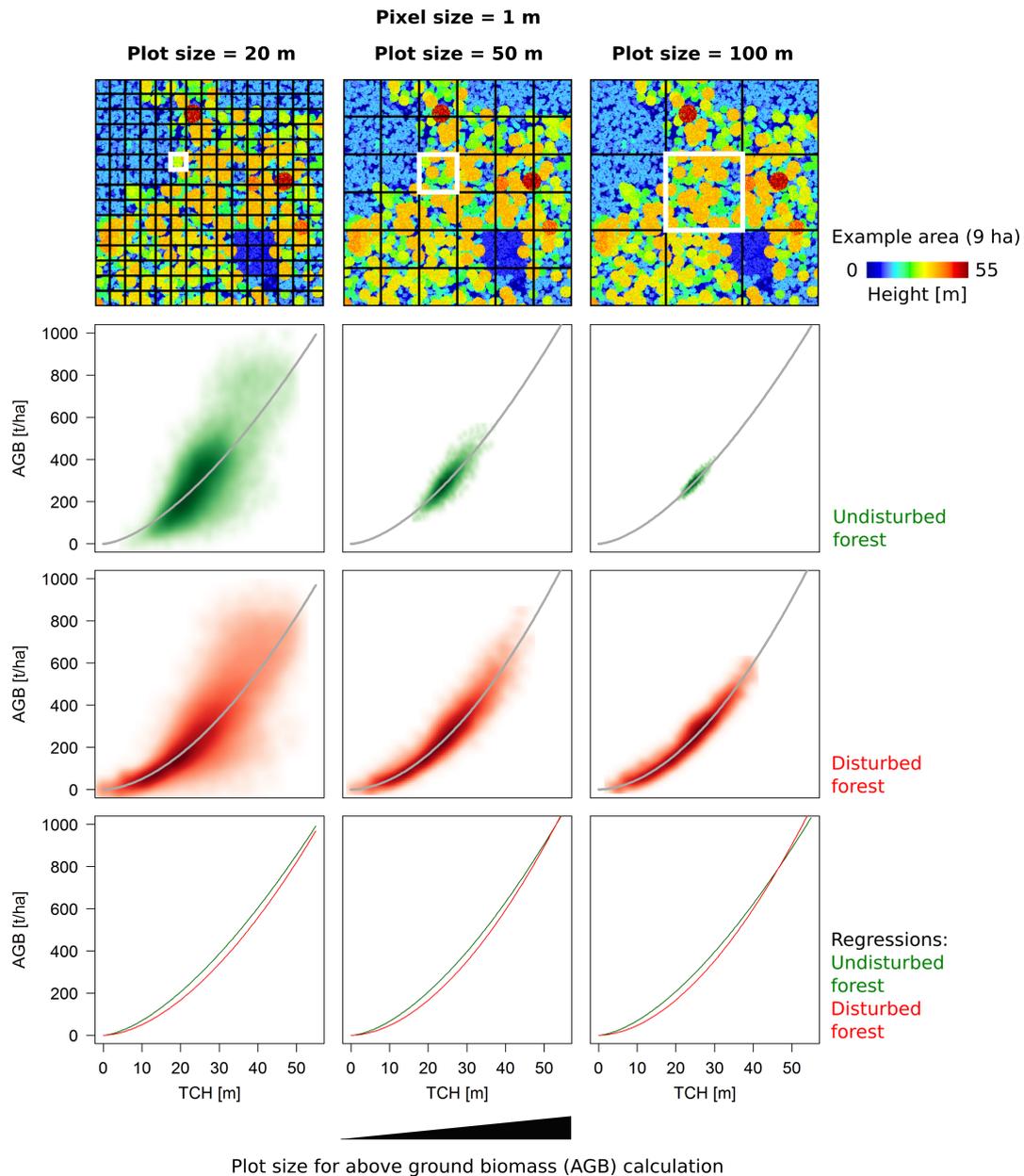


Figure 2.5: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) from 1-m pixel resolution (CHM) for different plot sizes. All data was derived from FORMIND and lidar simulations. 1) The first row demonstrates the sampling approach. Shown is a scene of 9 ha simulated forest with different stages of succession. The following rows show the TCH-to-AGB relationship with each record representing one 20-m, 50-m or 100-m plot, respectively, for 2) 1,400 ha of undisturbed simulated forest (green), 3) 1,400 ha of fire-disturbed and 1,400 ha of regularly logged simulated forest (red) and 4) the curves of the best power law fits.

metrics are all measures of forest height. Vegetation density metrics (e.g.,  $N_{AGR}/N_{GR}$  and FCC) and vertical heterogeneity metrics (e.g., SD and Shannon Index) were less accurate AGB predictors than height metrics. The best predictions at large plot scales were achieved by TCH (10 m) and TCH (5 m), whereas at small plot scales RH75, MCH, QMCH and TCH (1 m) were the most accurate predictors. We could not find any relationship between AGB and CV of height, profile skewness or P:H ratio. The

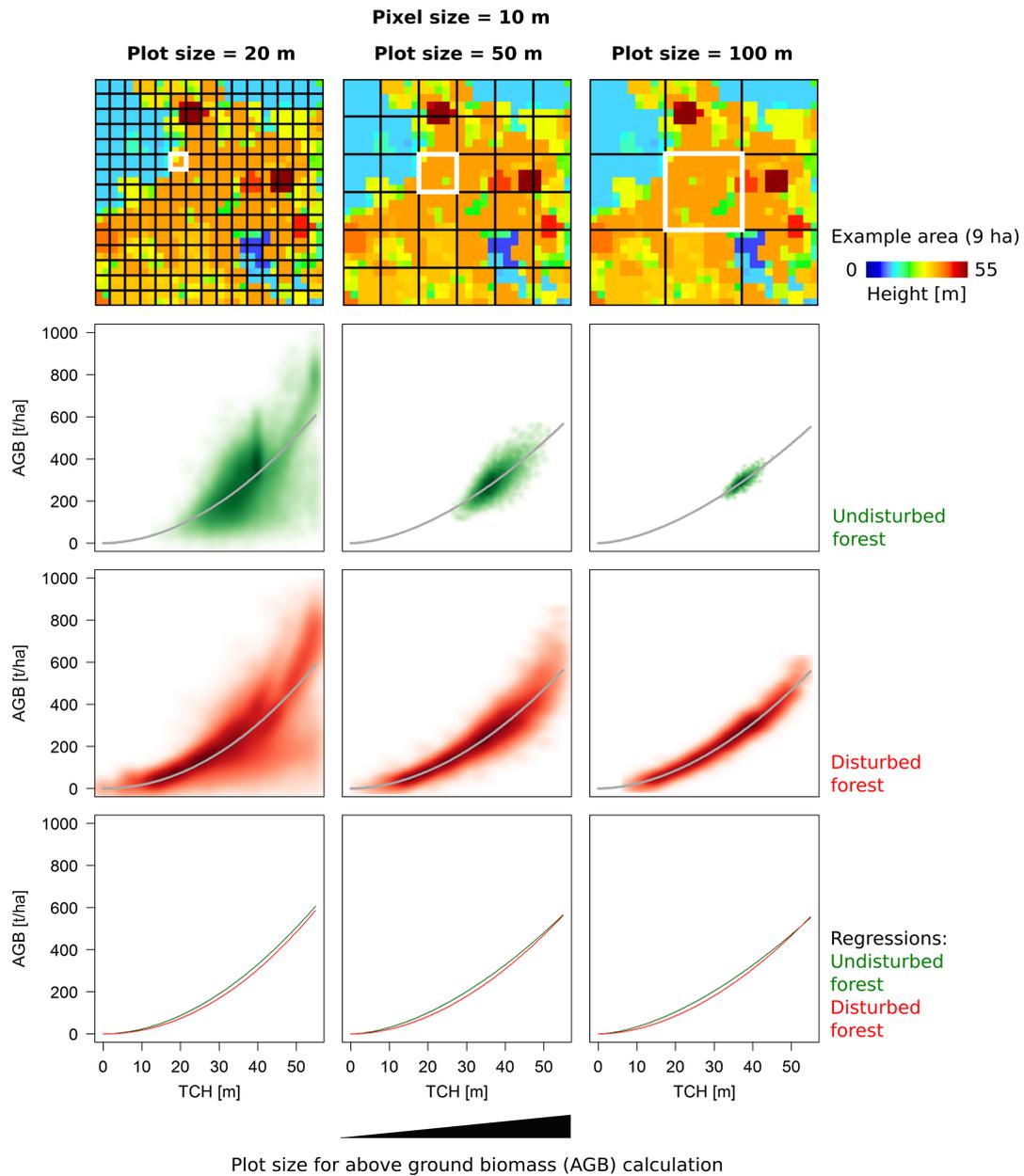


Figure 2.6: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) from 10-m pixel resolution (CHM) for different plot sizes. All data was derived from FORMIND and lidar simulations. 1) The first row demonstrates the sampling approach. Shown is a scene of 9 ha simulated forest with different stages of succession. The following rows show the TCH-to-AGB relationship with each record representing one 20-m, 50-m or 100-m plot, respectively, for 2) 1,400 ha of undisturbed simulated forest (green), 3) 1,400 ha of fire-disturbed and 1,400 ha of regularly logged simulated forest (red) and 4) the curves of the best power law fits.

Shannon Index of the profiles only showed a relationship with AGB for plot sizes  $\geq 50$  m. Scatter plots of a selection of metrics against AGB can be found in Figure A.12, nRMSE values for all metrics at all plot scales are displayed in Figure A.13 and detailed statistics and the coefficients of all fit power laws are listed in Table A.2.

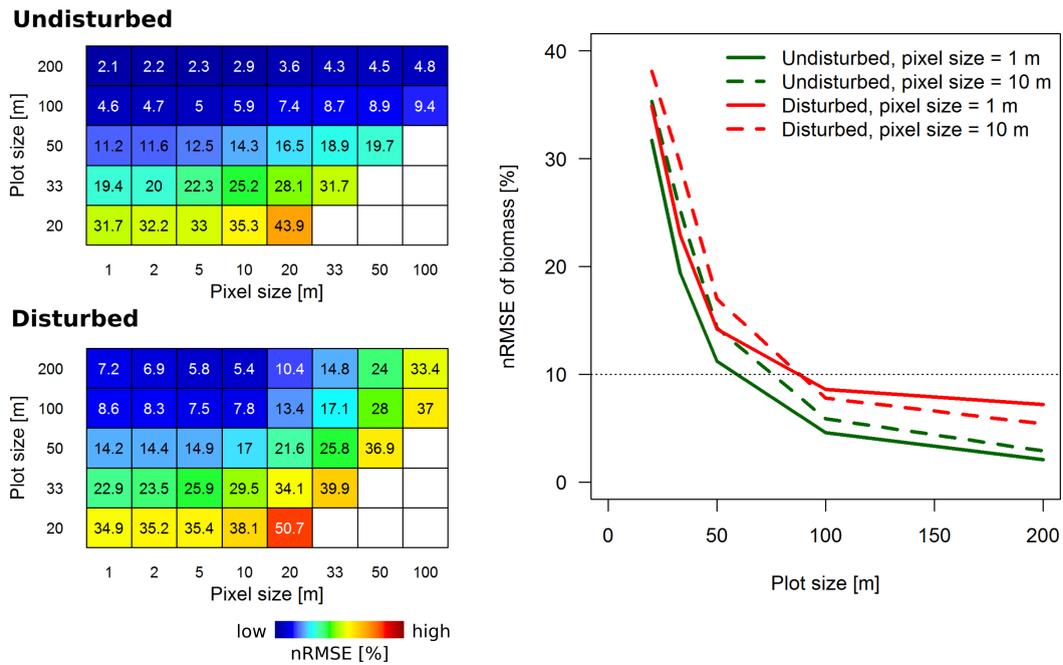


Figure 2.7: Normalized root mean square errors (nRMSE) [%] of power law models that describe the relationship between aboveground biomass (AGB) and top-of-canopy height (TCH) at different plot scales and different pixel resolutions for undisturbed and disturbed simulated forest. For pixel sizes of 1 and 10 m, the decrease in nRMSE with increasing plot size is shown on the right side.

## 2.5 DISCUSSION

This study demonstrated a new approach for simulating 3D lidar point clouds of forest stands and for investigating structural lidar metrics for their relationship with AGB of a tropical forest using forest simulations. We explored the accuracy of AGB predictions based on various lidar metrics, spatial scales and considering undisturbed and disturbed forest plots.

### 2.5.1 Lidar Simulations

Unlike other lidar simulation approaches that use detailed radiative transfer theory (Sun and Ranson, 2000; Ni-Meister et al., 2001; Kotchenova et al., 2003; Goodwin et al., 2007) or explicit 3D models of trees and ray tracing (Disney et al., 2010; Endo et al., 2012), our method requires only a minimal parameter set to efficiently compute synthetic lidar point clouds for large areas. Under simple assumptions, e.g., one DBH-to-height and DBH-to-crown-diameter allometry, a constant crown length proportion, cylindrical crowns shapes and a homogeneous leaf area density within crowns, the lidar model was able to reproduce the vertical lidar profiles of different 9-ha subplots within the 50-ha BCI megaplot to an overlap of 87%. An extinction factor  $k_{\text{NIR}}$  of approximately 0.2 was suggested by empirical measurements (Jones, 2013) and theoretical considerations (Campbell and Norman, 2012; Tang et al., 2012) and could be confirmed by our inverse modeling tests.

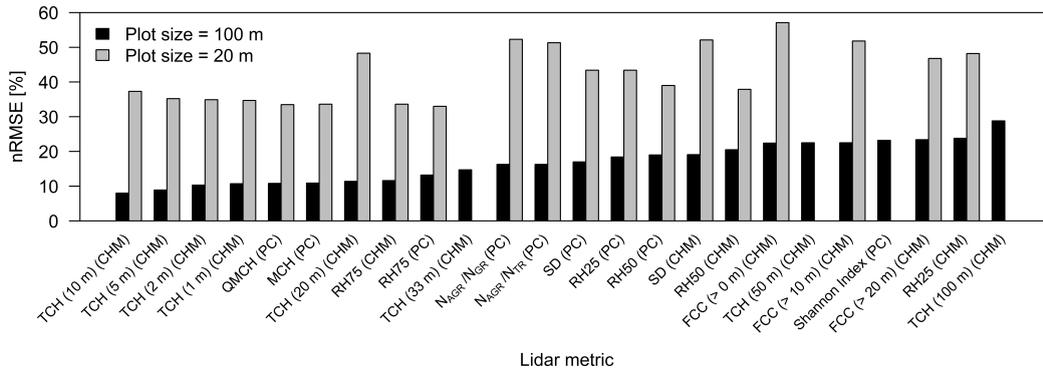


Figure 2.8: Normalized root mean square errors (nRMSE) [%] of power law models that describe the relationship between aboveground biomass (AGB) and various lidar metrics (for explanations of the abbreviations, please refer to the main text and Table 2.1) at plot scales of 100 and 20 m, respectively. From left to right, the metrics are sorted by increasing nRMSE at the 100-m plot size. Whether certain metrics were derived from point clouds (PC) or from canopy-height-models (CHM) is indicated in brackets. This analysis was based on pooled (undisturbed and disturbed) simulated forest data and lidar simulations. Missing bars indicate that no power law model could be fit at the 20-m plot size.

Airborne and simulated profiles for the 9-ha subplots matched well in general. They diverged most in the upper canopy, where the DBH-to-height allometry led to an overestimation of high trees. Frequencies of ground returns of simulated profiles were approximately 25% lower than for the airborne data, which could be adjusted by choosing another lidar return probability  $P_0$  for ground voxels. Because the exact size of the ground return peak does not affect most of the lidar metrics, we did not treat ground voxels differently than canopy voxels in this study. It should also be noted that simulated lidar profiles (inventory- and FORMIND-based) contain only returns from trees and ground. Non-woody vegetation such as shrubs and lianas may contribute to the airborne lidar profiles, particularly near ground, whereas they are absent in the simulations.

## 2.5.2 Biomass Prediction from Lidar Height

For the simulated BCI lidar dataset, TCH at various pixel resolutions performed better than any other lidar metric for biomass predictions. The lowest AGB prediction errors (< 10%) were found for large mapping units (plot sizes of 100 and 200 m) with TCH derived from CHMs with pixel sizes of 5 to 20 m. For the smaller mapping units of 50 m, 33 m and 20 m, the minimal achievable errors from any metric were 15%, 23% and 33%, respectively. At those scales, the high pixel resolution TCH, RH75 or point-cloud-based MCH and QMCH led to slightly smaller errors than TCH of medium pixel resolution. The finding that medium pixel resolution CHMs are sufficient to make highly accurate AGB predictions at the 1-ha scale is encouraging for spaceborne biomass mapping efforts on the global scale. The generation of high-resolution information (e.g., pixel size of 1 m) requires airborne laser scanning campaigns, whereas medium resolutions can be derived from satellites. The synthetic aperture radar satellite system TanDEM-X can provide forest heights closely correlated to TCH at a resolution of 10 m (referred to as H100 in the radar literature; Kugler et al., 2014;

Lee and Fatoyinbo, 2015). Future sensors, such as GEDI<sup>1</sup> and Tandem-L<sup>2</sup>, will provide data of similar horizontal resolution (20 to 50 m) and improved vertical resolution. Thus, TCH as well as MCH and RH75 of the vertical profiles are promising metrics for estimating AGB using these sensors. The analysis also showed that sensors that only provide maximum height at the coarse resolution of 100 m lead to AGB estimation errors of > 25%. It appears highly plausible that CHMs with pixels sizes around 10 m corresponding to the dimensions of the objects of interest, namely crowns of medium to large trees, which contribute most to the total AGB, are a good data source for AGB inference. High-resolution data such as 1-m pixel CHMs or the full point cloud have the advantage of providing detailed information on crown architecture and small gaps, but this information might only be additional noise in the signal for stand level AGB and may not be necessary for large-scale mapping.

### 2.5.3 *The Role of Structure Metrics*

Metrics of vertical heterogeneity (e.g., standard deviation or Shannon Index) and vegetation density (e.g.,  $N_{AGR}/N_{GR}$  or FCC) showed weaker relationships with AGB than most of the height metrics. Hence, these metrics might not be the optimal choice as single AGB predictors. However, considering vegetation structure in addition to mean height could potentially improve AGB estimations. Several approaches have been suggested to improve power-law-based lidar-to-AGB models by considering additional predictors. These predictors include horizontal and vertical structure indices (Tello et al., 2015) and texture metrics of the CHM (Abdullahi et al., 2016). Finally, when thinking beyond AGB stock prediction and towards the study of forest dynamics and disturbances based on remote sensing, structural metrics may become very important. The Shannon Index of the lidar profile has been previously associated with productivity and mortality (Stark et al., 2012), and gap fraction and size distribution may provide information about disturbances (Lobo and Dalling, 2014).

### 2.5.4 *Prediction Errors*

For all tested lidar metrics, we observed the tendency for the prediction errors to decrease with increasing plot scale. This pattern has been reported and quantified previously for MCH (Asner et al., 2010; Mascaro et al., 2011b), QMCH (Chen et al., 2016) and TCH (Köhler and Huth, 2010; Asner and Mascaro, 2014) and in general for the situation in which remote sensing footprints and ground plot extents do not fully match (Réjou-Méchain et al., 2014). In our analysis, the spatial locations and extents of ground plots and remote sensing data matched perfectly, because they were based on simulations. Also there was no displacement of crowns from stem locations. Thus, our dataset is free of geolocation errors and the observed residuals in the lidar-to-AGB relationships can be attributed to the following sources of uncertainty: 1) the highly clumped biomass distribution on the ground, i.e., the majority of biomass is localized in tree trunks at specific positions with empty space in between, whereas remote sensing signals capture the tree crowns, which are spread around the trunk positions; 2) edge effects of overhanging tree crowns with trunk positions and thus biomass being located outside the focal plot area; 3) the general variability among trees with

1 <http://science.nasa.gov/missions/gedi/>

2 <https://www.tandem-l.de/>

respect to their geometries and wood densities; and 4) the undergrowth vegetation that is obscured by the upper canopy and not detected by the remote sensing sensor. The error caused by 1) should decrease with increasing plot size due to the decrease in biomass variability (Figure 2.3) and the decreasing influence of single large trees. The error caused by 2) should decrease with increasing core area to edge length ratio. The error caused by 3) should decrease because differences at the individual tree level average out with increasing plot size. Only errors caused by 4) can be expected to be scale-independent. Using a crown-distributed instead of a stem-localized biomass distribution as ground truth has been shown to reduce estimation errors (Mascaro et al., 2011b). However, the actual biomass distribution in a forest is expected to be closer to being stem-localized than (uniformly) crown-distributed. Thus, reducing errors by assuming crown-distributed biomass does not necessarily lead to more accurate biomass maps. Our modeling approach may allow future studies to gain a closer look at the contributions of the separate error sources by switching them off one at a time. Different lidar metrics showed different changes in errors across scales: e.g., in moving from large to small plots, the errors of TCH<sub>20</sub>, TCH<sub>33</sub> and the Shannon Index increased much faster than for other metrics with similar errors at the 200-m scale (Figure 2.8 and A.13). For the Shannon Index, the relationship with AGB was entirely lost at scales smaller than 50 m.

#### 2.5.5 *Linking Remote Sensing with Dynamic Forest Models*

Despite the great potential of the proposed approach, relatively few studies have linked remote sensing and forest modeling. Applications include model initialization (Ranson et al., 2001; Hurtt et al., 2004), model parameterization (Falkowski et al., 2010), remote sensing calibration (Köhler and Huth, 2010; Palace et al., 2015), error quantification (Hurtt et al., 2010; Frazer et al., 2011) and the understanding of large-scale ecosystem patterns and processes (Shugart et al., 2015). Our study is the first to demonstrate how remote sensing simulations combined with a dynamic forest model can provide remote sensing metrics over the full range of disturbance-induced successional stages, which is particularly useful for tropical forests where available field data is limited. The lidar-to-AGB relationships can differ between disturbance types because one type (e.g., fire) might cause mosaics of surviving trees and bare ground, whereas another type (e.g., selective logging) might cause a height degradation throughout the entire study area. Horizontal heterogeneities, such as those caused by fires, are particularly problematic when lidar metrics are aggregated over larger areas. Thus, the disturbance regime of a region and the presence of the described phenomena should be taken into account when deciding which metric and resolution to choose for biomass mapping. Modeling can be one way to explore these effects in greater detail.

An important condition for combining a forest model and remote sensing is the structural realism of the model in the relevant aspects. Overall, our model was able to reproduce forest attributes and literature values well. Previous studies on BCI that linked AGB at the 1-ha scale to MCH derived from airborne lidar scans reported RMSE values of 17 t<sub>Carbon</sub> ha<sup>-1</sup> (Mascaro et al., 2011a) and 28.9 t<sub>AGB</sub> ha<sup>-1</sup> (Meyer et al., 2013) in agreement with the value of 27.1 t<sub>AGB</sub> ha<sup>-1</sup> we obtained for the pooled simulated dataset (Table A.2). A noteworthy deviation between the simulation data and reference data was that for comparable AGB values the simulated TCH was higher than the airborne TCH, particularly at the upper end of the AGB and TCH ranges

(described in detail in Appendix A.1.3). We believe that this deviation was primarily caused by the simple tree geometries used in the forest model. Using only one general DBH-to-height allometry for all trees might be suboptimal if the aim is to reproduce the natural height heterogeneity of the upper canopy at all scales. In our simulations, too many trees reached the maximum possible height of 55 m, which is an exceptional height on BCI observed for only one tree in the airborne lidar CHM. Hurtt et al. (2004) encountered a similar problem with large trees. In their case, model-derived canopy heights were restrained to a maximum, whereas observed lidar heights exceeded that limit. Therefore, one potential improvement for future model parameterizations would be to consider asymptotic instead of power law DBH-to-height allometries and allow for a certain plasticity of modeled heights and crown diameters. The sensitivity analysis about model assumptions showed that the alternative scenario using an asymptotic tree height allometry led to slight increases in  $R^2$  and decreases in nRMSE of the stand height to biomass relationship (Figures A.16 to A.18). Recent advances in individual tree delineation from airborne lidar (Duncanson et al., 2014; Ferraz et al., 2016) and terrestrial laser scanning (Raumonen et al., 2013) have the potential to improve our understanding of tree allometries and the structural realism of forest models. When models are able to reproduce observed patterns in the relationship between remote sensing metrics and static biomass stocks, we can move forward using the presented methodology to explore dynamic changes of biomass stocks.

## 2.6 CONCLUSION

This study introduced a novel approach for coupling remote sensing simulations with a dynamic forest model to derive structure-to-biomass relationships for a tropical forest, including disturbed stands. The lidar model was validated successfully with airborne and census reference data from Barro Colorado Island. The model proved its capacity for efficient and realistic lidar point cloud simulations for large, simulated forest stands. Virtual forest inventory datasets were generated with a forest model and sampled with the lidar simulation model. The results provide a comprehensive overview of biomass estimation errors for a wide range of lidar metrics and spatial scales and may guide decisions on which metric to choose for a given remote sensing data structure (e.g., point clouds, vertical profiles, canopy height models). It was found that height-to-biomass relationships were similar for undisturbed and disturbed forest, but errors were larger for the latter. Furthermore, we found that top-of-canopy height was an accurate biomass predictor even if pixel resolutions were only 10 to 20 m. Such resolutions could be derived at large scale from spaceborne sensors.

## 2.7 ACKNOWLEDGMENTS

We thank J. Dalling for providing the lidar data and the Smithsonian Tropical Research Institute for providing the census data for BCI. The BCI forest dynamics research project was founded by S.P. Hubbell and R.B. Foster and is now managed by R. Condit, S. Lao, and R. Perez under the Center for Tropical Forest Science and the Smithsonian Tropical Research in Panama. Numerous organizations have provided funding, principally the U.S. National Science Foundation, and hundreds of field workers have contributed. This study was conducted with funding by the German Federal Ministry for Economic Affairs and Energy (BMWi) under the funding reference

50EE1416. RF and AH were supported by the HGF-Helmholtz Alliance "Remote Sensing and Earth System Dynamics". We thank two reviewers for their constructive comments on our paper.

## MODEL-ASSISTED ESTIMATION OF TROPICAL FOREST BIOMASS CHANGE: A COMPARISON OF APPROACHES

---

### 3.1 ABSTRACT

Monitoring of changes in forest biomass requires accurate transfer functions between remote sensing derived changes in canopy height ( $\Delta H$ ) and the actual changes in aboveground biomass ( $\Delta AGB$ ). Different approaches can be used to accomplish this task: Direct approaches link  $\Delta H$  directly to  $\Delta AGB$ . Indirect approaches are based on deriving AGB stock estimates for two points in time and calculating the difference. In some studies direct approaches led to more accurate estimations, while in others indirect approaches led to more accurate estimations. It is unknown how each approach performs under different conditions and over the full range of possible changes. Here, we used a forest model (FORMIND) to generate a large dataset (> 28,000 ha) of natural and disturbed forest stands over time. Remote sensing of forest height was simulated on these stands to derive canopy height models for each time step. Three approaches for estimating  $\Delta AGB$  were compared: (1) the direct approach, (2) the indirect approach and (3) an enhanced direct approach (dir+tex), using  $\Delta H$  in combination with canopy texture. Total prediction accuracies of the three approaches measured as root mean squared errors were  $RMSE_{direct} = 18.7 \text{ t ha}^{-1}$ ,  $RMSE_{indirect} = 12.6 \text{ t ha}^{-1}$  and  $RMSE_{dir+tex} = 12.4 \text{ t ha}^{-1}$ . Further analyses revealed height-dependent biases in the  $\Delta AGB$  estimates of the direct approach, which did not occur with the other approaches. Finally, the three approaches were applied on radar-derived (TanDEM-X) canopy height changes on Barro Colorado Island (Panama). The study demonstrates the potential of forest modeling for improving the interpretation of changes observed in remote sensing data and for comparing different methodologies.

### 3.2 INTRODUCTION

Forests play a crucial role in the global carbon budget. Carbon stocks of forests worldwide are estimated to be around 350 to 600 Gt (Houghton et al., 2009; IPCC, 2007; Pan et al., 2011). Deforestation and forest degradation are estimated to cause an annual change of 1.1 Gt of carbon. However, these change estimations include large uncertainties, because they can only be derived indirectly from estimates of other carbon stocks and fluxes (Houghton et al., 2009). Passive optical remote sensing sensors are successfully used for monitoring the forest extent (Hansen et al., 2013). However, to quantify the aboveground biomass and thus carbon stocks of forests, passive optical sensors suffer from saturation and can be used only for forests with relatively low biomass (Zolkos et al., 2013; Lu et al., 2014). Active sensors such as light detection and ranging (lidar; Lefsky et al., 2002a; Wulder et al., 2012) and synthetic aperture radar (SAR; Treuhaft et al., 2015) enable measurements of the canopy height structure of forests, which can be used to derive information about the standing aboveground biomass (AGB). The height-to-biomass relationship at the stand level for area-based biomass estimations is a topic of many recent studies in different forest ecosystems and geographical regions. A multitude of remote sensing metrics (Lefsky et al., 2002b; Lu et al., 2014), spatial scales (Chen et al., 2016; Knapp et al., 2018a) and modeling approaches (Fassnacht et al., 2014) have been tested and compared in this context. It was found that often a single metric that captures canopy height can provide accurate biomass estimations based on equations derived from regression analysis (Drake et al., 2002). Among several possible metrics that describe average canopy height (e.g., height quantiles, mean profile height), the so called mean top-of-canopy

height (TCH; Asner and Mascaro, 2014) has become one of the most frequently used metrics (Marvin et al., 2014; Coomes et al., 2017; Getzin et al., 2017).

Country-wide (Asner et al., 2014) and even biome-wide maps (Saatchi et al., 2011; Baccini et al., 2012) have been published that build on the height-to-biomass relationship. Combining a baseline map of biomass stocks (Saatchi et al., 2011) with maps of forest cover change (Hansen et al., 2013) the carbon emissions due to tropical deforestation between 2000 and 2005 have been estimated to be  $0.81 \text{ Gt yr}^{-1}$  (Harris et al., 2012). Such an area-change-based estimation does not, however, account for the dynamics of degradation, disturbances and recovery within forested areas.

For the monitoring of changes in forest biomass, two basically different methods have been suggested in recent years, often referred to as the direct and the indirect approach (Cao et al., 2016). Both try to estimate the change in aboveground biomass ( $\Delta\text{AGB}$ ) or carbon of a forest during a certain time interval, based on two remote sensing data acquisitions, one before and one after the period of interest. In the indirect method the remotely sensed height information at the beginning and at the end of the time interval is used to estimate the standing biomass stocks before and after the interval. The biomass change is then calculated as the difference between the stocks at both points in time. Hence, the indirect approach relies on the height-to-biomass relationship. In contrast, the direct method tries to link the height change ( $\Delta\text{H}$ ) to the biomass change. Thus, first height change is calculated as the difference between height at the end and height at the start of the interval. Then, a  $\Delta\text{H}$ -to- $\Delta\text{AGB}$  relationship is used to estimate the change in biomass from the observed height change.

The direct and the indirect biomass change estimation have been applied in several studies (sometimes in direct comparison to each other), covering different forest types, from boreal (Bollandsås et al., 2013), temperate (Hudak et al., 2012; Zhao et al., 2018), subtropical (Cao et al., 2016) and tropical regions (Dubayah et al., 2010; Meyer et al., 2013). In some studies, the direct approach led to more accurate estimations, while in other studies the indirect approach led to more accurate estimations. It remains unclear, which approach performs better under which conditions. The studies dealing with the topic had a rather regional extent and were based on a limited set of forest inventory plots. Plots are usually small in size which results in large variation in the observed relationships. With new spaceborne lidar and SAR sensors becoming operational (e.g., TerraSAR-X-Add-on for Digital Elevation Measurements (TanDEM-X), Sentinel 1, Global Ecosystem Dynamics Investigation Lidar (GEDI), BIOMASS, Tandem-L) that enable large-scale measurements of forest structure, it is necessary to establish standardized approaches for how to estimate biomass changes from remote sensing.

There are advantages and disadvantages to both approaches. The advantage of the indirect approach is that it makes use of the well-studied H-to-AGB relationship which is applied at the beginning and end of the time interval. The direct approach requires the establishment of a  $\Delta\text{H}$ -to- $\Delta\text{AGB}$  relationship. The advantage of this approach is that with an established  $\Delta\text{H}$ -to- $\Delta\text{AGB}$  relationship, measurements of absolute forest heights are no longer required and net changes in canopy surface height are sufficient to estimate changes in AGB. This holds the potential to map  $\Delta\text{AGB}$  from changes observed in digital surface models (DSM; e.g., Shuttle Radar Topography Mission (SRTM) and TanDEM-X products), which represent only the canopy surface height. Knowing the terrain height (digital terrain model, DTM) underneath for deriving net vegetation height (canopy height model, CHM) would not be required (Solberg

et al., 2014; Solberg et al., 2015; Puliti et al., 2017). Such situations can arise if a certain technique is only capable of generating surface information (e.g., photogrammetry, SAR interferometry), or if the terrain information is of low precision due to dense vegetation or on hill slopes (e.g., large footprint lidar, polarimetric SAR interferometry).

Despite the lack of absolute vegetation height information, the DSMs do, however, contain valuable information on canopy structure. Crown and gap size distributions vary among forest stands, depending on their age and height structure and are reflected in the DSM texture. Textural information from optical remote sensing imagery has been used previously to map habitat heterogeneity (Tuanmu and Jetz, 2015) and forest biomass (Couteron et al., 2006; Proisy et al., 2012; Singh et al., 2015) and textural information from DSMs and CHMs has been used successfully in forest classification (Kennel et al., 2013) and estimation of average tree height, basal area and stem volume (Bohlin et al., 2012; Abdullahi et al., 2016). Thus, analyzing the DSM texture, i.e., local height variability among neighboring pixels, can provide a useful set of metrics to characterize a forest stand in the absence of absolute height information. This textural information could be used to enhance biomass change estimations in situations where a simple direct approach may fail. Each single texture metric alone may only show a weak relationship with canopy height, but using an ensemble of metrics and a machine learning algorithm (e.g., random forest), we expect to improve estimations of biomass change from DSM change compared to the simple direct approach. Comprehensive analyses of the relations between remote sensing metrics and ground-based metrics require a large number of ground-truth plots. The measurement effort becomes even larger when the goal is to analyze changes over time, which requires synchronized remote sensing and field campaigns at regular time intervals, ideally covering the full successional range of the forest. Biomass losses due to disturbances happen stochastically and even the most expansive field campaigns can hardly provide representative samples for them. For that reason, forest models have gained popularity to analyze simulated forests and explore relationships between remote sensing and ground-based metrics (Shugart et al., 2015).

Forest models have a long tradition in ecology and forestry. They have traditionally been used to understand processes in forest ecosystems and to test system behavior under the influence of management scenarios or disturbance events (Shugart, 2003). Starting from simple forest yield tables (Moser Jr, 1980), the incorporation of more ecological processes led to advanced types of forest models, including forest gap models (Botkin et al., 1972), and finally to the development of individual-based forest models (Huston et al., 1988). Forest ecosystem functions and forest structure emerge from individual trees and their interactions. Therefore, gap models are designed to work at local scale and represent forest structure and dynamics at a detailed level (Bugmann, 2001; Shugart et al., 2018). An architecture based on individual trees allows the modeling of structurally realistic concepts that are based on field measurements at different organizational levels (Grimm and Berger, 2016) – which was the key driver for the successful applications of gap models. As the impact of disturbance events or management in forests is mainly quantified at the tree level, an individual-based structure is a huge advantage for simulating disturbances.

The individual-based approach facilitates the linkage of forest models to remote sensing data. Previous applications have covered different aspects including mapping of biomass (Hurtt et al., 2004; Rödig et al., 2017b) and productivity (Rödig et al., 2018), understanding of height biomass relationships (Köhler and Huth, 2010; Palace et al., 2015), error quantification (Frazer et al., 2011) and monitoring of changes in forest

structure (Cazcarra-Bes et al., 2017). The individual-based forest model FORMIND (Fischer et al., 2016) has been used to analyze the relationship between many different lidar metrics and AGB for a tropical rainforest across scales and disturbance states (Knapp et al., 2018a). The simulated disturbances allowed to expand the range of stand structures and successional stages far beyond the range covered by the available inventory data of an old-growth forest.

In this study, we used FORMIND simulations to analyze and compare different ways of how to estimate biomass changes from height changes. We simulated time series of tropical forest stands (synthetic ground-truth) and synthetic remote sensing data. Tropical forests are of particular interest in the context of mapping forest biomass changes, due to their large share in the global vegetation carbon pool and their high deforestation rates. The methods developed on the basis of forest simulations were finally tested with TanDEM-X data acquired at two points in time.

The main research questions were as follows: (1) How can aboveground biomass change be estimated from canopy height change, if the goal is to cover the full range of tropical forest succession including disturbed forests? (2) Under which conditions do the different approaches work best? (3) Can canopy texture information improve estimates of biomass change in the absence of canopy height information?

### 3.3 MATERIALS AND METHODS

#### 3.3.1 *Study Area*

The study focusses on Barro Colorado Island (BCI), Panama (9.15° N, 79.85° W), a semi-deciduous tropical lowland rainforest site. Average daily maximum and minimum temperatures are 30.8 and 23.4 °C and the annual precipitation sum is 2,600 mm, with a dry season from January to April (Condit et al., 2001). BCI hosts a 50-ha (1,000 m × 500 m) rainforest observation plot, which has been continuously monitored for more than three decades, with every tree with a diameter at breast height (DBH) ≥ 1 cm being recorded at 5-yr intervals (Condit, 1998; Condit et al., 2012; Hubbell et al., 1999; Hubbell et al., 2005). This inventory dataset with its outstanding spatial and temporal dimensions, along with the large amount of research conducted around it, provides a rich source of information for forest model parameterization (Kazmierczak et al., 2014; Knapp et al., 2018a) as well as ground-truthing for remote sensing studies (Mascaro et al., 2011a; Meyer et al., 2013; Lobo and Dalling, 2014). In this study, a forest model parameterized with the census data was used and the census data of 2010 and 2015 (in combination with the same allometries as used in the forest model, Equations 3.1 to 3.3) served for ground-truthing of satellite-derived estimations of biomass change.

#### 3.3.2 *Forest Model Description*

The forest model FORMIND simulates the dynamic processes of establishment, growth, competition (for light and space) and mortality at the individual tree level. Species with similar ecological traits and growth characteristics are grouped into plant functional types (PFT). Biomass growth is mainly driven by light. Large trees receive most of the incoming radiation and shade smaller trees. The resulting biomass growth of each tree is determined by a physiology-based carbon balance including photosynthesis and

respiration. An increase in tree biomass results in stem diameter growth and through the use of allometric relationships also in a growth of tree height, stem volume and leaf area. A detailed description of the model processes can be found in Fischer et al. (2016). FORMIND was already fully parameterized for the study site (Knapp et al., 2018a). All species present in the BCI plot were grouped into four PFTs, according to stem diameter increment rates (slow, fast) and maximum tree height (small, tall). Allometric equations from the literature (Bohlman and O'Brien, 2006) were used to describe tree geometries (Knapp et al., 2018a). Tree height  $H_{\text{tree}}$  [m] and crown diameter  $CD_{\text{tree}}$  [m] are modelled as functions of  $DBH_{\text{tree}}$  [m] using Equations 3.1 and 3.2.

$$H_{\text{tree}} = 43.4 \cdot DBH_{\text{tree}}^{0.6} \quad (3.1)$$

$$CD_{\text{tree}} = 18.2 \cdot DBH_{\text{tree}}^{0.68} \quad (3.2)$$

PFTs can reach different maximal heights (20, 20, 40 and 55 m).  $AGB_{\text{tree}}$  is calculated from Equation 3.3, where  $F$  is the stem form factor, which accounts for the deviation from a cylindrical shape,  $\rho$  is the wood density [ $t_{\text{ODM}} \text{ m}^{-3}$ ] and  $\sigma$  is the stem-to-total AGB ratio of the tree. Parameter values are given in Knapp et al. (2018a).

$$AGB_{\text{tree}} = \pi \cdot \left(\frac{DBH_{\text{tree}}}{2}\right)^2 \cdot H_{\text{tree}} \cdot F \cdot \frac{\rho}{\sigma} \quad (3.3)$$

The parameterization has been shown to reproduce several patterns observed in the field (AGB, basal area, stem numbers and stem size distributions of the total plot and per PFT). Additionally, it has been used in combination with lidar simulations and could reproduce patterns of airborne lidar data (Knapp et al., 2018a). The temporal development of AGB and canopy height during primary succession is shown in Appendix B Figure B.1.

### 3.3.3 Simulations

FORMIND was used to simulate the development of a 16-ha area (400 m  $\times$  400 m) of the BCI forest for a long period (2,000 years). The first 200 years of spin-up were discarded from the analysis. In each year a full inventory table containing all trees with  $DBH \geq 3$  cm was stored and a virtual lidar scan of the area was sampled using the lidar simulation approach described by Knapp et al. (2018a) (Figure 3.1). In FORMIND, we simulated spatially explicit disturbances to frequently clear parts of the area. Disturbances were set to reoccur at random times and places within the area with an average time interval of 25 years and an average affected area of 50% (Fischer et al., 2016). These settings created a spatially heterogeneous mosaic of different forest regrowth stages. The scenario was designed to produce a dataset with maximum possible structural heterogeneity to cover the full range of possible biomass and height changes with the simulations.

### 3.3.4 Biomass, Height and Change Calculations

The simulated lidar data were processed to obtain a metric called  $TCH_{10}$  (mean top-of-canopy height with 10-m pixel resolution) at 1-ha scale. This metric has been

shown to yield a height-to-biomass relationship which produces accurate biomass stock estimates (RMSE = 19.8 t ha<sup>-1</sup>; nRMSE = 8%; R<sup>2</sup> = 0.96; Knapp et al., 2018a) To obtain TCH<sub>10</sub> from the simulated lidar point cloud a canopy height model (CHM) was produced by taking the height of the highest lidar return falling into each 10 m × 10 m area as pixel value. Next, all pixel values were averaged to obtain one TCH<sub>10</sub> value for each hectare. Aboveground biomass (AGB) of the simulated forest stands was aggregated at the 1-ha scale. Changes in observed AGB (called  $\Delta\text{AGB}_{\text{observed}}$ ) and changes in TCH<sub>10</sub> (called  $\Delta\text{TCH}_{10}$ ) over each 5-yr interval were calculated for each simulated hectare (Figure 3.1, see Appendix B for equivalent analysis for 10- and 25-yr intervals). The total number of analyzed 1-ha forest stands was 28,736.

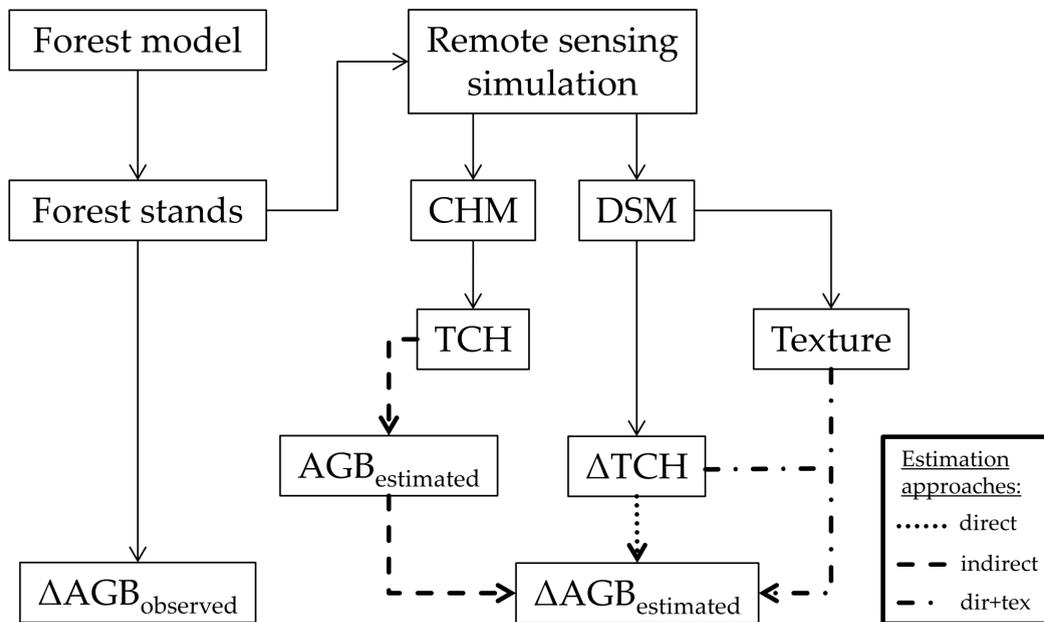


Figure 3.1: Technical flowchart of the analysis of the simulated data. The analysis of empirical data was conducted in the same way, just with different data sources: inventory data instead of forest model and TanDEM-X data instead of remote sensing simulation (section 3.3.8). Abbreviations: AGB = aboveground biomass, CHM = canopy height model, DSM = digital surface model, TCH = mean top-of-canopy height. Note that we distinguish between CHM and DSM. In the simulations CHM and DSM are the same, because the simulated forest stands are on flat terrain. However, CHMs are only required to derive TCH (indirect approach), while for  $\Delta\text{TCH}$  and texture calculations DSMs are sufficient (direct and dir+tex approach).

### 3.3.5 Texture Calculations

A variety of texture metrics were derived from the simulated CHM rasters (10-m resolution) to capture vegetation surface variability. We used a 3 × 3 pixel moving window (30 m × 30 m), only considering the direct neighbors for each pixel, and then averaged those textures over each hectare. First order texture metrics are independent of the spatial arrangement of pixel values within the moving window. The following first order texture metrics were derived (using the raster package in R (R Development Core Team, 2014): standard deviation, skewness, kurtosis, slope, topographic

ruggedness index (TRI), topographic position index (TPI) and roughness of the nine pixel values in the moving window. Second order texture metrics are based on the grey-level co-occurrence matrix (GLCM), i.e., they consider how often certain pixel values occur next to each other (Haralick et al., 1973). To obtain a limited set of discrete grey-levels for GLCM texture calculations, the original CHM values were rounded (5-m classes). The following GLCM metrics were calculated (using the `gcm` package in R): homogeneity, contrast, dissimilarity, entropy and angular second moment (ASM). Directionality in the co-occurrence pattern was not considered (circular version). It is important to note that all texture metrics are independent of the absolute pixel values. The texture metrics depend only on the differences in values between focal pixel and neighbor pixels. Thus, they can be derived from a CHM (terrain-normalized) or a DSM (not terrain-normalized) alike (assuming that the contribution of terrain variability is much smaller than the contribution of canopy surface variability to DSM variability at the given 30 m × 30 m moving window scale).

### 3.3.6 Biomass Change Estimation

Three different approaches for  $\Delta\text{AGB}$  estimation were tested with the simulated dataset: (1) the direct approach, (2) the indirect approach and (3) an enhanced direct approach involving canopy texture information (`dir+tex`; Figure 3.2).

#### 3.3.6.1 Direct Approach

A linear regression model with slope  $m$  and intercept  $n$  was fit between simulation-derived  $\Delta\text{AGB}_{\text{observed}}$  and  $\Delta\text{TCH}_{10}$  (Equation 3.4) and used for  $\Delta\text{AGB}_{\text{direct}}$  predictions.

$$\Delta\text{AGB}_{\text{direct}} = m \cdot \Delta\text{TCH}_{10} + n \quad (3.4)$$

#### 3.3.6.2 Indirect Approach

For each simulated forest stand (1 ha) and simulation year the AGB stock [ $\text{t ha}^{-1}$ ] was estimated using an established  $\text{TCH}_{10}$ -to-AGB power law relationship (Equation 3.5) with coefficients  $a = 0.4$  and  $b = 1.81$ , derived in a previous study (Knapp et al., 2018a).

$$\text{AGB} = a \cdot \text{TCH}_{10}^b \quad (3.5)$$

$\Delta\text{AGB}_{\text{indirect}}$  was calculated as the difference between AGB stocks at the beginning ( $t_1$ ) and end ( $t_2$ ) of the time interval (Equation 3.6).

$$\Delta\text{AGB}_{\text{indirect}} = \text{AGB}_{t_2} - \text{AGB}_{t_1} = a \cdot \text{TCH}_{10, t_2}^b - a \cdot \text{TCH}_{10, t_1}^b \quad (3.6)$$

#### 3.3.6.3 Enhanced Direct Approach (`dir+tex`)

We used the random forest machine learning algorithm (Breiman, 2001) to predict  $\Delta\text{AGB}_{\text{dir+tex}}$  from  $\Delta\text{TCH}_{10}$  (like in the direct approach) in combination with the CHM texture metrics. Random forest is an ensemble method based on regression trees. To train the algorithm, an ensemble of regression trees (here 1,000) is fit. Each single tree is trained using only a subset of the full training data and using only a subset

of available predictor variables. A prediction of a random forest is generated by averaging the predictions obtained from all single regression trees. We used the model selection procedure developed by Murphy et al. (2010) to obtain a parsimonious set of a few meaningful predictor variables (using the rfUtilities package in R). The goal of the procedure is to find the model that maximizes explained variability, minimizes the mean squared error (MSE) and needs the lowest possible number of predictor variables. In the procedure, random forest models are iteratively fit to the data, starting with all predictor variables. During each iteration, predictors are ranked by decreasing relative importance (contribution to MSE reduction) and predictors falling below an importance threshold are dropped before the next iteration. We used deciles from 0.1 to 1 as importance thresholds. Additionally, we chose a parsimony threshold of 0.05, meaning that in the end from all models for which MSE was only 5% or less above the MSE of the very best model, the model with the smallest number of predictors was chosen as the most parsimonious and hence best model. The model selection was applied using the whole dataset.

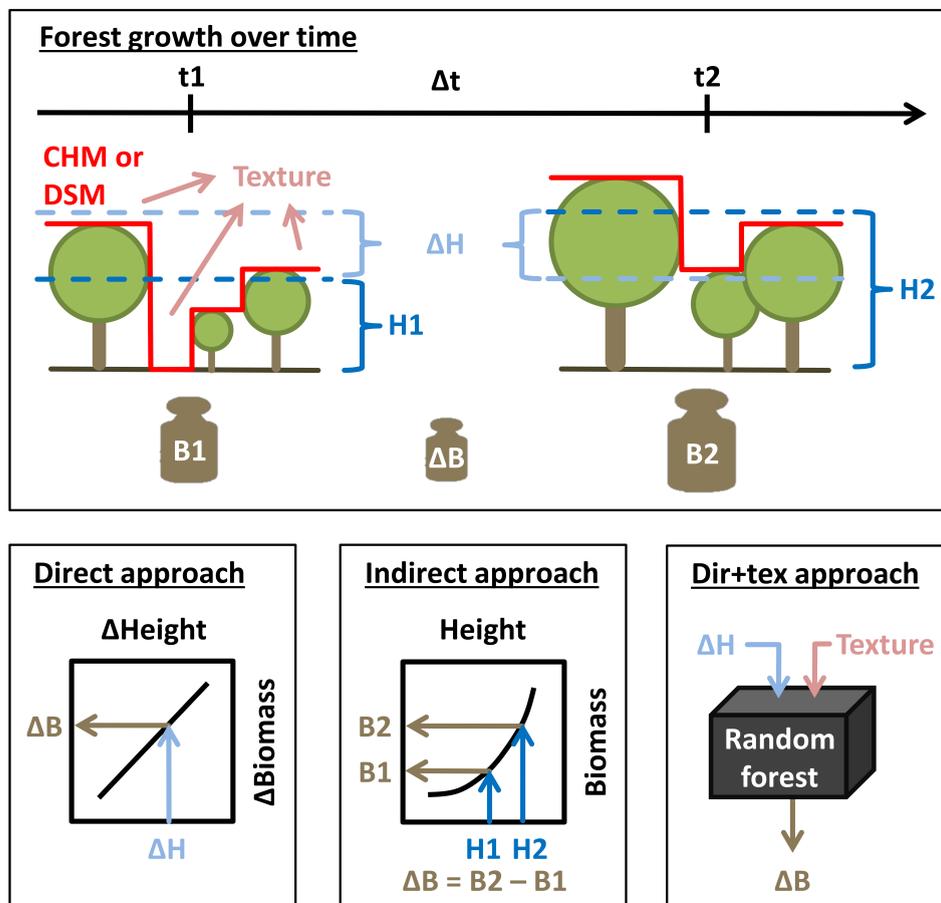


Figure 3.2: Inputs and principles of the three different approaches to derive biomass change ( $\Delta B$ ) over time from observed canopy heights ( $H$ ) or height change ( $\Delta H$ ) and canopy texture, respectively.  $H$  refers to mean top-of-canopy height, obtained by averaging the canopy height model (CHM, indicated by red line). The direct approach and the dir+tex approach do not require absolute canopy height information, thus digital surface models (DSM) can be used instead of CHMs.

The three different approaches for  $\Delta\text{AGB}$  prediction were evaluated based on how well predictions fit observations (i.e.,  $\Delta\text{AGB}_{\text{observed}}$  derived directly from the FORMIND output) and corresponding goodness-of-fit statistics such as  $R^2$ , root mean squared error (RMSE) and bias (mean residual value). To avoid overfitting of the machine learning model in the dir+tex approach, the dataset was split into five similarly sized groups and five random forest models were trained with 80% of the data, respectively. The predictions were then made for the remaining 20% test data, which were not part of the training dataset (5-fold cross validation).

### 3.3.7 Evaluation Experiments by Bootstrapping

To assess how prediction accuracy depends on forest height, we conducted a bootstrapping analysis of the simulated dataset. We iterated through forest height classes  $i$  from 5 to 50 m in 1-m steps. For each height class  $i$  the simulated dataset was subset, selecting all stands for which the stand height at the initial year ( $\text{TCH}_{10, t1}$ ) was inside the height class  $i$  with a class width of 10 m (i.e.,  $i - 5 \text{ m} < \text{TCH}_{10, t1} < i + 5 \text{ m}$ ). From all available stands in each height class  $i$ , 1,000 stands were randomly sampled with replacement. Estimations of  $\Delta\text{AGB}$  were derived for the 1,000 sample stands using (1) the indirect, (2) the direct and (3) the dir+tex approach. These  $\Delta\text{AGB}$ -estimates were compared against the actual  $\Delta\text{AGB}_{\text{observed}}$  values of each stand to calculate and record  $R^2$ , RMSE and bias. The whole sampling procedure was repeated 100 times for each height class.

### 3.3.8 Application on TanDEM-X Data

To finally test the derived models on real world data, we used imagery derived from the TanDEM-X satellites over BCI in the years 2011 and 2015. TanDEM-X is a radar interferometer with synthetic aperture (SAR) providing single pass interferometric (Graham, 1974; Bamler and Hartl, 1998) and polarimetric interferometric (Cloude and Papathanassiou, 1998) data at X-band (Krieger et al., 2013). The data have been acquired in a bistatic configuration, with a practically zero temporal baseline and a spatial baseline corresponding to a height of ambiguity of about 47 m for the 2011 and 70 m for the 2015 acquisition. Canopy height was calculated from interferometric TanDEM-X data at HH polarization in combination with a lidar DTM (acquired in 2009; Knapp et al., 2018a; Lobo and Dalling, 2014) needed as reference for the terrain topography below the trees. The single inversion steps are described in detail in Kugler et al. (2014). Here, the Single-Pol Inversion was applied. The forest layer was modeled as a random volume (random distribution of scatterers) over a ground layer as described by Attema and Ulaby (1978) and Treuhaft et al. (1996). This random volume over ground model (RVoG) assumes that the backscattering at X-band along forest height, can then be described by an exponential backscatter function in which the backscattered power decreases from the tree tops to the ground. Furthermore, Kugler et al. (2014) assumed that the interferometric measurement at X-band contains only backscattering from the vegetation layer (no backscattering from the ground below the forest). With these preconditions and with the ground information from the lidar DTM, we constructed CHMs of 10-m resolution from the TanDEM-X data for the years 2011 and 2015.  $\text{TCH}_{10}$  was derived at 1-ha scale for the 50-ha inventory plot for both years.  $\Delta\text{AGB}$  was estimated from  $\Delta\text{TCH}_{10}$  using the three approaches. As ground-

truth, AGB at 1-ha scale was calculated from DBH measurements of the BCI census data of 2010 and 2015 using the same tree allometries as in FORMIND (Equation 3.3). Inventory-based  $\Delta\text{AGB}$  was compared against the three TanDEM-X-based  $\Delta\text{AGB}$  estimates.

### 3.4 RESULTS

#### 3.4.1 Simulation Results

In total, the simulated forest dataset consisted of 28,736 ha. For each of these 1-ha stands, the changes in biomass and height were recorded over a 5-yr interval. Biomass increased on 72.6% (20,868 ha) and decreased on 27.4% (7,868 ha) of the stands. Height increased on 74.9% (21,522 ha), decreased on 25% (7,186 ha) and stayed equal on 0.1% (28 ha) of the stands. The maximal biomass gain was  $83.7 \text{ t ha}^{-1}$  and the maximal biomass loss was  $-428.6 \text{ t ha}^{-1}$ . The maximal height gain was 10.3 m and the maximal height loss was  $-42.6 \text{ m}$ .

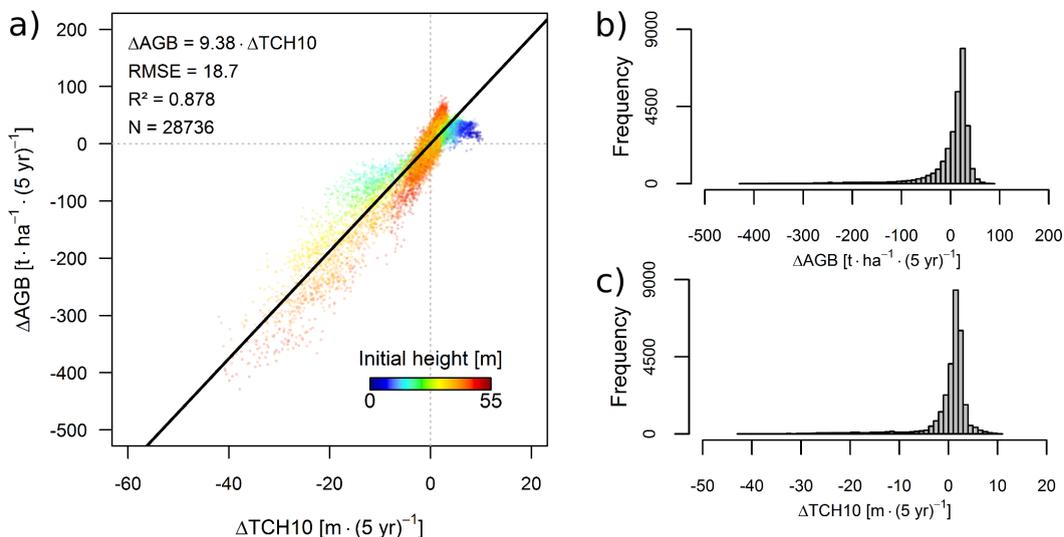


Figure 3.3: Simulated data showing (a) the  $\Delta\text{TCH}_{10}$ -to- $\Delta\text{AGB}$  relationship for a 5-yr time interval between first and second measurement and (b + c) the frequency distributions of both variables. Each point represents a 1-ha forest stand. Colors indicate the initial height ( $\text{TCH}_{10}$ ) of each stand. The black line represents the linear regression model.

To establish a direct relationship between height change  $\Delta\text{TCH}_{10}$  and biomass change  $\Delta\text{AGB}$ , a linear regression model was fit between the two variables (Figure 3.3) with a significant slope term of 9.38 ( $p < 0.001$ ) and an insignificant intercept of 0.14 ( $p = 0.218$ ). This matches the expectation that a change of zero in forest height should, on average, also result in zero change in biomass. Forcing the regression through the origin did not change the slope term. The  $R^2$  was 0.878 and the RMSE was  $18.7 \text{ t ha}^{-1} (5 \text{ yr})^{-1}$ . Figure 3.3 shows the scatterplot with colors indicating the  $\text{TCH}_{10}$  of each stand at the initial year. Low, early successional stands (blue) tend to show large  $\Delta\text{TCH}_{10}$  values associated with small  $\Delta\text{AGB}$  values, whereas high, old-growth forests (red) tend to show small  $\Delta\text{TCH}_{10}$  values associated with large  $\Delta\text{AGB}$  values.

Results for the analyses of 10- and 25-yr time intervals can be found in Appendix B (Figure B.2).

The derived  $\Delta TCH_{10}$ -to- $\Delta AGB$  relationship was used to estimate  $\Delta AGB$  from  $\Delta TCH_{10}$  following the direct approach. Alternatively,  $\Delta AGB$  was estimated for each hectare following the indirect approach and the dir+tex approach. Statistics were derived for the linear relationship of prediction vs. observation. The direct approach resulted in  $R^2 = 0.878$  (RMSE = 18.7 t ha<sup>-1</sup> (5 yr)<sup>-1</sup>). The indirect approach resulted in a  $R^2 = 0.945$  (RMSE = 12.61 t ha<sup>-1</sup> (5 yr)<sup>-1</sup>). The cross-validation of the dir+tex approach resulted in a  $R^2 = 0.947$  (RMSE = 12.42 t ha<sup>-1</sup> (5 yr)<sup>-1</sup>). Figure 3.4 shows the 1:1-plots for the different approaches.

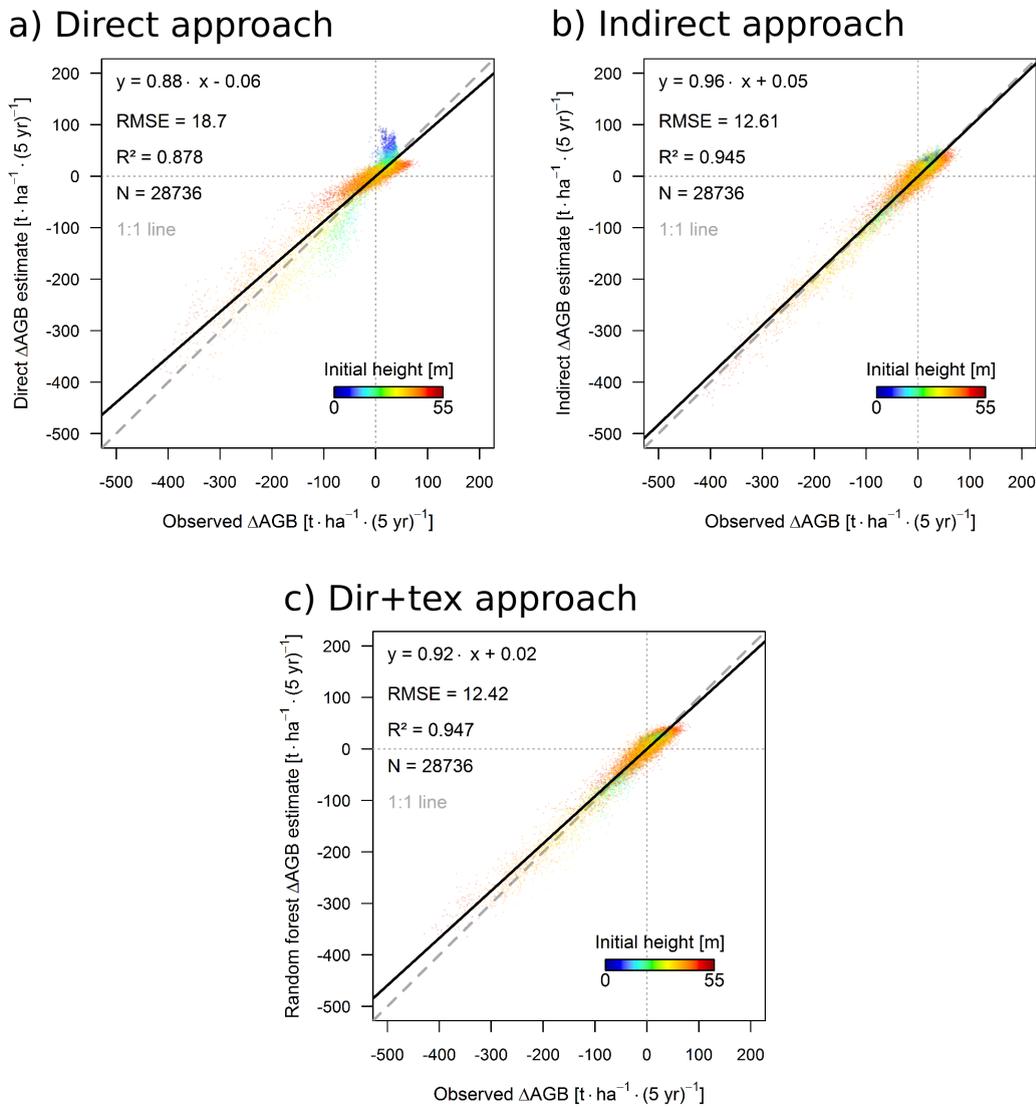


Figure 3.4: 1:1-plots of estimated  $\Delta AGB$  versus observed  $\Delta AGB$  following (a) the direct (b) the indirect and (c) the dir+tex approach. Each point represents a 1-ha forest stand. Colors indicate initial height (TCH<sub>10</sub>) of each stand. Black lines represent linear regression models.

In the random forest model selection procedure for the dir+tex approach nine predictor variables were selected to form the most parsimonious model (Figure B.3).

$\Delta TCH_{10}$  was identified clearly as the most important predictor variable, followed by first order texture variables skewness, topographic position index and kurtosis. Second order (GLCM-based) texture metrics were of minor importance.

The derived models for  $\Delta AGB$  estimation were evaluated for forests with different heights (bootstrapping analysis). We explored height gains and losses separately (Figure 3.5). According to the  $R^2$  values (Figure 3.5a-c), the linear trend between prediction and observation is weak for stands with low canopy heights ( $TCH_{10} < 10$  m) but the  $R^2$  reach values around 0.9 for all stands with  $TCH_{10} \geq 10$  m. However, if losses are regarded separately, stands need to have an initial height  $\geq 25$  m to show such high  $R^2$  values. If gains are regarded separately, the linear trend is generally less pronounced with maximal observed  $R^2$  values being around 0.5. Overall, the  $R^2$  patterns are similar for the direct, the indirect and the dir+tex approach with a slight increase from low to high stands.

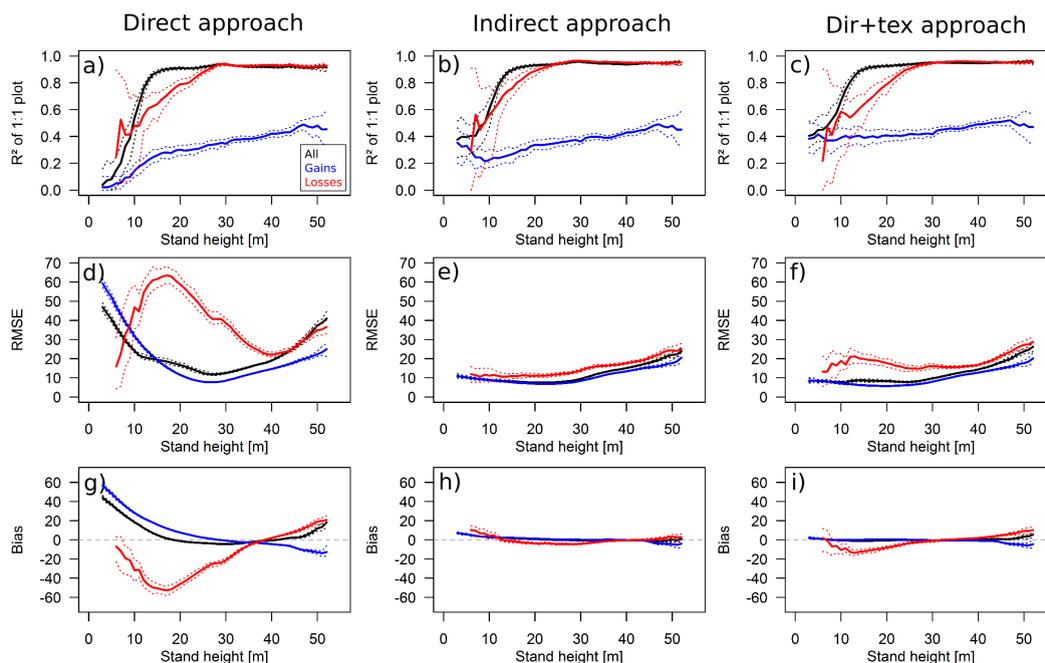


Figure 3.5: Prediction statistics of the three approaches plotted over stand heights.  $R^2$ , RMSE and Bias were calculated from 1,000 samples in each stand height class. Solid lines represent the mean and dashed lines represent the minima and maxima of 100 bootstrapping replicates. Sampling was done using the full dataset (“All”, black) and using exclusively stands with positive (“Gains”, blue) or negative (“Losses”, red) height changes, respectively.

In contrast to the  $R^2$  patterns, the prediction error (RMSE) patterns are quite distinct between the approaches. RMSE for the direct approach (Figure 3.5d) are large for low and high stands and have a minimum at intermediate stand heights around 25 m (40 m in case of losses). For the indirect approach (Figure 3.5e) RMSE stay constant around 10 t for stands with heights  $< 30$  m and increase slightly for stands with heights  $\geq 30$  m. For the dir+tex approach (Figure 3.5f) the RMSE pattern for gains is similar to the indirect approach. For losses, the pattern is similar to the direct approach but with smaller RMSE values.

Finally, when we look at the systematic biases of the  $\Delta\text{AGB}$  predictions, the indirect approach does not show any pronounced bias over the entire height range (Figure 3.5h). The direct approach leads to considerable systematic biases when estimating  $\Delta\text{AGB}$  from height gains and losses (Figure 3.5g). Gains are overestimated by up to 60 t for the lowest forest stands and underestimated by up to -20 t for the highest stands. Losses are negatively biased by up to -50 t for low stands and positively biased by up to 20 t for high stands. The biases are around zero for stands around 35 m height. The biases for the dir+tex approach are higher than the ones for the indirect approach but much lower than the ones for the direct approach (Figure 3.5i). Gains are unbiased for stands up to 40 m height. Biases of losses show a similar but less pronounced pattern to the direct approach.

### 3.4.2 Theoretical Considerations about the $\Delta H$ -to- $\Delta\text{AGB}$ Relationship

For a better understanding of the observed biases, in particular the height-dependent biases of the direct approach, we made the following theoretical considerations. For simplicity, we replaced  $\text{TCH}_{10}$  with  $H$ . The problem with assuming a linear relationship between  $\Delta H$  and  $\Delta\text{AGB}$  is that the underlying  $H$ -to- $\text{AGB}$  relationship is non-linear and can be described, e.g., by a power law (Equation 3.7 with parameters  $a$  and  $b$  from Equation 3.5). The relationship between height change and biomass change at any given height can be derived from the first derivative of Equation 3.7 (see Equation 3.8 and Equation 3.9).

$$\text{AGB} = a \cdot H^b \quad (3.7)$$

$$\frac{d\text{AGB}}{dH} = a \cdot b \cdot H^{b-1} \quad (3.8)$$

$$\Delta\text{AGB} = a \cdot b \cdot H^{b-1} \cdot \Delta H \quad (3.9)$$

Equation 3.9 describes the relationship between  $\Delta\text{AGB}$  and  $\Delta H$ , which depends also on  $H$  (which is unknown in case of missing DTM). Thus, the magnitude of  $\Delta\text{AGB}$  is not only dependent on  $\Delta H$  but also on the absolute height  $H$  of a stand. For any given stand height  $H$ , the  $\Delta H$ -to- $\Delta\text{AGB}$  relationship can be expressed with a linear model, but the slope of the linear model is different for each possible  $H$  (Figure 3.6a). If we assume that  $H$  lies inside the range from 0 to 55 m, the marked area in the plot encloses all possible  $\Delta H$ - $\Delta\text{AGB}$  combinations (black envelope resembling the shape of a propeller). Stands of a given initial height move along power law-shaped trajectories over time (e.g., open circles for 10 m and filled circles for 30 m initial height in Figure 3.6a). The possible gains for a certain time interval are constrained by the forest's growth rate, while losses can be large even within short time periods, due to the stochastic occurrence of mortality and disturbance events. For 5-, 10- and 25-yr intervals, the simulations show the subspace of observable  $\Delta H$ - $\Delta\text{AGB}$  combinations (colored envelopes in Figure 3.6b). The simulation data exceed the black propeller, particularly around the coordinate origin. The explanation is that the black propeller envelope only applies under the assumption of an exact power law relationship between  $H$  and  $\text{AGB}$  without variability in the  $H$ -to- $\text{AGB}$  relationship.

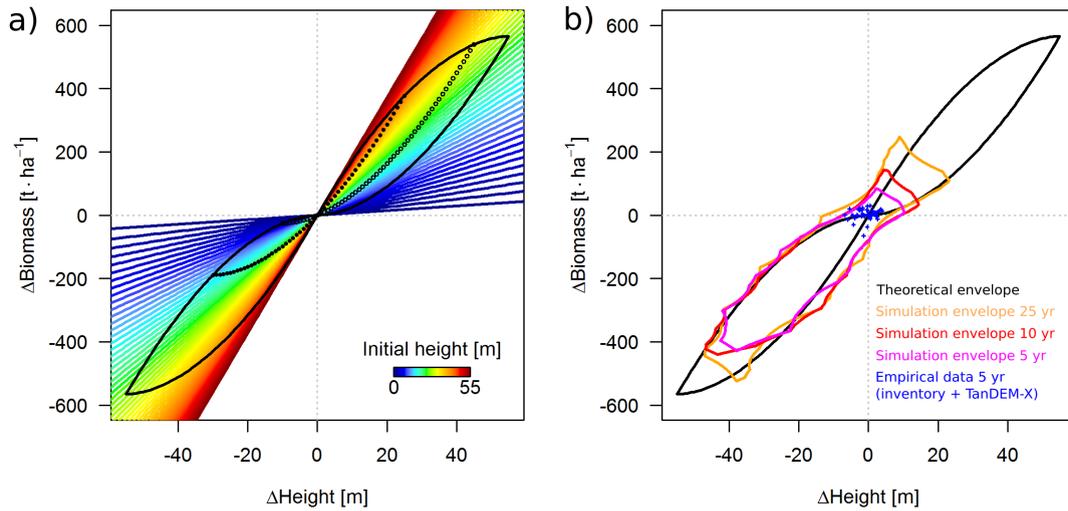


Figure 3.6: Theoretical considerations on the relationship between height change and biomass change. (a) shows the slopes of the relationship for different initial stand heights. The black envelope covers all possible change combinations if the initial height is restricted to the range of 0 to 55 m and the height-to-biomass relationship is an exact power law. The (open and closed) circles mark trajectories of different forest stands with a given start height (10 and 30 m). (b) shows the envelopes of simulated data for different time intervals and the empirical data from the BCI 50-ha plot.

### 3.4.3 Results for the 50-ha Plot

The CHMs of the BCI 50-ha plot derived from TanDEM-X data of 2011 and 2015 (Figure B.4, Appendix B) served to test the three  $\Delta\text{AGB}$  estimation models which had been fit with simulation data. At the 1-ha scale, the mean  $\text{TCH}_{10}$  over the 50 ha was 31.0 m ( $\pm 3.4$  m SD) in 2011 and 30.3 m ( $\pm 3.9$  m SD) in 2015. Distributions of canopy height ( $\text{TCH}_{10}$ ) in both years and canopy height change ( $\Delta\text{TCH}_{10}$ ) within the four years are given in Figure 3.7a-c. The distribution of biomass change ( $\Delta\text{AGB}$ ) calculated from the inventory data is given in Figure 3.7d. There was a slight loss in average canopy height (mean  $\Delta\text{TCH}_{10} = -0.6$  m  $\pm 2.6$  m SD), but a slight gain in average biomass (mean  $\Delta\text{AGB} = 3.3$  t  $\text{ha}^{-1}$   $\pm 16.6$  t  $\text{ha}^{-1}$  SD). When plotted against each other  $\Delta\text{AGB}$  and  $\Delta\text{TCH}_{10}$  of the BCI plot scatter closely around the coordinate origin in Figure 3.6b (blue points), showing neither strong gains nor losses.

Table 3.1:  $R^2$  values for the 1:1 relationships of  $\Delta\text{AGB}$  predictions from TanDEM-X data following the three different approaches vs. inventory-based  $\Delta\text{AGB}$  and against each other.

	Direct	Indirect	Dir+tex
Inventory	0.006	0.007	0.003
Dir+tex	0.966	0.946	-
Indirect	0.991	-	-

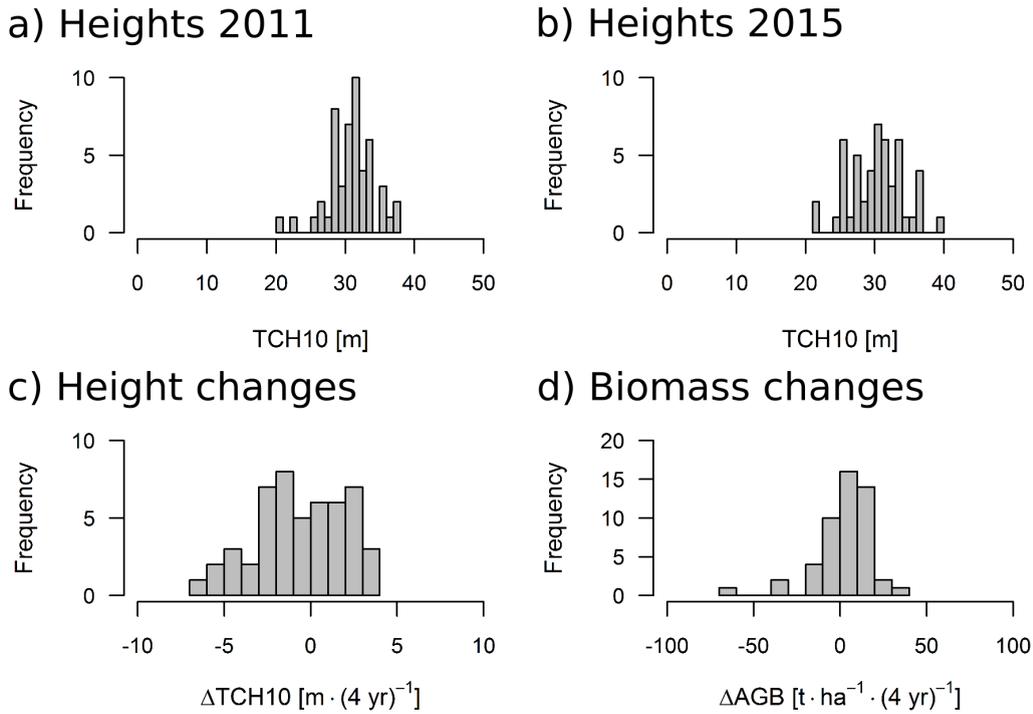


Figure 3.7: TanDEM-X-derived distributions of (a+b) mean top-of-canopy height ( $TCH_{10}$ ), (c) height change ( $\Delta TCH_{10}$ ) and (d) inventory-derived biomass change at 1-ha scale within the 4-yr interval.

The TanDEM-X-derived  $\Delta AGB$  estimates from all three approaches did not show any significant correlation with the observed  $\Delta AGB$  derived from the forest inventory data from censuses in 2010 and 2015, with  $R^2$  values close to zero (Figure 3.8 and Table 3.1). The estimates of  $\Delta AGB$  from the three approaches were, however, closely correlated among each other with  $R^2$  values  $\geq 0.96$  (Table 3.1). Thus, within the narrow range of height changes observed in the BCI plot over the short time interval of 4 years, the three different approaches produced very similar predictions. However, the observed height changes do not reflect the changes in AGB adequately, hence, none of the approaches resulted in  $\Delta AGB$  estimates that could be confirmed by the ground-truth data.

Figure 3.9 shows  $ABG$  stocks plotted over  $TCH_{10}$  at 1-ha scale for both years. The arrows illustrate the changes between 2011 and 2015 in both attributes. There is no clear link in the sense that an increase or decrease in AGB is accompanied with an increase or decrease of canopy height, respectively. On 23 hectares AGB and canopy height changed in the same direction, whereas on the remaining 27 hectares they changed in opposite directions. This explains the difficulties in estimating  $\Delta AGB$ .

### 3.5 DISCUSSION

Our results provide insight into the relationship between canopy height and biomass change over time for a tropical lowland rainforest. Understanding this relationship is crucial to quantify forest carbon losses and gains with remote sensing data. Using forest simulations, including disturbances, we could produce a large dataset covering a wide range of stand structures and successional stages. With this dataset, we were

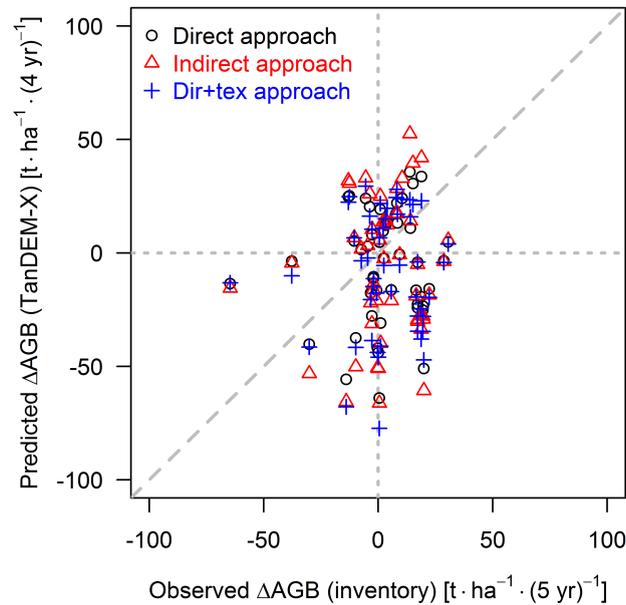


Figure 3.8: 1:1 plot of predicted  $\Delta\text{AGB}$  (based on TanDEM-X-derived  $\Delta\text{TCH}_{10}$ ) versus observed  $\Delta\text{AGB}$  (based on forest inventory data) at 1-ha scale for the BCI 50-ha plot and the time interval between 2011 and 2015. Colors represent the direct (black), the indirect (red) and the dir+tex (blue) approach.

able to compare the performance of three different approaches to estimate biomass change.

### 3.5.1 Performance of the Approaches

In comparison, the indirect approach performed better than the direct approach. The former led to more precise and unbiased  $\Delta\text{AGB}$  estimations over the full range of possible height changes. Nevertheless, looking only at the statistics for the whole dataset, both approaches showed high  $R^2$  values. However, the colors indicating the initial height in the 1:1-plot (Figure 3.4a) and the bootstrapping results (Figure 3.5) reveal that the direct approach only worked well in a window of forest heights around 30 m, while for low and high stands it produced strongly biased results. Analyzing gains and losses further revealed the asymmetry in the  $\Delta\text{H}$ - $\Delta\text{AGB}$  relationship, due to slow, continuous growth and abrupt, stochastic mortality, and its effect on statistics that quantify prediction accuracy. The biases of the direct approach for low and high stands can be explained by the non-linearity of the H-AGB relationship, leading to different slopes in the  $\Delta\text{H}$ - $\Delta\text{AGB}$  relationship, depending on stand height. The development of an enhanced direct approach, which avoids such bias and allows accurate estimations, even in the absence of information about H, was the goal behind the dir+tex approach. The textural information should compensate for the missing information about H. It was shown that estimations of  $\Delta\text{AGB}$  from the dir+tex approach were nearly unbiased and of similar accuracy as the ones from the indirect approach.

In this study, we did not test the robustness of the approaches with regard to the time interval, i.e. whether it is possible to calibrate an approach based on observations over a certain time interval and then apply it for predictions over a different time

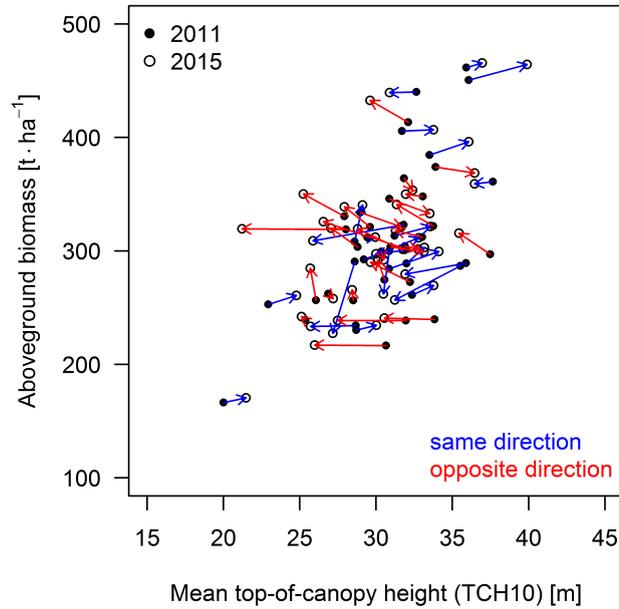


Figure 3.9: Aboveground biomass (AGB) plotted over canopy height ( $TCH_{10}$ ) for the two different years. Each point represents 1-ha of the BCI 50-ha plot. Arrows represent the change in both attributes over the 4-yr interval. Colors of the arrows indicate whether AGB and  $TCH_{10}$  changed in the same (blue) or opposite (red) direction.

interval. For the indirect approach the interval is irrelevant, because it is based on two independent biomass stock estimates. For the direct approach estimates depend on the slope of the  $\Delta H$ - $\Delta AGB$  relationship, which was similar for 5-, 10- and 25-yr intervals (9.38, 9.99, 10.51). In the dir+tex approach, the only variable depending on the interval is  $\Delta H$  (like in the direct approach), while all texture variables only describe the state of the forest at the initial point of the interval. Hence, we expect all three approaches to be fairly robust for variable time intervals. The length of the time interval does, however, influence the frequency distribution of observable changes, with large gains being only possible over longer time intervals and large losses being more frequent over longer time intervals (Figure B.2).

### 3.5.2 Comparison with Other Studies

In an earlier study comparing the indirect and direct approach on BCI, Meyer et al. (2013) used data acquired in 1998 with the Laser Vegetation Imaging Sensor (LVIS, large footprint) and discrete return small footprint lidar data from 2009 (11-yr interval). They could not find a direct  $\Delta H$ - $\Delta AGB$  relationship at the 1-ha scale ( $R^2 < 0.1$ ) and using the indirect approach they report large estimation uncertainties. They considered the different sensor types as a major reason for the missing direct  $\Delta H$ - $\Delta AGB$ -relationship. In another study using data from two successive LVIS campaigns in 1998 and 2005 (7-yr interval) at La Selva, Costa Rica, Dubayah et al. (2010) found a direct  $\Delta H$ - $\Delta AGB$  relationship ( $R^2 = 0.65$ ,  $RSE = 10.54 \text{ t ha}^{-1}$  or excluding secondary forests  $R^2 = 0.5$ ,  $RSE = 8.86 \text{ t ha}^{-1}$ ) using a two-predictor model based on changes in height quantiles  $\Delta RH_{50}$  and  $\Delta RH_{100}$ . The statistics of that relationship are in the order of magnitude of what was observed for gains in our simulations. Several comparative

studies in boreal (Bollandsås et al., 2013; Næsset et al., 2013) and subtropical (Cao et al., 2016) forests report  $\Delta$ AGB estimations to be more accurate when obtained from the direct approach as opposed to the indirect approach. Others, covering tropical Asia (Englhart et al., 2013) or temperate North America (Hudak et al., 2012), only tested the indirect approach. In two studies, we found scatterplots that indicate a different slope in the  $\Delta$ H- $\Delta$ AGB relationship for secondary (disturbed) sites than for old-growth (undisturbed) sites (Figure 8 in Dubayah et al., 2010 and Figure 9F in Hudak et al., 2012). This is in agreement with our simulation results and the explanation given in Figure 3.6.

As soon as the goal is to cover a wide range of successional stages including disturbed stands, the non-linear H-AGB relationship leads to H-dependent biases in the linear  $\Delta$ H- $\Delta$ AGB relationship. The condition for an unbiased direct approach would be that the underlying H-AGB relationship is linear. In past studies, the direct approach has worked better in boreal sites (coniferous forests) and worse in tropical sites (broadleaved forests). Hence, the differences in crown shapes and forest structure between those forest types might lead to a more linear H-AGB relationship in coniferous forests, making them more suitable for the direct approach.

Recently, there have been attempts to generalize the height-to-biomass relationships accounting for differences between geographical regions and forest types (Asner and Mascaro, 2014; Bouvier et al., 2015). These relationships are, however, not linear, but power laws. Thus, the linear relationship that has been found in some studies in the past might be an effect of either small plot sizes, causing large variability (Solberg et al., 2015; Solberg et al., 2017), or a limited sampling of successional stages, not covering the full spectrum of H-AGB combinations.

A linear H-AGB relationship is a strict criterion for the application of a method. Many forests may not fulfill this criterion and for large-scale mapping of AGB changes (from height changes) a method independent of such a requirement should be preferred. With the proposed dir+tex approach, we found a method that works independently of knowledge about absolute canopy height. Biomass changes are estimated based on height change and local canopy texture at the initial point in time. Predictions showed similar accuracies to the indirect approach. Biases were somewhat larger than for the indirect but much smaller than for the direct approach. This texture-assisted approach could therefore be applied to multi-temporal DSM data for which no exact underlying DTM is available and hence CHM creation is impossible. This applies to global DSMs, originating from different spaceborne sensor systems (SRTM, TanDEM-X) (Puliti et al., 2017; Solberg et al., 2014; Solberg et al., 2015) and to DSMs derived photogrammetrically from digital stereo imagery (Bohlin et al., 2012; Vastaranta et al., 2013).

### 3.5.3 Outcome of TanDEM-X Application

Despite the good performance of the three approaches on the simulation dataset, none of them was able to produce accurate predictions of biomass change for the BCI 50-ha plot, based on canopy height change derived from TanDEM-X data. The main difference between the simulated data and the TanDEM-X data was that the forest in the simulations was exposed to a disturbance regime causing dramatic changes in canopy height and biomass, while no such disturbances occurred on the BCI plot within the observed time interval. Thus, the changes observed on the BCI plot are

only a subset of possible changes for a 5-yr interval (Figure 3.6b). BCI only represents old-growth dynamics with small changes in canopy height and biomass, while the simulations cover the entire range of possible changes. We suspect that there are two main reasons for the difficulties in predicting  $\Delta\text{AGB}$  from TanDEM-X-derived  $\Delta\text{TCH}_{10}$  on the BCI plot: (1) the variability unexplained by the prediction models and (2) the uncertainty in canopy height estimates.

Regarding the unexplained variability (1), the bootstrapping results tell us that for an old-growth forest, like on the BCI plot, with an average height  $\text{TCH}_{10} = 30$  m, the average estimation error for  $\Delta\text{AGB}$  over a 5-yr interval is around  $10 \text{ t ha}^{-1}$  for all three approaches (e.g., for the indirect approach  $\text{RMSE} = 9.5 \text{ t ha}^{-1}$ ). This is a high uncertainty given that the standard deviation of  $\Delta\text{AGB}$  from the BCI inventory was only  $16.6 \text{ t ha}^{-1}$  and the mean  $\Delta\text{AGB}$  was only  $3.3 \text{ t ha}^{-1}$ . The high  $R^2$  values in the direct  $\Delta\text{H}-\Delta\text{AGB}$  relationship of the simulated data were caused by the large losses. Since such losses did not occur on BCI, the  $R^2$  for the BCI plot should be expected to be much lower. In the simulations, the average  $R^2$  for gains only was 0.35 for a 30 m high stand. Even this value can be expected to be too optimistic for the BCI plot, as the simulated gains include areas of recovery after disturbances, which were also not present in the BCI plot. Thus, the intrinsic variability in canopy height and biomass alone leads to a very weak  $\Delta\text{H}-\Delta\text{AGB}$  relationship for old growth forests. The simulations did not account for the uncertainty in canopy height estimates from remote sensing, which is yet another factor influencing the results from the BCI plot.

Uncertainty in X-band-derived tropical forest heights (2) has been reported to lie in a range from 1 to 6 m (Kugler et al., 2014; Treuhaft et al., 2017). The precise numbers may vary depending on the sensor configuration, the height metrics and the spatial scale considered. In our study, the main limitation of the TanDEM-X derived heights arose from the sub-optimum spatial baselines available, that were practically too large for the forest heights within the study site, leading to a low interferometric coherence (Figure B.5) and to rather inaccurate forest heights. Note that the limited penetration capability at the X-band was circumvented by using lidar-derived topographic information. The observed height changes on the BCI plot were small (standard deviation of 2.6 m for  $\Delta\text{TCH}_{10}$  at ha-scale and mean  $\Delta\text{TCH}_{10}$  of -0.6 m) in comparison to the large uncertainty.

Thus, the two reasons, (1) variability in the  $\Delta\text{H}-\Delta\text{AGB}$  relationship and (2) uncertainty in canopy height estimates, lead to a low signal-to-noise ratio and explain why the presented approaches are insufficient for the detection of small changes in AGB in an undisturbed old-growth tropical rainforest. Further sources of uncertainty are the variability in tree geometry at the individual level which is simplified in the forest model by allometric equations, as well as temporal mismatch (between census and remote sensing acquisition), errors in the census data and geolocation errors.

### 3.5.4 Perspectives

The presented approaches to estimate biomass changes at 1-ha scale based on 10-m resolution vegetation height changes are best suited for monitoring changes due to disturbances, degradation and growth. They could serve to map biomass changes caused by these processes from TanDEM-X data at regional scale. They cannot detect small changes associated with old-growth forest dynamics. Future research efforts should focus on reducing the uncertainty in biomass change estimates. Uncertainty

in the prediction models can be reduced, e.g., by considering data at finer horizontal resolution, the full vertical signal distribution, the multitude of variables that can describe forest structure and individual-tree-based approaches (Zhao et al., 2018). The uncertainty in height estimations due to the variability in the remote sensing data can be reduced by fitting trends to dense time series of multiple acquisitions (Treuhart et al., 2017). Future research should also try to disentangle the processes and compartments behind biomass dynamics (including also belowground biomass) by quantifying the contributions of wood productivity, mortality, foliage and fine root turnover (He et al., 2012). Upcoming spaceborne missions (e.g., GEDI, BIOMASS, Tandem-L) and the increasing data collections from airborne campaigns (lidar, SAR, photogrammetry) will provide the measurements to advance in this direction. Data fusion will help to use the available data most effectively with each individual mission contributing to a different aspect, e.g., by providing detailed vertical foliage distributions and ground elevation at points in space (GEDI), or wall-to-wall measurements of SAR backscatter from woody biomass components (BIOMASS, Tandem-L). Forest models will help to understand and interpret new remote sensing observations.

### 3.6 CONCLUSION

In this study, we analyzed different approaches for estimating aboveground biomass change in a tropical rainforest from observed changes in mean top-of-canopy height. With forest simulations, it was possible to generate and analyze data covering a much wider range of changes than usually available from field data. Goodness-of-fit statistics ( $R^2$ , RMSE and bias) for the  $\Delta$ AGB estimations were computed over the full range of possible stand heights, also demonstrating an asymmetry between gains and losses. It was found that a direct, linear  $\Delta$ H-to- $\Delta$ AGB relationship only provides accurate predictions under limited conditions and can lead to large prediction biases when applied over a wide range of stand heights. The indirect approach, which builds on the H-to-AGB stock relationship, can be used to avoid such biases, but it is dependent on accurate measurements of canopy height. A third approach, based on random forest machine learning, was found to provide similar prediction accuracies to the indirect approach. The latter did not require canopy height as input. Instead, canopy height change in combination with a set of canopy texture metrics served as predictors. The simulation-derived approaches were not sensitive enough to detect small biomass changes based on TanDEM-X data for a 50-ha plot in Panama. Further research is required to improve our ability to detect even small changes in biomass with remote sensing. The presented approaches with their accuracies of around  $\pm 13 \text{ t ha}^{-1}$  are nevertheless well suited for monitoring of forest biomass changes at large scales. In summary, the study demonstrated the potential of using a forest model for improving the understanding and interpretation of multi-temporal remote sensing data and for evaluating different approaches.

### 3.7 ACKNOWLEDGMENTS

We thank the Smithsonian Tropical Research Institute for providing the census data for BCI. The BCI forest dynamics research project was founded by S.P. Hubbell and R.B. Foster and is now managed by R. Condit, S. Lao, and R. Perez under the Center for Tropical Forest Science and the Smithsonian Tropical Research in Panama.

Numerous organizations have provided funding, principally the U.S. National Science Foundation, and hundreds of field workers have contributed. We further thank the German Aerospace Center (DLR) for providing TanDEM-X data under OTHER6890 and V. Cazcarra-Bes for his help with data processing. We thank J. Dalling for providing airborne lidar data for BCI. This study was conducted with funding by the German Federal Ministry for Economic Affairs and Energy (BMWi) under the funding reference 50EE1416. R.F., K.P. and A.H. were supported by the HGF-Helmholtz Alliance "Remote Sensing and Earth System Dynamics" HA-310 under the funding reference RA37012. We thank three reviewers and the editor for their constructive comments on our paper.

FOREST STRUCTURE METRICS TO GENERALIZE BIOMASS  
ESTIMATION FROM LIDAR ACROSS CONTINENTS

---

#### 4.1 ABSTRACT

Forest aboveground biomass is a key variable in remote sensing based forest monitoring. Active sensor systems, such as lidar, can generate detailed canopy height products. Relationships between canopy height and biomass are commonly established via regression analysis using information from ground-truth plots. In this way, many site-specific height-biomass relationships have been proposed in the literature and applied for mapping in regional contexts. However, such relationships are only valid within the specific forest type for which they were calibrated. A generalized relationship would facilitate biomass estimation across forest types and regions. In this study, a combination of structural descriptors is proposed as an approach for generalization between forest types. Each descriptor is supposed to quantify a different aspect of forest structure, i.e., mean canopy height, maximum canopy height, maximum stand density, vertical heterogeneity and wood density. Lidar data covering 194 ha of forest inventory plots from five different sites including temperate and tropical forests from Africa, Europe, North, Central and South America were used. Biomass predictions using the best general model (nRMSE = 12.4%,  $R^2 = 0.74$ ) were found to be almost as accurate as predictions using five site-specific models (nRMSE = 11.7%,  $R^2 = 0.77$ ). The results further allow interpretation about the importance of the employed structure descriptors in the biomass estimation and the mechanisms behind the relationships. Understanding the relationship between canopy structure and aboveground biomass and being able to generalize it across forest types are important steps towards consistent large scale biomass mapping and monitoring using airborne and spaceborne platforms.

#### 4.2 INTRODUCTION

Quantifying global carbon stocks of forests as well as their changes over time requires spatially explicit measurements and monitoring (Harris et al., 2012). The primary variable of interest hereby is forest aboveground biomass (AGB). Thus, there is a growing amount of literature about estimating forest biomass from canopy height metrics derived from light detection and ranging (lidar), synthetic aperture radar or photogrammetry (Goetz and Dubayah, 2011; Zolkos et al., 2013; Asner and Mascaro, 2014; Lu et al., 2014). The majority of these studies investigated data from specific forest sites with the goal to find the best prediction model, i.e. maximizing explained variability (e.g.,  $R^2$ ), minimizing prediction error (e.g., RMSE) and minimizing systematic bias, while using the most parsimonious set of predictor variables (Zolkos et al., 2013). Different statistical approaches have been used including multiple linear regression models and machine learning methods (Fassnacht et al., 2014). As a result various site- and forest type-specific relations for biomass estimation have been proposed and applied successfully for biomass mapping at regional scale.

Comparatively few studies have tried to seek for generalization in the estimation approaches (e.g., Lefsky et al., 2002b; Asner et al., 2012b; Magnussen et al., 2012; Asner and Mascaro, 2014; Vincent et al., 2014; Bouvier et al., 2015). For consistent global mapping of forest biomass, however, it would be desirable to have more generic relationships which are applicable across different forest types and biogeographic regions. Such an approach would contribute to a better understanding of how structural attributes differ between forest types and how they are related to biomass. It would

also facilitate biomass mapping across the globe. Given the variety of metrics that can be derived from lidar data (Næsset, 2002), it would further be desirable to have a minimum set of meaningful metrics, describing different aspects of forest structure, to avoid problems with multicollinearity and extensive model selection procedures (Bouvier et al., 2015). A widely used general approach is the one proposed by Asner et al. (2012b) and modified by Asner and Mascaro (2014) for pan-tropical application. The function is inspired by individual tree allometry, where tree AGB can be modeled as a multiplicative power law of tree diameter at breast height (DBH; or tree basal area BA), tree height and species-specific wood density (Chave et al., 2014). Hence, as a stand level equivalent for area-based AGB estimation a power law of stand BA sum, mean top-of-canopy height (TCH) and average wood density is used. It further assumes a linear relationship between BA and TCH, which may differ between regions, and regional differences in average wood density can be considered. This approach has been established using data from different tropical regions (Hawaii, Panama, Peru, Madagascar) and has been applied successfully in other tropical regions, e.g., Colombia (Asner et al., 2012a), Malaysia (Coomes et al., 2017) and Tanzania (Getzin et al., 2017).

In a recent study, Bouvier et al. (2015) suggested a different model for generalized AGB estimation. They used several a priori defined lidar-based metrics that captured different aspects of forest structure. Their model was able to produce accurate AGB estimations for different forest types in France. However, site-specific coefficients led to higher prediction accuracies for each site, compared to using only one set of coefficients across all sites.

In this study, we attempted to find a model that is generally applicable throughout different forest types and even different biomes by including structural information on forest stands. Tropical forests in Panama, French Guiana and Gabon were analyzed along with temperate forests in the United States of America and Germany. We tried to estimate BA and AGB at the 1-ha scale. Usually, the two are closely correlated. However, BA is a simple inventory-derived metric, which is easily comparable among sites and studies. Inventory-based AGB, on the other hand, is more complex to compute. Assumptions about allometric relationships and wood density values are required to derive single tree AGB. It can lead to considerable differences in stand AGB if different assumptions are chosen for the same stand (Duncanson et al., 2017). Therefore, both variables were chosen in this analysis – BA for its robustness and comparability and AGB as the major variable of interest in forest carbon mapping efforts.

We hypothesized that the following structural forest attributes may contribute to explaining stand BA and AGB: a) mean canopy height, b) maximum possible stand density c) maximum possible tree height, d) vertical canopy heterogeneity, and for AGB, additionally, e) average wood density. Most of these structural attributes can be quantified in several alternative ways. Thus, for one attribute there may be a set of several candidate metrics. In this analysis, data from 194 ha of temperate and tropical forest from five megaplot sites were combined with the following goals: 1) to find a generic approach for BA and AGB estimation that can be applied across all sites without causing prediction bias at any individual site and 2) to investigate the contributions of the different structural attributes.

### 4.3 MATERIALS AND METHODS

#### 4.3.1 *Study Sites*

Data from five forest sites covering different forest types and biogeographical zones were used (Table 4.1). Four study sites are part of the ForestGEO megaplot network (Anderson-Teixeira et al., 2015) and thus they have been inventoried according to a standard protocol. The data structure for the fifth site, Paracou, is similar to the one of the other sites. For each tree the diameter at breast height (DBH), spatial position and species identity were recorded. In this study, all trees with a DBH  $\geq 10$  cm were considered and all given numbers refer to trees above this size threshold. In the following the five sites are briefly described. 1) Barro Colorado Island (BCI), Panama, is a Central American lowland tropical moist forest site with an annual precipitation of 2,580 mm and an average temperature of 27.1 °C. The census on the 50-ha plot was conducted in 2010 (Condit, 1998; Hubbell et al., 1999; Hubbell et al., 2005) and comprised 22,084 trees, which belonged to 223 species. 2) Paracou, French Guiana, is a South American lowland tropical rainforest with an annual precipitation of 3,040 mm and an average temperature of 26 °C. There are 16 large plots with plots 1 to 15 having an extent of 250 m  $\times$  250 m each and plot 16 having an extent of 500 m  $\times$  500 m. In 1986 and 1987, selective logging with different treatment intensities (timber logging, fuelwood logging, thinning) has been conducted on some of the plots, while others have served as control plots (Hérault and Piponiot, 2018). Since our analysis was conducted for 100 m  $\times$  100 m units, only subareas of 200 m  $\times$  200 m measured from the south-western corners of plots 1 to 15 were used. In total 85 ha from Paracou were analyzed. Censuses were conducted in 2015 and comprised 53,501 trees of 713 species. 3) Rabi, Gabon, is a Central African lowland tropical rainforest site with an annual precipitation of 2,300 mm and an average temperature of 26 °C. The census on the 25-ha plot was conducted from 2010 to 2012 (Labrière et al., 2018) and comprised 12,019 trees, which belonged to 235 species. 4) The Smithsonian Environmental Research Center (SERC) plot, United States of America, is a North American deciduous broadleaved temperate forest site with an annual precipitation of 1,070 mm and an average temperature of 13.2 °C. The census on the 16-ha plot was conducted in 2014 (McMahon and Parker, 2015; Král et al., 2016) and comprised 4,719 trees, which belonged to 39 species. 5) Traunstein, Germany, is a Central European managed mixed temperate forest site, which includes conifer and broadleaf plantations. It has an annual precipitation of 1,240 mm and an average temperature of 7.6 °C. The census on the 25-ha plot was conducted in 2016. Due to the shape of the plot a rectangular 18-ha subarea was selected for the analysis. It comprised 7,182 trees, which belonged to 25 species.

#### 4.3.2 *Inventory Data*

The inventory data was processed to calculate AGB of each individual tree. Based on species identity wood density values were assigned to each tree using the ForestGEO wood density database<sup>1</sup> and in the case of Paracou the Global Wood Density Database<sup>2</sup> (Chave et al., 2009; Zanne et al., 2009). If a wood density value was not available at

1 <http://ctfs.si.edu/Public/Datasets/CTFSWoodDensity/>

2 <https://datadryad.org/resource/doi:10.5061/dryad.234>

Table 4.1: Information on the study sites.

Site	Forest type	Size [ha]	Location	Year of inventory	Year of lidar scan	Basal area [m <sup>2</sup> ha <sup>-1</sup> ]
BCI	Neotropical moist	50	9.15° N 79.85° W	2010	2009	17.3 - 38.5
Paracou	Neotropical wet	85	5.27° N 52.92° W	2015	2015	24.8 - 38.7
Rabi	Afrotropical wet	25	1.92° S 9.88° E	2010 - 2012	2015	20.8 - 36.7
SERC	Nearctic temperate broadleaf	16	38.89° N 76.56° W	2014	2017	26.2 - 43.5
Traunstein	Palaearctic temperate mixed	18	47.94° N 12.67° E	2016	2016	7 - 44.6

species level, the median value of available species at higher taxonomic levels (genus, family) was used. For the few cases in which no related species were present, the overall median wood density at the site was assigned. From DBH the height of each tree was calculated using site specific asymptotic allometric relationships. These relationships were derived by fitting regression models of the Michaelis-Menten type (Equation 4.1) to the diameter height dataset from Jucker et al. (2017), grouped by biogeographical region and forest type.

$$H = \frac{h_{\max} \cdot D}{d_{1/2} + D} \quad (4.1)$$

This equation describes tree height  $H$  [m] as a function of DBH  $D$  [m] with two parameters:  $h_{\max}$ , which is the asymptotic maximal possible tree height, and  $d_{1/2}$ , which is the DBH of a tree that has reached a height of half of  $h_{\max}$ . A verification whether the derived models describe the DBH-height relations at each site reasonably well was done by plotting the curves together with the maximal observed DBH and maximal (lidar-derived) height of each hectare (Figure C.1). This showed that the relationships match the observed values for the BCI, Paracou, Rabi and Traunstein plot, but strongly underestimate the tree heights at SERC. Thus, for SERC we discarded the diameter-height relationship obtained from the dataset and instead obtained parameters by directly fitting a regression model to the data points in Figure C.1, representing the lidar-derived maximal heights on each hectare. All height allometry parameters used are listed in Table 4.2.

Aboveground biomass (AGB [t]) of each tree was calculated according to the general allometric equation suggested by Chave et al. (2014) (called "Model 5" in the original publication; Equation 4.2) with DBH  $D$  [m], height  $H$  [m] and wood density  $WD$  [t m<sup>-3</sup>].

$$AGB = 0.559 \cdot D^2 \cdot H \cdot WD \quad (4.2)$$

### 4.3.3 Lidar Data

Small footprint discrete return lidar data was collected in BCI in August 2009 using an Optech ALTM Gemini sensor (Lobo and Dalling, 2014), in Paracou in October 2015

Table 4.2: Parameters for the different diameter-height relationships modeled with a Michaelis-Menten equation.

Site	Species group	$h_{\max}$	$d_{1/2}$
BCI	All	57.4	0.43
Paracou	All	57.4	0.43
Rabi	All	59.9	0.48
SERC (discarded)	All	37	0.22
SERC (used instead)	All	54.7	0.27
Traunstein	Angiosperms	48.8	0.25
Traunstein	Gymnosperms	68.9	0.5

using an Riegl LMS Q 780, in Rabi in 2015 using a Riegl VQ-480i sensor (Labrière et al., 2018), at SERC in July 2017 using a Riegl VQ-480i sensor (Cook et al., 2013) and in Traunstein in August 2016 using a Riegl LMS Q 680i sensor. The lidar point clouds were terrain-normalized using LAStools (Isenburg, 2011) and rasterized to canopy height models (CHM) with 1-m resolution, by taking the height of the highest return in each 1-m<sup>2</sup>-cell (Figure 4.1). No interpolation was used and cells with no return were filled with value zero (ground height).

#### 4.3.4 Forest Structure Metrics

All four megaplots were divided into square-shaped subplots of 1-ha size each. At 1-ha scale a variety of structural metrics was calculated from 1) the inventory data and 2) the lidar data. Inventory-based metrics included basal area sum (BA), number of stems per ha (N), quadratic mean tree diameter (at breast height,  $D_g$ ), maximum DBH per ha ( $D_{\max}$ ), mean wood density weighted for tree basal area ( $WD_{BA}$ ) or tree aboveground volume ( $WD_{AGV}$ ) and stand density index (SDI, Equation 4.3), which is a standardized metric for stocking (Reineke, 1933).

$$SDI = N \cdot \left(\frac{25}{D_g}\right)^{-1.605} \quad (4.3)$$

Additionally to those metrics derived at 1-ha-level, we derived a set of metrics at site-level. Those site-level metrics included maximum basal area sum ( $BA_{s\max}$ ) and maximum SDI ( $SDI_{s\max}$ ) of all the 1-ha plots at each site  $s$  and basal area-weighted mean wood density ( $WD_{sBA}$ ) and aboveground volume-weighted mean wood density ( $WD_{sAGV}$ ) across all trees at each site  $s$ .

Lidar-based metrics were maximum canopy height per 1-ha plot  $H_{\max}$  and per site  $H_{s\max}$ , mean top-of-canopy height from CHMs of two different resolutions (1-m and 10-m pixels called  $TCH_1$  and  $TCH_{10}$ ), standard deviation of the 1-m CHM ( $SD_{CHM}$ ), coefficient of variation of the 1-m CHM ( $CV_{CHM}$ ) and Gini index of the 1-m CHM ( $Gini_{CHM}$ ). The vertical foliage profile (VFP) was derived from the vertical profile of the 1-m CHM following the approach described by Harding et al. (2001) (Equation 4.4).

$$VFP(h_i) = \frac{1}{k \cdot \Delta h} \cdot \ln\left(\frac{GP(h_i)}{GP(h_{i+1})}\right) \quad (4.4)$$

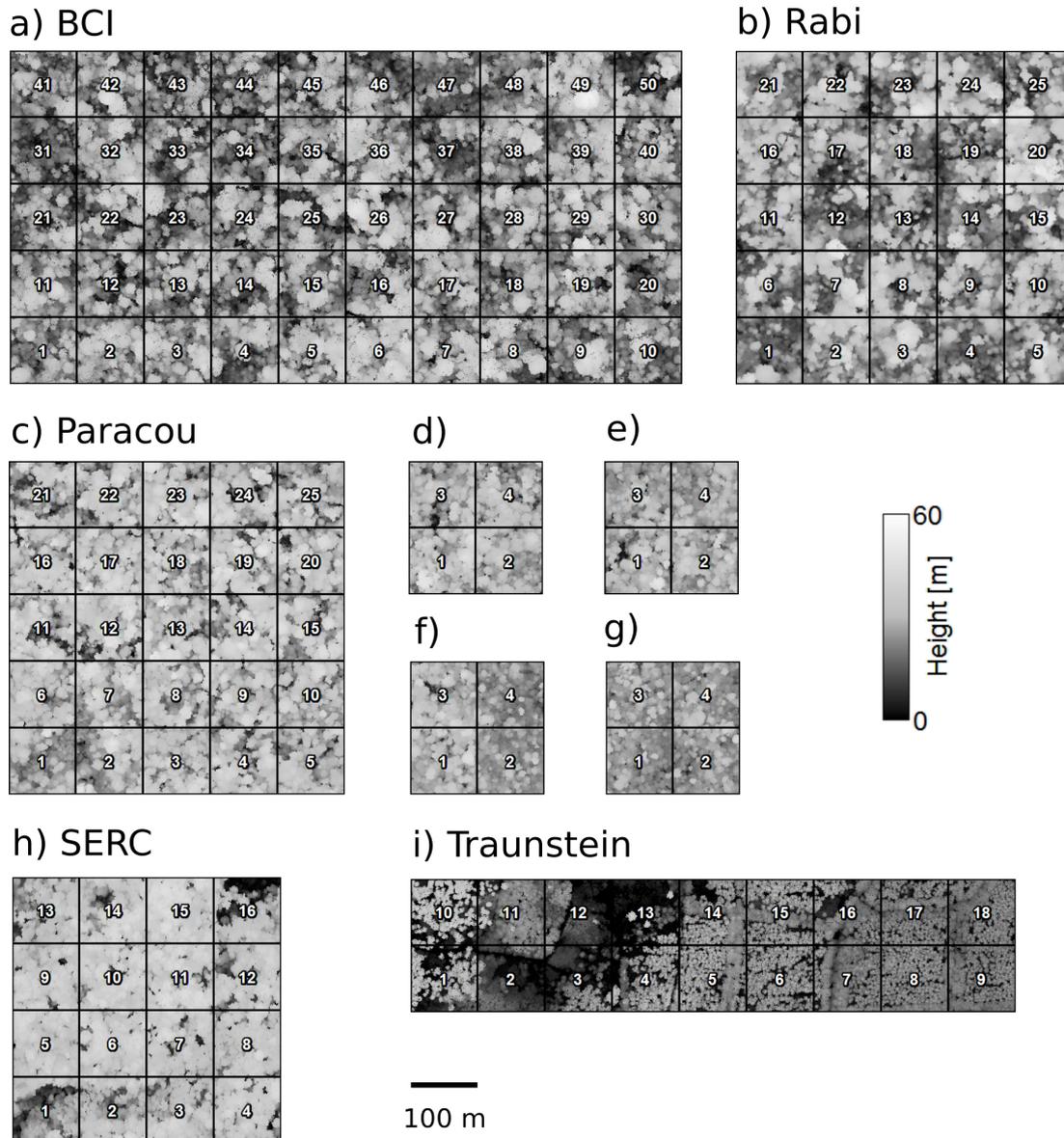


Figure 4.1: Canopy height models of the five study sites: a) Barro Colorado Island, b) Rabi, c) Paracou Plot 16 (biodiversity plot), d) Paracou Plot 1 (control plot), e) Paracou Plot 2 (selective logging), f) Paracou Plot 3 (selective logging and timber stand improvement), g) Paracou Plot 4 (selective logging, timber stand improvement and fuelwood collection), h) Smithsonian Environmental Research Center and i) Traunstein. The black grids and numbers represent the 1-ha subplots with each ha representing one record in the analysis.

with  $k$  being the light extinction coefficient,  $\Delta h$  the width of one height bin (here 1 m) and  $GP(h_i)$  the gap probability (value of the cumulative CHM profile) in height bin  $h_i$  (Ni-Meister et al., 2001). All pixels below 5 m height were regarded as ground and  $k$  was set to 0.3. The parameter  $k$  can be described as the quotient of a projection coefficient  $G$ , which is 0.5 for a random leaf angle distribution, and a clumping index  $C$ , which on average is 1.58 for different forest types (Tang et al., 2012). A value of  $k = 0.3$  has been shown to result in good LAI estimations (Getzin et al., 2017). The same vertical distribution metrics as for the CHMs were derived from the VFP, namely  $SD_{VFP}$ ,  $CV_{VFP}$  and  $Gini_{VFP}$ .

#### 4.3.5 Multivariate Regression Analysis

Regression analysis was conducted to find the best relationship and set of predictor variables for BA and AGB estimation with the main objective to minimize overall root mean squared error (RMSE) across sites. The regression models had the functional form of multivariate power laws. Each predictor was supposed to capture a different structural aspect of the forest. Several candidate metrics were grouped into sets of potential predictors and tested in different combinations. By categorizing metrics into different structural aspects it was avoided to test all possible metric combinations including redundant ones that capture the same structural aspect. For BA a four predictor equation was used (Equation 4.5). For AGB a five predictor equation was used (Equation 4.6).

$$BA = a_0 \cdot P_h^{a_h} \cdot P_d^{a_d} \cdot P_m^{a_m} \cdot P_v^{a_v} \quad (4.5)$$

$$AGB = b_0 \cdot P_h^{b_h} \cdot P_d^{b_d} \cdot P_m^{b_m} \cdot P_v^{b_v} \cdot P_w^{b_w} \quad (4.6)$$

Each  $P_x$  represents a predictor for a certain structural aspect  $x$  with  $a_x$  being the coefficients for BA estimation and  $b_x$  being the coefficients for AGB estimation. The predictor sets were defined as follows: The predictor for mean canopy height  $P_h$  contained  $TCH_1$  and  $TCH_{10}$  as possible variables. The maximum possible canopy height  $P_m$  was exclusively represented by  $H_{smax}$ . The predictor for maximum density (stocking) contained  $BA_{smax}$  and  $SDI_{smax}$ . For vertical heterogeneity of the canopy  $P_v$   $SD_{CHM}$ ,  $CV_{CHM}$ ,  $Gini_{CHM}$ ,  $SD_{VFP}$ ,  $CV_{VFP}$  and  $Gini_{VFP}$  have been explored. Average wood density  $P_w$  was only included in the AGB estimation and contained  $WD_{sBA}$  and  $WD_{sAGV}$ .

Maximum likelihood parameter estimation in R (R Development Core Team, 2014) was used to derive the coefficients for Equations 4.5 and 4.6. All possible metric combinations were tested including all possible subsets discarding one or several predictors  $P_x$ . The goodness-of-fit was evaluated based on linear regression of predictions against observations of the dependent variable to quantify  $R^2$ , RMSE and nRMSE (normalized RMSE by dividing it by the mean observed value). Wilcoxon tests were performed to check whether the mean prediction residual at each site deviated significantly from zero, with the goal of identifying unbiased prediction models. For each predictor combination 1,000 bootstrapping replicates were performed, by resampling the dataset randomly with replacement. Site-level metrics  $BA_{smax}$ ,  $SDI_{smax}$ ,  $H_{smax}$  and  $WD_{sAGV}$  were recalculated based on the resampled dataset, i.e., if the plot with the largest  $H_{max}$  of site  $s$  was not in the resampled dataset,  $H_{smax}$  was set to the largest  $H_{max}$  of any plots from site  $s$  present in the resampled dataset (likewise for the other metrics). Mean bootstrapped statistics ( $RMSE_b$ ,  $nRMSE_b$  and  $R^2_b$ ) served to evaluate the different models.

#### 4.3.6 Site-specific Reference Regression Models

Site-specific reference models were required to assess the performance of the derived general, site-independent, structure-based multi predictor regression models. For this purpose, single predictor regression models were fit. As predictors for these reference

models  $P_h$ , i.e.,  $TCH_1$  or  $TCH_{10}$ , were used. These metrics have been most widely used for this purpose (Asner and Mascaro, 2014; Knapp et al., 2018a). The models given by Equations 4.7 and 4.8 were fit by splitting the dataset into five subsets with each subset containing only records from one site  $s$  and using the same fitting procedure as described above.

$$BA = a_{0,s} \cdot P_h^{a_{h,s}} \quad (4.7)$$

$$AGB = b_{0,s} \cdot P_h^{b_{h,s}} \quad (4.8)$$

## 4.4 RESULTS

### 4.4.1 Forest Structure at Different Sites

The different structure attributes at 1-ha scale varied within and among the five sites (Figure 4.2 and Figure C.2). BA values ranged from 7 to 44.6  $\text{m}^2 \text{ha}^{-1}$  with a mean of 29.8  $\text{m}^2 \text{ha}^{-1}$ . AGB values ranged from 76 to 638  $\text{t ha}^{-1}$  with a mean of 354  $\text{t ha}^{-1}$ . Mean top-of-canopy height ranged from 5.6 to 38 m when calculated from 1 m  $\times$  1 m pixels ( $TCH_1$ ) and from 15.8 to 41.8 m when calculated from 10 m  $\times$  10 m pixels ( $TCH_{10}$ ). In both cases the distributions for the tropical sites were similar while TCH were on average higher at SERC and lower at Traunstein. Mean wood densities per hectare were calculated on a BA-weighted and on an AGV-weighted basis. Both were similar in their distributions and mean wood densities calculated for each site across the entire megaplot weighted by either BA or AGV were almost identical with the largest difference being 0.02  $\text{t m}^{-3}$  in the case of BCI ( $WD_{sBA}$ : BCI: 0.51  $\text{t m}^{-3}$ , Paracou: 0.69  $\text{t m}^{-3}$ , Rabi: 0.66  $\text{t m}^{-3}$ , SERC: 0.48  $\text{t m}^{-3}$ , Traunstein: 0.5  $\text{t m}^{-3}$ ;  $WD_{sAGV}$ : BCI: 0.49  $\text{t m}^{-3}$ , Paracou: 0.69  $\text{t m}^{-3}$ , Rabi: 0.66  $\text{t m}^{-3}$ , SERC: 0.47  $\text{t m}^{-3}$ , Traunstein: 0.5  $\text{t m}^{-3}$ ). Due to this similarity, only AGV-weighted wood density at site level  $WD_{sAGV}$  was considered in the further analysis. Mean wood densities at Paracou and Rabi exceeded the values from all other sites strongly with  $WD_{AGV}$  at the 1-ha scale ranging from 0.6 to 0.74  $\text{t m}^{-3}$  at Paracou and Rabi and from 0.42 to 0.55  $\text{t m}^{-3}$  at all other sites. Stand density index values ranged from 138 to 778 with the lowest values occurring in recently managed parts of Traunstein. The maximal SDIs are proxies for the highest possible stocking density in the different forest types (BCI: 683, Paracou: 749, Rabi: 703, SERC: 708, Traunstein: 778). The maximum canopy height covered a wide range from 27.1 m to 54.7 m.  $H_{\max}$  at SERC only covered a very narrow range falling inside the range of the tropical sites, while  $H_{\max}$  at Traunstein were much lower. The maximum canopy heights per site  $H_{s\max}$  were 54.7 m at BCI, 50 m at Paracou, 52.6 m at Rabi, 46.2 m at SERC and 40.3 m at Traunstein.

The vertical heterogeneity was measured in several different ways using standard deviation, coefficient of variation and Gini index of the canopy height model and the vertical foliage profile, respectively. CHM- and VFP-based vertical structure metrics showed quite different distribution patterns (Figure C.2). For the CHM-based metrics, Paracou and SERC showed the lowest values, due to a homogenous canopy surface, BCI and Rabi showed intermediate values, due to their rough canopy surface with large trees alternating with gaps and Traunstein showed (at least for CV and Gini index) the highest values, due to its heterogeneous structure composed of old and

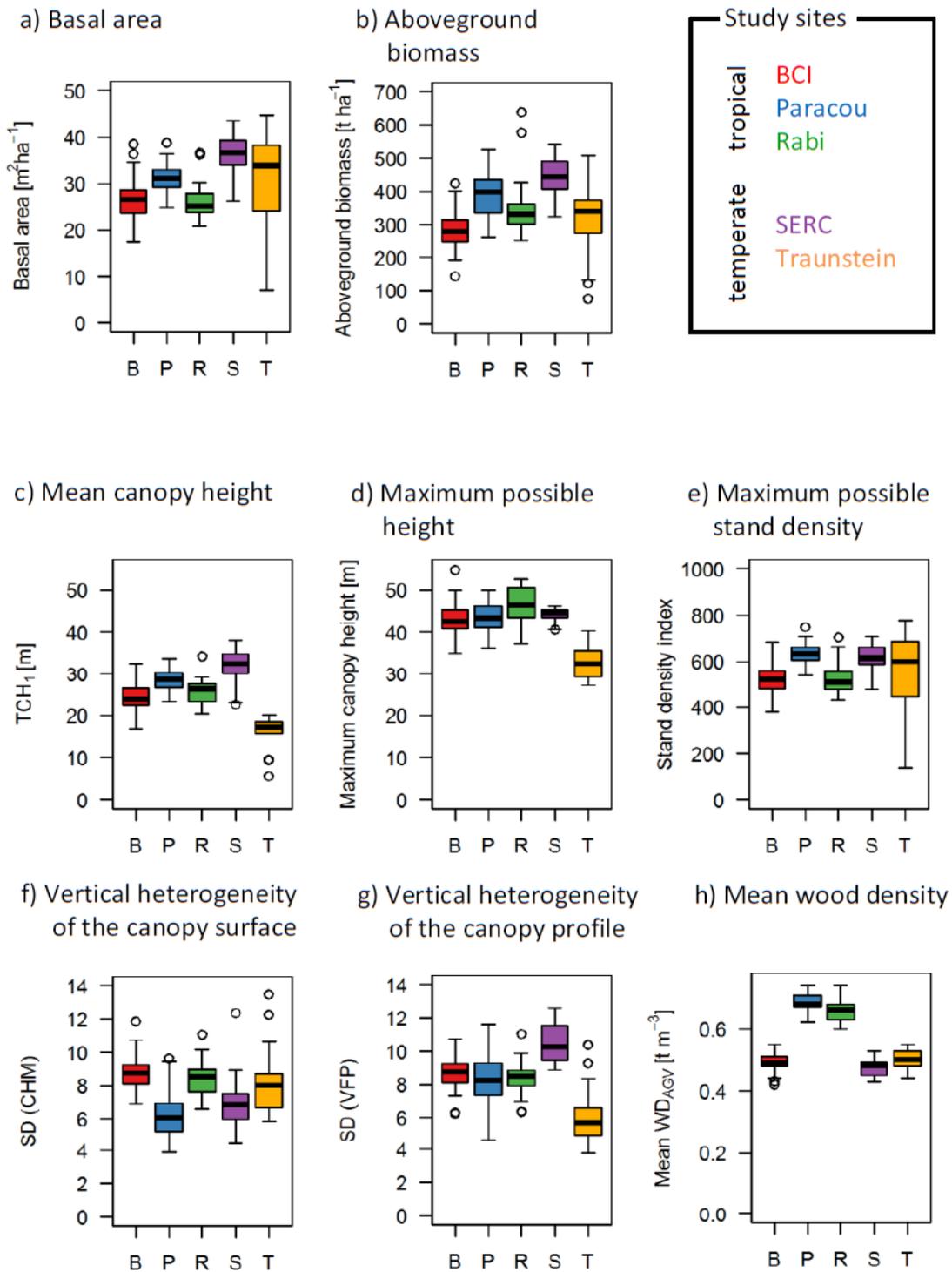


Figure 4.2: Boxplots of the distributions of a selection of forest structure metrics at 1-ha scale across the five study sites Barro Colorado Island (B), Paracou (P), Rabi (R), Smithsonian Environmental Research Center (S) and Traunstein (T). Graphics a) and b) depict the two target variables basal area and aboveground biomass. Graphics c) to h) depict six possible predictor variables.

young stands interrupted by forest roads. For the VFP- based vertical structure metrics, Traunstein showed the lowest values, which is in accordance with the fact that large parts of the plot are single-layered stands of different age, while the other sites showed higher values, indicating a more complex, multi-layered canopy.

#### 4.4.2 Site-specific Basal Area Estimation

Here, basal area was estimated from lidar using a single structural descriptor of stand height  $P_h$ . Mean top-of-canopy height at 1- and 10-m pixel resolution ( $TCH_1$  and  $TCH_{10}$ ) were tested as  $P_h$  to derive site-specific power law coefficients ( $a_{0,site}$  and  $a_{h,site}$ ). Coefficients for each site are listed in Table C.1 and a scatterplot with site-specific curves is displayed in Figure 4.3. The following goodness-of-fit statistics were derived across all sites using the site-specific relationships: The  $TCH_1$ -based basal area predictions resulted in  $RMSE = 2.5 \text{ m}^2 \text{ ha}^{-1}$  (8.3%) and  $R^2 = 0.79$ . The  $TCH_{10}$ -based basal area predictions resulted in  $RMSE = 2.8 \text{ m}^2 \text{ ha}^{-1}$  (9.5%) and  $R^2 = 0.73$ . In both cases, the mean residuals were not significantly different from zero at any site (Wilcoxon tests with the smallest p-value of all sites being  $p = 0.54$ ).

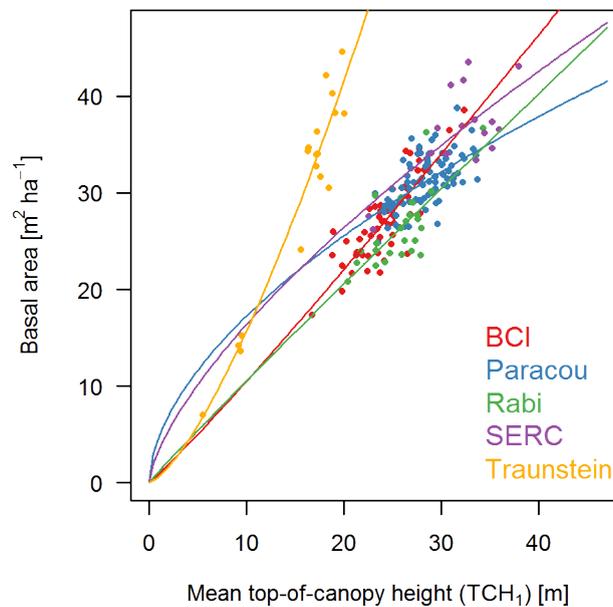


Figure 4.3: Site-specific relationships (power laws) between basal area and  $TCH_1$  with each point representing 1 ha.

#### 4.4.3 Generalized Basal Area Estimation

Here, basal area was estimated using several structural descriptors from lidar, which were supposed to capture different aspects of forest structure (Equation 4.5). In total, 125 models consisting of different descriptors and metrics were analyzed. The best models found are listed in Table 4.3. The models are ranked according to increasing mean bootstrapped  $RMSE_b$ . The listed models represent the best overall model and

the best models with certain structural descriptors  $P_x$  removed. For the different  $P_x$  in most cases the same metrics were selected. For  $P_h$  mostly  $TCH_{10}$  was selected and  $TCH_1$  only occurred once. For  $P_d$  always the site specific maximal basal area  $BA_{smax}$  was selected. For  $P_m$  the site specific maximal tree height  $H_{smax}$  was the only available metric. For  $P_v$ , however, four different metrics appear in the list of best models, namely  $SD_{VFP}$ ,  $Gini_{VFP}$ ,  $CV_{VFP}$  and  $CV_{CHM}$ . The overall best model was one using all four structural descriptors (Equation 4.9; nRMSE = 9.8%).

$$BA = 9.2 \cdot TCH_{10}^{1.3} \cdot BA_{smax}^{0.359} \cdot H_{smax}^{-1.03} \cdot SD_{VFP}^{-0.305} \quad (4.9)$$

The goodness-of-fit decreased only marginally if information on maximal possible height ( $P_m$ ) was excluded from the predictors (nRMSE = 10%). The third best model was a two-predictor model using only mean current and maximal possible canopy height ( $P_h$  and  $P_m$ , nRMSE = 10.9%). Hence, there was no other three-predictor model that could exceed this two predictor model in accuracy. It was followed by a model using mean canopy height and maximal possible stand density ( $P_h$  and  $P_d$ , nRMSE = 11.5%), and one using mean canopy height and vertical heterogeneity ( $P_h$  and  $P_v$ , nRMSE = 12.2%). At the lower end, the best model making no use of current mean canopy height (no  $P_h$ ) was somewhat better (nRMSE = 14.2%) than the one using exclusively mean canopy height ( $P_h = TCH_{10}$ , nRMSE = 14.6%). Thus, adding any structural descriptor decreased nRMSE by at least 2.4% compared to a model purely based on canopy height.

The additional goal was finding a relationship that is unbiased across all sites. According to the Wilcoxon tests predictions of the best found model were slightly but significantly biased for BCI ( $p = 0.0036$ ) and Paracou ( $p = 0.05$ ), whereas the second best model did not show any significant bias for any of the sites (Table C.2). Figure 4.4 shows the 1:1 plots for site-specific  $TCH_1$ -based predictions (a), the generalized  $TCH_{10}$ -based predictions (c) and the predictions using the best model based on structural descriptors (e). Figure 4.4 also shows the residual distributions resulting from each of the three predictions for the different sites (b, d, f).

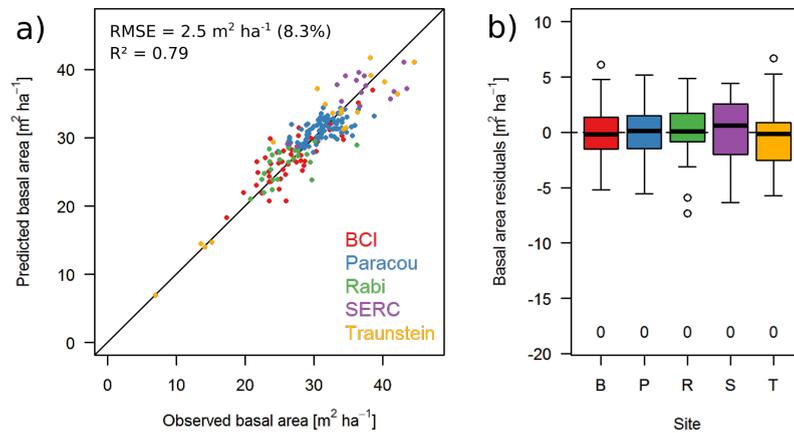
#### 4.4.4 Site-specific Aboveground Biomass Estimation

Analogous to BA, AGB was modeled as a power law function of mean canopy height  $P_h$ , using either  $TCH_1$  or  $TCH_{10}$  by fitting site-specific coefficients ( $b_{0,site}$  and  $b_{h,site}$ ; Table C.1). Applying these site-specific relationships, the following goodness-of-fit statistics were derived across all sites: The  $TCH_1$ -based AGB predictions resulted in RMSE = 41 t ha<sup>-1</sup> (11.6%) and  $R^2 = 0.78$ . The  $TCH_{10}$ -based AGB predictions resulted in RMSE = 41.8 t ha<sup>-1</sup> (11.8%) and  $R^2 = 0.77$ . In both cases, the mean prediction residuals were not significantly different from zero at all sites (Wilcoxon tests with the smallest p-value of all sites being  $p = 0.39$ ).

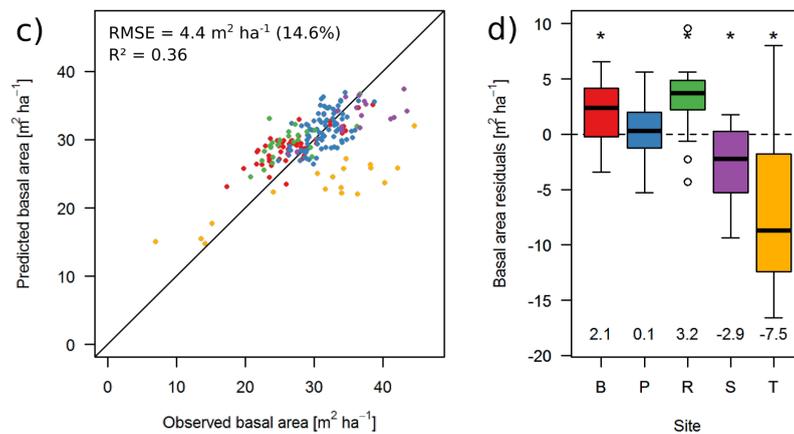
#### 4.4.5 Generalized Aboveground Biomass Estimation

To derive a generalized AGB estimation model the same structural descriptors as for basal area were used. Additionally, a fifth descriptor for the average wood density  $P_w$  was introduced, which resulted in 251 models in total. Table 4.4 lists the best models found for different combinations of structural descriptors. The models are ranked

## Site-specific single predictor model



## General single predictor model



## General multi predictor model

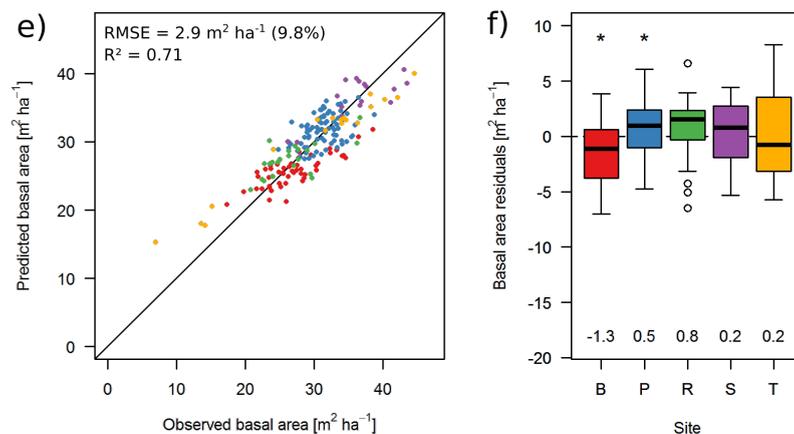


Figure 4.4: Scatterplots of predicted basal area against observed basal area using a) a site-specific single predictor model, c) a general single predictor model (based on  $TCH_{10}$ ) and e) the best general multi predictor model (Equation 4.9). The boxplots on the right hand side show the distribution (quartiles) of prediction residuals at each site with numbers below displaying the mean residual value (bias) and asterisks above indicating whether the means deviate significantly from zero (b, d, f).

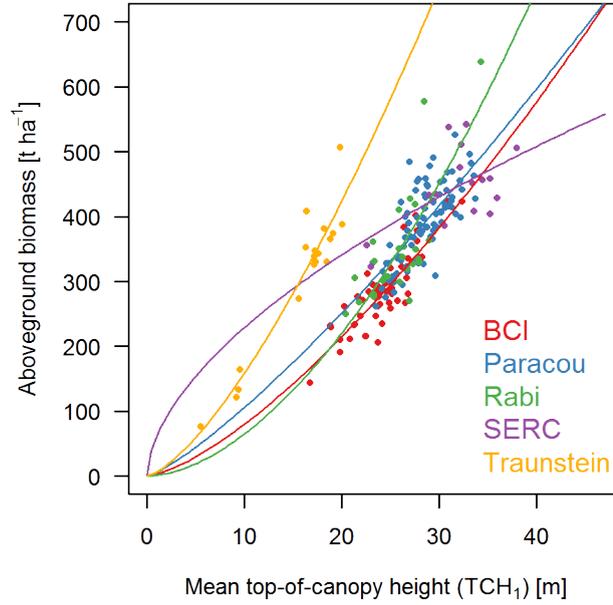


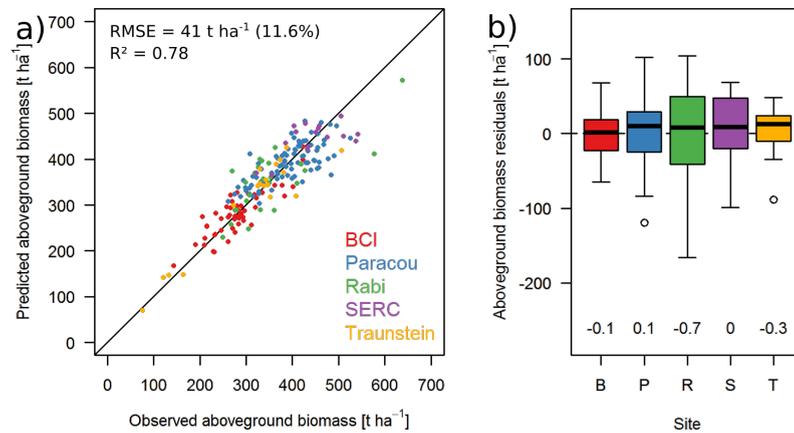
Figure 4.5: Site-specific relationships (power laws) between aboveground biomass and  $TCH_1$  with each point representing 1 ha.

according to increasing  $RMSE_b$  (derived from bootstrapping). The listed models represent the best overall model and the best models with a reduced number of structural descriptors  $P_x$ . Compared to the basal area estimation more descriptor combinations are possible due to the additional parameter  $P_w$ .  $TCH_{10}$  was selected in most cases for  $P_h$ . For  $P_d$  maximal basal area  $BA_{smax}$  and maximal stand density index  $SDI_{smax}$  per site do appear in the models. For  $P_v$  three different metrics have been selected:  $SD_{CHM}$ ,  $SD_{VFP}$  and  $CV_{VFP}$ . In the case of aboveground biomass estimation the best model was the one using all five available structural descriptors (Equation 4.10;  $nRMSE = 12.4\%$ ).

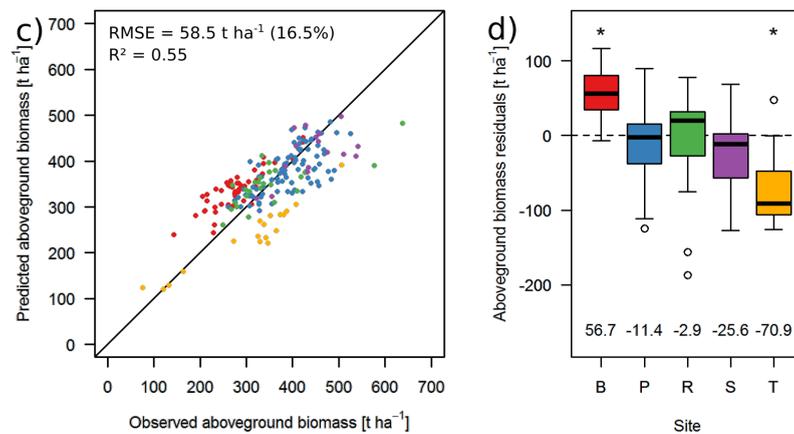
$$AGB = 1.92 \cdot TCH_1^1 \cdot SDI_{smax}^{0.979} \cdot H_{smax}^{-1.24} \cdot SD_{CHM}^{0.212} \cdot WD_{sAGV}^{0.0838} \quad (4.10)$$

Leaving either  $P_v$ ,  $P_w$  or both aside increased the  $nRMSE$  by around 1%. The best model that did not rely on any site related ground-based information was the one using only  $P_h$  and  $P_m$  ( $nRMSE = 13.5\%$ ). Table 4.4 documents the results for other descriptor combinations. E.g., a naïve single predictor model based on  $TCH_{10}$  only had  $nRMSE = 16.5\%$ .  $P_d$  and  $P_m$  were more important than  $P_v$  (according to their presence in the best models), which is different from the basal area estimation where  $P_v$  was more important than  $P_d$  and  $P_m$ . The best model without  $P_h$  had a  $nRMSE = 17.8\%$ .

## Site-specific single predictor model



## General single predictor model



## General multi predictor model

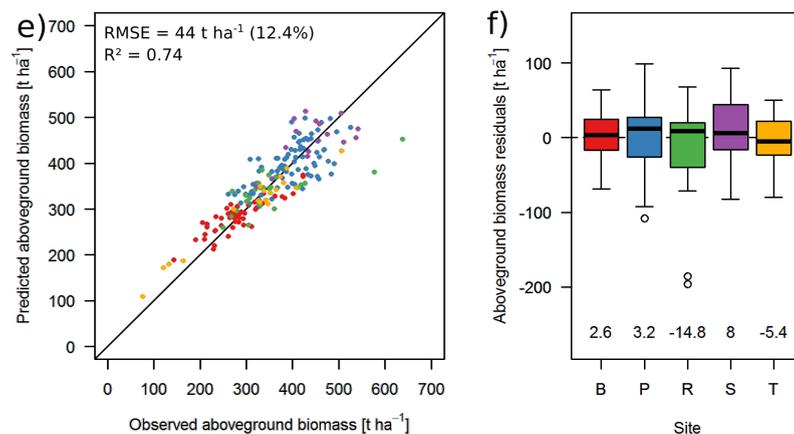


Figure 4.6: Scatterplots of predicted aboveground biomass against observed aboveground biomass using a) a site-specific single predictor model, c) a general single predictor model (based on  $TCH_{10}$ ) and e) the best general multi predictor model (Equation 4.10). The boxplots on the right hand side show the distribution (quartiles) of prediction residuals at each site with numbers below displaying the mean residual value (bias) and asterisks above indicating whether the means deviate significantly from zero (b, d, f).

Table 4.3: The best basal area estimation models for different predictor combinations ranked for increasing mean bootstrapping root mean squared error (RMSE<sub>b</sub>). For explanation of the columns names and variable names please refer to the main text.

Mean canopy height $P_h$	Max. stand density $P_d$	Max. canopy height $P_m$	Vertical hetero-geneity $P_v$	RMSE	nRMSE	$R^2$	RMSE <sub>b</sub>	nRMSE <sub>b</sub>	$R^2_b$	Bias
TCH <sub>10</sub>	BA <sub>smax</sub>	H <sub>smax</sub>	SD <sub>VFP</sub>	2.9	9.8%	0.71	2.9	9.6%	0.72	yes
TCH <sub>10</sub>	BA <sub>smax</sub>	-	Gini <sub>VFP</sub>	3.0	10.0%	0.70	3.1	10.4%	0.67	no
TCH <sub>1</sub>	-	H <sub>smax</sub>	-	3.3	10.9%	0.64	3.1	10.4%	0.67	yes
TCH <sub>10</sub>	BA <sub>smax</sub>	-	-	3.4	11.5%	0.60	3.6	12.0%	0.56	yes
TCH <sub>10</sub>	-	-	CV <sub>VFP</sub>	3.6	12.2%	0.55	3.6	12.0%	0.56	yes
-	BA <sub>smax</sub>	-	CV <sub>CHM</sub>	4.2	14.2%	0.40	4.2	14.1%	0.39	no
TCH <sub>10</sub>	-	-	-	4.4	14.6%	0.36	4.3	14.5%	0.36	yes

Table 4.4: The best aboveground biomass estimation models for different predictor combinations ranked for increasing mean bootstrapping root mean squared error (RMSE<sub>b</sub>). For explanation of the columns names and variable names please refer to the main text.

Mean canopy height $P_h$	Max. stand density $P_d$	Max. canopy height $P_m$	Vertical hetero-geneity $P_v$	Mean wood density $P_w$	RMSE	nRMSE	$R^2$	RMSE <sub>b</sub>	nRMSE <sub>b</sub>	$R^2_b$	Bias
TCH <sub>1</sub>	SDI <sub>smax</sub>	H <sub>smax</sub>	SD <sub>CHM</sub>	WD <sub>sAGV</sub>	44.0	12.4%	0.74	46.1	13.0%	0.71	no
TCH <sub>10</sub>	BA <sub>smax</sub>	H <sub>smax</sub>	SD <sub>VFP</sub>	-	47.4	13.4%	0.70	47.9	13.5%	0.69	yes
TCH <sub>10</sub>	SDI <sub>smax</sub>	H <sub>smax</sub>	-	WD <sub>sAGV</sub>	45.6	12.9%	0.73	48.5	13.7%	0.69	no
TCH <sub>10</sub>	SDI <sub>smax</sub>	H <sub>smax</sub>	-	-	46.0	13.0%	0.72	48.5	13.7%	0.69	yes
TCH <sub>1</sub>	-	H <sub>smax</sub>	-	-	47.7	13.5%	0.70	48.9	13.8%	0.68	yes
TCH <sub>10</sub>	BA <sub>smax</sub>	-	CV <sub>VFP</sub>	WD <sub>sAGV</sub>	51.2	14.5%	0.67	49.7	14.0%	0.67	yes
TCH <sub>10</sub>	BA <sub>smax</sub>	-	-	WD <sub>sAGV</sub>	60.1	17.0%	0.55	50.6	14.3%	0.66	yes
TCH <sub>10</sub>	SDI <sub>smax</sub>	-	-	-	50.9	14.4%	0.66	50.7	14.3%	0.66	yes
TCH <sub>10</sub>	-	-	CV <sub>VFP</sub>	-	54.3	15.3%	0.61	53.9	15.2%	0.61	yes
TCH <sub>10</sub>	-	-	-	WD <sub>sAGV</sub>	57.2	16.1%	0.57	57.2	16.1%	0.57	yes
TCH <sub>10</sub>	-	-	-	-	58.5	16.5%	0.55	58.3	16.4%	0.55	yes
-	BA <sub>smax</sub>	-	SD <sub>VFP</sub>	-	63.0	17.8%	0.48	62.2	17.5%	0.49	yes

With regard to the goal of finding a relationship that is unbiased across all sites, the Wilcoxon tests identified two models for which the mean of residuals at none of the single sites differed significantly from zero (Table C.3): the models in rows 1 and 3 in Table 4.4. For all other models predictions were biased for at least one site. Figure 4.6 shows the 1:1 plots for site-specific TCH<sub>1</sub>-based predictions (a), the generalized TCH<sub>10</sub>-based predictions (c) and the structure guided predictions using the best model (e). Figure 4.6 also shows the residual distributions resulting from each of the three predictions for the different sites (b, d, f).

#### 4.4.6 Comparison of Results

Overall, achieved relative errors in BA estimation were somewhat lower than the ones for AGB estimation. The exclusion of different structural descriptors led to an increase in estimation errors. Figure 4.7 shows the obtained nRMSE for different sets of  $P_x$  in comparison to the nRMSE of site-specific estimations. For BA estimation, the best unbiased generic model required four coefficients and resulted in a nRMSE of 10.4%, which is 2.1% higher than the nRMSE of 8.3% obtained from five site-specific models, requiring ten (two per site) coefficients. For AGB estimation, the best unbiased generic model required six coefficients and resulted in a nRMSE of 13%, which is 1.4% higher than the nRMSE of 11.6% obtained from five site-specific models, requiring ten coefficients.

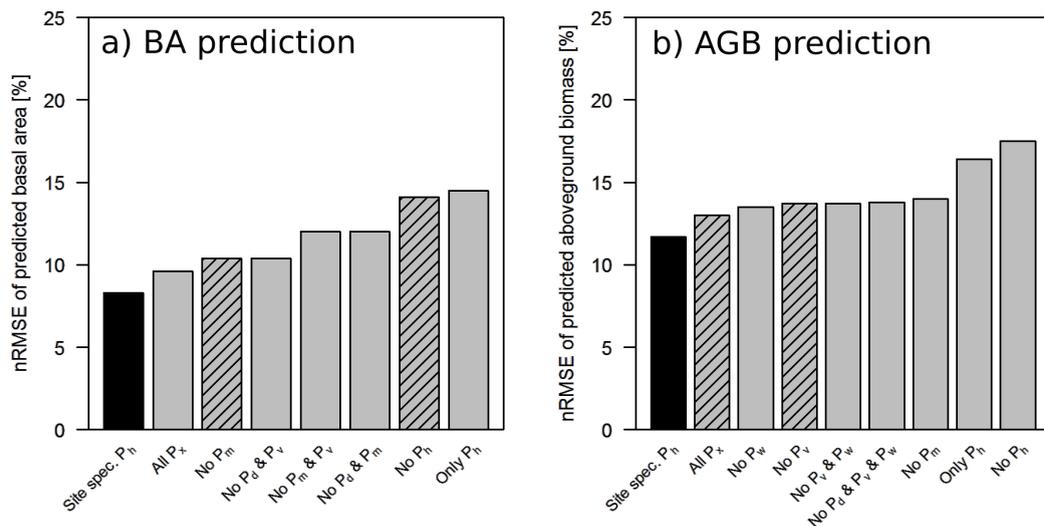


Figure 4.7: Summary of how the exclusions of certain structural descriptors  $P_x$  influence the normalized root mean squared error (nRMSE) of basal area (a) and aboveground biomass estimation (b). The black bar represents the site-specific reference model. The grey bars represent mean bootstrapping nRMSE of the different generic models. The striped bars mark the models which produce unbiased predictions at all sites. For the meaning of the indices of the predictors please refer to the main text.

## 4.5 DISCUSSION

The goal of this study was to determine a set of forest structure metrics that can be used for BA and AGB estimation from CHMs at very distinct forest sites, which belong to different biomes. It could be shown that a combination of four metrics capturing mean canopy height, maximal stand density, maximal canopy height and vertical heterogeneity could estimate BA using a generic model across all sites with a high accuracy, which was almost as good as the accuracy achieved by site-specific models. The accuracy for AGB estimation was slightly weaker than the one for BA estimation.

### 4.5.1 *The Role of Mean Canopy Height*

It was found that the mean canopy height ( $P_h$ ), represented here by  $TCH_1$  and  $TCH_{10}$ , was the most important predictor variable, which is in support of its wide use in previous studies (Lefsky et al., 2002b; Asner and Mascaro, 2014; Duncanson et al., 2015). It was important in BA and AGB estimation, with accuracies decreasing considerably when  $P_h$  was dropped from the models. Despite the mathematical simplicity of TCH (the mean height of all CHM pixels) it is a quite comprehensive metric capturing much of the forest structure in a single number. It is influenced by the heights and crown sizes of the trees (which contribute to the CHM) and therefore closely related to Lorey's height (BA-weighted mean tree height; Asner and Mascaro, 2014). However, TCH also provides information on horizontal vegetation density, if ground pixels, e.g., in canopy gaps, are included in its computation (Lu et al., 2014). There have been studies that tried to separate the "height" and "density" aspect of TCH by calculating mean canopy height only from canopy pixels (and excluding ground pixels) and capturing horizontal vegetation density as fractional canopy cover, i.e., the relative proportion of canopy pixels above an arbitrary height threshold, or its inverse, the gap fraction (Bouvier et al., 2015). It has also been shown that fractional canopy cover alone can predict AGB in tropical forests quite well over a range of canopy height thresholds (Meyer et al., 2018).

It was found that  $TCH_{10}$  derived from a rough 10-m-pixel CHM often performed better than  $TCH_1$  derived from 1-m pixels. This has been observed also in an earlier study at BCI using TCH in single predictor models (Knapp et al., 2018a). It might be explained by the ability of  $TCH_{10}$  to capture the canopy structure of the large trees, which contribute most to BA and AGB, and the larger gaps where such trees are missing.  $TCH_1$  includes more detail and is influenced by the structure of individual tree crowns and small gaps within and between crowns, which might not be relevant or even counterproductive for estimating BA and AGB. In particular, in the context of generalization between different forest types it might be beneficial that  $TCH_{10}$  "ignores" differences in crown shapes.

### 4.5.2 *The Role of Stand Density*

Maximal stand density per site ( $P_d$ ) was of high importance for AGB and BA estimation. Only one in the best eight AGB models did not contain  $P_d$ . Asner and Mascaro (2014) pointed out that for many sites BA shows a linear relationship with TCH, but with considerable differences in the slopes, which was therefore an important term in their AGB estimation model. Differences in this relationship can be expected because

at different sites different tree species may occur, which have different geometries, in particular regarding the relationships between DBH and height and DBH and crown diameter. Of two stands with the same canopy height, one may contain trees with slender crowns and has a much higher stocking than the other one containing trees with wide crowns. We tried to reduce the necessary information about density as much as possible by only using the maximum observed value per site. As this parameter is not derived from remote sensing data, there is either inventory data or expert knowledge required. As metrics for  $P_d$  either  $BA_{smax}$  or  $SDI_{smax}$  were used. The two are independent from each other: The highest SDI identifies the stand with the highest stocking according to the self-thinning rule (Reineke, 1933), which is not necessarily the stand with the highest current BA (Figure C.3). Among the five sites investigated, the tropical sites had lower  $BA_{smax}$  (38.5, 38.7, and 36.7  $m^2 ha^{-1}$ ) than the temperate sites (43.5 and 44.6  $m^2 ha^{-1}$ ).  $SDI_{smax}$ , however, was similar at BCI, Rabi and SERC (683, 703, 708) and somewhat higher at Paracou (749) and Traunstein (778). As shown, either of the two metrics could improve the AGB estimation in comparison to the case of missing  $P_d$ .

#### 4.5.3 *The Role of Maximum Height*

An inclusion of maximum possible height ( $P_m$ ) was expected to improve estimation models. The reason behind is the same as for stand density, namely the possibility of regionally different DBH-height relationships of trees, that lead to differences in the maximum possible canopy height.  $P_m$  can be easily extracted from remote sensing data (in contrast to  $P_d$ ). Here, the maximum observed canopy height  $H_{smax}$  (in the CHM) was used under the assumption that the plots are large enough to be representative for the maximum possible tree height in the respective forest types. Maximum height showed no relevance in BA estimation, but prediction errors for AGB increased from 11.8% to 14.8% if it was dropped from the model.  $H_{smax}$  might act as a standardization for TCH. Site-specific relationships between TCH and AGB (Figure 4.5) show that, e.g., the forest at Traunstein reaches high AGB values at much lower mean canopy heights than other sites. By additionally providing the information that also  $H_{smax}$  at Traunstein is lower than elsewhere, the TCH values are put into the perspective of how high is the forest now and how high could it possibly become. This standardization role of  $H_{smax}$  is supported by the fact that all selected models have negative coefficients ( $a_m, b_m$ ), commonly close to -1 for  $H_{smax}$ , and positive coefficients ( $a_h, b_h$ ), commonly close to +1 for TCH, i.e., the ratio  $TCH / H_{smax}$  is used in the predictions. Models based on  $TCH_1$  and  $H_{smax}$  also were the best two-predictor models in BA and AGB estimation, respectively.

#### 4.5.4 *The Role of Vertical Heterogeneity*

Vertical heterogeneity ( $P_v$ ) was after mean canopy height ( $P_h$ ) the second descriptor derived at individual plot-level rather than site-level. It was included in the best BA and AGB models, however dropping it increased the nRMSE by less than 1%. From the six candidate metrics for  $P_v$ , various were chosen in different models.

The calculation of the vertical metrics was either based on the canopy height model or on the vertical foliage profile. As visible in Figure C.2, the distributions of SD, CV and Gini index differed strongly depending on whether they were CHM- or VFP-

based. CHM-based variability metrics describe the heterogeneity of the canopy surface, including ground pixels, i.e., canopy gaps. VFP-based variability metrics describe the vertical layering of the reconstructed foliage profile, which does not contain any ground component, but up-weights profile parts in the lower heights to compensate for the occlusion by high trees. Hence, their contributions to BA and AGB estimation might be different: CHM-based metrics rather characterize forests in the spectrum from smooth canopy surfaces, as observed for young, dense stands, to rough canopy surfaces, as observed for old or disturbed stands. VFP-based metrics rather account for the overseen trees in the lower canopy. Other studies have also identified the vertical heterogeneity as a component in prediction models. Magnussen et al. (2012) proposed a two-predictor model based on 1) TCH and 2) the variance of the CHM divided by TCH, which is closely related to  $CV_{CHM}$  used here. Bouvier et al. (2015) considered two vertical metrics in their four-predictor model: 1) variance of the CHM and 2) CV of the leaf area density in the VFP. To conclude, there is a wide variety of metrics that characterize vertical heterogeneity and they may in fact capture quite different aspects of forest structure. They do contribute in the improvement and generalization of BA and AGB estimation. Future analyses should try to achieve a better understanding of how the different metrics are related to ground-based metrics of forest structure, and whether a combination of several of them could further improve estimation results.

#### 4.5.5 *The Role of Wood Density*

Regional differences in average wood density have been suspected to be a main reason behind differences in the height-to-biomass relationship of forests (Asner and Mascaro, 2014; Vincent et al., 2014; Meyer et al., 2018). In our analysis, however, dropping the wood density parameter ( $P_w$ ) led only to a slight increase in nRMSE of less than 1% for AGB estimations. The values of  $WD_{sAGV}$  were very similar for BCI, SERC and Traunstein, but considerably higher for Paracou and Rabi. If region-specific estimates on average wood density are available they should definitely be considered in AGB estimation models. Nevertheless, our results suggest that compared to other parameters wood density is of minor importance for a generalized AGB estimation. With regard to the question, how average wood density should be calculated, Vincent et al. (2014) argued to use AGV instead of BA as a weighting variable, as AGV of the trees in the ground-truth plots has to be calculated anyway to derive tree AGB, and AGV is the structurally more appropriate weighting variable compared to BA. In this study, for all five study sites  $WD_{sAGV}$  and  $WD_{sBA}$  were found to be very similar. Thus, only  $WD_{sAGV}$  was further used in the analyses.

#### 4.5.6 *Generalization and Outlook*

With the identified structural variables and the fitted coefficients we propose general prediction models for BA and AGB estimation in different forest types. Having such models and also understanding the contribution of different forest structural aspects is important for consistent large scale mapping of forest carbon stocks (Lefsky et al., 2002b). This is particularly relevant for upcoming spaceborne missions such as GEDI (Stavros et al., 2017), BIOMASS (Le Toan et al., 2011) or Tandem-L (Moreira et al., 2015) which will provide consistent forest height measurements across very different forest types, not all of which are represented sufficiently in ground-truth datasets.

As a next step, the proposed relationships need to be tested at other forest sites to either confirm or, if necessary, adapt them. Further, the influence of spatial scale needs to be investigated, as different sensors produce measurements at different scales. Finally, methods need to be developed for acquiring more of the structural variables entirely from remote sensing and becoming independent from any ground-based input. Individual tree delineation from high resolution canopy height data can be applied to derive stand density information directly from remote sensing (Duncanson et al., 2015; Ferraz et al., 2016). Average wood density can be estimated based on forest type or even species classification using passive optical remote sensing (Fassnacht et al., 2016). These technologies have to be combined to derive very detailed estimates from airborne acquisitions at landscape scale. The estimates can then serve as training areas for wall-to-wall mapping using spaceborne products.

#### 4.6 CONCLUSION

Data from temperate and tropical forest plots were combined to develop a general equation for biomass (and basal area) estimation based on a set of forest structure metrics from remote sensing. The different structural dimensions were a priori defined. The results provided insight in the relative importance of mean and maximal canopy height, stand density, vertical heterogeneity and wood density for biomass estimation. Not all of those forest attributes can be derived from lidar data. For maximal stand density and mean wood density field-based information is required at the site level. Alternatively, a model without those attributes can be chosen from the list of models, at the expense of slightly lower prediction accuracies. The found relationships should provide guidance towards a standardized workflow for estimating aboveground biomass for forest carbon mapping and monitoring from remote sensing.

#### 4.7 ACKNOWLEDGMENTS

This study would not have been possible without the work of large number of people collecting data in the field and conducting airborne lidar campaigns. We thank all of them for providing access to their data and making this comparative study possible. In particular we thank the Smithsonian Tropical Research Institute for providing the census data for BCI. The BCI forest dynamics research project was founded by S.P. Hubbell and R.B. Foster and is now managed by R. Condit, S. Lao, and R. Perez under the Center for Tropical Forest Science and the Smithsonian Tropical Research in Panama. Numerous organizations have provided funding, principally the U.S. National Science Foundation, and hundreds of field workers have contributed. We thank J. Dalling for providing the lidar data from BCI. We thank B. Hérault, G. Derroire and A. Dourdain for providing census data and G. Vincent for providing the lidar data from Paracou. The Rabi data sets were collected during the AfriSAR campaign and we thank A. Alonso and H. Memiaghe for providing the census data and S. Saatchi for providing the lidar data. The Rabi plot is a collaborative project of CENAREST, the Center for Conservation and Sustainability (CCS) and the Smithsonian Institution Forest Global Earth Observatory (ForestGEO). We thank G. Parker, S. McMahon and S. Davis for providing census data from SERC. These data were gathered as part of forest ecology studies at the Smithsonian Environmental Research Center. SERC is a participant in the ForestGEO network. We thank B. Cook and D. Morton for providing

the lidar data for SERC, which was collected during a G-LiHT campaign. We thank H. Pretzsch, P. Biber, M. Heym and the numerous field workers for providing the census data from Traunstein. Traunstein is a participant in the ForestGEO network. We thank K. Papathanassiou, F. Kugler, V. Cazcarra-Bes, M. Tello and M. Pardini for providing the lidar data from Traunstein. We further thank the biosphere members within the HGF-Helmholtz Alliance “Remote Sensing and Earth System Dynamics” for many fruitful discussions on forest structure. This study was conducted with funding by the German Federal Ministry for Economic Affairs and Energy (BMWi) under the funding reference 50EE1416. R.F., K.P. and A.H. were supported by the HGF-Helmholtz Alliance “Remote Sensing and Earth System Dynamics” HA-310 under the funding reference RA37012.

## SYNTHESIS AND OUTLOOK

---

## 5.1 MAIN RESULTS AND LIMITATIONS

As a dominant component of global land cover and the carbon budget forests play a significant role in the earth and climate system. Today, they face multiple threats, primarily due to human influence, but also increasingly due to climate change. If we are to conserve global forest covers, it is imperative to develop the tools to better analyze and understand their dynamics. Field inventories, i.e., sample based approaches, alone cannot fulfill the requirements to investigate all aspects. In this thesis methods from the disciplines of remote sensing and forest modeling were combined and new approaches developed. These novel approaches are contributing to the overall goals of enabling large scale measurements of forest attributes and improving the mechanistic understanding behind observations of forest structure.

The challenges which were met in this thesis were the following: 1) development of a lidar simulator and its integration into a forest model; 2) use of lidar and forest simulations to explore biomass stock estimation at different scales and under different disturbance regimes; 3) simulation of forest dynamics to explore possibilities for remote sensing of biomass change over time and 4) combination of forest structure descriptors for generalizing biomass estimations across different forest types from different continents. In the following the main results of each of these tasks are summarized and discussed including also limitations.

### 5.1.1 *Integrating Remote Sensing Simulations in a Forest Model*

Forest gap models like FORMIND produce large datasets. They can cover hundreds of hectares over hundreds of time steps and consist of millions of trees. Simulating lidar point clouds for such datasets requires storage capacities comparable to real airborne lidar campaigns at regional scale ( $\sim 50$  to  $1,000$  MB  $\text{km}^{-2}$ ) and an efficient algorithm and code. Hence, detailed ray-tracing algorithms that act at the leaf and branch level (Disney et al., 2010) are unsuitable for simulations of such large forest areas. The voxel-based lidar simulation approach developed in this thesis has a level of complexity to capture all the important patterns obtained from airborne lidar data while being quickly applicable for large forest stands. Producing point clouds which mimic airborne lidar data ensures the flexibility to aggregate the data along different dimensions (e.g., vertical profiles, canopy height models, arbitrary distribution metrics) at any desired spatial resolution. This is a great advantage compared to earlier studies, which used canopy height as a forest model output at a model-given spatial resolution (Hurtt et al., 2004; Köhler and Huth, 2010). Additionally, the simulated lidar products can be compared one-to-one to their real-world equivalents.

The benefit of linking the use of forest and lidar simulations for understanding and interpreting remote sensing observations has been demonstrated in this thesis. An aspect which did not get as much attention in this work, but bears enormous potential is the use of forward remote sensing simulations in calibration and validation of forest models. Individual-based forest models require a long list of parameters, not all of which are easy to obtain from field data or literature. Hence, calibration and validation based on multiple patterns are common practice and often essential for obtaining accurate parameterizations (Wiegand et al., 2003). Quite a range of patterns can be obtained from field data (i.e., basal area, stand density, stem size distributions, tree positions), but patterns describing the vertical dimension (canopy structure)

are insufficiently covered in the field. Lidar data provides a multitude of metrics describing canopy structure. In forest model calibration approaches, these patterns can now be derived from simulations and be compared with those derived from airborne data (e.g., like in the FORMIND calibration described in Appendix A.1.2), with the potential to improve parameterizations of forest models with regard to structural realism in the future.

The lidar simulator was implemented in C++ and integrated in the FORMIND source code as an optional module. For more application flexibility, i.e., using it in combination with other forest models or simulating lidar data based on field inventory data, the lidar simulator was also implemented in R. The performance of the R version is comparable to the compiled C++ version, due to the integration of the R packages `data.table` and `parallel`, which speed up the processing considerably. It has been published (together with a set of analysis functions) in an R package on GitHub<sup>1</sup>.

### 5.1.2 *Effects of Scale and Disturbance States on Biomass Estimation*

The spatial scale has been identified as a major influence on biomass estimation accuracy. Biomass can be mapped accurately for large area units ( $\geq 1$  ha), but with decreasing size of the mapping units down to the size of common field sampling plots (~10 to 30 m) the errors increase strongly. Surprisingly, scale effects have only lately become a focus of interest, in particular the problems with very small plots also found in this thesis (Mascaro et al., 2011b; Zolkos et al., 2013; Réjou-Méchain et al., 2014). In Chapter 2 it was shown that the variability in canopy height (condensed to a metric of mean canopy height) is an important factor for biomass estimation and the maximal observed height as a single number is a weak predictor of biomass.

Being aware of the effects of spatial scale is relevant for planning of remote sensing missions, to design the sensor resolution accordingly. The small footprint size of the GEDI mission (~20 m) is expected to cause a rather high uncertainty in biomass estimation at the single footprint level (Dubayah, personal communication). Parts of the uncertainty will average out when upscaling to a 1-ha resolution product. The BIOMASS mission, as another example, will provide P-band SAR-based estimates of canopy height and biomass at 200-m resolution. The question here is rather if the horizontal resolution of 200-m is sufficient for accurate biomass estimates. Simulations similar to the ones conducted for airborne lidar in this thesis could be used to evaluate the estimation accuracies of the proposed spaceborne missions.

Differences in the derived height-to-biomass relationships between undisturbed and disturbed forests were small. With the lack of field data covering the successional gradient after disturbance events, these simulation findings cannot be confirmed yet for the real forest. If such differences existed, a likely reason would be differences in tree allometries and wood densities among early and late successional species. The species-level variability of these attributes is lost in the PFT approach of FORMIND, but the general trend of lower and higher wood density of early and late successional species, respectively, is included in the PFT approach. Hence, the model could show small differences in the height-to-biomass relationships, but to which degree these exist in reality requires further investigation.

The main benefit of using disturbance simulations was rather the possibility to generate heterogeneous, mixed-state plots, where more than one successional stage

<sup>1</sup> <https://github.com/niknap/slidaRtools>

occur next to each other. Such cases are common in real landscapes and are crucial to be present in the model, if the model is supposed to be able to represent the full range of stand structures observed in nature. In Chapter 2 the potential of combining a forest model with disturbance simulations and lidar simulations was shown, for generating and investigating the full range of possible stand structures. In Chapter 3 the focus was on the specific advantage of simulating time series of forest and remote sensing data.

### 5.1.3 *Detecting Biomass Changes over Time with Remote Sensing*

Repeated synchronous collections of airborne remote sensing and ground-truth data have been rarely realized. Simulated time series of remote sensing and forest inventory data have a high potential for filling knowledge gaps in change detection. Because of data scarcity, few studies have analyzed forest biomass changes with lidar so far. One could argue that mapping biomass stocks at different points in time also enables change calculation. However, it is important to point out that the relative errors associated with biomass stock estimates are usually in the range of 10 to 20% (at 1-ha scale). These errors are much larger for biomass change estimates (> 30%), due to the fact that changes are smaller than stocks. Especially, the small gains caused by forest growth seem to be often below the detection limits. The simulations in Chapter 3 have served to analyze this aspect and compare different approaches of change estimation, also considering the asymmetry of gains and losses. Further research is needed to improve the precision of change estimates.

Indirect change estimation based on biomass stock estimates at two time points is only possible if information on canopy height above ground (i.e., also terrain height) is available. Otherwise a direct approach, i.e., a direct height-change-to-biomass-change relationship is required. A systematic bias in a direct height-change-to-biomass-change relationship has been identified in this study. Biases could be removed in an improved direct approach that involved canopy texture analysis and random forest machine learning (dir+tex). Such an approach bears the potential for large scale biomass change detection, based on the global TanDEM-X digital surface model product (Solberg et al., 2014). Since it involves machine learning, it requires a large training dataset, which was only possible to generate using the linked forest and remote sensing simulations. Further research is required for a mechanistic understanding of the relationships between textural metrics, forest structure (Abdullahi et al., 2016) and terrain variability, and simulated data sets are an ideal test bed for this. They will allow experiments on how tree crown sizes affect canopy texture (Proisy et al., 2007) and how sensitive texture metrics are with regard to uneven terrain. The application of textural metrics also requires a thorough investigation of the influence of the moving window size, i.e., the considered pixel neighborhood from which texture is calculated, which comes in as a third scale component between the pixel and the plot scale of the data.

### 5.1.4 *Generalization of Structural Relationships Across Forest Types*

A general equation that links forest height and remote sensing-derived structure to biomass, regardless of the specific forest type, would be a big achievement in the field of forest mapping. It could reduce the efforts of ground-truth plot sampling and would lead to comparable remote sensing products for forest carbon monitoring worldwide.

Despite the good performance of TCH as single biomass predictor in numerous studies since its proposal (Asner and Mascaro, 2014), it is obvious that one single lidar metric is not sufficient for generalization across forest types due to the differences in forest structures and tree allometries (Bouvier et al., 2015). In Chapter 4, a combination of structural descriptors was identified which allowed unbiased biomass predictions across five distinct forests from four different continents, representing the temperate and tropical biomes. The estimation errors using the derived general equation were only marginally higher compared to using site-specific equations.

Please note the dataset used in the study is far from representing the full range of possible forest types and structures. Fortunately, forest plot networks such as ForestGEO make comparison and synthesis studies like this possible. If more lidar campaigns were conducted over further inventory sites, these could serve to validate and improve the general equation for biomass estimation. Ideally, as many forest types as possible from all biogeographical regions (e.g., a combination of the 8 biogeographic realms with the 14 biomes or the 867 ecoregions identified by Olson et al., 2001) and across different successional stages should be represented in the database. An important point is also the accessibility of data. Many datasets are being collected but not shared with the scientific community. Data policies like the ones of the ForestGEO network and remote sensing missions such as G-LiHT, which provide easy and free of charge access to their datasets, are positive examples fostering scientific progress.

In lack of empirical data, simulated data could be useful in this context. The parameterizations of FORMIND for different forest types across the world could be used to simulate lidar data and test the general equation. With simulated data it would be possible to manipulate the tree allometries and test the robustness of the biomass equation for extreme cases. A truly robust general equation should even be able to deal with extreme forest structures, which might go beyond what is observable in nature.

## 5.2 IDEAS FOR FUTURE RESEARCH

In the following section, possible future research directions based on the results of this thesis are outlined. Some of the proposed ideas are already part of ongoing research activities. All ideas have the common goal, to exploit more of the potentials, which emerge from the link between remote sensing and forest modeling.

### 5.2.1 *Disentangling Uncertainty Using Simulations*

Any remote sensing product contains uncertainties. Some of them can be reduced by refinement of technology and methods, while others are inherent. In either case, it is important to understand the mechanisms behind uncertainties in derived remote sensing products and to be able to quantify them. Forest and remote sensing simulations not only enable to derive relations between different forest attributes, but also allow detailed uncertainty analyses. E.g., in the biomass estimation procedure, various sources of uncertainty are contained. These can be categorized into 1) allometric variability, 2) structural ambiguity, 3) border effects, 4) geolocation errors and 5) instrument errors. Disentangling the contributions of each of these error sources is difficult with empirical data, but can be achieved with simulated remote sensing data of forest stands.

**Allometric variability** refers to the anatomic differences between trees. While allometric equations describe the average relationships between tree DBH and other morphological attributes, each tree differs to a certain degree from this average relationship. Thus, there is some uncertainty in the relationship between tree height and crown diameter (remote sensing observables) and DBH from which the biomass is calculated. Furthermore, there is variability in the wood density among species and individuals. The uncertainty that is introduced by this variability at tree level can be quantified by comparing simulations that explicitly account for the variability with simulations where the variability is switched off, i.e., all trees follow the exact same allometry.

**Structural ambiguity** refers to the fact that different forest stand structures can result in identical remote sensing signals. Hence, the signals (e.g., lidar profiles) are ambiguous with respect to the variable of interest (e.g., biomass). This is, to a certain degree, caused by allometric variability, as described before, but also the spatial positions of trees inside the plot area are important, because the contribution of smaller trees to the signal can be very different depending on whether they are standing freely visible for the sensor or are covered by crowns of larger trees. In a currently running project, the lidar simulator developed during this thesis was adapted to simulate large-footprint waveforms (e.g., with the technical specifications of the ICESat GLAS instrument with its 65 m footprint diameter). Lidar sampling was simulated for the whole Amazon rainforest using the regionalized FORMIND version (Rödig et al., 2017b) with the goal to quantify the structural ambiguity (regarding biomass estimates) behind each real ICESat GLAS waveform collected over the Amazon (Figure 5.1; Rödig, Knapp, et al., in prep.<sup>2</sup>). For each GLAS waveform the 100 most similar simulated waveforms were derived from FORMIND. For some GLAS profiles the biomass distribution of the 100 stands behind these simulated profiles was narrow (low uncertainty, Figure 5.1b<sub>2</sub>), while for others it was wide (high uncertainty, Figure 5.1a<sub>2</sub>). Unlike most previous studies, which estimated biomass based on a few lidar metrics, this approach takes the full vertical profile information into account. Future analyses need to find out whether there are specific profile characteristics associated with low and high structural ambiguity.

**Border effects** refer to the problem that crowns of trees near the plot borders can cross these borders. While the biomass estimate for a plot is based on all trees that have their stem foot position inside the plot area (an alternative approach is the concept of crown-distributed biomass suggested by Mascaro et al., 2011b), the remote sensing signal includes contributions of all tree crowns that occupy the canopy space above the plot. Trees which are inside the plot often have parts of their crowns outside the plot area. On the other hand, there are trees which are located outside the plot but close to its border and parts of their crowns are inside the plot area. Hence, parts of the crowns of trees inside the plot are lost in the remote sensing signal, while parts of the crowns of trees outside the plot have an unwanted contribution. Border effects are strong in cases of a large border-to-area ratio, i.e., for small plots (e.g., 20 m × 20 m), and are expected to become small for large enough plots. It would be valuable to investigate such border effects in more detail to improve sampling methods. In a current project, the magnitude of border effects in biomass estimation at different spatial scales is investigated using simulated lidar data. One possible approach is to

<sup>2</sup> Rödig, E., Knapp, N., Fischer, R., Bohn, F., Dubayah, R., Tang, H., Huth, A. (in prep.). The state of the Amazon rainforest derived from forest modeling and full lidar profiles.

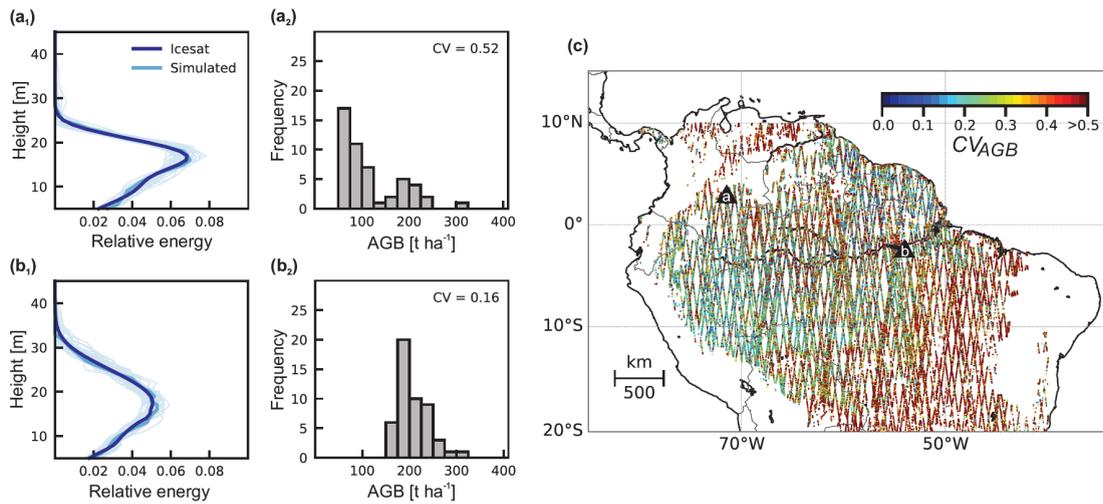


Figure 5.1: Two ICESat waveforms (dark blue) and for each the 100 most similar simulated waveforms (light blue, a<sub>1</sub>, b<sub>1</sub>) along with the aboveground biomass distributions of the 100 FORMIND stands behind the simulated waveforms, respectively (a<sub>2</sub>, b<sub>2</sub>). The map shows the ambiguity (CV = coefficient of variation) of biomass estimates for 1 million ICESat footprints across Amazonia (c; Rödiger, Knapp, et al., in prep.).

introduce periodic boundary conditions in the simulations. That means crowns which leave the plot on one side reappear on the opposite side and no crowns of other trees can enter the plot, i.e., border effects are being “switched off” (Figure 5.2, Knapp et al., in prep.<sup>3</sup>).

**Geolocation errors** refer to the uncertainty caused by weak co-registration of remote sensing and ground-truth data. The stronger the spatial mismatch between the two, the larger is the error. Spatial shifts are commonly in the cm range, but can also reach several meters. In simulated datasets geolocation errors do not exist, but can be introduced to systematically investigate them (Frazer et al., 2011).

**Instrument errors** refer to errors caused by the technical specifications of the sensor. An understanding of the noise patterns of a sensor system is important in simulating realistic remote sensing signals, e.g., for pre-launch assessment of mission requirements (Hancock et al., 2017). It involves testing the influence of different possible inclination angles and sensor-target distances, which are given by the flight path and viewing direction of the sensor platform and the earth’s topography. It further involves scattering processes in the atmosphere, possible interactions with clouds and the influence of environmental background radiation. It finally requires knowledge about the magnitude of stochastic noise introduced at the level of sensor electronics (Hancock, personal communication).

The main potential for forest models in disentangling remote sensing uncertainty lies in improving the understanding of the interplay between border effects, structural ambiguity and allometric variability in the overall observed uncertainty. Instrument errors fall into the domain of the mission engineers and require expert knowledge in earth observation, radiative transfer and sensor electronics. Geolocation errors are rather easy to understand and need to be avoided as much as possible.

<sup>3</sup> Knapp, N., Fischer, R., Rödiger, E., Taubert, F., Dantas de Paula, M., Huth, A. (in prep.). Tree crowns and plot borders: quantifying the contribution of border effects in area-based biomass estimations from remote sensing.

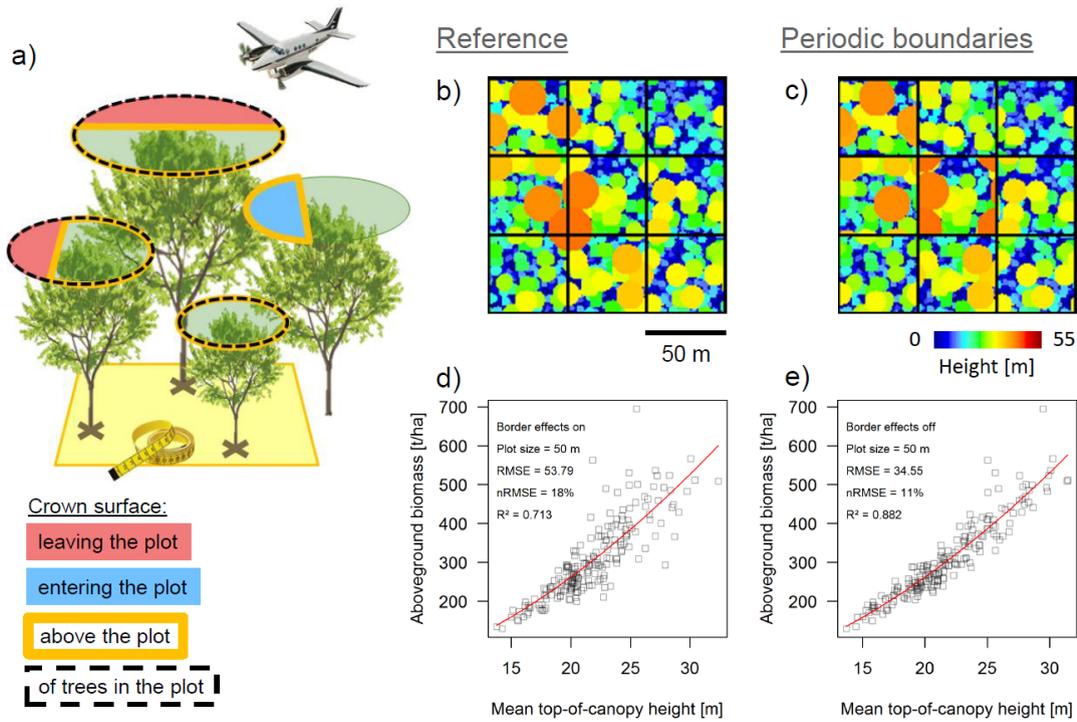


Figure 5.2: The border effect is caused by tree crowns leaving and entering the plot area (a). Its influence on biomass estimation can be analyzed by enforcing periodic boundary conditions on simulated canopy height models. To illustrate this, an example canopy height model of nine 50 m × 50 m plots is shown under normal conditions (b) and with periodic boundaries for each plot (c). The regression analysis of 200 plots shows that the overall error of 18% (d) would reduce to 11% if border effects were non-existent (e; Knapp et al., in prep.).

### 5.2.2 Remote Sensing of Forest Productivity

Besides standing biomass, there is a growing demand to estimate forest productivity from remote sensing, in order to understand the climate vegetation interaction. Coarse estimates of vegetation productivity are based on the fraction of absorbed photosynthetic radiation and can be derived from MODIS at 1-km resolution (Zhao et al., 2005). For more fine scale productivity estimates the forest structure has to be considered. **Productivity** may refer to the gross primary productivity (GPP), net primary productivity (NPP, which is GPP minus plant respiration) or net ecosystem productivity (NEP, which is NPP minus losses due to mortality and consumption by heterotrophic organisms). The challenge is that unlike biomass, which can be calculated from inventory data via allometric equations, productivity is more difficult to measure in field plots. Repeated censuses provide estimates on tree growth for deriving net primary productivity, but usually with intervals of several years. Eddy-flux towers measure carbon fluxes above the canopy near-real time, but are complex and costly systems, which only enable investigations at a few selected points (Rödig et al., 2017a).

Process-oriented forest models enable detailed analyses of the relationship between stand structure and productivity. A recent study using FORMIND and a stand structure generator called ‘forest factory’ has shown how tree **height heterogeneity** can

explain differences in productivity for stands of similar density (basal area; Bohn and Huth, 2017). In a follow-up study, lidar simulations of the forest factory stands were used to identify remote sensing metrics which are closely correlated to field-based metrics that describe **horizontal** and **vertical forest structure** (Figure 5.3; Fischer, Knapp et al., in prep.<sup>4</sup>). Based on this information, aboveground biomass (AGB) and aboveground woody productivity (AWP) were derived as functions of mean top-of-canopy height (TCH) and the standard deviation of canopy height (SDH).

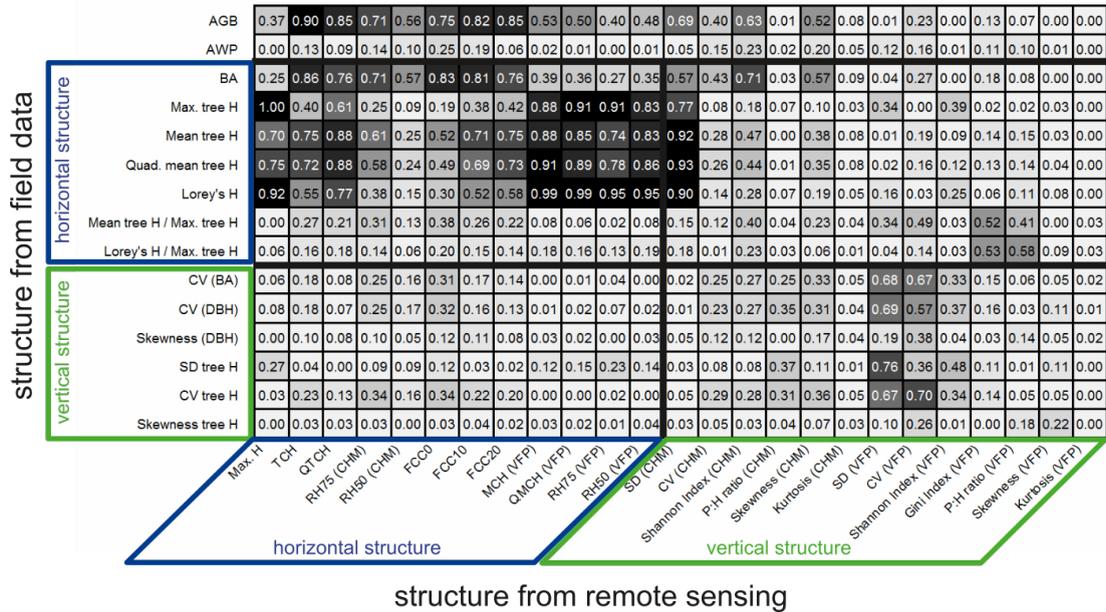


Figure 5.3: Matrix of R<sup>2</sup> values of relationships between field-based and remote sensing-derived descriptors of horizontal and vertical forest structure (Fischer, Knapp et al., in prep.).

In a current pilot project, these relationships are used to estimate forest productivity across Germany. As no wall-to-wall lidar data for Germany was available, lidar data was simulated for the 47,000 plots of the German forest inventory (Bundeswaldinventur) and derived lidar metrics were interpolated in space to generate maps of horizontal and vertical structure, AGB and AWP (Figure 5.4; Fischer, Knapp et al., in prep.; Fischer et al., 2019).

### 5.2.3 Model Initialization at Individual Tree Level

Many ecological processes in forests act at the individual tree level or are emergent properties caused by the interactions among individuals. The strength of individual-based models to represent these processes appropriately is what makes forest gap models so useful and warrants their wide application. Remote sensing products have been used for a long time to describe these processes at an aggregated, area-based scale. With increasing resolution of remote sensing data also the information retrieval at individual tree level has become possible. In the author's opinion, it is a logical next step to link the approaches of individual-based remote sensing and individual-based modeling. However, there has been little progress in this direction so far. In

<sup>4</sup> Fischer, R, Knapp, N, Bohn, F, Shugart, HH, Huth, A (in prep.). The relevance of horizontal and vertical structure for forest biomass and productivity - new perspectives for remote sensing.

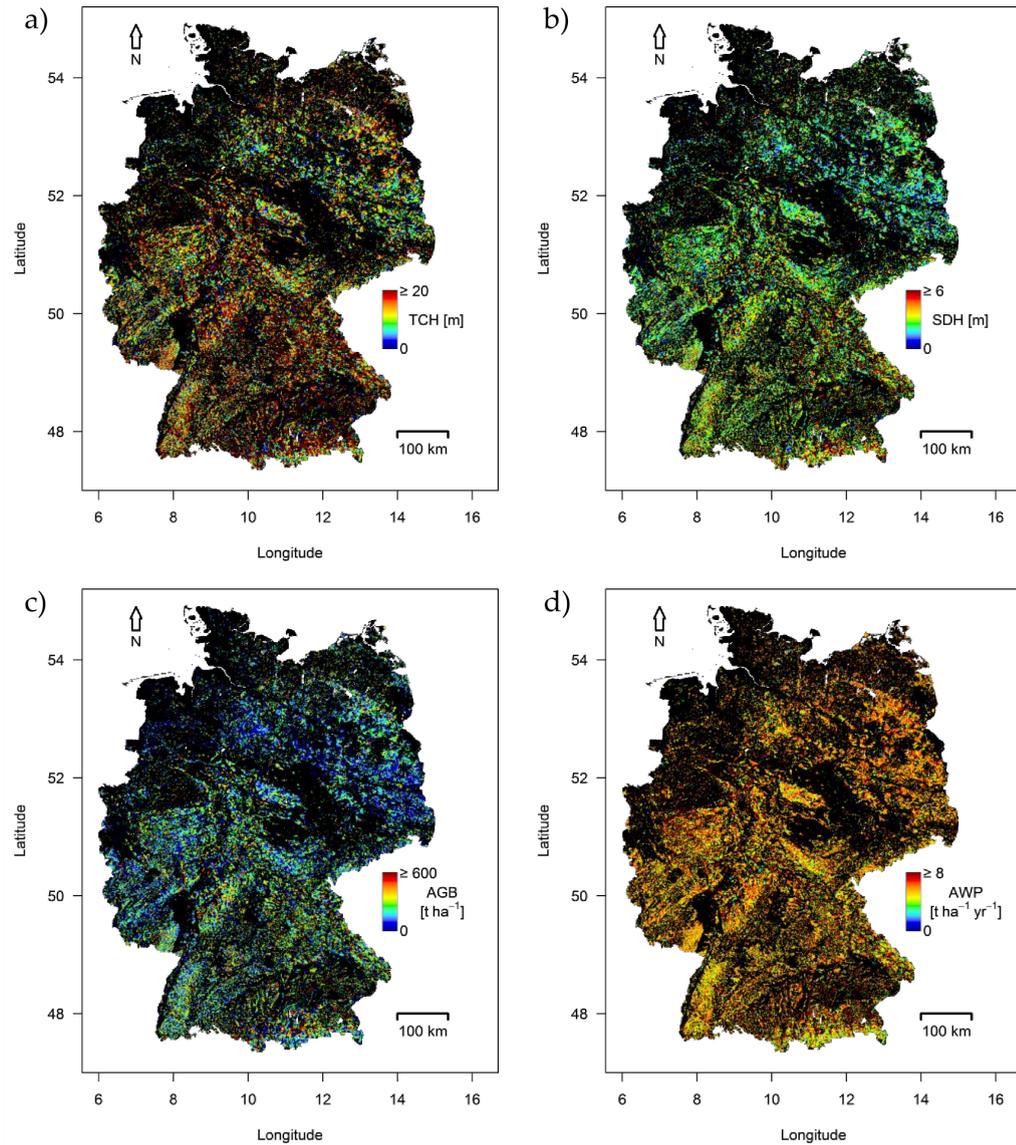


Figure 5.4: Maps of Germany showing mean top-of-canopy height (TCH), standard deviation of height (SDH), aboveground biomass (AGB) and aboveground woody productivity (AWP) based on the German forest inventory and lidar simulations (Fischer, Knapp et al., in prep.).

the course of this thesis, several Master student projects have been supervised, with the goal of applying **individual tree crown delineation (ITCD)** algorithms to lidar data and derive information for forest model parameterization and initialization. The two biggest challenges in the field of ITCD are 1) the detection of understory trees in multilayered stands and 2) the processing time. Many ITCD algorithms operate on canopy height models and are built to detect trees in the upper canopy layer (Lindberg and Holmgren, 2017). One algorithm called **adaptive mean shift 3D (AMS3D)** operates on full point clouds and has shown promising results even in a complex tropical rainforest (Ferraz et al., 2016). This algorithm was re-implemented in a Master student projects at the UFZ with the goal of optimizing processing speed (Falke, 2017). In another Master student project a workflow was developed for time efficient parallel

processing of lidar point clouds at state scale (entire Thuringia; Nguyen, 2017). By combining the methods developed in both projects, it may soon be possible to count and measure every individual tree at country scale using airborne lidar data.

Ongoing Master student projects are dealing with the questions on how to use ITCD to derive allometric relationships for model parameterization and to initialize forest simulations with the precise spatial and structural configuration as found in the field (Fujimoto, in prep.<sup>5</sup>; Naya-Geiger, in prep.<sup>6</sup>). Several project proposals have been prepared for establishing a processing chain that uses lidar-based ITCD for initialization of dynamic forest simulations to estimate productivity and development under variable environmental conditions at country scale. Meanwhile, the AMS3D re-implementation has been applied successfully on several different datasets (Figure 5.5; Vincent personal communication) and the code was recently published as an R package on GitHub<sup>7</sup>.

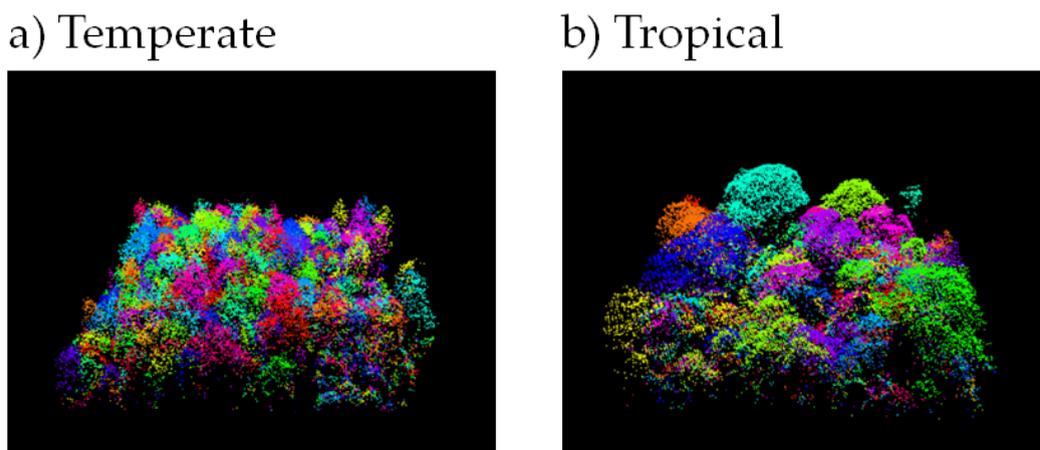


Figure 5.5: Lidar point clouds of 1 ha of temperate (Traunstein) and tropical (BCI) forest with each color representing one tree crown cluster identified by the adaptive mean shift 3D algorithm.

### 5.3 VISION

This thesis has laid the foundations for coupling remote sensing and individual-based forest modeling with benefits for both disciplines. In combining the methods described here with data from upcoming missions such as GEDI (precise height and vertical structure), EnMAP (hyperspectral signatures revealing functional traits and taxonomic identities) and Tandem-L (wall-to-wall horizontal and vertical structure at subannual revisit intervals) we are not far from being able to localize, identify and measure all trees at continental scale and predict their future fate based on simulations. Even if individual tree crowns cannot be delineated as directly as in the

5 Fujimoto A. (in prep.). Currently conducted Master's thesis at the University of Osaka on the topic of using UAV-based structure from motion to conduct ITCD and initialize FORMIND for estimating growth of a Japanese coniferous forest.

6 Naya-Geiger A. (in prep.). Currently conducted Master's thesis at UFZ in cooperation with HNE Eberswalde on the topic of using lidar and AMS3D to derive allometric relationships of trees and initialize FORMIND for estimating productivity of the Traunstein megaplot.

7 <https://github.com/niknap/MeanShiftR>

case of airborne lidar, satellite derived vegetation profiles will allow the reconstruction of tree size distributions (e.g., Stark et al., 2015). The species information will come from a combination of multispectral, hyperspectral and textural data in combination with machine learning algorithms (e.g., Fassnacht et al., 2016). Hence, in the near future **individual-based forest model applications at continental scale**, which will be constrained by detailed remote sensing products of high resolution, will become possible. To be clear, the goal should be to develop methods which provide the option to simulate any tree in any forest on earth where needed, which does not mean that the model has to explicitly contain representations of all 3 trillion trees on earth (Crowther et al., 2015). The **fusion of remote sensing and forest modeling** will blur the boundaries between model initialization, calibration, validation and remote sensing interpretation sensu Plummer (2000) and is maybe best described with the term assimilation. This means for such a system to work smoothly the remote sensing interpretation needs to be well established and the model needs to be flexible to integrate remote sensing input at the starting point and for updates during runtime. There is also a need for producing synthetic remote sensing output for data comparisons.

A forest model, which could be initialized by remote sensing with the current state and structure of the forest at any place in the world to run simulations under different climate, disturbance and management scenarios would offer interesting new possibilities for **application**. It could support climate change mitigation and resource planning through monitoring of forest disturbances and quantification of the impacts of afforestation projects, e.g., in the REDD+ context. It would help forest managers with optimizing silvicultural treatments and harvesting operations while considering multiple ecosystem services of forests. Further, it would facilitate climate change adaptation by predicting possible changes in optimal species compositions at any site. In contrast to frequently used statistical habitat suitability models, a process-oriented model would allow mechanistic simulations on species range shifts and migration rates due to climate change. Predictions could be validated, based on a range of remote sensing patterns, and if necessary adjusted, based on the continuous assimilation of remote sensing observations.

The development of such an **integrated global forest information system** could be realized with existing technology within the next five to ten years. Meanwhile, basic research in forest ecology, modeling and remote sensing has to find answers and solutions to open questions and problems. On the ecology and modeling side, one big challenge is to understand which degree of **complexity** and structural realism is required to accurately predict the variables of interest. Individual-based forest gap models are at a level of intermediate complexity. Tree growth can be represented (and with terrestrial laser scanning also measured; Disney, 2018) in far higher detail than realized in gap models. **Functional-structural plant models** that explicitly model plant organs such as branches and leaves (e.g., represented as L-systems) can help to study morphological plasticity and the environmental conditions causing it (Wang et al., 2018). This level of detail is good for understanding how processes like competition lead to observed structures. It is however, not practicable in large scale simulations and therefore higher order rules (e.g., allometric relationships), which account for different conditions (e.g., competition), need to be derived to simplify processes at aggregated levels. **Aggregation** can go far beyond the individual tree level and the dynamics of individual-based forest models can also be approximated with differential equations systems (Tietjen and Huth, 2006). Such stock-based approaches can be

computationally more efficient and may also be easier to link to coarse scale remote sensing observations.

On the remote sensing side one big challenge is to make the best use out of all the data which is already available (Pettorelli et al., 2014). Current trends in the earth observation community are about the provision of data access and **exploitation platforms** (data cubes and cloud computing; Baumann et al., 2018). This new infrastructure will facilitate **big data** processing and enable the application of computationally intensive machine learning or physically based inversion methods for large areas. For forest remote sensing this could enable the implementation of automated disturbance detection and classification and the surveillance of local weather and microclimate for an improved risk assessment.

The fate of the world's forests in the 21st century is critical if we want to preserve a livable environment for present and future generations. With remote sensing and forest modeling we own the technology that could enable us to monitor and manage forests in a sustainable way. If we learn to apply it properly it will hopefully provide guidance to cope with global change.



APPENDIX OF CHAPTER 2:  
LINKING LIDAR AND FOREST MODELING TO ASSESS BIOMASS  
ESTIMATION ACROSS SCALES AND DISTURBANCE STATES

---

A.1 SUPPLEMENTARY INFORMATION

A.1.1 *Lidar Model Parameterization*

The light extinction model contains two parameters,  $P_0$  and  $k$ . The value of the surface return probability  $P_0$  affects the point density of simulated lidar point clouds and must be chosen to match densities of airborne lidar data. The extinction coefficient  $k$  affects the shape of the vertical distribution of simulated lidar returns. The literature suggests that  $k$  values for near infrared (NIR) wavelengths, as commonly found in airborne lidar systems, are lower than those for the photosynthetically active radiation (PAR) spectrum due to the lower absorptivity of leaves. Jones (2013) reported values of  $k_{\text{NIR}} = 0.17$  and  $k_{\text{PAR}} = 0.5$  in crop fields. Calculations following Campbell and Norman (2012) led to  $k_{\text{NIR}} = 0.22$  and  $k_{\text{PAR}} = 0.45$ , assuming absorptivities of  $\alpha_{\text{NIR}} = 0.2$  and  $\alpha_{\text{PAR}} = 0.8$ , a vertical incidence angle of  $0^\circ$  and a spherical leaf angle distribution.

To find a good  $k$  for our lidar simulations, we ran the lidar model with varying  $k$  and compared the outputs to airborne lidar data. Two independent datasets of the BCI 50-ha plot were used for this purpose: 1) the airborne lidar point cloud acquired in 2009 and 2) the census data containing positions and diameter at breast height (DBH) of 246,903 trees with a  $\text{DBH} \geq 1$  cm acquired in 2010. The collection times of the two datasets were considered to be in close proximity such that changes in the forest structure between the two dates could be neglected. The census data were used as input to the lidar simulations. The airborne data were used as a reference, which should be approximated using simulations. Based on the DBH [m] for each tree, the height  $H$  [m] and crown diameter  $CD$  [m] were calculated according to given allometric equations (Bohlman and O'Brien, 2006).

$$H = 43.4 \cdot \text{DBH}^{0.6} \tag{A.1}$$

$$CD = 18.16 \cdot \text{DBH}^{0.68} \tag{A.2}$$

These relationships are based on measurements from BCI. Equation A.1 represents the average allometry over different functional groups of species, whereas Equation A.2 is based on subcanopy species. We tested different crown diameter allometries by simulating canopy height models (CHMs) based on the inventory data and comparing them to the airborne CHM. Visual inspection along with quantification of canopy gap fraction (area proportion below 20 m height) per ha led us to the decision to choose the allometry that results in the largest possible crown diameter for a given DBH, which was the one called "subcanopy" in Bohlman and O'Brien (2006). Overall, mean gap fractions were 0.32 (airborne), 0.34 (simulated with "subcanopy" allometry)

and 0.47 (simulated with the average "all" allometry). Regarding the height allometry, we had to set a maximal height limit of 40 m. Most forests and tree species show an asymptote in height as diameter gets larger (Jucker et al., 2017) and the power law would overestimate height of very large trees. Additional parameters required for the lidar simulations were crown length, crown shape and LAI. Crown lengths were assumed to be a constant proportion of  $0.4 \cdot H$  for all trees. Crown shapes were modeled as cylinders. LAI was set to 2.0 for all trees. All these allometric equations and parameters were retained in the FORMIND simulations.

A 9-ha subplot of the 50-ha plot was selected. Synthetic lidar point clouds of these 9 ha were generated by applying different values for  $k$  and  $P_0$  from 0.1 to 1 in steps of 0.1. The pulse density for the simulations was set to  $4 \text{ m}^{-2}$ . From the airborne and simulated lidar point clouds, normalized vertical profiles were derived by counting the points in each 1-m height bin and dividing by the total number of points. The relative overlap (intersection area divided by union area) between the airborne and the simulated profile served as a measure for evaluating each  $k$ . Based on this approach,  $k = 0.1$  led to the best profile overlap (88.6%), slightly ahead of  $k = 0.2$  (87.3%) (Figure 2.4a). Because 0.2 is in better agreement with the literature values (Campbell and Norman, 2012; Jones, 2013), it was chosen for all further lidar simulations. The value was validated on another 9-ha subset for which  $k = 0.2$  led to an equally good overlap (87.5%) between simulated and airborne lidar profile (Figure 2.4b). The value of  $P_0$  did not affect the shapes of the simulated profiles. We chose a value of  $P_0 = 0.2$ , because resulting return densities of the simulated point clouds were most similar to the airborne point cloud ( $4 \text{ m}^{-2}$ ). The 9 ha used for calibration and validation covered the  $300 \text{ m} \times 300 \text{ m}$  squares in the southwestern and northeastern corners of the megaplot, respectively. Those two areas were selected, because of the maximal possible spatial distance within the plot and because they cover different habitat types with the southwest being mainly classified as low plateau, stream and swamp and the northeast being mainly classified as high plateau, slope and young forest (Harms et al., 2001). The lidar simulations worked equally well for both plots (Figure 2.4).

#### A.1.2 FORMIND Model Parameterization

The full model description has been published in Fischer et al. (2016) and no changes to the model structure have been made. The parameterization of FORMIND for BCI was originally developed and described by Kazmierczak et al. (2014) and is based on 50 ha census data (years 1990 to 2010). Kazmierczak et al. (2014) showed that grouping the 323 tree species that are found at the site into four plant functional types (PFT) provides sufficient variability between trees to reproduce many patterns observed in the field data, while keeping model complexity and parametrization effort within a feasible range. The aim of this study was to additionally reproduce remote sensing patterns; thus, the original 4-PFT parametrization was revised to incorporate the allometric relationships of tree heights and crown sizes described in Equations A.1 and A.2. The four PFTs were defined as follows: 1) small, slow growing; 2) large, slow growing; 3) small, fast growing; and 4) large, fast growing species. Fast growing species were those for which the 75-percentile of annual DBH increment exceeded 5 mm. Large species were those with a maximal DBH  $> 27.5 \text{ cm}$  (corresponding to a maximal height  $> 20 \text{ m}$ , according to Equation A.1). All four PFTs were modeled with the same height and crown diameter allometries. The maximal heights of the

PFTs were set to (1) 20 m, (2) 40 m, (3) 20 m and (4) 55 m. Allowing trees of one group to reach a higher maximum height was necessary to obtain the typical canopy pattern of BCI, with scattered emergent trees. Since in our species grouping the largest species show higher DBH increments than 5 mm yr<sup>-1</sup> (75-percentile) based on field data, we decided for them (PFT 4) to grow higher than the slow growing large trees (PFT 2). Crown lengths were modeled as 0.4 times H. The wood specific gravity of each PFT was calculated as the basal area weighted mean of the wood specific gravities of all contained species. Species-specific wood specific gravity values were derived from the database of the CTFS-ForestGEO network (Anderson-Teixeira et al., 2015). Aboveground biomass (AGB) of trees was calculated using Equation A.3:

$$AGB = \pi \cdot \left(\frac{DBH}{2}\right)^2 \cdot H \cdot F \cdot \frac{\rho}{\sigma} \quad (A.3)$$

where F is a stem form factor which accounts for the deviation from a cylindrical shape and is calculated as a power law from  $f_0 \cdot DBH^{f_1}$ ,  $\rho$  is the wood density (t<sub>ODM</sub> m<sup>-3</sup>) and  $\sigma$  is the stem-to-total AGB ratio of the tree. Parameter values are listed in Table A.1.

Annual diameter increments were calculated for all individual trees in the BCI inventory using the CTFS R package (<http://ctfs.si.edu/Public/CTFSRPackage/>). To fit diameter growth curves for each PFT, the records were grouped into ten equidistant DBH classes spanning the full DBH range of the PFT. Within each DBH class, the 90-percentile of DBH increment was calculated and a 3rd order polynomial model was fit to describe DBH increment (90-percentile) as a dependent variable of DBH.

A few parameters were calibrated inversely by varying values within a reasonable range and comparing emergent patterns of the simulation output to their equivalents from field data. Calibrated parameters were annual seed ingrowth, background mortality and parameters of the light response curve for photosynthesis. Patterns used for calibration were AGB, basal area, stem number and stem size distribution of each PFT after 500 yr of simulation. Additionally, simulated lidar patterns, i.e., the vertical lidar profile, mean canopy profile height and the CHM, were considered and compared with airborne data. Calibration was performed using a dynamically dimensioned search (DDS) optimization algorithm (Lehmann and Huth, 2015). All parameters used for the FORMIND simulations can be found in Table A.1.

### A.1.3 Comparison of Simulation-Derived Lidar-to-Biomass Relations with Reference Data

We quantified the bias between n reference observations (BCI inventory and airborne lidar data) and predictions based on the power laws that were fit to the simulated data and normalized it to the mean AGB observed in the reference data to obtain nBias (Equation A.4).

$$nBias = \frac{\sum_{i=1}^n (predAGB_i - refAGB_i)}{n} \cdot \frac{1}{mean(refAGB)} \quad (A.4)$$

A shift between simulated and reference field data was observed. For TCH-based biomass predictions the nBias was always negative. For the undisturbed scenario, nBias ranged from -6 to -28% (Figure A.10); for the disturbed scenario, it ranged

Table A.1: Model parameters to simulate the forest on Barro Colorado Island (BCI) using FORMIND with four plant functional types (PFT). The parameters are described in Fischer et al. (2016).

Process	Parameter	Unit	Plant functional type (PFT)				Reference
			1	2	3	4	
geometry	$H_{\max}$	m	20	40	20	55	7
	$h_0$				43.4		1
	$h_1$				0.6		1
	$cl_0$				0.4		10
	$cd_0$				18.16		1
	$cd_1$				0.68		1
	$cd_2$				0		1
	$\rho$	$t_{\text{ODM}} \text{ m}^{-3}$	0.62	0.56	0.52	0.39	7
	$\sigma$				0.7		9
	$f_0$				0.49		3
	$f_1$				-0.1		3
	$l_0$				2		7, 9
$l_1$				0		7, 9	
recruitment	$N_{\text{seed}}$	$\text{ha}^{-1} \text{ yr}^{-1}$	80	78	80	70	3
	$I_{\text{seed}}$		0.1	0.1	0.4	0.25	3
	$D_{\text{min}}$	m			0.01		3
mortality	$M_B$	$\text{yr}^{-1}$	0.025	0.022	0.092	0.013	7, 9
	$f_{\text{fall}}$				0.3		3, 4
photosynt.	$I_0$	$\mu\text{mol}_{\text{photon}} \text{ m}^{-2} \text{ s}^{-1}$			850		8
	$k$				0.5		5
	$l_{\text{day}}$	h			12		8
	$\phi_{\text{act}}$	d			365		8
	$P_{\text{max}}$	$\mu\text{mol}_{\text{CO}_2} \mu\text{mol}_{\text{photon}}^{-1}$	1.7	1.6	9.9	7.6	9
	$\alpha$	$\mu\text{mol}_{\text{CO}_2} \text{ m}^{-2} \text{ s}^{-1}$	0.4	0.5	0.3	0.3	9
growth	$g_0$		0.001297	0.001303	0.011046	0.01685	7
	$g_1$		0.004319	0.02878	0.024529	-0.022556	7
	$g_2$		0.135525	-0.022137	-0.14802	0.03249	7
	$g_3$		-0.544071	0.003709	-0.188849	-0.010326	7
fire	$\lambda$	yr			25		11
	$\beta$				0.5		11
	$S_{\text{fire}}$				0.55		9
	$e_{\text{moisture}}$				0.4		9
	$\text{min}_{\text{fuelload}}$				0.25		9
	$\text{tol}_{\text{fire}}$				2		6

References: 1) Bohlman and O'Brien (2006), 2) Dislich et al. (2009), 3) Kazmierczak et al. (2014), 4) Brokaw (1985), 5) Campbell and Norman (2012), 6) Busing and Solomon (2006), 7) field data, 8) climate data, 9) calibrated, 10) estimated, 11) Cochrane et al. (1999).

from -20 to -33% (Figure A.11). In both cases, the largest biases were observed for pixel resolutions of 10 m, whereas biases decreased towards higher and lower pixel resolutions. Plot size had no effect on nBias.

We also fit power laws for the TCH-to-AGB-relationship to the reference field data (Figure A.4) and to the simulated data under the different disturbance regimes (fire and logging) separately (Figures A.7 and A.8). The curves of the reference data fit the simulation-based curves best for small CHM pixel sizes and medium to large plot sizes (Figure A.9). For larger pixel sizes, the reference curve bends more strongly upwards than the simulation curves do. Thus, the simulation-based power laws tend to mostly underestimate real AGB, particularly for large TCH values.

The greatest differences between simulation scenarios can be observed in the case in which only the maximum lidar height in a plot is taken as TCH (plot size = pixel size in Figure A.4 and following). In those cases, the undisturbed curves have rather flat slopes, indicating low variability of maximal tree height with AGB at the plot scale. The lower lying curves of the fire-disturbed forest result from cases in which single surviving trees in otherwise burned plots cause high maximum tree heights along with low total AGB stocks on the plots (Figure A.7). Such cases do not occur in the logged forest scenario, in which large trees are taken out homogeneously across the whole plot (Figure A.8).

As for TCH, for most other metrics the model derived power laws underestimated the reference AGB. The nBias for all metrics at all scales is plotted in Figure A.14. For some metrics, RH25, RH50 and FCC (> 20 m), the nBias was below 10%. For certain vegetation density metrics (e.g.,  $N_{AGR}/N_{GR}$  and FCC), nBias increased with plot size, whereas for the majority of metrics no clear trend with plot size was detected.

#### A.1.4 Sensitivity of Lidar Simulations to Alternative Tree Representations

The lidar simulations are based on assumptions about tree geometries. For some of those assumptions, there exist alternatives, which could possibly improve the model's quality. Therefore, we analyzed how alternative assumptions about the leaf area distribution, the crown shape, the diameter-height relation and the crown length proportion influence the derived metrics from the lidar simulations. This analysis was done by simulating lidar data for the 50-ha inventory data and varying the assumptions.

Figure A.15 shows the obtained vertical profiles in comparison to the airborne profile for the 50-ha plot. Figures A.16 to A.18 show the relations between three different lidar metrics (MCH and TCH from 1- and 10-m pixels) and AGB at 1-ha scale. Case (a) shows the base scenario (geometric assumptions as explained in the main text). In (b), we changed the assumption of a fixed LAI of 2 for each tree with the assumption of a constant leaf area volume density (LAD). A value of  $LAD = 0.29 \text{ m}^2 \text{ m}^{-3}$  was chosen to preserve the overall average  $LAI = 5.4 \text{ m}^2 \text{ m}^{-2}$  of the 50-ha plot. In (c), we replaced the assumption of cylindrical tree crowns by ellipsoid tree crowns, preserving the crown diameters and lengths. In (d), we replaced the DBH-height relationship of the power law type (with a maximal possible height of 40 m) by an asymptotic function of the Michaelis-Menten type (Equation A.5):

$$H = \frac{55 \cdot DBH}{0.5 + DBH} \quad (\text{A.5})$$

In (e), we calculated crown lengths  $CL$  of trees using the "All" crown depth function from Bohlman and O'Brien (2006) for which Equation A.6 contains the parameters in meter units:

$$CL = 17.5 \cdot DBH^{0.7} \quad (\text{A.6})$$

In (f), all the alternative assumptions were combined. The LAD in this case was  $0.44 \text{ m}^2 \text{ m}^{-3}$ , due to the altered crown geometries.

All alternative assumptions did not alter the simulated lidar profiles much (Figure A.15). Compared to our base scenario we observed slight decreases in the overlap with the airborne reference profile and marginal changes in the MCH and ground return peaks. Using ellipsoid crown shapes and an asymptotic height curve smooths the profiles at the upper canopy limit around 40 m. The scatterplots of MCH (Figure A.16),  $TCH_1$  (Figure A.17) and  $TCH_{10}$  (Figure A.18) against AGB for the different tree geometric assumptions do not provide arguments for or against a certain scenario. Different geometric assumptions produce similar results. Using an asymptotic instead of a power law height allometry led to slight increases in  $R^2$  and decreases in nRMSE of the stand height to biomass relationships.

Table A.2: This table can be found as Table S2 in the digital supplementary material of Knapp et al. (2018a) as an Excel file<sup>1</sup>. It contains model parameters and statistics for all 672 cases in which it was possible to fit a power law model between a lidar metric and aboveground biomass (AGB). Columns: 1) Data set: simulation scenario or field reference; 2) Lidar metric (based on point cloud PC or canopy height model CHM); 3) Scale: plot size [m]; 4) a: pre-exponential factor; 5) b: scaling exponent; 6) Mean: mean observed AGB [t], 7) RMSE: root mean squared error of AGB estimation [t] quantified within the data set itself; 8) Bias: bias of AGB estimation [t] quantified within the data set itself; 9) RMSE (all): quantified against the pooled simulation data from different scenarios; 10) Bias (all): quantified against the pooled simulation data from different scenarios; 11) Bias (ref): quantified against the field reference data; 12)  $R^2$ : coefficient of determination for the linear fit between predicted and observed AGB; 13) Intercept: of the linear fit; 14) Slope: of the linear fit.

<sup>1</sup> <https://www.sciencedirect.com/science/article/abs/pii/S0034425717305679>

## A.2 SUPPLEMENTARY GRAPHICS

## A.2.1 Graphics Concerning FORMIND Output Patterns

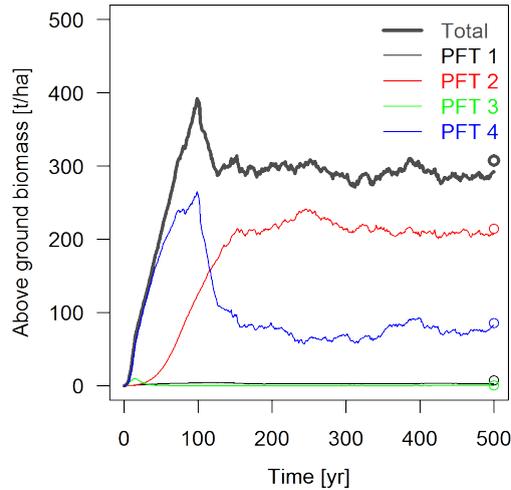


Figure A.1: Aboveground biomass (AGB) succession of the four plant functional types (PFT) on an area of 16 ha over 500 years of undisturbed FORMIND simulation. The circles on the right side represent the field reference data from the BCI 50-ha plot.

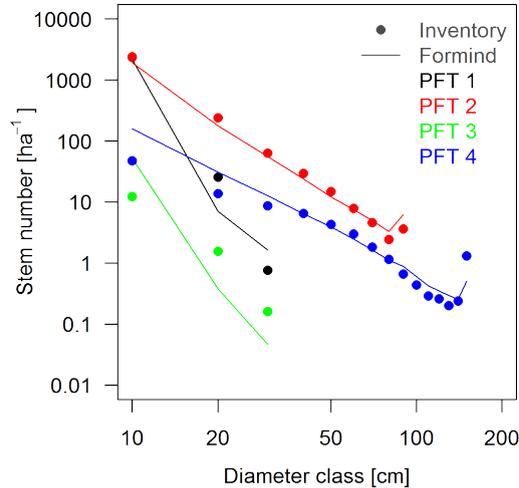


Figure A.2: Stem diameter distributions of the four plant functional types (PFT) with 10-cm class width on a double logarithmic scale. Different symbols represent the BCI 50-ha plot (Inventory) and a 16-ha FORMIND simulation after 500 years. Note that PFT 2 and 4 can only reach maximal heights of 40 and 55 m, respectively, in FORMIND, which in turn limits their diameters to maxima of 87 and 148 cm, respectively. This explains why the largest diameter classes of these PFTs have higher stem numbers than the next smaller classes. All trees in the field inventory with diameters exceeding the maximal possible diameters in the model were assigned to the largest diameter class.

## A.2.2 Graphics Concerning Biomass Prediction from Top-of-Canopy Height

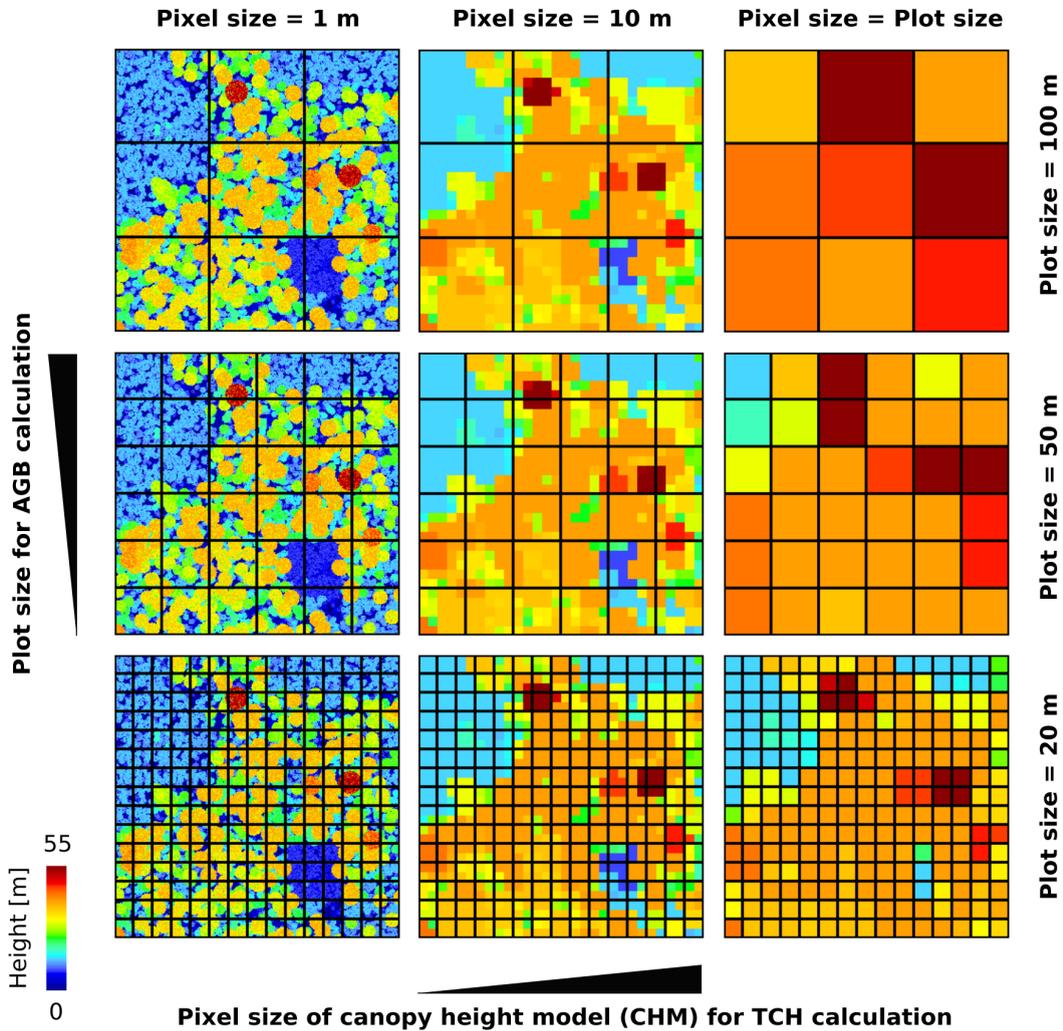


Figure A.3: Example of the sampling approach. The canopy-height-models (CHM) show a scene of 9 ha of simulated forest with different parts of the area in different stages of succession after fire disturbances. Aboveground biomass (AGB) and top-of-canopy height (TCH) were calculated for each plot using different plot sizes (black grids). CHMs of different pixel sizes were derived from the underlying lidar point cloud by taking the height of the highest return within the extent of each pixel of the respective resolution. TCH was calculated from the CHM by averaging the pixel values in each plot. In the extreme case, pixel size equals plot size such that TCH equals the maximal lidar height found in a plot (right column). Plot sizes of 20 m, 50 m and 100 m side lengths (rows) and pixel sizes of 1 m, 10 m and the maximum possible (columns) are shown. The arrangement of resolutions corresponds to the arrangement of scatterplots displayed in Figures A.4 to A.9.

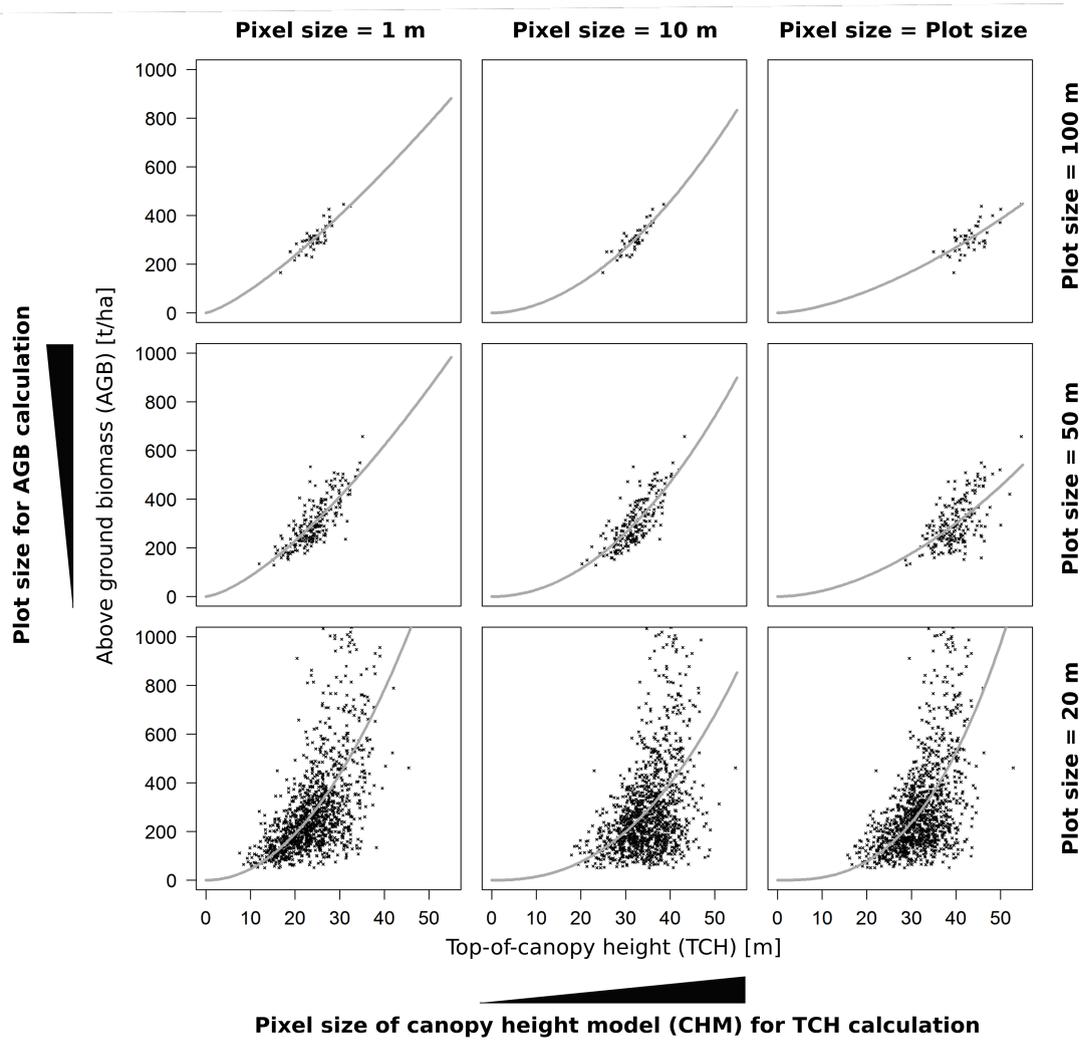


Figure A.4: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The black symbols represent the data from the **50-ha BCI plot**. The grey lines represent the best power law fits. Note that the Y-axis always ends at  $1000 \text{ t ha}^{-1}$  to be consistent with Figures A.5 to A.9. However, at the 20-m plot scale some plots had AGB values beyond  $1000 \text{ t ha}^{-1}$  (see also Figure 2.3).

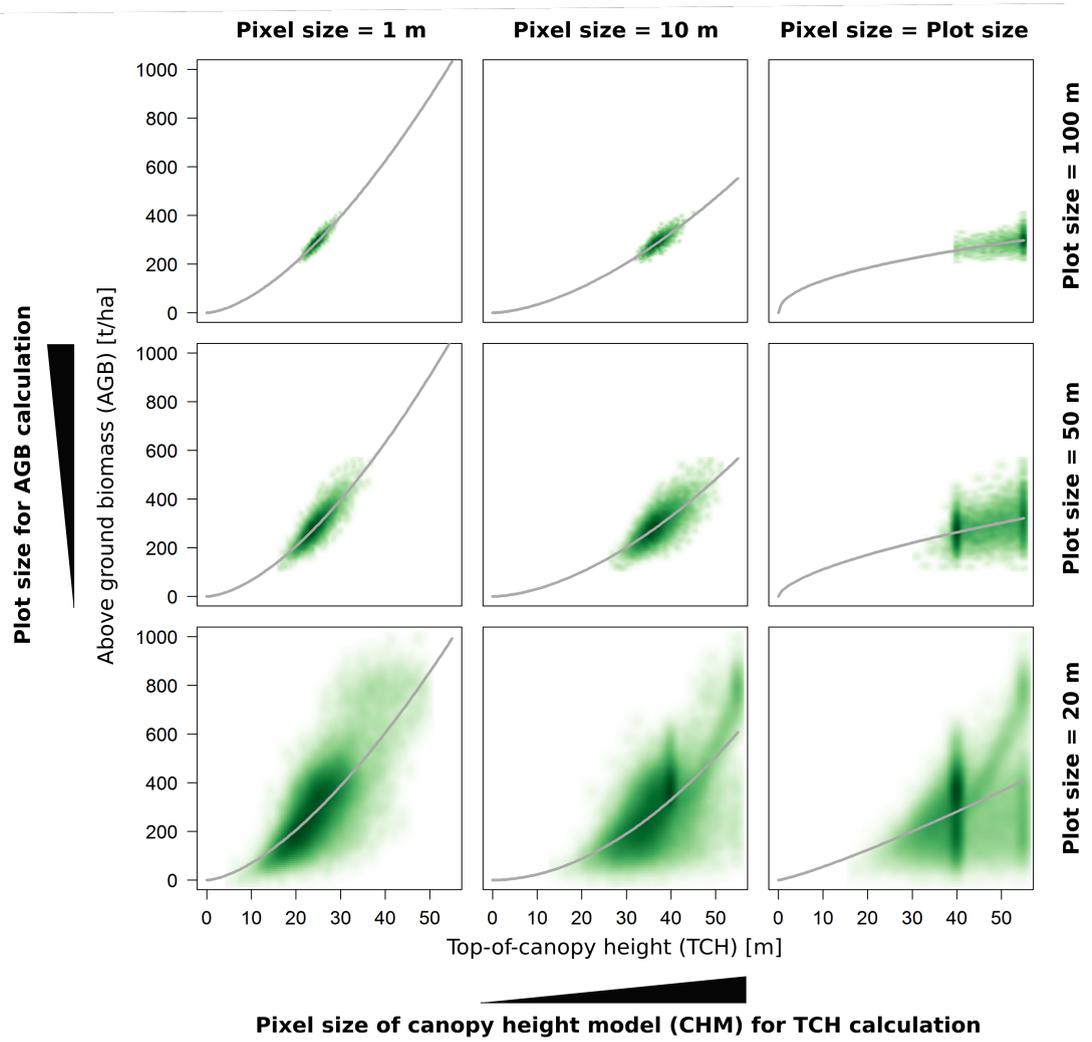


Figure A.5: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The green shading represents 1,400 ha of **undisturbed** forest simulated with FORMIND. The grey lines represent the best power law fits.

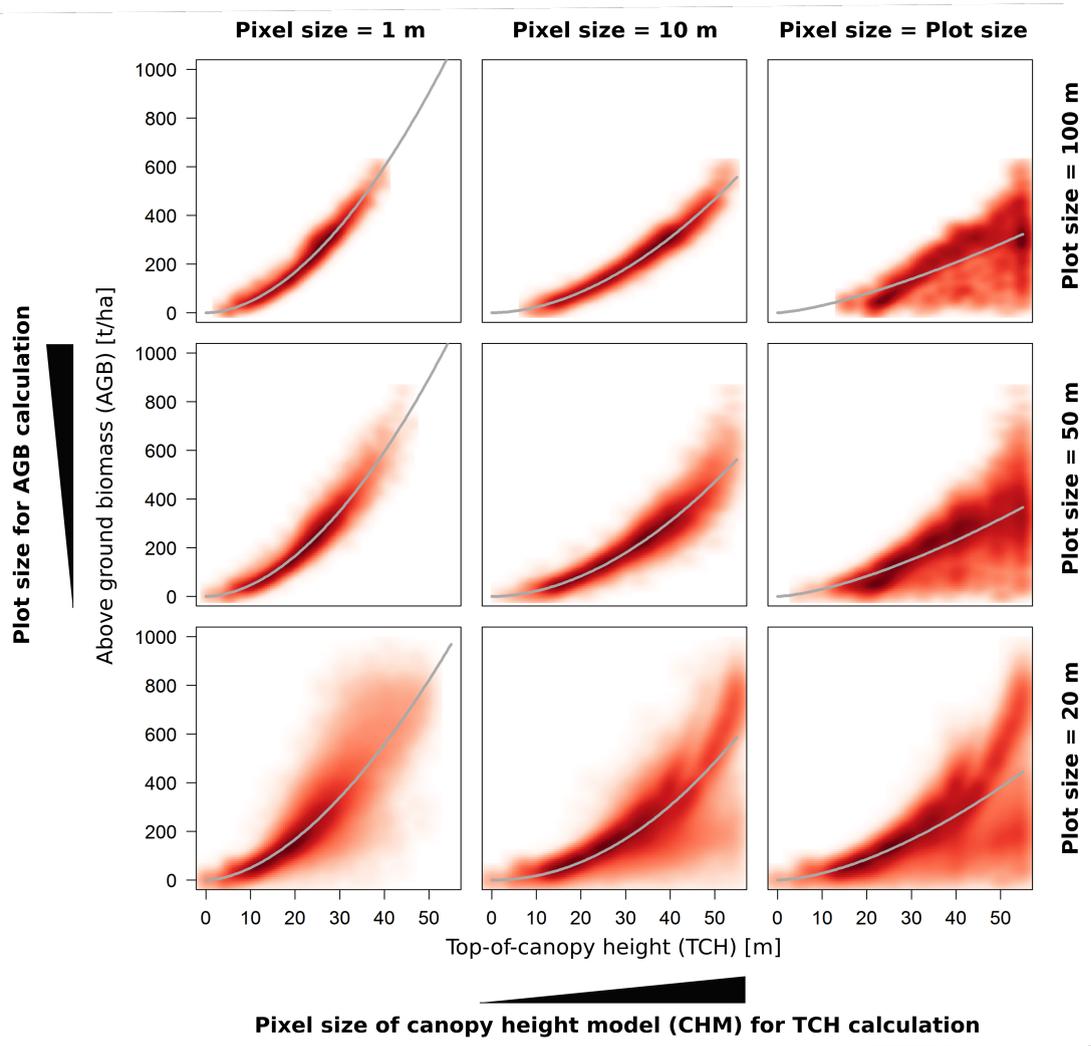


Figure A.6: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The red shading represents 1,400 ha of **fire-disturbed** and 1,400 ha of regularly **logged** forest simulated with FORMIND. The grey lines represent the best power law fits.

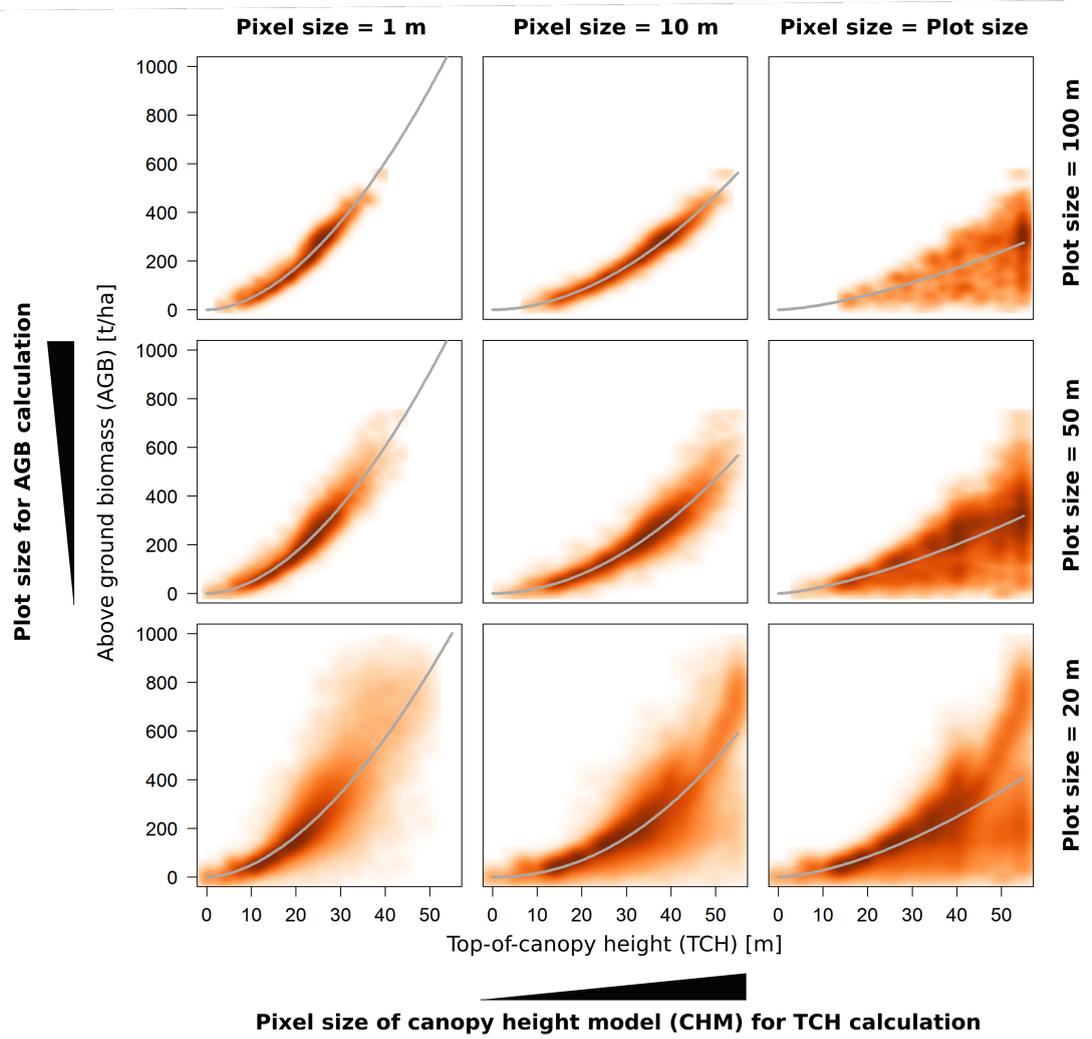


Figure A.7: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The orange shading represents 1,400 ha of **fire-disturbed** forest simulated with FORMIND. The grey lines represent the best power law fits.

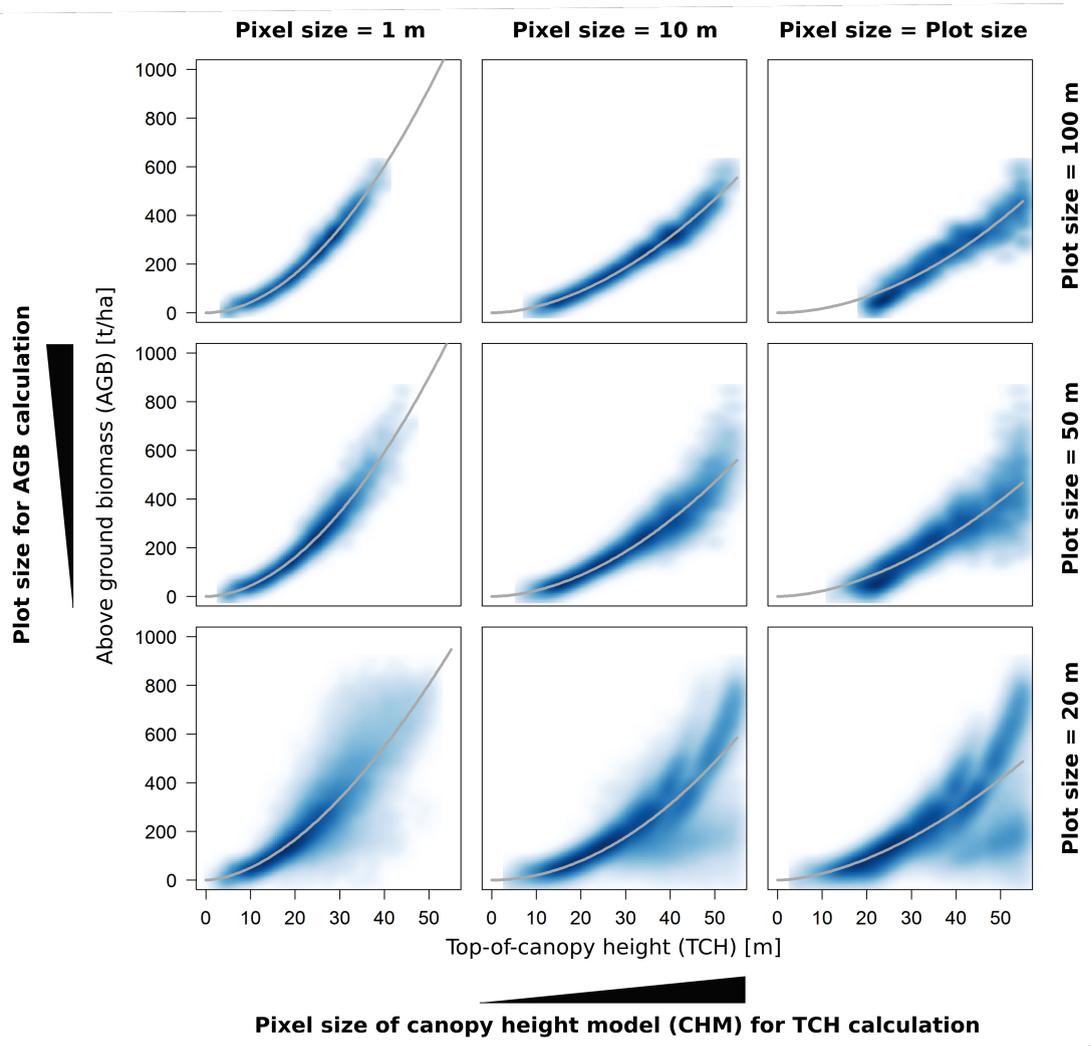


Figure A.8: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The blue shading represents 1,400 ha of regularly **logged** forest simulated with FORMIND. The grey lines represent the best power law fits.

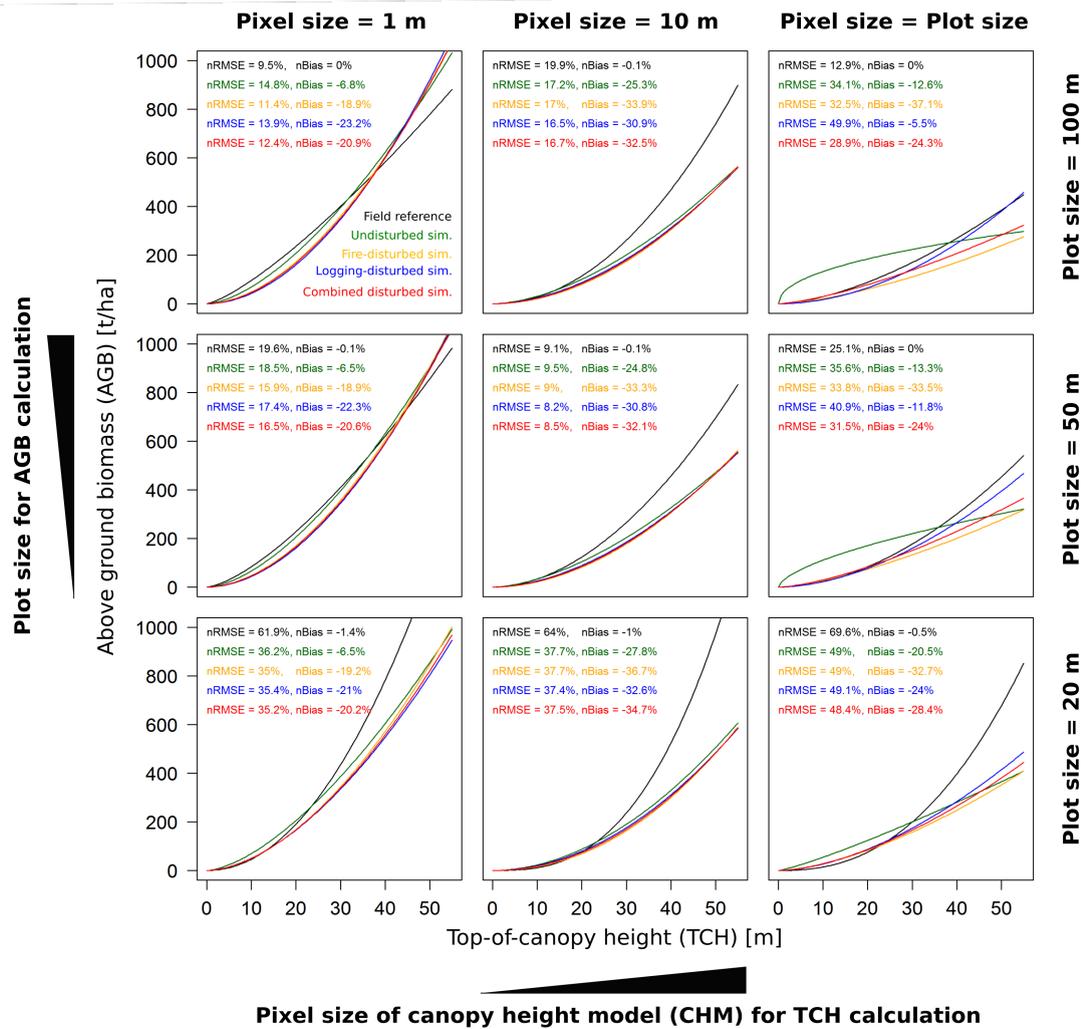


Figure A.9: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The curves represent the best power law fits for the different data sets, represented by the following colors: 1) black = field reference; 2) green = undisturbed; 3) orange = fire-disturbed; 4) blue = logging-disturbed; 5) red = combined-disturbed. To calculate nRMSE, the pooled simulated data set with all scenarios combined was used, except for nRMSE of the field reference data, where only the field data itself was used. The nBias values always refer to the bias between predictions derived from each model and the corresponding field reference observations.

Plot size [m]								
200	-5.7	-14.5	-22	-23.5	-21.9	-16.5	-12.9	-13.1
100	-6.8	-16	-22.6	-24.8	-21.6	-15.7	-13.3	-12.6
50	-6.5	-15.8	-22.8	-25.3	-23.2	-17.5	-13.3	
33	-8.2	-17.1	-22.7	-25.1	-23.2	-14.7		
20	-6.5	-15.1	-23.2	-27.8	-20.5			
	1	2	5	10	20	33	50	100
	Pixel size [m]							

Figure A.10: Normalized biases (nBias) [%] of power law models that describe the relationship between aboveground biomass (AGB) and top-of-canopy height (TCH) at different plot scales and different CHM pixel resolutions for undisturbed simulated forest. The nBias quantifies the difference between model prediction based on FORMIND-derived data and field observations from the BCI reference data.

Plot size [m]								
200	-19.6	-28.2	-31.9	-31.3	-29	-28.4	-25.2	-25.5
100	-20.9	-29.2	-32.9	-32.1	-29.2	-28.5	-24.4	-24.3
50	-20.6	-28.8	-33	-32.5	-29.8	-28.8	-24	
33	-22.2	-30	-33	-33.1	-30.4	-28		
20	-20.2	-27.9	-33.5	-34.7	-28.4			
	1	2	5	10	20	33	50	100
	Pixel size [m]							

Figure A.11: Normalized biases (nBias) [%] of power law models that describe the relationship between aboveground biomass (AGB) and top-of-canopy height (TCH) at different plot scales and different CHM pixel resolutions for disturbed simulated forest. The nBias quantifies the difference between model prediction based on FORMIND-derived data and field observations from the BCI reference data.

## A.2.3 Graphics Concerning Biomass Prediction from Various Lidar Metrics

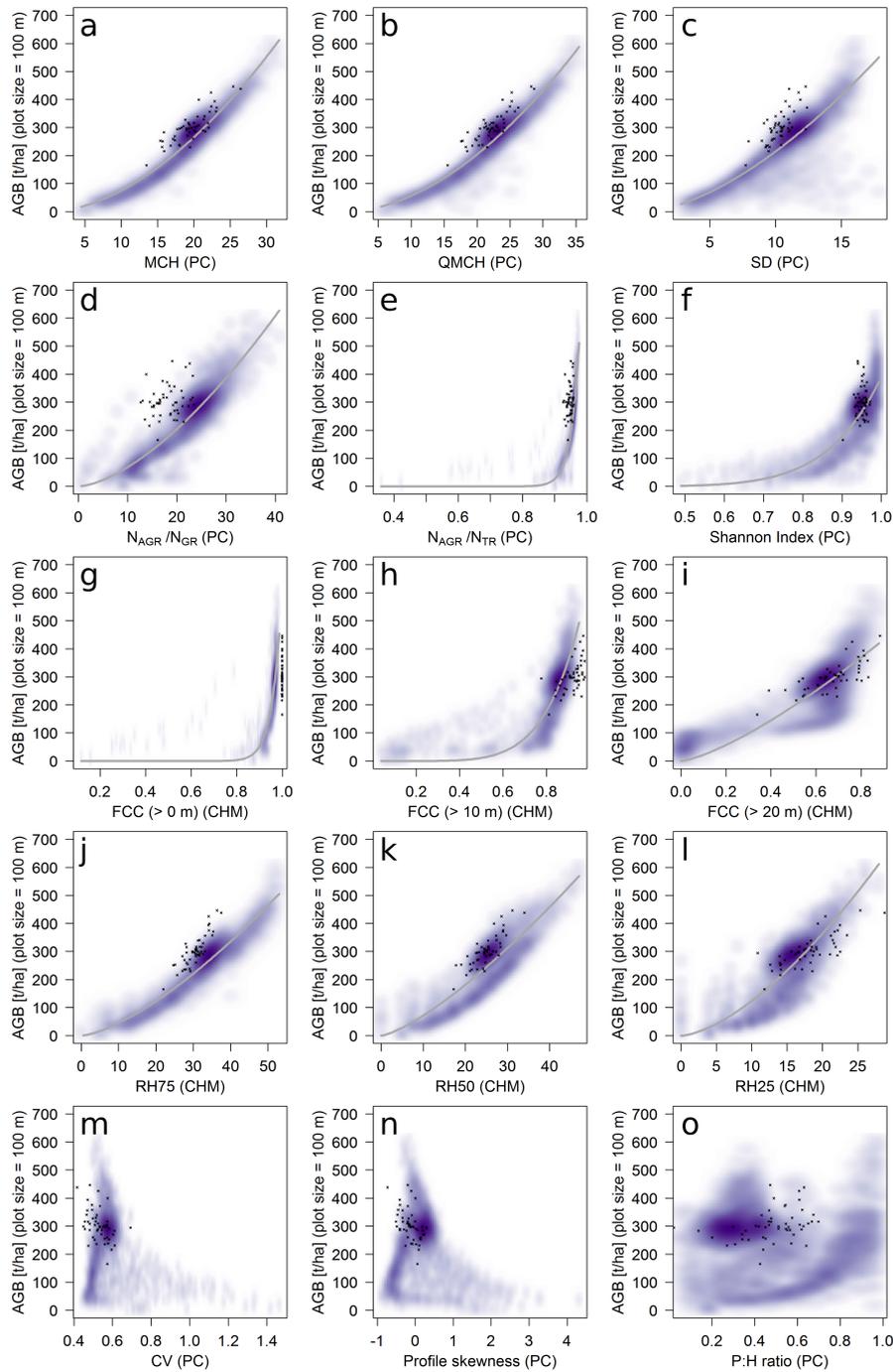


Figure A.12: Aboveground biomass (AGB) as a function of miscellaneous lidar metrics at the 1-ha scale. The violet shading represents 4,200 ha of forest simulated with FORMIND (undisturbed, fire-disturbed and regularly logged, 1,400 ha each). The black symbols represent reference data from the 50-ha BCI plot. The grey lines represent the best power law fits for the simulated data. Whether certain metrics were derived from point clouds (PC) or from canopy-height-models (CHM) is indicated in brackets. For explanations of the abbreviations, please refer to the main text.

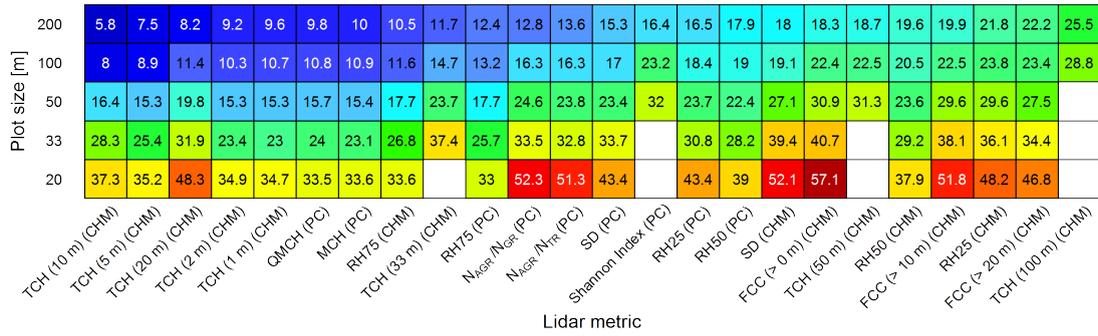


Figure A.13: Normalized root mean square errors (nRMSE) [%] of power law models that describe the relationship between aboveground biomass (AGB) and various lidar metrics (for explanations of the abbreviations, please refer to the main text and Table 2.1) at different plot scales. Whether certain metrics were derived from point clouds (PC) or from canopy-height-models (CHM) is indicated in brackets. This analysis was based on pooled (undisturbed and disturbed) simulated forest data. Empty cells indicate that no power law model could be fit.

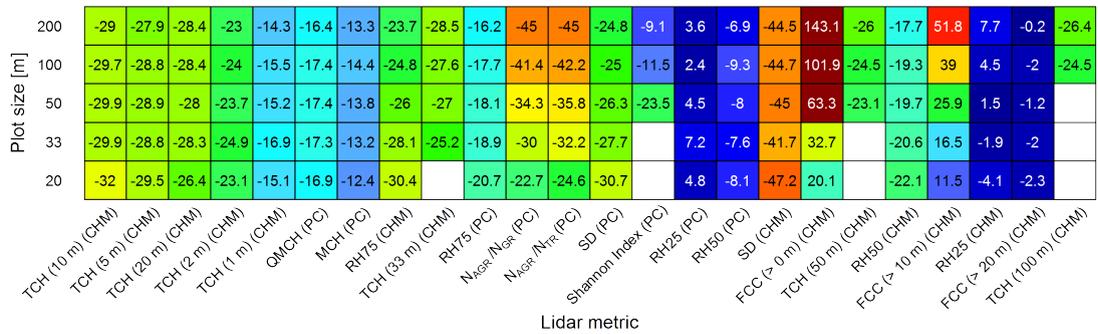


Figure A.14: Normalized biases (nBias) [%] of power law models that describe the relationship between aboveground biomass (AGB) and various lidar metrics (for explanations of the abbreviations please refer to the main text) at different plot scales. The nBias quantifies the difference between model prediction based on FORMIND-derived data and field observations from the BCI reference data. Whether certain metrics were derived from point clouds (PC) or from canopy-height-models (CHM) is indicated in brackets. This analysis was based on pooled (undisturbed and disturbed) simulated forest data. Empty cells indicate that no power law model could be fit.

A.2.4 Graphics Concerning the Sensitivity of Lidar Simulations to Alternative Tree Representations

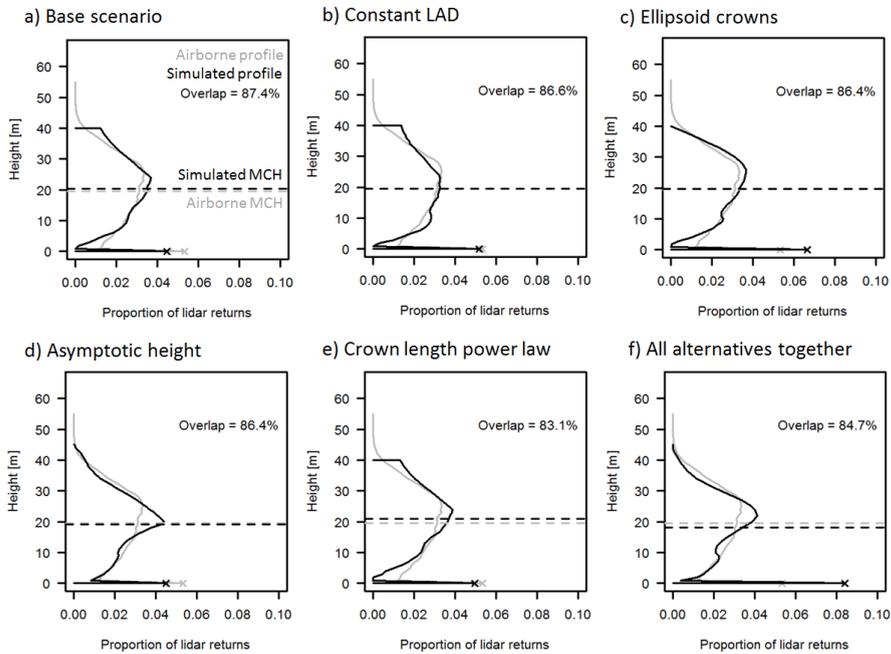


Figure A.15: Simulated lidar profiles of the BCI 50-ha plot under alternative assumptions about the tree geometry and leaf area volume density (LAD). Dashed lines indicate mean canopy profile height (MCH).

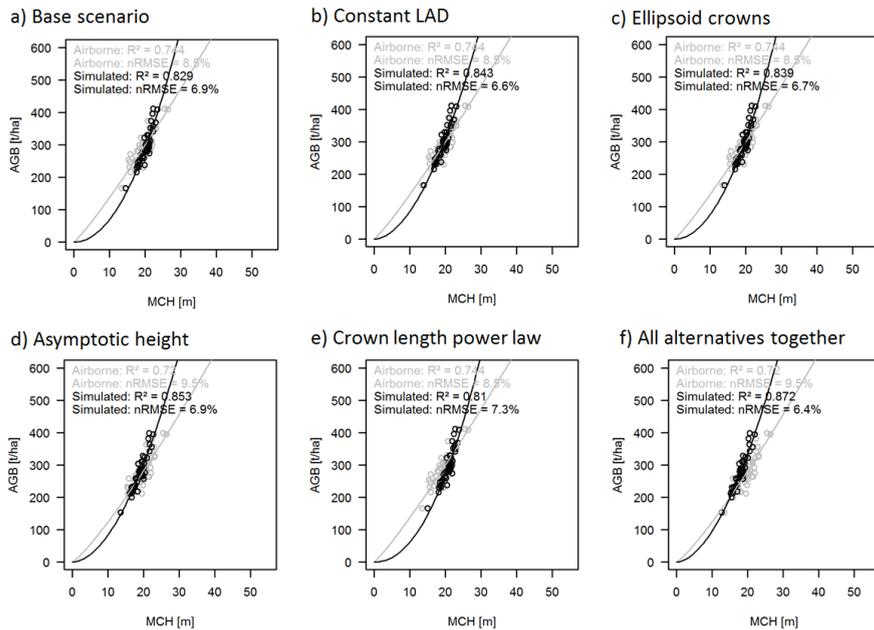


Figure A.16: Relationship between aboveground biomass (AGB) and mean canopy profile height (MCH). Points represent values at the scale of one hectare. Each graphic shows results from lidar simulations under alternative assumptions about the tree geometry.

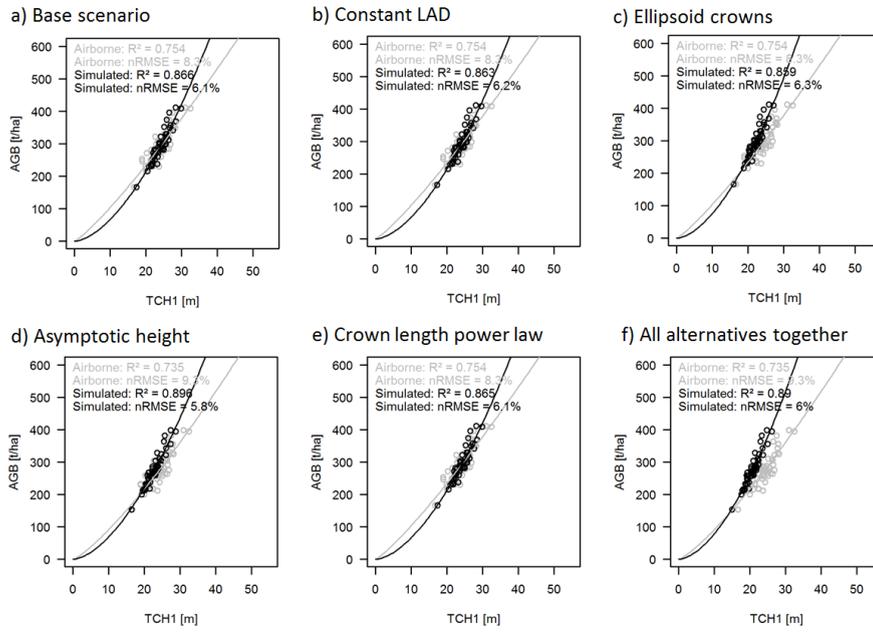


Figure A.17: Relationship between aboveground biomass (AGB) and mean top-of-canopy height derived from 1-m pixels ( $TCH_1$ ). Points represent values at the scale of one hectare. Each graphic shows results from lidar simulations under alternative assumptions about the tree geometry.

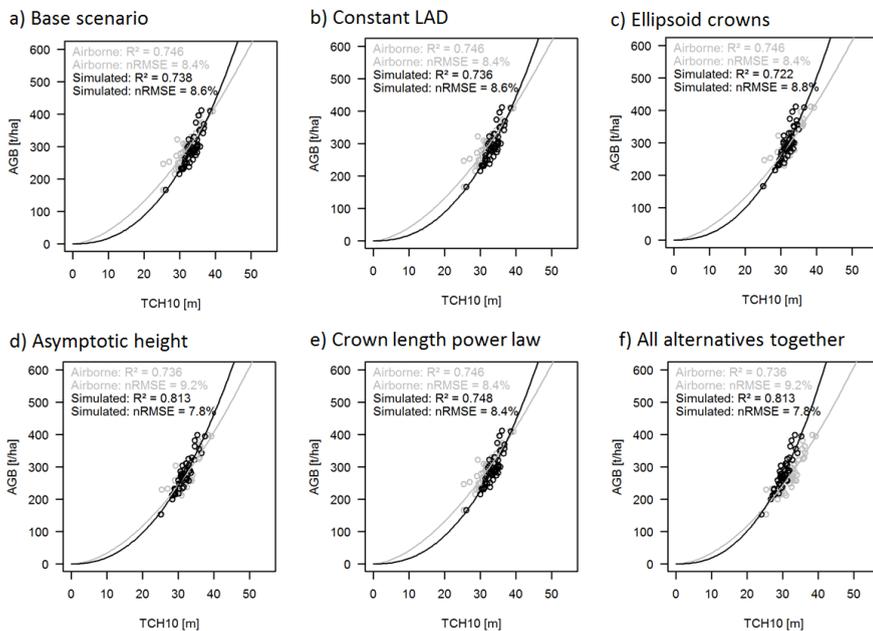


Figure A.18: Relationship between aboveground biomass (AGB) and mean top-of-canopy height derived from 1-m pixels ( $TCH_{10}$ ). Points represent values at the scale of one hectare. Each graphic shows results from lidar simulations under alternative assumptions about the tree geometry.



## APPENDIX OF CHAPTER 3: MODEL-ASSISTED ESTIMATION OF TROPICAL FOREST BIOMASS CHANGE: A COMPARISON OF APPROACHES

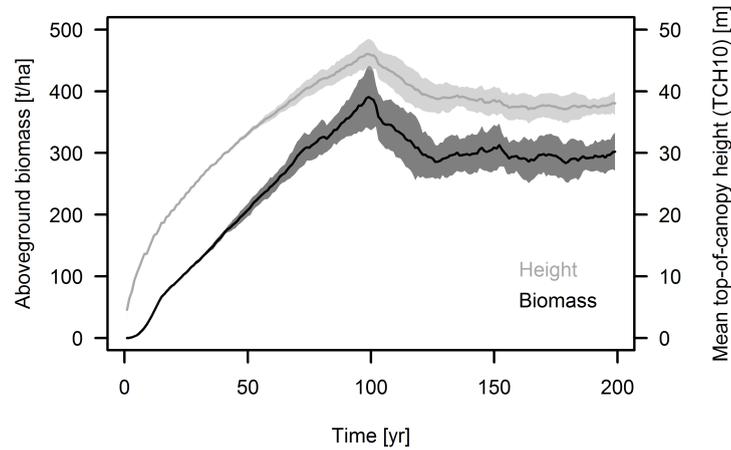


Figure B.1: Development of aboveground biomass and mean top-of-canopy height over time during primary succession in a FORMIND run using the Barro Colorado Island parameterization. Lines represent the mean of 16 ha and envelopes the standard deviations at 1-ha scale.

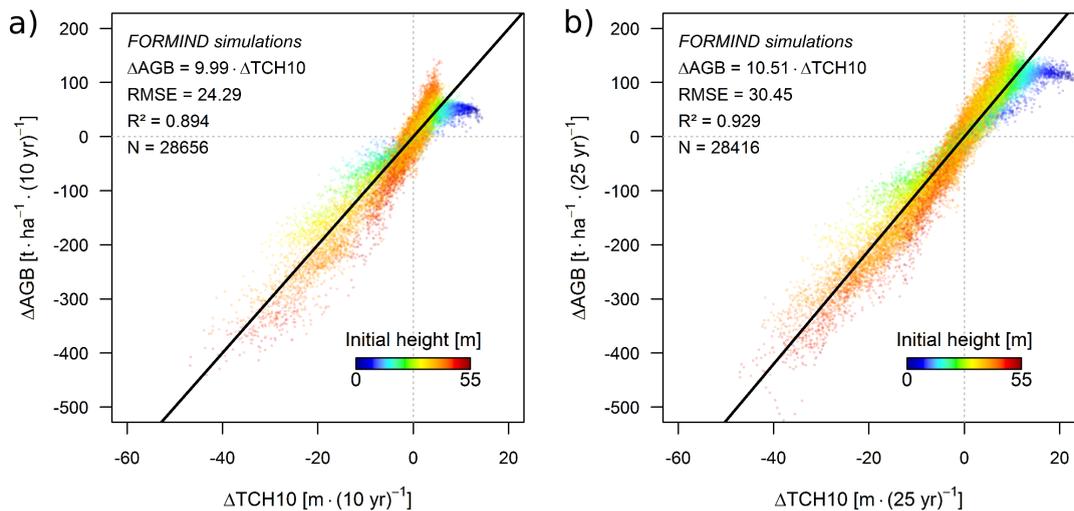


Figure B.2: Simulated data showing the  $\Delta TCH_{10}$ -to- $\Delta AGB$  relationship for (a) a 10-yr and (b) a 25-yr time interval between first and second measurement. Each point represents a 1-ha forest stand. Colors indicate the initial height ( $TCH_{10}$ ) of each stand. The black line represents the linear regression model.

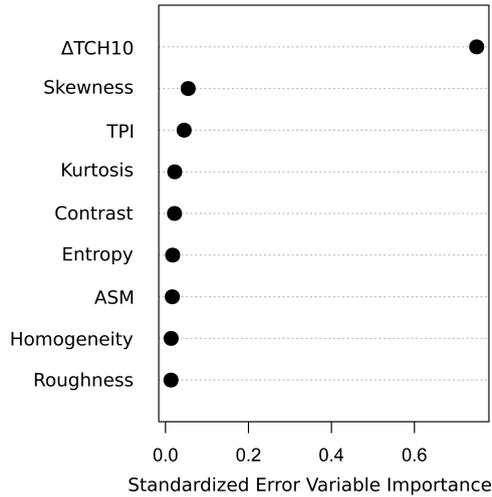
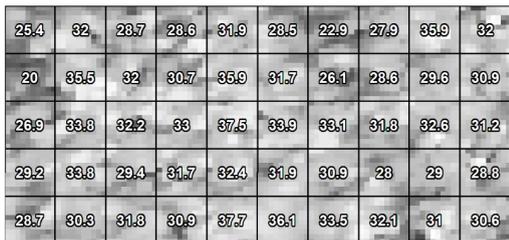


Figure B.3: The nine predictor variables, contributing to the random forest predictions in the dir+tex approach, ranked by decreasing standardized importance. Canopy height change ( $\Delta TCH_{10}$ ) was the most important predictor, followed by eight canopy texture metrics. For explanation of the abbreviations please refer to Section 3.3.5.

a) CHM 2011



b) CHM 2015

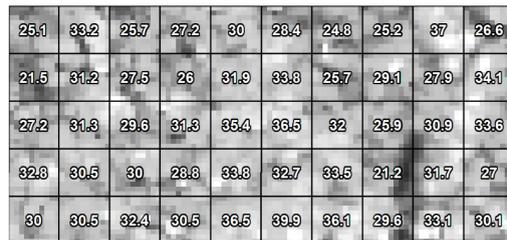
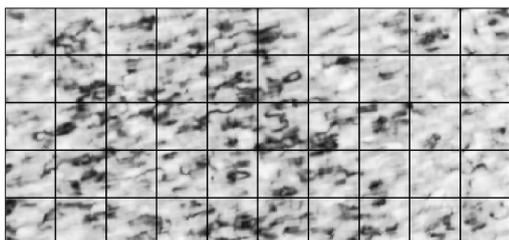


Figure B.4: TanDEM-X-derived canopy height models (CHM) for the years 2011 (a) and 2015 (b) on the BCI 50-ha plot. CHMs have 10-m pixel resolution and mean top-of-canopy height ( $TCH_{10}$ ) [m] was calculated at 1-ha scale (labels on maps).

a) Coherence 2011



b) Coherence 2015

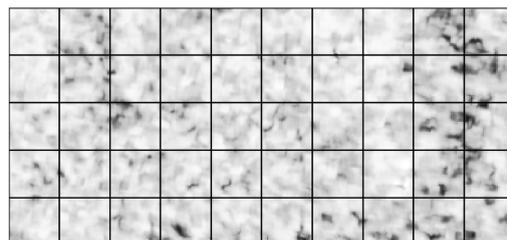


Figure B.5: TanDEM-X-derived interferometric coherence for the years 2011 (a) and 2015 (b) on the BCI 50-ha plot.

APPENDIX OF CHAPTER 4:  
FOREST STRUCTURE METRICS TO GENERALIZE BIOMASS  
ESTIMATION FROM LIDAR ACROSS CONTINENTS

---

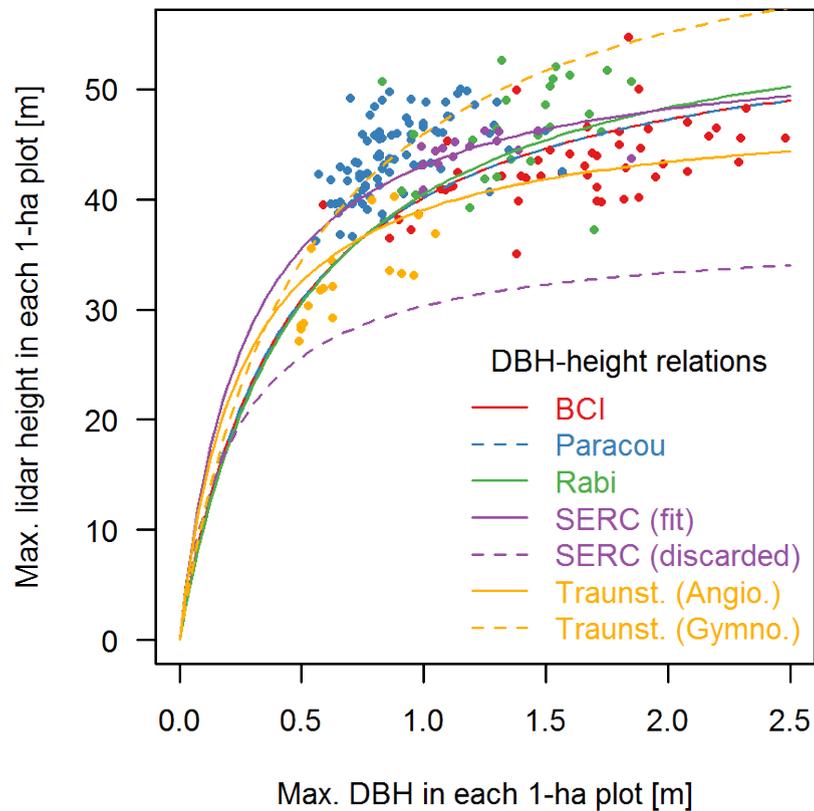


Figure C.1: Maximal lidar height in each 1-ha plot plotted over maximal diameter at breast height (DBH) in each plot. The curves represent the DBH-height allometries derived from dataset of Jucker et al. (2017). The derived curve for SERC (dashed violet) did not match the data, hence a regression was fit to this dataset (solid violet) and used as DBH-height allometry instead.

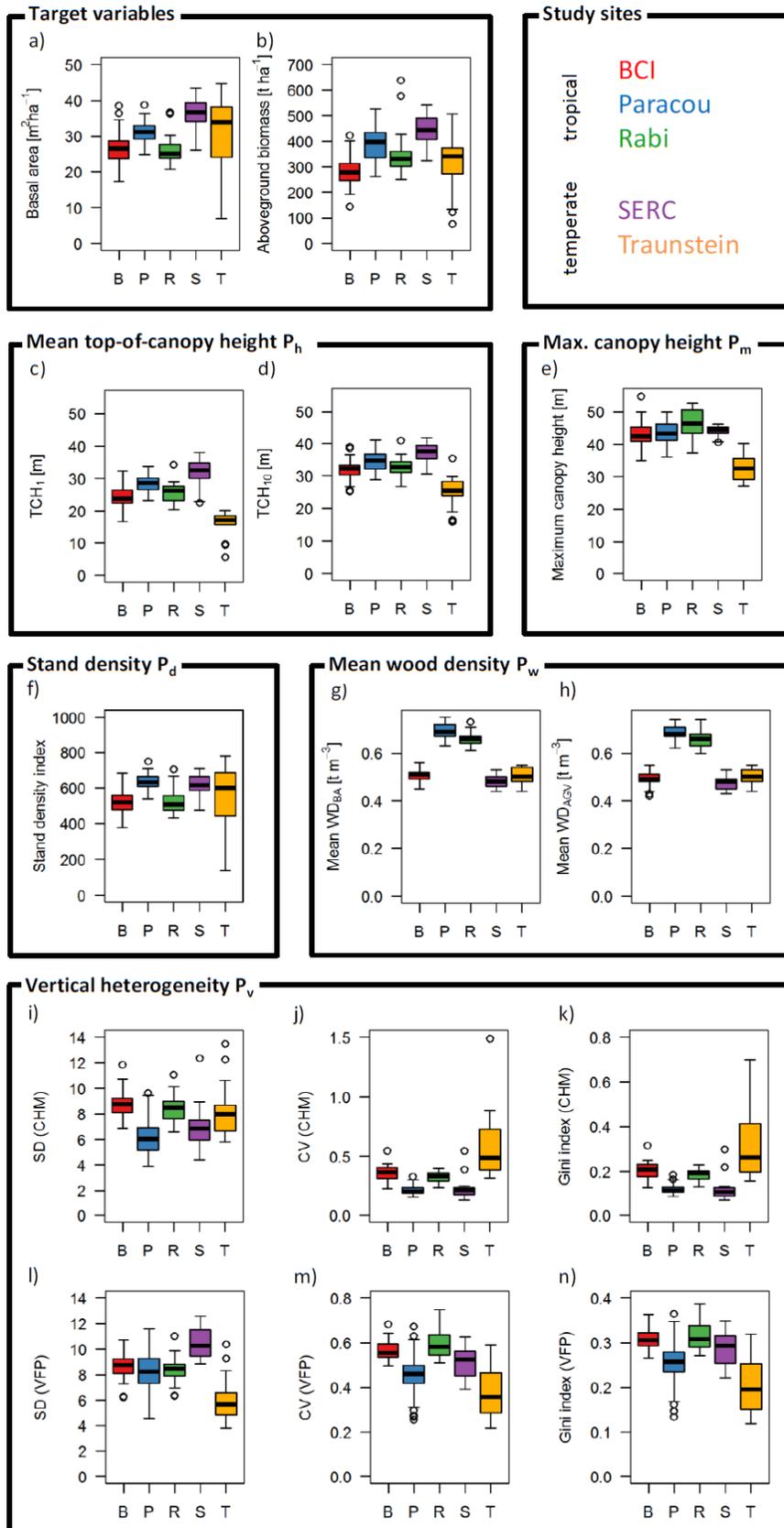


Figure C.2: Boxplots of the distributions of the different forest structure metrics at 1-ha scale across all study sites Barro Colorado Island (B), Paracou (P), Rabi (R), Smithsonian Environmental Research Center (S) and Traunstein (T).

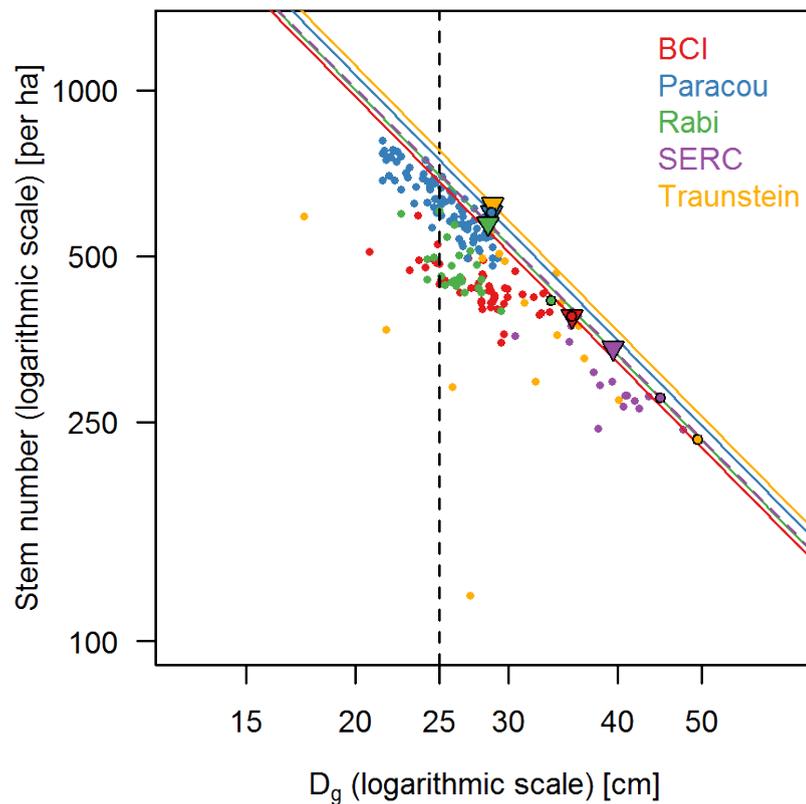


Figure C.3: Relationship between quadratic mean stem diameter  $D_g$  and stem number on log-log-scale. Each point represents one hectare with colors indicating the sites. Lines represent the self-thinning trajectories for each site according to Reinecke's rule with a slope of  $-1.605$  and an intercept based on the highest stand density index (SDI) observed at each site. The SDI can be graphically derived as the  $y$ -values (stem numbers) corresponding to the intersection points of the lines with the dashed vertical line, which represents the standard  $D_g$  of 25 cm. Black circles mark the plots with the highest basal area per site and black triangles mark the plots with the highest SDI per site, respectively. At BCI and Paracou these two maxima coincide in the same plot.

Table C.1: Site-specific reference models for basal area (BA) and aboveground biomass (AGB) estimation based on either TCH<sub>1</sub> or TCH<sub>10</sub> as single predictors.

Site	Target variable	Predictor P <sub>h</sub>	a <sub>0,site</sub> / b <sub>0,site</sub>	a <sub>h,site</sub> / b <sub>h,site</sub>	RMSE <sub>site</sub>	nRMSE <sub>site</sub>	R <sup>2</sup> <sub>site</sub>	RMSE <sub>all</sub>	nRMSE <sub>all</sub>	R <sup>2</sup> <sub>all</sub>	Wilcoxon p-value
BCI	BA	TCH <sub>1</sub>	0.885	1.07	2.4	8.8%	0.69				0.84
Paracou	BA	TCH <sub>1</sub>	4.65	0.569	2.1	6.8%	0.36				0.8
Rabi	BA	TCH <sub>1</sub>	1.13	0.969	2.7	10.1%	0.52	2.5	8.3%	0.79	0.71
SERC	BA	TCH <sub>1</sub>	3.33	0.691	3.2	8.9%	0.52				0.86
Traunstein	BA	TCH <sub>1</sub>	0.613	1.41	3.1	10.3%	0.91				0.93
BCI	BA	TCH <sub>10</sub>	0.163	1.47	2.6	9.5%	0.64				0.9
Paracou	BA	TCH <sub>10</sub>	4.84	0.526	2.2	7%	0.31				0.75
Rabi	BA	TCH <sub>10</sub>	0.487	1.14	2.8	10.5%	0.48	2.8	9.5%	0.73	0.54
SERC	BA	TCH <sub>10</sub>	0.839	1.04	3.4	9.4%	0.47				0.9
Traunstein	BA	TCH <sub>10</sub>	0.268	1.47	4.9	16.1%	0.79				0.9
BCI	AGB	TCH <sub>1</sub>	2.98	1.43	29.6	10.5%	0.74				0.98
Paracou	AGB	TCH <sub>1</sub>	5.93	1.25	40.1	10.3%	0.54				0.6
Rabi	AGB	TCH <sub>1</sub>	1.1	1.77	58.2	16.6%	0.57	41	11.6%	0.78	0.89
SERC	AGB	TCH <sub>1</sub>	60.6	0.576	48.4	10.9%	0.35				0.86
Traunstein	AGB	TCH <sub>1</sub>	6.04	1.42	36.6	11.8%	0.89				0.39
BCI	AGB	TCH <sub>10</sub>	0.316	1.96	28.7	10.2%	0.75				0.91
Paracou	AGB	TCH <sub>10</sub>	4.65	1.25	42.8	11%	0.48				0.57
Rabi	AGB	TCH <sub>10</sub>	0.638	1.81	57.8	16.5%	0.57	41.8	11.8%	0.77	0.73
SERC	AGB	TCH <sub>10</sub>	12.5	0.988	49.8	11.2%	0.31				0.86
Traunstein	AGB	TCH <sub>10</sub>	0.331	2.12	32.7	10.6%	0.91				0.73

Table C.2: The best basal area estimation models for each possible predictor ( $P_x$ ) combination ranked for increasing mean bootstrapping root mean squared error ( $RMSE_b$ ). The regression coefficients  $a_x$  for each model are given along with goodness of fit statistics. The Wilcoxon test p-values indicate for each site (B = Barro Colorado Island, P = Paracou, R = Rabi, S = Smithsonian Environmental Research Center, T = Traunstein) whether predictions are biased, i.e., mean prediction residuals are significantly different ( $p \leq 0.05$ ) from zero. For explanation of the columns names and variable names please refer to the main text.

Predictor variables				Regression coefficients					Single model statistics			Bootstrapping statistics			Wilcoxon test p-values per site				
$P_h$	$P_d$	$P_m$	$P_v$	$a_0$	$a_h$	$a_d$	$a_m$	$a_v$	RMSE	nRMSE	$R^2$	$RMSE_b$	n $RMSE_b$	$R^2_b$	B	P	R	S	T
TCH <sub>10</sub>	BA <sub>smax</sub>	H <sub>smax</sub>	SD <sub>VFP</sub>	9.2	1.3	0.359	-1.03	-0.305	2.9	9.8%	0.71	2.9	9.6%	0.72	0.0036	0.05	0.09	0.86	0.72
TCH <sub>10</sub>	BA <sub>smax</sub>	-	Gini <sub>VFP</sub>	4.37E-03	1.02	1.33	-	-0.292	3	10%	0.7	3.1	10.4%	0.67	0.71	0.74	0.65	0.06	0.3
TCH <sub>1</sub>	-	H <sub>smax</sub>	-	812	0.614	-	-1.35	-	3.3	10.9%	0.64	3.1	10.4%	0.67	5.30E-04	0.022	0.011	0.12	0.55
TCH <sub>10</sub>	BA <sub>smax</sub>	-	-	3.63E-03	0.879	1.62	-	-	3.4	11.5%	0.6	3.6	12%	0.56	0.0073	0.0092	0.41	0.011	0.26
TCH <sub>10</sub>	-	-	CV <sub>VFP</sub>	0.766	0.967	-	-	-0.394	3.6	12.2%	0.55	3.6	12%	0.56	0.26	0.0029	0.052	0.0017	0.0034
-	BA <sub>smax</sub>	-	CV <sub>CHM</sub>	0.138	-	1.37	-	-0.262	4.2	14.2%	0.4	4.2	14.1%	0.39	0.58	0.9	0.96	0.56	0.77
TCH <sub>10</sub>	-	-	-	1.78	0.807	-	-	-	4.4	14.6%	0.36	4.3	14.5%	0.36	3.20E-05	0.77	6.40E-05	0.0017	0.0023

Table C.3: The best aboveground biomass estimation models for each possible predictor ( $P_x$ ) combination ranked for increasing mean bootstrapping root mean squared error ( $RMSE_b$ ). The regression coefficients  $b_x$  for each model are given along with goodness of fit statistics. The Wilcoxon test p-values indicate for each site (B = Barro Colorado Island, P = Paracou, R = Rabi, S = Smithsonian Environmental Research Center, T = Traunstein) whether predictions are biased, i.e., mean prediction residuals are significantly different ( $p \leq 0.05$ ) from zero. For explanation of the columns names and variable names please refer to the main text.

Predictor variables					Regression coefficients						Single model statistics			Bootstrapping statistics			Wilcoxon test p-values per site				
$P_h$	$P_d$	$P_m$	$P_v$	$P_w$	$b_0$	$b_h$	$b_d$	$b_m$	$b_v$	$b_w$	RMSE	nRMSE	$R^2$	$RMSE_b$	n $RMSE_b$	$R^2_b$	B	P	R	S	T
TCH <sub>1</sub>	SDI <sub>smax</sub>	H <sub>smax</sub>	SD <sub>CHM</sub>	WD <sub>sAGV</sub>	1.92	1	0.979	-1.24	0.212	0.0838	44	12.4%	0.74	46.1	13%	0.71	0.44	0.3	0.94	0.56	0.64
TCH <sub>10</sub>	BA <sub>smax</sub>	H <sub>smax</sub>	SD <sub>VFP</sub>	-	1.52	2	0.601	-0.753	-0.383	-	47.4	13.4%	0.7	47.9	13.5%	0.69	7.90E-05	0.79	0.017	0.25	0.47
TCH <sub>10</sub>	SDI <sub>smax</sub>	H <sub>smax</sub>	-	WD <sub>sAGV</sub>	0.793	1.45	0.785	-1.05	-	0.0936	45.6	12.9%	0.73	48.5	13.7%	0.69	0.099	0.17	0.33	0.74	1
TCH <sub>10</sub>	SDI <sub>smax</sub>	H <sub>smax</sub>	-	-	0.08	1.46	1.05	-0.921	-	-	46	13%	0.72	48.5	13.7%	0.69	0.016	0.29	0.08	0.86	0.8
TCH <sub>1</sub>	-	H <sub>smax</sub>	-	-	3520	0.959	-	-1.38	-	-	47.7	13.5%	0.7	48.9	13.8%	0.68	0.031	0.82	0.56	0.025	0.47
TCH <sub>10</sub>	BA <sub>smax</sub>	-	CV <sub>VFP</sub>	WD <sub>sAGV</sub>	0.406	1.22	0.692	-	-0.156	0.284	51.2	14.5%	0.67	49.7	14%	0.67	1.40E-06	0.33	0.94	0.013	0.0066
TCH <sub>10</sub>	BA <sub>smax</sub>	-	-	WD <sub>sAGV</sub>	3.62E-03	0.549	2.7	-	-	0.665	60.1	17%	0.55	50.6	14.3%	0.66	0.061	0.7	0.063	0.46	2.50E-04
TCH <sub>10</sub>	SDI <sub>smax</sub>	-	-	-	1.55E-05	1.56	1.74	-	-	-	50.9	14.4%	0.66	50.7	14.3%	0.66	7.60E-05	0.022	0.34	0.013	6.70E-04
TCH <sub>10</sub>	-	-	CV <sub>VFP</sub>	-	1.15	1.57	-	-	-0.312	-	54.3	15.3%	0.61	53.9	15.2%	0.61	6.60E-09	0.98	0.37	0.065	5.30E-05
TCH <sub>10</sub>	-	-	-	WD <sub>sAGV</sub>	3.12	1.38	-	-	-	0.201	57.2	16.1%	0.57	57.2	16.1%	0.57	5.50E-09	0.8	0.31	0.0034	7.60E-05
TCH <sub>10</sub>	-	-	-	-	2.3	1.44	-	-	-	-	58.5	16.5%	0.55	58.3	16.4%	0.55	1.10E-09	0.063	0.56	0.074	1.10E-04
-	BA <sub>smax</sub>	-	SD <sub>VFP</sub>	-	0.87	-	1.38	-	0.645	0.782	63	17.8%	0.48	62.2	17.5%	0.49	5.90E-04	0.81	0.12	0.013	0.5



## LIST OF FIGURES

---

Figure 1.1	Ranges of different forest types . . . . .	2
Figure 1.2	Stocks and fluxes in the global carbon cycle . . . . .	3
Figure 1.3	Global map of the main drivers of forest cover loss . . . . .	5
Figure 1.4	Carbon stocks and fluxes of compartments in a tropical rainforest	8
Figure 1.5	Wavelengths and frequencies of remote sensing sensor types . .	10
Figure 1.6	Lidar signal generation for one laser beam . . . . .	13
Figure 1.7	Active remote sensing techniques (lidar and SAR) for canopy height measurements . . . . .	15
Figure 1.8	Lidar point clouds of a temperate and a tropical forest and different aggregation steps. . . . .	17
Figure 1.9	Flowchart of the main processes simulated in the individual- based forest model FORMIND. . . . .	22
Figure 2.1	Workflow of the study. . . . .	28
Figure 2.2	Principle of the lidar model. . . . .	30
Figure 2.3	Relative frequency distributions of aboveground biomass (AGB).	35
Figure 2.4	Vertical lidar profiles. . . . .	36
Figure 2.5	Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) from 1-m pixel resolution (CHM) for different plot sizes. . . . .	37
Figure 2.6	Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) from 10-m pixel resolution (CHM) for different plot sizes. . . . .	38
Figure 2.7	Normalized root mean square errors (nRMSE) [%] of power law models. . . . .	39
Figure 2.8	Normalized root mean square errors (nRMSE) [%] of power law models that describe the relationship between aboveground biomass (AGB) and various lidar metrics. . . . .	40
Figure 3.1	Technical flowchart of the analysis of the simulated data. . . . .	51
Figure 3.2	Inputs and principles of the three different approaches. . . . .	53
Figure 3.3	Simulated data showing the $\Delta TCH_{10}$ -to- $\Delta AGB$ relationship for a 5-yr time interval. . . . .	55
Figure 3.4	1:1-plots of estimated $\Delta AGB$ versus observed $\Delta AGB$ . . . . .	56
Figure 3.5	Prediction statistics of the three approaches plotted over stand heights. . . . .	57
Figure 3.6	Theoretical considerations on the relationship between height change and biomass change. . . . .	59
Figure 3.7	TanDEM-X-derived distributions. . . . .	60
Figure 3.8	1:1 plot of predicted $\Delta AGB$ (based on TanDEM-X-derived $\Delta TCH_{10}$ ) versus observed $\Delta AGB$ (based on forest inventory data). . . . .	61
Figure 3.9	Aboveground biomass (AGB) plotted over canopy height ( $TCH_{10}$ ) for the two different years. . . . .	62
Figure 4.1	Canopy height models of the five study sites. . . . .	73

Figure 4.2	Boxplots of the distributions of a selection of forest structure metrics. . . . .	76
Figure 4.3	Site-specific relationships (power laws) between basal area and TCH <sub>1</sub> . . . . .	77
Figure 4.4	Scatterplots of predicted basal area against observed basal area.	79
Figure 4.5	Site-specific relationships (power laws) between aboveground biomass and TCH <sub>1</sub> . . . . .	80
Figure 4.6	Scatterplots of predicted aboveground biomass against observed aboveground biomass. . . . .	81
Figure 4.7	Summary of how the exclusions of certain structural descriptors P <sub>x</sub> influence the normalized root mean squared error. . . . .	83
Figure 5.1	Two ICESat waveforms (dark blue) and for each the 100 most similar simulated waveforms (light blue) along with the aboveground biomass distributions . . . . .	95
Figure 5.2	The border effect is caused by tree crowns leaving and entering the plot area . . . . .	96
Figure 5.3	Matrix of R <sup>2</sup> values of relationships between field-based and remote sensing-derived descriptors of horizontal and vertical forest structure . . . . .	97
Figure 5.4	Maps of Germany showing mean top-of-canopy height (TCH), standard deviation of height (SDH), aboveground biomass (AGB) and aboveground woody productivity (AWP) . . . . .	98
Figure 5.5	Lidar point clouds of 1 ha of temperate (Traunstein) and tropical (BCI) forest with each color representing one tree crown cluster identified by the adaptive mean shift 3D algorithm. . . . .	99
Figure A.1	Aboveground biomass (AGB) succession of the four plant functional types (PFT) on an area of 16 ha over 500 years of undisturbed FORMIND simulation. . . . .	109
Figure A.2	Stem diameter distributions of the four plant functional types (PFT) with 10-cm class width on a double logarithmic scale. . .	109
Figure A.3	Example of the sampling approach. The canopy-height-models (CHM) show a scene of 9 ha of simulated forest with different parts of the area in different stages of succession after fire disturbances. . . . .	110
Figure A.4	Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. . . . .	111
Figure A.5	Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The green shading represents 1,400 ha of undisturbed forest simulated with FORMIND. . . . .	112
Figure A.6	Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The red shading represents 1,400 ha of fire-disturbed and 1,400 ha of regularly logged forest simulated with FORMIND. . . . .	113

Figure A.7	Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The orange shading represents 1,400 ha of fire-disturbed forest simulated with FORMIND. . . . .	114
Figure A.8	Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The blue shading represents 1,400 ha of regularly logged forest simulated with FORMIND. . . . .	115
Figure A.9	Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) for different plot sizes and different CHM pixel resolutions. The curves represent the best power law fits for the different data sets. . . . .	116
Figure A.10	Normalized biases (nBias) of power law models that describe the relationship between aboveground biomass (AGB) and top-of-canopy height (TCH) at different plot scales and different CHM pixel resolutions for undisturbed simulated forest. . . . .	117
Figure A.11	Normalized biases (nBias) of power law models that describe the relationship between aboveground biomass (AGB) and top-of-canopy height (TCH) at different plot scales and different CHM pixel resolutions for disturbed simulated forest. . . . .	117
Figure A.12	Aboveground biomass (AGB) as a function of miscellaneous lidar metrics at the 1-ha scale. . . . .	118
Figure A.13	Normalized root mean square errors (nRMSE) of power law models that describe the relationship between aboveground biomass (AGB) and various lidar metrics at different plot scales. . . . .	119
Figure A.14	Normalized biases (nBias) of power law models that describe the relationship between aboveground biomass (AGB) and various lidar metrics at different plot scales. . . . .	119
Figure A.15	Simulated lidar profiles of the BCI 50-ha plot under alternative assumptions about the tree geometry and leaf area volume density (LAD). . . . .	120
Figure A.16	Relationship between aboveground biomass (AGB) and mean canopy profile height (MCH). . . . .	120
Figure A.17	Relationship between aboveground biomass (AGB) and mean top-of-canopy height derived from 1-m pixels (TCH <sub>1</sub> ). . . . .	121
Figure A.18	Relationship between aboveground biomass (AGB) and mean top-of-canopy height derived from 1-m pixels (TCH <sub>10</sub> ). . . . .	121
Figure B.1	Development of aboveground biomass and mean top-of-canopy height over time during primary succession in a FORMIND run. . . . .	123
Figure B.2	Simulated data showing the $\Delta TCH_{10}$ -to- $\Delta AGB$ relationship for (a) a 10-yr and (b) a 25-yr time interval. . . . .	123
Figure B.3	The nine predictor variables, contributing to the random forest predictions in the dir+tex approach. . . . .	124
Figure B.4	TanDEM-X-derived canopy height models (CHM) for the years 2011 (a) and 2015 (b) on the BCI 50-ha plot. . . . .	124
Figure B.5	TanDEM-X-derived interferometric coherence for the years 2011 (a) and 2015 (b) on the BCI 50-ha plot. . . . .	124
Figure C.1	Maximal lidar height in each 1-ha plot plotted over maximal diameter at breast height (DBH) in each plot. . . . .	125

Figure C.2	Boxplots of the distributions of the different forest structure metrics at 1-ha scale across all study sites. . . . .	126
Figure C.3	Relationship between quadratic mean stem diameter $D_g$ and stem number on log-log-scale. . . . .	127

## LIST OF TABLES

---

Table 1.1	Wavelength ranges of the different regions and bands in the electromagnetic spectrum. . . . .	11
Table 2.1	List of the lidar metrics and the underlying data. . . . .	34
Table 3.1	$R^2$ values for the 1:1 relationships. . . . .	59
Table 4.1	Information on the study sites. . . . .	71
Table 4.2	Parameters for the different diameter-height relationships modeled with a Michaelis-Menten equation. . . . .	72
Table 4.3	The best basal area estimation models for different predictor combinations. . . . .	82
Table 4.4	The best aboveground biomass estimation models for different predictor combinations. . . . .	82
Table A.1	Model parameters to simulate the forest on Barro Colorado Island (BCI) using FORMIND with four plant functional types (PFT). . . . .	106
Table A.2	This table can be found as Table S2 in the digital supplementary material of Knapp et al. (2018a) as an Excel file. . . . .	108
Table C.1	Site-specific reference models for basal area (BA) and aboveground biomass (AGB) estimation. . . . .	128
Table C.2	The best basal area estimation models for each possible predictor ( $P_x$ ) combination. . . . .	129
Table C.3	The best aboveground biomass estimation models for each possible predictor ( $P_x$ ) combination. . . . .	129

## ACRONYMS

---

Acronym	Explanation
ACD	Aboveground carbon density
AGB	Aboveground biomass
AGR	Aboveground returns
AGV	Aboveground volume
AMS3D	Adaptive mean shift 3D algorithm
ASM	Angular second moment
AWP	Aboveground woody productivity
B	Biomass
BA	Basal area
BCI	Barro Colorado Island
BWI	Bundeswaldinventur
CAO-2	Carnegie airborne observatory 2
CD	Crown diameter
CHM	Canopy height model
CO <sub>2</sub>	Carbon dioxide
CV	Coefficient of variation
D	Diameter (stem)
DBH	Diameter at breast height (stem)
DDS	Dynamically dimensioned search
DEM	Digital elevation model
DGVM	Dynamic global vegetation model
dir+tex	Direct and texture
DLR	Deutsches Luft- und Raumfahrtzentrum
DSM	Digital surface model
DTM	Digital terrain model
EBV	Essential biodiversity variables
ECV	Essential climate variables
EnMAP	Environmental mapping and analysis program
FAO	Food and agriculture organization
FAPAR	Fraction of absorbed photosynthetic active radiation
FCC	Fractional canopy cover
FORMIND	Forest model individual-based
FSC	Forest stewardship council

Acronym	Explanation
GCOS	Global climate observation system
GEDI	Global ecosystem dynamics investigation lidar
GEOBON	Group on earth observations biodiversity observation network
GF	Gap fraction
GLAS	Geoscience laser altimeter system
GLCM	Grey-level co-occurrence matrix
G-LiHT	Goddard's lidar, hyperspectral and thermal airborne imager
GNSS	Global navigation satellite system
GP	Gap probability
GPP	Gross primary productivity
GR	Ground returns
H	Height
HH	Horizontal/horizontal polarisation
HOME	Height of median energy
HV	Horizontal/vertical polarisation
IBM	Individual-based model
ICESat	Ice, cloud, and land elevation satellite
InSAR	Synthetic aperture radar interferometry
IPCC	Intergovernmental panel on climate change
IR	Infrared
ISS	International space station
ITCD	Individual tree crown delineation
LAD	Leaf area volume density
LAI	Leaf area index
LIDAR	Light detection and ranging
LM	Lidar metric
LVIS	Land, vegetation and ice sensor
MCH	Mean canopy height
MODIS	Moderate-resolution imaging spectroradiometer
MSE	Mean squared error
NASA	National aeronautics and space administration
nDSM	Normalized digital surface model
NDVI	Normalized difference vegetation index
NEE	Net ecosystem exchange
NEP	Net ecosystem productivity
NIR	Near infrared
NPP	Net primary productivity
nRMSE	Normalized root mean square error

Acronym	Explanation
PAR	Photosynthetically active radiation
PC	Point cloud
PEFC	Programme for the endorsement of forest certification schemes
PFT	Plant functional type
PollnSAR	Polarimetric synthetic aperture radar interferometry
QMCH	Quadratic mean canopy height
QTCH	Quadratic mean top-of-canopy height
RADAR	Radio detection and ranging
REDD+	Reduction of emissions from deforestation and forest degradation and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries
RH	Relative height quantile
RMSE	Root mean squared error
RVoG	Random volume over ground
SAR	Synthetic aperture radar
SD	Standard deviation
SDH	Standard deviation of height
SDI	Stand density index
SERC	Smithsonian environmental research center
SLICER	Scanning lidar imager of canopies by echo recovery
SRTM	Shuttle radar topography mission
TanDEM-X	TerraSAR-X add-on for digital elevation measurement
TCH	Mean top-of-canopy height
TPI	Topographic position index
TR	Total returns
TRI	Topographic ruggedness index
UAV	Unmanned aerial vehicle
UFZ	Helmholtz Centre for Environmental Research
UNFCCC	United nations framework convention on climate change
VCF	Vegetation continuous field
VFP	Vertical foliage profile
VH	Vertical/horizontal polarisation
VIS	Visible light
VOD	Vegetation optical depth
VV	Vertical/vertical polarisation
WD	Wood density



## BIBLIOGRAPHY

---

- Abdullahi, S, Kugler, F, and Pretzsch, H (2016). "Prediction of stem volume in complex temperate forest stands using TanDEM-X SAR data." In: *Remote Sensing of Environment* 174, pp. 197–211. ISSN: 00344257. DOI: 10.1016/j.rse.2015.12.012. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425715302339>.
- Achard, F, Eva, HD, Stibig, HJ, Mayaux, P, Gallego, J, Richards, T, and Malingreau, JP (2002). "Determination of deforestation rates of the world's humid tropical forests." English. In: *Science* 297.5583, pp. 999–1002. DOI: 10.1126/science.1070656.
- Anderegg, WR, Konings, AG, Trugman, AT, Yu, K, Bowling, DR, Gabbitas, R, Karp, DS, Pacala, S, Sperry, JS, Sulman, BN, and Zenes, N (2018). "Hydraulic diversity of forests regulates ecosystem resilience during drought." In: *Nature* 561.7724, pp. 538–541. ISSN: 14764687. DOI: 10.1038/s41586-018-0539-7. URL: <http://dx.doi.org/10.1038/s41586-018-0539-7>.
- Anderson-Teixeira, KJ et al. (2015). "CTFS-ForestGEO: A worldwide network monitoring forests in an era of global change." In: *Global Change Biology* 21.2, pp. 528–549. ISSN: 13652486. DOI: 10.1111/gcb.12712. URL: <http://www.ncbi.nlm.nih.gov/pubmed/25258024>.
- Anderson-Teixeira, KJ, Wang, MMH, Mcgarvey, JC, and Lebauer, DS (2016). "Carbon dynamics of mature and regrowth tropical forests derived from a pantropical database (TropForC-db)." In: *Global Change Biology* 22.5, pp. 1690–1709. ISSN: 13652486. DOI: 10.1111/gcb.13226.
- Armstrong, A, Fischer, R, Huth, A, Shugart, H, and Fatoyinbo, T (2018). "Simulating forest dynamics of lowland rainforests in Eastern Madagascar." In: *Forests* 9.4, pp. 1–17. ISSN: 19994907. DOI: 10.3390/f9040214.
- Aryal, RR, Latifi, H, Heurich, M, and Hahn, M (2017). "Impact of slope, aspect, and habitat-type on LiDAR-derived digital terrain models in a near natural, heterogeneous temperate forest." In: *PGF - Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 85.4, pp. 243–255. ISSN: 25122819. DOI: 10.1007/s41064-017-0023-2.
- Asner, GP, Clark, JK, Mascaró, J, Galindo García, GA, Chadwick, KD, Navarrete Encinales, DA, Paez-Acosta, G, Cabrera Montenegro, E, Kennedy-Bowdoin, T, Duque, Á, Balaji, A, Von Hildebrand, P, Maatoug, L, Phillips Bernal, JF, Yepes Quintero, AP, Knapp, DE, García Dávila, MC, Jacobson, J, and Ordóñez, MF (2012a). "High-resolution mapping of forest carbon stocks in the Colombian Amazon." In: *Biogeosciences* 9.7, pp. 2683–2696. ISSN: 17264170. DOI: 10.5194/bg-9-2683-2012.
- Asner, GP, Martin, RE, Knapp, DE, Tupayachi, R, Anderson, CB, Sinca, F, Vaughn, NR, and Llactayo, W (2017). "Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation." In: *Science* 355.6323, pp. 385–389. ISSN: 0036-8075. DOI: 10.1126/science.aaj1987.
- Asner, GP (2009). "Automated mapping of tropical deforestation and forest degradation: CLASlite." In: *Journal of Applied Remote Sensing* 3.1, p. 033543. ISSN: 1931-3195. DOI: 10.1117/1.3223675. URL: <http://remotesensing.spiedigitallibrary.org/article.aspx?doi=10.1117/1.3223675>.
- Asner, GP and Martin, RE (2009). "Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests." In: *Frontiers in Ecology and the Environment* 7.5, pp. 269–276. ISSN: 1540-9295. DOI: 10.1890/070152. URL: <http://www.esajournals.org/doi/abs/10.1890/070152>.
- Asner, GP and Mascaró, J (2014). "Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric." In: *Remote Sensing of Environment* 140, pp. 614–624. ISSN: 00344257. DOI: 10.1016/j.rse.2013.09.023. URL: <http://linkinghub.elsevier.com/retrieve/pii/S003442571300360X>.
- Asner, GP, Flint Hughes, R, Varga, TA, Knapp, DE, and Kennedy-Bowdoin, T (2009). "Environmental and biotic controls over aboveground biomass throughout a tropical rain forest."

## BIBLIOGRAPHY

- In: *Ecosystems* 12.2, pp. 261–278. ISSN: 14329840. DOI: 10.1007/s10021-008-9221-5. URL: <http://link.springer.com/10.1007/s10021-008-9221-5>.
- Asner, GP, Powell, GV, Mascaro, J, Knapp, DE, Clark, JK, Jacobson, J, Kennedy-Bowdoin, T, Balaji, A, Paez-Acosta, G, Victoria, E, Secada, L, Valqui, M, and Hughes, RF (2010). “High-resolution forest carbon stocks and emissions in the Amazon.” In: *Proceedings of the National Academy of Sciences of the United States of America* 107.38, pp. 16738–16742. ISSN: 1091-6490. DOI: 10.1073/pnas.1004875107. arXiv: arXiv:1408.1149. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2944749&tool=pmcentrez&rendertype=abstract>.
- Asner, GP, Hughes, RF, Mascaro, J, Uowolo, AL, Knapp, DE, Jacobson, J, Kennedy-Bowdoin, T, and Clark, JK (2011). “High-resolution carbon mapping on the million-hectare Island of Hawaii.” In: *Frontiers in Ecology and the Environment* 9.8, pp. 434–439. ISSN: 1540-9295. DOI: 10.1890/100179. URL: <http://www.esajournals.org/doi/abs/10.1890/100179>.
- Asner, GP, Mascaro, J, Muller-Landau, HC, Vieilledent, G, Vaudry, R, Rasamoelina, M, Hall, JS, and Breugel, M van (2012b). “A universal airborne LiDAR approach for tropical forest carbon mapping.” In: *Oecologia* 168.4, pp. 1147–1160. ISSN: 00298549. DOI: 10.1007/s00442-011-2165-z. URL: <http://www.ncbi.nlm.nih.gov/pubmed/22033763>.
- Asner, GP, Knapp, DE, Boardman, J, Green, RO, Kennedy-Bowdoin, T, Eastwood, M, Martin, RE, Anderson, C, and Field, CB (2012c). “Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion.” In: *Remote Sensing of Environment* 124, pp. 454–465. ISSN: 00344257. DOI: 10.1016/j.rse.2012.06.012. URL: <http://dx.doi.org/10.1016/j.rse.2012.06.012>.
- Asner, GP, Knapp, DE, Martin, RE, Tupayachi, R, Anderson, CB, Mascaro, J, Sinca, F, Chadwick, KD, Sousan, S, Higgins, M, Farfan, W, Silman, MR, Augusto, W, León, L, and Neyra Palomino, AF (2014). *The High-Resolution Carbon Geography of Perú*. Tech. rep. Stanford, USA: Carnegie Institution for Science, p. 65. URL: <ftp://dgc.stanford.edu/pub/asner/carbonreport/CarnegiePeruCarbonReport-English.pdf>.
- Attema, EPW and Ulaby, FT (1978). “Vegetation modeled as a water cloud.” In: *Radio science* 13.2, pp. 357–364. URL: 10.1029/RS013i002p00357.
- Auld, G, Gulbrandsen, LH, and McDermott, CL (2008). “Certification Schemes and the Impacts on Forests and Forestry.” In: *Annual Review of Environment and Resources* 33.1, pp. 187–211. ISSN: 1543-5938. DOI: 10.1146/annurev.environ.33.013007.103754. URL: <http://www.annualreviews.org/doi/10.1146/annurev.environ.33.013007.103754>.
- Avitabile, V et al. (2016). “An integrated pan-tropical biomass map using multiple reference datasets.” In: *Global Change Biology* 22.4, pp. 1406–1420. ISSN: 13652486. DOI: 10.1111/gcb.13139.
- Baccini, A, Goetz, SJ, Walker, WS, Laporte, NT, Sun, M, Sulla-Menashe, D, Hackler, J, Beck, PSA, Dubayah, R, Friedl, MA, Samanta, S, and Houghton, RA (2012). “Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps.” In: *Nature Climate Change* 2.3, pp. 182–185. ISSN: 1758-678X. DOI: 10.1038/nclimate1354. URL: <http://dx.doi.org/10.1038/nclimate1354>.
- Bamler, R and Hartl, P (1998). “Synthetic aperture radar interferometry.” In: *Inverse Problems* 14.4, p. 55. ISSN: 0266-5611. DOI: 10.1088/0266-5611/14/4/001. URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=838084>{\%}5Cnhttp://stacks.iop.org/0266-5611/14/i=4/a=001?key=crossref.8f6cbe410d55076593d70a4a99ec4bc6.
- Bannari, A, Morin, D, Bonn, F, and Huete, AR (1995). “A review of vegetation indices.” In: *Remote Sensing Reviews* 13.1-2, pp. 95–120. ISSN: 02757257. DOI: 10.1080/02757259509532298.
- Baumann, P, Rossi, AP, Bell, B, Clements, O, Evans, B, Hoenig, H, Hogan, P, Kakalettris, G, Koltsida, P, Mantovani, S, Marco Figuera, R, Merticariu, V, Misev, D, Pham, HB, Siemen, S, and Wagemann, J (2018). “Fostering Cross-Disciplinary Earth Science Through Datacube Analytics.” In: *Earth Observation Open Science and Innovation. ISSI Scientific Report Series, vol 15*. Ed. by PP Mathieu and C Aubrecht. Cham: Springer, pp. 91–119. ISBN: 978-3-319-65633-5. DOI: 10.1007/978-3-319-65633-5\_5. URL: <http://link.springer.com/10.1007/978-3-319-65633-5>{\\_}5.
- Beech, E, Rivers, M, Oldfield, S, and Smith, PP (2017). “GlobalTreeSearch: The first complete global database of tree species and country distributions.” In: *Journal of Sustainable Forestry*

- 36.5, pp. 454–489. ISSN: 1540756X. DOI: 10.1080/10549811.2017.1310049. URL: <https://doi.org/10.1080/10549811.2017.1310049><http://dx.doi.org/10.1080/10549811.2017.1310049>.
- Bitterlich, W (1952). "Die Winkelzählprobe." In: *Forstwissenschaftliches Centralblatt* 71.7-8, pp. 215–225.
- Blair, JB, Coyle, DB, Bufton, JL, and Harding, DJ (1994). "Optimization of an airborne laser altimeter for remote sensing of vegetation and tree canopies." In: *Geoscience and Remote Sensing Symposium, 1994. IGARSS'94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International*. Vol. 2. IEEE, pp. 939–941. DOI: 10.1109/IGARSS.1994.399307.
- Blair, JB, Rabine, DL, and Hofton, MA (1999). "The Laser Vegetation Imaging Sensor: A medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography." In: *ISPRS Journal of Photogrammetry and Remote Sensing* 54.2-3, pp. 115–122. ISSN: 09242716. DOI: 10.1016/S0924-2716(99)00002-7.
- Bohlin, J, Wallerman, J, and Fransson, JE (2012). "Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM." In: *Scandinavian Journal of Forest Research* 27.7, pp. 692–699. ISSN: 02827581. DOI: 10.1080/02827581.2012.686625.
- Bohlman, S and O'Brien, S (2006). "Allometry, adult stature and regeneration requirement of 65 tree species on Barro Colorado Island, Panama." In: *Journal of Tropical Ecology* 22.02, pp. 123–136. ISSN: 0266-4674. DOI: 10.1017/S0266467405003019. URL: [http://www.journals.cambridge.org/abstract{\\\_}S0266467405003019](http://www.journals.cambridge.org/abstract{\_}S0266467405003019).
- Bohn, FJ, May, F, and Huth, A (2017). "Species composition and forest structure explain the temperature sensitivity patterns of productivity in temperate forests." In: *Biogeosciences Discussions* 2017, pp. 1–26. DOI: 10.5194/bg-2017-335. URL: <https://www.biogeosciences-discuss.net/bg-2017-335/>.
- Bohn, FJ and Huth, A (2017). "The importance of forest structure to biodiversity–productivity relationships." In: *Royal Society Open Science* 4.1. ISSN: 20545703. DOI: 10.1098/rsos.160521.
- Bohn, FJ, Frank, K, and Huth, A (2014). "Of climate and its resulting tree growth: Simulating the productivity of temperate forests." In: *Ecological Modelling* 278, pp. 9–17. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2014.01.021. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0304380014000647>.
- Bollandsås, OM, Gregoire, TG, Næsset, E, and Øyen, BH (2013). "Detection of biomass change in a Norwegian mountain forest area using small footprint airborne laser scanner data." In: *Statistical Methods and Applications* 22.1, pp. 113–129. ISSN: 16182510. DOI: 10.1007/s10260-012-0220-5.
- Bonan, G (2008). "Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests." In: *Science* 320.5882, pp. 1444–1449. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1155121. URL: <http://science.sciencemag.org/content/320/5882/1444.abstract>.
- Botkin, DB, Janak, JF, and Wallis, JR (1972). "Some Ecological Consequences of a Computer Model of Forest Growth." In: *The Journal of Ecology* 60.3, pp. 849–872. ISSN: 00220477. DOI: 10.2307/2258570.
- Bouvier, M, Durrieu, S, Fournier, RA, and Renaud, JP (2015). "Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data." In: *Remote Sensing of Environment* 156, pp. 322–334. ISSN: 00344257. DOI: 10.1016/j.rse.2014.10.004. URL: <http://dx.doi.org/10.1016/j.rse.2014.10.004>.
- Breiman, L (2001). "Random Forests." In: *Machine Learning* 45, pp. 5–32. URL: <http://link.springer.com/article/10.1023/A:1010933404324>.
- Brinck, K, Fischer, R, Groeneveld, J, Lehmann, S, Paula, MDD, Pütz, S, Sexton, JO, Song, D, and Huth, A (2017). "High resolution analysis of tropical forest fragmentation and its impact on the global carbon cycle." In: *Nature Communications, Published online: 17 March 2017; | doi:10.1038/ncomms14855* 115, pp. 445–459. ISSN: 2041-1723. DOI: 10.1038/NCOMMS14855. URL: <http://www.nature.com/articles/ncomms14855>.

- Brokaw, NVL (1985). "Gap-Phase Regeneration in a Tropical Forest." In: *Ecology* 66.3, pp. 682–687. ISSN: 0012-9658. DOI: 10.2307/1940529. URL: <http://dx.doi.org/10.2307/1940529>.
- Bugmann, H (2001). "A review of forest gap models." In: *Climatic Change* 51.3-4, pp. 259–305. ISSN: 01650009. DOI: 10.1023/A:1012525626267. URL: <http://link.springer.com/article/10.1023/A:1012525626267>.
- Busing, RT and Solomon, AM (2006). *Modeling the effects of fire frequency and severity on forests in the northwestern United States*. Scientific Investigations Report 2006-5061, US Geological Survey. URL: <http://pubs.usgs.gov/sir/2006/5061/>.
- Calders, K, Newnham, G, Burt, A, Murphy, S, Raunonen, P, Herold, M, Culvenor, D, Avitabile, V, Disney, M, Armston, J, and Kaasalainen, M (2015). "Nondestructive estimates of above-ground biomass using terrestrial laser scanning." In: *Methods in Ecology and Evolution* 6.2, pp. 198–208. ISSN: 2041210X. DOI: 10.1111/2041-210X.12301.
- Campbell, GS and Norman, JM (2012). *An Introduction to Environmental Biophysics*. Berlin Heidelberg: Springer Science & Business Media. ISBN: 978-1-461-21626-1.
- Campbell, JB and Wynne, RH (2011). *Introduction to remote sensing*. 5th ed. New York, USA: Guilford Press, p. 667.
- Cao, L, Coops, NC, Innes, JL, Sheppard, SRJ, Fu, L, Ruan, H, and She, G (2016). "Estimation of forest biomass dynamics in subtropical forests using multi-temporal airborne LiDAR data." In: *Remote Sensing of Environment* 178, pp. 158–171. ISSN: 00344257. DOI: 10.1016/j.rse.2016.03.012. URL: <http://dx.doi.org/10.1016/j.rse.2016.03.012>.
- Carlowitz, HC von (1713). *Sylvicultura oeconomica, oder, Hausswirthliche Nachricht und naturmassige Anweisung zur wilden Baumzucht*.
- Cazcarra-Bes, V, Tello-Alonso, M, Fischer, R, Heym, M, and Papathanassiou, K (2017). "Monitoring of Forest Structure Dynamics by Means of L-Band SAR Tomography." In: *Remote Sensing* 9.12, p. 1229. ISSN: 2072-4292. DOI: 10.3390/rs9121229. URL: <http://www.mdpi.com/2072-4292/9/12/1229>.
- Chao, S (2012). *Forest peoples: numbers across the world*. Tech. rep., 24 pp. URL: [http://www.forestpeoples.org/sites/fpp/files/publication/2012/05/forest-peoples-numbers-across-world-final\\_0.pdf](http://www.forestpeoples.org/sites/fpp/files/publication/2012/05/forest-peoples-numbers-across-world-final_0.pdf).
- Chave, J, Andalo, C, Brown, S, and Cairns, M (2005). "Tree allometry and improved estimation of carbon stocks and balance in tropical forests." In: *Oecologia* 145.1, pp. 87–99. URL: <http://link.springer.com/article/10.1007/s00442-005-0100-x>.
- Chave, J, Condit, R, Lao, S, Caspersen, JP, Foster, RB, and Hubbell, SP (2003). "Spatial and temporal variation of biomass in a tropical forest: results from a large census plot in Panama." In: *Journal of Ecology* 91.2, pp. 240–252. ISSN: 00220477. DOI: 10.1046/j.1365-2745.2003.00757.x. URL: <http://doi.wiley.com/10.1046/j.1365-2745.2003.00757.x>.
- Chave, J, Coomes, D, Jansen, S, Lewis, SL, Swenson, NG, and Zanne, AE (2009). "Towards a worldwide wood economics spectrum." In: *Ecology Letters* 12.4, pp. 351–366. ISSN: 1461023X. DOI: 10.1111/j.1461-0248.2009.01285.x.
- Chave, J, Réjou-Méchain, M, Búrquez, A, Chidumayo, E, Colgan, MS, Delitti, WBC, Duque, A, Eid, T, Fearnside, PM, Goodman, RC, Henry, M, Martínez-Yrizar, A, Mugasha, WA, Muller-Landau, HC, Mencuccini, M, Nelson, BW, Ngomanda, A, Nogueira, EM, Ortiz-Malavassi, E, Pélissier, R, Ploton, P, Ryan, CM, Saldarriaga, JG, and Vieilledent, G (2014). "Improved allometric models to estimate the aboveground biomass of tropical trees." In: *Global Change Biology* 20.10, pp. 3177–3190. ISSN: 13652486. DOI: 10.1111/gcb.12629. URL: <http://www.ncbi.nlm.nih.gov/pubmed/24817483>.
- Chazdon, RL, Brancalion, PHS, Laestadius, L, Bennett-Curry, A, Buckingham, K, Kumar, C, Moll-Rocek, J, Vieira, ICG, and Wilson, SJ (2016). "When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration." In: *Ambio* 45.5, pp. 538–550. ISSN: 0044-7447. DOI: 10.1007/s13280-016-0772-y. URL: <http://link.springer.com/10.1007/s13280-016-0772-y>.
- Chen, Q (2010). "Retrieving vegetation height of forests and woodlands over mountainous areas in the Pacific Coast region using satellite laser altimetry." In: *Remote Sensing of Environment* 114.7, pp. 1610–1627. ISSN: 00344257. DOI: 10.1016/j.rse.2010.02.016. URL: <http://dx.doi.org/10.1016/j.rse.2010.02.016>.

- Chen, Q (2013). "Lidar remote sensing of vegetation biomass." In: *Remote sensing of natural resources*. Vol. 399, pp. 399–420. ISBN: 978-1-4665-5692-8. DOI: doi:10.1201/b15159-28.
- Chen, Q (2015). "Modeling aboveground tree woody biomass using national-scale allometric methods and airborne lidar." In: *ISPRS Journal of Photogrammetry and Remote Sensing* 106, pp. 95–106. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2015.05.007. URL: <http://dx.doi.org/10.1016/j.isprsjprs.2015.05.007>.
- Chen, Q, McRoberts, RE, Wang, C, and Radtke, PJ (2016). "Forest aboveground biomass mapping and estimation across multiple spatial scales using model-based inference." In: *Remote Sensing of Environment* 184, pp. 350–360. ISSN: 00344257. DOI: 10.1016/j.rse.2016.07.023. URL: <http://dx.doi.org/10.1016/j.rse.2016.07.023>.
- Cloude, SR and Papathanassiou, KP (1998). "Polarimetric SAR interferometry." In: *IEEE Transactions on Geoscience and Remote Sensing* 36.5, pp. 1551–1565. DOI: 10.1109/36.718859.
- Cochrane, MA, Alencar, A, Schulze, M, Souza, C, Nepstad, D, Lefebvre, P, and Davidson, E (1999). "Positive Feedbacks in the Fire Dynamic of Closed Canopy Tropical Forests." In: *Science* 284.5421, pp. 1832–1835. ISSN: 00368075. DOI: 10.1126/science.284.5421.1832. URL: <http://www.sciencemag.org/cgi/doi/10.1126/science.284.5421.1832>.
- Condit, R (1998). *Tropical forest census plots*. Berlin, Germany and George Town, Texas: Springer-Verlag and R. G. Landes Company, p. 211. ISBN: 3540641440. DOI: 10.1007/978-3-662-03664-8.
- Condit, R, Lao, S, Pérez, R, Dolins, S, Foster, R, and Hubbell, S (2012). *Barro Colorado Forest Census Plot Data*. DOI: 10.5479/data.bci.20130603.
- Condit, R, Robinson, WD, Ibáñez, R, Aguilar, S, Sanjur, A, Martínez, R, Stallard, RF, García, T, Angehr, GR, Petit, L, Wright, SJ, Robinson, TR, and Heckadon, S (2001). "The Status of the Panama Canal Watershed and Its Biodiversity at the Beginning of the 21st Century." In: *BioScience* 51.5, pp. 389–398. DOI: [https://doi.org/10.1641/0006-3568\(2001\)051\[0389:TS0TPC\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0389:TS0TPC]2.0.CO;2).
- Cook, BD, Corp, La, Nelson, RF, Middleton, EM, Morton, DC, McCorkel, JT, Masek, JG, Ranson, KJ, Ly, V, and Montesano, PM (2013). "NASA goddard's LiDAR, hyperspectral and thermal (G-LiHT) airborne imager." In: *Remote Sensing* 5.8, pp. 4045–4066. ISSN: 20724292. DOI: 10.3390/rs5084045.
- Coomes, DA, Dalponte, M, Jucker, T, Asner, GP, Banin, LF, Burslem, DFRP, Lewis, SL, Nilus, R, Phillips, OL, Phua, MH, and Qie, L (2017). "Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data." In: *Remote Sensing of Environment* 194, pp. 77–88. ISSN: 00344257. DOI: 10.1016/j.rse.2017.03.017. URL: <http://dx.doi.org/10.1016/j.rse.2017.03.017>.
- Couteron, P, Barbier, N, and Gautier, D (2006). "Textural ordination based on fourier spectral decomposition: A method to analyze and compare landscape patterns." In: *Landscape Ecology* 21.4, pp. 555–567. ISSN: 09212973. DOI: 10.1007/s10980-005-2166-6.
- Crowther, TW et al. (2015). "Mapping tree density at a global scale." In: *Nature* 525.7568, pp. 201–205. ISSN: 14764687. DOI: 10.1038/nature14967. arXiv: 9605103 [cs].
- Curtis, PG, Slay, CM, Harris, NL, Tyukavina, A, and Hansen, MC (2018). "Classifying drivers of global forest loss." In: *Science* 361.6407, pp. 1108–1111. ISSN: 10959203. DOI: 10.1126/science.aau3445. arXiv: 1709.03428. URL: <http://www.sciencemag.org/lookup/doi/10.1126/science.aau3445>.
- Dantas de Paula, M, Groeneveld, J, and Huth, a (2015). "Tropical forest degradation and recovery in fragmented landscapes — Simulating changes in tree community, forest hydrology and carbon balance." In: *Global Ecology and Conservation* 3, pp. 664–677. ISSN: 23519894. DOI: 10.1016/j.gecco.2015.03.004. URL: <http://linkinghub.elsevier.com/retrieve/pii/S235198941500030X>.
- De Sy, V, Herold, M, Achard, F, Asner, GP, Held, A, Kellndorfer, J, and Verbesselt, J (2012). "Synergies of multiple remote sensing data sources for REDD+ monitoring." In: *Current Opinion in Environmental Sustainability* 4.6, pp. 696–706. ISSN: 18773435. DOI: 10.1016/j.cosust.2012.09.013. URL: <http://linkinghub.elsevier.com/retrieve/pii/S1877343512001200>.
- Dengler, J, Jansen, F, Glöckler, F, Peet, RK, Cáceres, M de, Chytrý, M, Ewald, J, Oldeland, J, Lopez-Gonzalez, G, Finckh, M, Mucina, L, Rodwell, JS, Schaminée, JHJ, and Spencer,

- N (2011). "The Global Index of Vegetation-Plot Databases (GIVD): A new resource for vegetation science." In: *Journal of Vegetation Science* 22.4, pp. 582–597. ISSN: 11009233. DOI: 10.1111/j.1654-1103.2011.01265.x.
- Diamond, J (2005). *Collapse: How societies choose to fail or succeed*. Viking Press. ISBN: 978-0670033379.
- Dislich, C, Günter, S, Homeier, J, Schröder, B, and Huth, A (2009). "Simulating forest dynamics of a tropical montane forest in South Ecuador." In: *Erdkunde* 63.4, pp. 347–364. ISSN: 00140015. DOI: 10.3112/erdkunde.2009.04.05. URL: <http://www.erdkunde.uni-bonn.de/archive/2009/simulating-forest-dynamics-of-a-tropical-montane-forest-in-south-ecuador>.
- Disney, MI, Kalogirou, V, Lewis, P, Prieto-Blanco, A, Hancock, S, and Pfeifer, M (2010). "Simulating the impact of discrete-return lidar system and survey characteristics over young conifer and broadleaf forests." In: *Remote Sensing of Environment* 114.7, pp. 1546–1560. ISSN: 00344257. DOI: 10.1016/j.rse.2010.02.009.
- Disney, M (2018). "Terrestrial LiDAR: a 3D revolution in how we look at trees." In: *New Phytologist*. ISSN: 0028646X. DOI: 10.1111/nph.15517. URL: <http://doi.wiley.com/10.1111/nph.15517>.
- Drake, JB, Dubayah, RO, Clark, DB, Knox, RG, Blair, JB, Hofton, MA, Chazdon, RL, Weishampel, JF, and Prince, S (2002). "Estimation of tropical forest structural characteristics, using large-footprint lidar." In: *Remote Sensing of Environment* 79.2-3, pp. 305–319. ISSN: 00344257. DOI: 10.1016/S0034-4257(01)00281-4. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425701002814>.
- Drusch, M, Del Bello, U, Carlier, S, Colin, O, Fernandez, V, Gascon, F, Hoersch, B, Isola, C, Laberinti, P, Martimort, P, Meygret, A, Spoto, F, Sy, O, Marchese, F, and Bargellini, P (2012). "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services." In: *Remote Sensing of Environment* 120, pp. 25–36. ISSN: 00344257. DOI: 10.1016/j.rse.2011.11.026.
- Dubayah, RO, Sheldon, SL, Clark, DB, Hofton, MA, Blair, JB, Hurtt, GC, and Chazdon, RL (2010). "Estimation of tropical forest height and biomass dynamics using lidar remote sensing at la Selva, Costa Rica." In: *Journal of Geophysical Research: Biogeosciences* 115.2, pp. 1–17. ISSN: 01480227. DOI: 10.1029/2009JG000933. URL: <http://doi.wiley.com/10.1029/2009JG000933>.
- Duncanson, LI, Cook, BD, Hurtt, GC, and Dubayah, RO (2014). "An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems." In: *Remote Sensing of Environment* 154, pp. 378–386. ISSN: 00344257. DOI: 10.1016/j.rse.2013.07.044. URL: <http://dx.doi.org/10.1016/j.rse.2013.07.044>.
- Duncanson, LI, Dubayah, RO, Cook, BD, Rosette, J, and Parker, G (2015). "The importance of spatial detail: Assessing the utility of individual crown information and scaling approaches for lidar-based biomass density estimation." In: *Remote Sensing of Environment* 168, pp. 102–112. ISSN: 00344257. DOI: 10.1016/j.rse.2015.06.021. URL: <http://dx.doi.org/10.1016/j.rse.2015.06.021>.
- Duncanson, L, Huang, W, Johnson, K, Swatantran, A, McRoberts, RE, and Dubayah, R (2017). "Implications of allometric model selection for county-level biomass mapping." In: *Carbon Balance and Management* 12.1. ISSN: 17500680. DOI: 10.1186/s13021-017-0086-9.
- Elias, M and Potvin, C (2003). "Assessing inter- and intra-specific variation in trunk carbon concentration for 32 neotropical tree species." In: *Canadian Journal of Forest Research* 33.6, pp. 1039–1045. ISSN: 0045-5067. DOI: 10.1139/x03-018. URL: <http://www.nrcresearchpress.com/doi/abs/10.1139/x03-018>.
- Endo, T, Sawada, Y, Kobayashi, T, and Sawada, H (2012). "Developing a 3D Waveform Lidar Simulator for Forest." In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XXXIX-B8. September. Melbourne, Australia, pp. 399–402. DOI: 10.5194/isprsarchives-XXXIX-B8-399-2012. URL: <http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XXXIX-B8/399/2012/isprsarchives-XXXIX-B8-399-2012.pdf>.
- Englhart, S, Jubanski, J, and Siegert, F (2013). "Quantifying Dynamics in Tropical Peat Swamp Forest Biomass with Multi-Temporal LiDAR Datasets." In: *Remote Sensing* 5.5, pp. 2368–2388.

- ISSN: 2072-4292. DOI: 10.3390/rs5052368. URL: <http://www.mdpi.com/2072-4292/5/5/2368/>.
- FAO (2015). *Global Forest Resources Assessment 2015*. Rome, Italy, p. 56. ISBN: 9789251088210. URL: <http://www.fao.org/forestry/fra2005/en/>.
- Falke, J (2017). *3D Adaptive Mean Shift Algorithm: Single Tree Detection in Tropical Forests from Airborne LiDAR Data Implemented in R*. Tech. rep. Leipzig: Helmholtz Centre for Environmental Research, p. 16. URL: <https://www.hnee.de/en/Programmes/Master-degree/Forest-Information-Technology/students-results/projects-results/summer-term-2017/Student-Research-Projects-2017-E9256.htm>.
- Falkowski, MJ, Hudak, AT, Crookston, NL, Gessler, PE, Uebler, EH, and Smith, AM (2010). "Landscape-scale parameterization of a tree-level forest growth model: a k-nearest neighbor imputation approach incorporating LiDAR data." In: *Canadian Journal of Forest Research* 40.2, pp. 184–199. ISSN: 0045-5067. DOI: 10.1139/X09-183. URL: <http://www.nrcresearchpress.com/doi/abs/10.1139/x09-183{\#}.VpVr3fkrKUL>.
- Fassnacht, FE, Hartig, F, Latifi, H, Berger, C, Hernández, J, Corvalán, P, and Koch, B (2014). "Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass." In: *Remote Sensing of Environment* 154.1, pp. 102–114. ISSN: 00344257. DOI: 10.1016/j.rse.2014.07.028.
- Fassnacht, FE, Latifi, H, Stereńczak, K, Modzelewska, A, Lefsky, M, Waser, LT, Straub, C, and Ghosh, A (2016). "Review of studies on tree species classification from remotely sensed data." In: *Remote Sensing of Environment* 186, pp. 64–87. ISSN: 00344257. DOI: 10.1016/j.rse.2016.08.013. arXiv: arXiv:1011.1669v3.
- Ferraz, A, Saatchi, S, Mallet, C, and Meyer, V (2016). "Lidar detection of individual tree size in tropical forests." In: *Remote Sensing of Environment* 183, pp. 318–333. ISSN: 00344257. DOI: 10.1016/j.rse.2016.05.028. URL: <http://dx.doi.org/10.1016/j.rse.2016.05.028>.
- Fine, P and Ree, R (2006). "Evidence for a Time-Integrated Species-Area Effect on the Latitudinal Gradient in Tree Diversity." In: *The American Naturalist* 168.6, pp. 796–804. ISSN: 0003-0147. DOI: 10.1086/508635. URL: <http://www.journals.uchicago.edu/doi/10.1086/508635>.
- Fischer, R (2013). "Modellierung der Dynamik afrikanischer Tropenwälder. Analyse des Einflusses von Störungen auf tropische Wälder mit Hilfe des Waldmodells FORMIND." PhD thesis. Dissertation, Universität Osnabrück, Germany.
- Fischer, R, Armstrong, A, Shugart, HH, and Huth, A (2014). "Simulating the impacts of reduced rainfall on carbon stocks and net ecosystem exchange in a tropical forest." In: *Environmental Modelling and Software* 52, pp. 200–206. ISSN: 13648152. DOI: 10.1016/j.envsoft.2013.10.026. URL: <http://dx.doi.org/10.1016/j.envsoft.2013.10.026>.
- Fischer, R, Ensslin, A, Rutten, G, Fischer, M, Schellenberger Costa, D, Kleyer, M, Hemp, A, Paulick, S, and Huth, A (2015). "Simulating Carbon Stocks and Fluxes of an African Tropical Montane Forest with an Individual-Based Forest Model." In: *Plos One* 10.4, e0123300. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0123300. URL: <http://dx.plos.org/10.1371/journal.pone.0123300>.
- Fischer, R, Bohn, F, Dantas de Paula, M, Dislich, C, Groeneveld, J, Gutiérrez, AG, Kazmierczak, M, Knapp, N, Lehmann, S, Paulick, S, Pütz, S, Rödig, E, Taubert, F, Köhler, P, and Huth, A (2016). "Lessons learned from applying a forest gap model to understand ecosystem and carbon dynamics of complex tropical forests." In: *Ecological Modelling* 326, pp. 124–133. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2015.11.018. URL: <http://www.sciencedirect.com/science/article/pii/S0304380015005505>.
- Fischer, R, Rödig, E, and Huth, A (2018). "Consequences of a Reduced Number of Plant Functional Types for the Simulation of Forest Productivity." In: *Forests* 9.8, p. 460. ISSN: 13973142. DOI: 10.1111/j.1399-3046.2005.00332.x. URL: <http://www.mdpi.com/1999-4907/9/8/460>.
- Fischer, R, Knapp, N, Bohn, F, and Huth, A (2019). "Remote Sensing Measurements of Forest Structure Types for Ecosystem Service Mapping." In: *Atlas of Ecosystem Services: Drivers, Risks, and Societal Responses*. Ed. by M Schröter, A Bonn, S Klotz, R Seppelt, and C Baessler. 1st ed. Springer International Publishing, p. 430. ISBN: 978-3-319-96229-0. DOI: 10.1007/978-3-319-96229-0.

- Fisher, RA, Koven, CD, Anderegg, WR, Christoffersen, BO, Dietze, MC, Farrior, CE, Holm, JA, Hurtt, GC, Knox, RG, Lawrence, PJ, Lichstein, JW, Longo, M, Matheny, AM, Medvigy, D, Muller-Landau, HC, Powell, TL, Serbin, SP, Sato, H, Shuman, JK, Smith, B, Trugman, AT, Viskari, T, Verbeeck, H, Weng, E, Xu, C, Xu, X, Zhang, T, and Moorcroft, PR (2018). "Vegetation demographics in Earth System Models: A review of progress and priorities." In: *Global Change Biology* 24.1, pp. 35–54. ISSN: 13652486. DOI: 10.1111/gcb.13910. arXiv: 0608246v3 [arXiv:physics].
- Frazer, GW, Magnussen, S, Wulder, MA, and Niemann, KO (2011). "Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass." In: *Remote Sensing of Environment* 115.2, pp. 636–649. ISSN: 00344257. DOI: 10.1016/j.rse.2010.10.008. URL: <http://dx.doi.org/10.1016/j.rse.2010.10.008>.
- Getzin, S, Fischer, R, Knapp, N, and Huth, A (2017). "Using airborne LiDAR to assess spatial heterogeneity in forest structure on Mount Kilimanjaro." In: *Landscape Ecology* 32.9, pp. 1881–1894. ISSN: 15729761. DOI: 10.1007/s10980-017-0550-7. URL: <http://link.springer.com/10.1007/s10980-017-0550-7>.
- Gibbs, HK, Brown, S, Niles, JO, and Foley, JA (2007). "Monitoring and estimating tropical forest carbon stocks: making REDD a reality." In: *Environmental Research Letters* 2.2007, p. 045023. ISSN: 1748-9326. DOI: 10.1088/1748-9326/2/4/045023. URL: <http://stacks.iop.org/1748-9326/2/i=4/a=045023?key=crossref.4118e8af5a9a3ac02c1bb32f8a92c50f>.
- Goetz, S and Dubayah, R (2011). "Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change." In: *Carbon Management* 2.April, pp. 231–244. ISSN: 1758-3004. DOI: 10.4155/cmt.11.18. URL: <http://www.future-science.com/doi/pdf/10.4155/cmt.11.18>.
- Goodwin, NR, Coops, NC, and Culvenor, DS (2007). "Development of a simulation model to predict LiDAR interception in forested environments." In: *Remote Sensing of Environment* 111.4, pp. 481–492. ISSN: 00344257. DOI: 10.1016/j.rse.2007.04.001. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425707001496>.
- Graham, LC (1974). "Synthetic interferometer radar for topographic mapping." In: *Proceedings of the IEEE* 62.6, pp. 763–768. DOI: 10.1109/PROC.1974.9516.
- Grimm, V and Berger, U (2016). "Structural realism, emergence, and predictions in next-generation ecological modelling: Synthesis from a special issue." In: *Ecological Modelling* 326, pp. 177–187. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2016.01.001. URL: <http://dx.doi.org/10.1016/j.ecolmodel.2016.01.001>.
- Groeneveld, J, Alves, L, Bernacci, L, Catharino, E, Knogge, C, Metzger, J, Pütz, S, and Huth, a (2009). "The impact of fragmentation and density regulation on forest succession in the Atlantic rain forest." In: *Ecological Modelling* 220.19, pp. 2450–2459. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2009.06.015. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0304380009004153>.
- Guanter, L et al. (2015). "The EnMAP spaceborne imaging spectroscopy mission for earth observation." In: *Remote Sensing* 7.7, pp. 8830–8857. ISSN: 20724292. DOI: 10.3390/rs70708830.
- Guo, Q, Su, Y, Hu, T, Zhao, X, Wu, F, Li, Y, Liu, J, Chen, L, Xu, G, Lin, G, Zheng, Y, Lin, Y, Mi, X, Fei, L, and Wang, X (2017). "An integrated UAV-borne lidar system for 3D habitat mapping in three forest ecosystems across China." In: *International Journal of Remote Sensing* 20, pp. 1–19. ISSN: 0143-1161. DOI: 10.1080/01431161.2017.1285083. URL: <https://www.tandfonline.com/doi/full/10.1080/01431161.2017.1285083>.
- Hancock, S, Armston, J, Tang, H, Patterson, PL, Healey, SP, Marselis, S, Duncanson, L, Hofton, MA, Kellner, JR, Luthcke, SB, Sun, X, Blair, JB, and Dubayah, R (2017). "The GEDI Performance Tool." In: *AGU Fall Meeting Abstracts*.
- Hansen, MC, DeFries, RS, Townshend, JR, Sohlberg, R, Dimiceli, C, and Carroll, M (2002). "Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data." In: *Remote Sensing of Environment* 83.1-2, pp. 303–319. ISSN: 00344257. DOI: 10.1016/S0034-4257(02)00079-2.
- Hansen, MC, Potapov, PV, Moore, R, Hancher, M, Turubanova, Sa, Tyukavina, a, Thau, D, Stehman, SV, Goetz, SJ, Loveland, TR, Kommareddy, a, Egorov, a, Chini, L, Justice, CO, and

- Townshend, JRG (2013). "High-resolution global maps of 21st-century forest cover change." In: *Science (New York, N.Y.)* 342.2013, pp. 850–3. ISSN: 1095-9203. DOI: 10.1126/science.1244693. arXiv: 1011.1669v3. URL: <http://www.ncbi.nlm.nih.gov/pubmed/24233722>.
- Hansen, MC, Krylov, A, Tyukavina, A, Potapov, PV, Turubanova, S, Zutta, B, Ifo, S, Margono, B, Stolle, F, and Moore, R (2016). "Humid tropical forest disturbance alerts using Landsat data." In: *Environmental Research Letters* 11.3, p. 034008. ISSN: 1748-9326. URL: <http://stacks.iop.org/1748-9326/11/i=3/a=034008>.
- Haralick, RM, Shanmugam, K, and Dinstein, I (1973). "Textural Features for Image Classification." In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC-3.6, pp. 610–621. ISSN: 0018-9472. DOI: 10.1109/TSMC.1973.4309314. URL: <http://ieeexplore.ieee.org/document/4309314/>.
- Harding, DJ, Lefsky, MA, Parker, GG, and Blair, JB (2001). "Laser altimeter canopy height profiles methods and validation for closed-canopy, broadleaf forests." In: *Remote Sensing of Environment* 76.3, pp. 283–297. ISSN: 00344257. DOI: 10.1016/S0034-4257(00)00210-8.
- Harms, KE, Condit, R, Hubbell, SP, and Foster, RB (2001). "Habitat association of trees and shrubs in a 50-ha neotropical forest plot." In: *Journal of Ecology* 89.6, pp. 947–959. DOI: 10.1111/j.1365-2745.2001.00615.x.
- Harris, NL, Brown, S, Hagen, SC, Saatchi, SS, Petrova, S, Salas, W, Hansen, MC, Potapov, PV, and Lotsch, a (2012). "Baseline Map of Carbon Emissions from Deforestation in Tropical Regions." In: *Science* 336.6088, pp. 1573–1576. ISSN: 0036-8075. DOI: 10.1126/science.1217962. URL: <http://www.sciencemag.org/cgi/doi/10.1126/science.1217962>.
- He, L, Chen, JM, Pan, Y, Birdsey, R, and Kattge, J (2012). "Relationships between net primary productivity and forest stand age in U.S. forests." In: *Global Biogeochemical Cycles* 26.3, pp. 1–19. ISSN: 08866236. DOI: 10.1029/2010GB003942.
- Héroult, B and Piponirot, C (2018). "Key drivers of ecosystem recovery after disturbance in a neotropical forest." In: *Forest Ecosystems* 5.1, p. 2. ISSN: 2197-5620. DOI: 10.1186/s40663-017-0126-7. URL: <https://forestecosyst.springeropen.com/articles/10.1186/s40663-017-0126-7>.
- Hiltner, U, Bräuning, A, Gebrekirstos, A, Huth, A, and Fischer, R (2016). "Impacts of precipitation variability on the dynamics of a dry tropical montane forest." In: *Ecological Modelling* 320, pp. 92–101. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2015.09.021. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0304380015004524>.
- Hiltner, U, Huth, A, Bräuning, A, Héroult, B, and Fischer, R (2018). "Simulation of succession in a neotropical forest: High selective logging intensities prolong the recovery times of ecosystem functions." In: *Forest Ecology and Management* 430. June, pp. 517–525. ISSN: 03781127. DOI: 10.1016/j.foreco.2018.08.042. URL: <https://doi.org/10.1016/j.foreco.2018.08.042>.
- Hoch, G and Körner, C (2005). "Growth, demography and carbon relations of *Polylepis* trees at the world's highest treeline." In: *Functional Ecology* 19.6, pp. 941–951. ISSN: 02698463. DOI: 10.1111/j.1365-2435.2005.01040.x.
- Hollmann, R, Merchant, CJ, Saunders, R, Downy, C, Buchwitz, M, Cazenave, A, Chuvieco, E, Defourny, P, De Leeuw, G, Forsberg, R, Holzer-Popp, T, Paul, F, Sandven, S, Sathyendranath, S, Van Roozendaal, M, and Wagner, W (2013). "The ESA climate change initiative: Satellite data records for essential climate variables." In: *Bulletin of the American Meteorological Society* 94.10, pp. 1541–1552. ISSN: 00030007. DOI: 10.1175/BAMS-D-11-00254.1.
- Houghton, RA, Hall, F, and Goetz, SJ (2009). "Importance of biomass in the global carbon cycle." In: *Journal of Geophysical Research: Biogeosciences* 114.3, pp. 1–13. ISSN: 01480227. DOI: 10.1029/2009JG000935.
- Houghton, RA, Baccini, A, and Walker, WS (2018). "Where is the residual terrestrial carbon sink?" In: *Global Change Biology* 24.8, pp. 3277–3279. ISSN: 13652486. DOI: 10.1111/gcb.14313. URL: <http://doi.wiley.com/10.1111/gcb.14313>.
- Hubbell, S, Foster, R, O'Brien, S, Harms, K, Condit, R, Wechsler, B, Wright, S, and Loo de Lao, S (1999). "Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest." In: *Science* 283.5401, pp. 554–557. URL: <http://www.sciencemag.org/content/283/5401/554.short>.

- Hubbell, SP, Condit, R, and Foster, RB (2005). *Barro Colorado Forest Census Plot Data*. URL: <http://ctfs.si.edu/webatlas/datasets/bci>.
- Hudak, AT, Strand, EK, Vierling, La, Byrne, JC, Eitel, JUH, Martinuzzi, S, and Falkowski, MJ (2012). "Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys." In: *Remote Sensing of Environment* 123, pp. 25–40. ISSN: 00344257. DOI: 10.1016/j.rse.2012.02.023. URL: <http://dx.doi.org/10.1016/j.rse.2012.02.023>.
- Hurttt, GC, Fisk, J, Thomas, RQ, Dubayah, R, Moorcroft, PR, and Shugart, HH (2010). "Linking models and data on vegetation structure." In: *Journal of Geophysical Research* 115, pp. 1–11. ISSN: 0148-0227. DOI: 10.1029/2009JG000937.
- Hurttt, GC, Dubayah, R, Drake, J, Moorcroft, PR, Pacala, SW, Blair, JB, and Fearon, MG (2004). "Beyond Potential Vegetation: Combining Lidar Data and a Height-Structured Model for Carbon Studies." In: *Ecological Applications* 14.3, pp. 873–883. ISSN: 1051-0761. DOI: 10.1890/02-5317. URL: <http://doi.wiley.com/10.1890/02-5317>.
- Huston, M, DeAngelis, D, and Post, W (1988). "New Models Unify Computer be explained by interactions among individual organisms." In: *Bioscience* 38.10, pp. 682–691.
- Huth, A (1999). *Modellierung des Wachstums und der Nutzung von tropischem Regenwald*, p. 318.
- Huth, A, Drechsler, M, and Köhler, P (2004). "Multicriteria evaluation of simulated logging scenarios in a tropical rain forest." In: *Journal of Environmental Management* 71.4, pp. 321–333. ISSN: 03014797. DOI: 10.1016/j.jenvman.2004.03.008. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0301479704000568>.
- IPCC (2007). *Climate Change 2007: The Physical Science Basis*. Tech. rep. Intergovernmental Panel on Climate Change, p. 996.
- IPCC (2013). *Climate Change 2013: The Physical Science Basis, Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by TF Stocker, D Qin, GK Plattner, MMB Tignor, SK Allen, J Boschung, A Nauels, Y Xia, V Bex, and PM Midgley. New York, USA: Cambridge University Press.
- Immitzer, M, Atzberger, C, and Koukal, T (2012). "Tree species classification with Random forest using very high spatial resolution 8-band worldView-2 satellite data." In: *Remote Sensing* 4.9, pp. 2661–2693. ISSN: 20724292. DOI: 10.3390/rs4092661.
- Isenburg, M (2011). *LAStools - efficient tools for LiDAR processing*. URL: <http://lastools.org>.
- Jacquemoud, S, Verhoef, W, Baret, F, Bacour, C, Zarco-Tejada, PJ, Asner, GP, François, C, and Ustin, SL (2009). "PROSPECT + SAIL models: A review of use for vegetation characterization." In: *Remote Sensing of Environment* 113.SUPPL. 1, S56–S66. ISSN: 00344257. DOI: 10.1016/j.rse.2008.01.026. URL: <http://dx.doi.org/10.1016/j.rse.2008.01.026>.
- Jetz, W, Cavender-Bares, J, Pavlick, R, Schimel, D, Davis, FW, Asner, GP, Guralnick, R, Kattge, J, Latimer, AM, Moorcroft, P, Schaepman, ME, Schildhauer, MP, Schneider, FD, Schrodtt, F, Stahl, U, and Ustin, SL (2016). "Monitoring plant functional diversity from space." In: *Nature Plants* 2.3, pp. 1–5. ISSN: 20550278. DOI: 10.1038/NPLANTS.2016.24. URL: <http://dx.doi.org/10.1038/nplants.2016.24>.
- Jing, L, Hu, B, Noland, T, and Li, J (2012). "An individual tree crown delineation method based on multi-scale segmentation of imagery." In: *ISPRS Journal of Photogrammetry and Remote Sensing* 70, pp. 88–98. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2012.04.003. URL: <http://dx.doi.org/10.1016/j.isprsjprs.2012.04.003>.
- Jones, HG (2013). *Plants and Microclimate - A Quantitative Approach to Environmental Plant Physiology*. Cambridge: Cambridge University Press. ISBN: 978-1-107-51163-7.
- Jubanski, J, Ballhorn, U, Kronseder, K, Franke, J, and Siegert, F (2013). "Detection of large above-ground biomass variability in lowland forest ecosystems by airborne LiDAR." In: *Biogeosciences* 10.6, pp. 3917–3930. ISSN: 17264170. DOI: 10.5194/bg-10-3917-2013. URL: <http://www.biogeosciences.net/10/3917/2013/>.
- Jucker, T et al. (2017). "Allometric equations for integrating remote sensing imagery into forest monitoring programmes." In: *Global Change Biology* 23.1, pp. 177–190. ISSN: 13652486. DOI: 10.1111/gcb.13388.
- Juffe-Bignoli, D, Burgess, ND, Bingham, H, Belle, EMS, De Lima, MG, Deguignet, M, Bertzky, B, Milam, AN, Martinez-Lopez, J, Lewis, E, and Others (2014). "Protected planet report 2014."

- In: *UNEP-WCMC* 11, p. 80. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.735.3685&rep=rep1&type=pdf>.
- Justice, CO, Townsend, JRG, Vermote, EF, Masuoka, E, Wolfe, RE, Saleous, N, Roy, DP, and Morisette, JT (2002). "An overview of MODIS Land data processing and products." In: *Remote Sensing of Environment* 83, pp. 3–15. DOI: 10.1016/S0034-4257(02)00084-6.
- Kammesheidt, L, Köhler, P, and Huth, A (2001). "Sustainable Timber Harvesting in Venezuela: A Modelling Approach." In: *Journal of Applied Ecology* 38.4, pp. 756–770. DOI: 10.1046/j.1365-2664.2001.00629.x.
- Kazmierczak, M, Wiegand, T, and Huth, A (2014). "A neutral vs. non-neutral parametrizations of a physiological forest gap model." In: *Ecological Modelling* 288, pp. 94–102. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2014.05.002.
- Kennel, P, Tramon, M, Barbier, N, and Vincent, G (2013). "Canopy height model characteristics derived from airborne laser scanning and its effectiveness in discriminating various tropical moist forest types." In: *International Journal of Remote Sensing* 34.24, pp. 8917–8935. ISSN: 0143-1161. DOI: 10.1080/01431161.2013.858846. URL: <http://www.tandfonline.com/doi/abs/10.1080/01431161.2013.858846>.
- Knapp, N, Fischer, R, and Huth, A (2018a). "Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states." In: *Remote Sensing of Environment* 205, pp. 199–209. ISSN: 00344257. DOI: 10.1016/j.rse.2017.11.018. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425717305679>.
- Knapp, N, Huth, A, Kugler, F, Papathanassiou, K, Condit, R, Hubbell, SP, and Fischer, R (2018b). "Model-assisted estimation of tropical forest biomass change: A comparison of approaches." In: *Remote Sensing* 10.5, pp. 1–23. ISSN: 20724292. DOI: 10.3390/rs10050731. URL: <https://www.mdpi.com/2072-4292/10/5/731>.
- Köhler, P and Huth, A (2010). "Towards ground-truthing of spaceborne estimates of above-ground life biomass and leaf area index in tropical rain forests." In: *Biogeosciences* 7.8, pp. 2531–2543. ISSN: 17264170. DOI: 10.5194/bg-7-2531-2010. URL: <http://www.biogeosciences.net/7/2531/2010/>.
- Köhler, P and Huth, A (1998). "The effects of tree species grouping in tropical rainforest modelling: Simulations with the individual-based model Formind." In: *Ecological Modelling* 109.3, pp. 301–321. ISSN: 03043800. DOI: 10.1016/S0304-3800(98)00066-0.
- Kohyama, T (1992). "Size-structured multi-species model of rain forest trees." In: *Functional Ecology*, pp. 206–212. DOI: 10.2307/2389756.
- Kotchenova, SY, Shabanov, NV, Knyazikhin, Y, Davis, AB, Dubayah, RO, and Myneni, RB (2003). "Modeling lidar waveforms with time-dependent stochastic radiative transfer theory for remote estimations of forest structure." In: *Journal Of Geophysical Research-Atmospheres* 108.D15, p. 4484. ISSN: 0148-0227. DOI: 10.1029/2002JD003288. URL: <http://www.agu.org/pubs/crossref/2003/2002JD003288.shtml>{\%}5Cnpapers2://publication/doi/10.1029/2002JD003288.
- Král, K, Shue, J, Vrška, T, Gonzalez-Akre, EB, Parker, GG, McShea, WJ, and McMahon, SM (2016). "Fine-scale patch mosaic of developmental stages in Northeast American secondary temperate forests: the European perspective." In: *European Journal of Forest Research* 135.5, pp. 981–996. ISSN: 16124669. DOI: 10.1007/s10342-016-0988-1.
- Krieger, G, Moreira, A, Fiedler, H, Hajnsek, I, Werner, M, Younis, M, and Zink, M (2007). "TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry." In: *IEEE Trans. Geosci. Remote Sens.* 45.11, pp. 3317–3341. ISSN: 0196-2892. DOI: 10.1109/TGRS.2007.900693. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp={&}arnumber=4373373>.
- Krieger, G, Zink, M, Bachmann, M, Bräutigam, B, Schulze, D, Martone, M, Rizzoli, P, Steinbrecher, U, Antony, JW, De Zan, F, and Others (2013). "TanDEM-X: A radar interferometer with two formation-flying satellites." In: *Acta Astronautica* 89, pp. 83–98. DOI: 10.1016/j.actaastro.2013.03.008.
- Kugler, F, Schulze, D, Hajnsek, I, Pretzsch, H, and Papathanassiou, KP (2014). "TanDEM-X Pol-InSAR performance for forest height estimation." In: *IEEE Transactions on Geoscience and Remote Sensing* 52.10, pp. 6404–6422. ISSN: 01962892. DOI: 10.1109/TGRS.2013.2296533.

- Kuusik, A (2018). "Canopy Radiative Transfer Modeling." In: *Comprehensive Remote Sensing V. 3 Remote Sensing of Terrestrial Ecosystems*. Ed. by S Liang. 1988. Elsevier, pp. 9–22. DOI: 10.1016/B978-0-12-409548-9.10534-2. URL: <http://linkinghub.elsevier.com/retrieve/pii/B9780124095489105342>.
- Labrière, N, Tao, S, Chave, J, Scipal, K, Toan, TL, Abernethy, K, Alonso, A, Barbier, N, Bissengou, P, Casal, T, Davies, SJ, Ferraz, A, Hérault, B, Jaouen, G, Jeffery, KJ, Kenfack, D, Korte, L, Lewis, SL, Malhi, Y, Memiaghe, HR, Poulsen, JR, Réjou-Méchain, M, Villard, L, Vincent, G, White, LJ, and Saatchi, S (2018). "In Situ Reference Datasets From the TropiSAR and AfriSAR Campaigns in Support of Upcoming Spaceborne Biomass Missions." In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1–11. ISSN: 21511535. DOI: 10.1109/JSTARS.2018.2851606.
- Laurance, WF (1999). "Reflections on the tropical deforestation crisis." In: *Biological Conservation* 91.2-3, pp. 109–117. ISSN: 00063207. DOI: 10.1016/S0006-3207(99)00088-9.
- Lausch, A, Erasmi, S, King, DJ, Magdon, P, and Heurich, M (2016). "Understanding forest health with remote sensing-Part I-A review of spectral traits, processes and remote-sensing characteristics." In: *Remote Sensing* 8.12, pp. 1–45. ISSN: 20724292. DOI: 10.3390/rs8121029.
- Le Quéré, C et al. (2016). "Global Carbon Budget 2016." In: *Earth System Science Data* 8.2, pp. 605–649. ISSN: 18663516. DOI: 10.5194/essd-8-605-2016.
- Le Toan, T, Quegan, S, Davidson, M, Balzter, H, Paillou, P, Papathanassiou, K, Plummer, S, Rocca, F, Saatchi, S, Shugart, H, and Ulander, L (2011). "The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle." In: *Remote Sensing of Environment* 115.11, pp. 2850–2860. ISSN: 00344257. DOI: 10.1016/j.rse.2011.03.020. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425711001362>.
- Lee, JS, Grunes, MR, and Pottier, E (2001). "Quantitative comparison of classification capability: Fully polarimetric versus dual and single-polarization SAR." In: *IEEE Transactions on Geoscience and Remote Sensing* 39.11, pp. 2343–2351. ISSN: 01962892. DOI: 10.1109/36.964970.
- Lee, S-k and Fatoyinbo, TE (2015). "TanDEM-X Pol-InSAR Inversion for Mangrove Canopy Height Estimation." In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8.7, pp. 1–11. ISSN: 1939-1404. DOI: 10.1109/JSTARS.2015.2431646.
- Lefsky, MA, Cohen, WB, Acker, SA, Parker, GG, Spies, TA, and Harding, D (1999). "Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests." In: *Remote Sensing of Environment* 70.3, pp. 339–361. ISSN: 00344257. DOI: 10.1016/S0034-4257(99)00052-8. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425799000528>.
- Lefsky, MA, Cohen, WB, Parker, GG, and Harding, DJ (2002a). "Lidar remote sensing for ecosystem studies." In: *BioScience* 52.1, pp. 19–30. ISSN: ISSN 0006-3568. DOI: 10.1641/0006-3568(2002)052[0019:LRSFES]2.0.CO;2.
- Lefsky, MA, Cohen, WB, Harding, DJ, Parker, GG, Acker, SA, and Gower, ST (2002b). "Lidar remote sensing of above-ground biomass in three biomes." In: *Global Ecology and Biogeography* 11.5, pp. 393–399. ISSN: 1466-822X. DOI: 10.1046/j.1466-822x.2002.00303.x.
- Lehmann, S and Huth, A (2015). "Fast calibration of a dynamic vegetation model with minimum observation data." In: *Ecological Modelling* 301, pp. 98–105. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2015.01.013. URL: <http://dx.doi.org/10.1016/j.ecolmodel.2015.01.013>.
- Lenton, TM, Held, H, Kriegler, E, Hall, JW, Lucht, W, Rahmstorf, S, and Schellnhuber, HJ (2008). "Tipping elements in the Earth's climate system." In: *Proceedings of the National Academy of Sciences* 105.6, pp. 1786–1793. ISSN: 0027-8424. DOI: 10.1073/pnas.0705414105. arXiv: science.1200807 [10.1126]. URL: <http://www.pnas.org/cgi/doi/10.1073/pnas.0705414105>.
- Lindberg, E and Holmgren, J (2017). "Individual Tree Crown Methods for 3D Data from Remote Sensing." In: *Current Forestry Reports* 3.1, pp. 19–31. ISSN: 2198-6436. DOI: 10.1007/s40725-017-0051-6. URL: <http://link.springer.com/10.1007/s40725-017-0051-6>.
- Liu, YY, De Jeu, RAM, McCabe, MF, Evans, JP, and Van Dijk, AIJM (2011). "Global long-term passive microwave satellite-based retrievals of vegetation optical depth." In: *Geophysical Research Letters* 38.18, pp. 1–6. ISSN: 00948276. DOI: 10.1029/2011GL048684.

- Lobo, E and Dalling, JW (2014). "Spatial scale and sampling resolution affect measures of gap disturbance in a lowland tropical forest: implications for understanding forest regeneration and carbon storage." In: *Proceedings of the Royal Society B: Biological Sciences* 281.1778, pp. 20133218–20133218. ISSN: 0962-8452. DOI: 10.1098/rspb.2013.3218. URL: <http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2013.3218>.
- Lopez-Gonzalez, G, Lewis, SL, Burkitt, M, and Phillips, OL (2011). "ForestPlots.net: A web application and research tool to manage and analyse tropical forest plot data." In: *Journal of Vegetation Science* 22.4, pp. 610–613. ISSN: 11009233. DOI: 10.1111/j.1654-1103.2011.01312.x.
- Lorey, T (1878). "Die mittlere Bestandeshöhe." In: *Allgemeine Forst- und Jagdzeitung* 54, pp. 149–155.
- Los, SO, Rosette, JA, Kljun, N, North, PR, Chasmer, L, Suárez, JC, Hopkinson, C, Hill, RA, Van Gorsel, E, Mahoney, C, and Berni, JA (2012). "Vegetation height and cover fraction between 60° S and 60° N from ICESat GLAS data." In: *Geoscientific Model Development* 5.2, pp. 413–432. ISSN: 1991959X. DOI: 10.5194/gmd-5-413-2012.
- Lu, D, Chen, Q, Wang, G, Moran, E, Batistella, M, Zhang, M, Vaglio Laurin, G, and Saah, D (2012). "Aboveground Forest Biomass Estimation with Landsat and LiDAR Data and Uncertainty Analysis of the Estimates." In: *International Journal of Forestry Research* 2012.1, pp. 1–16. ISSN: 1687-9368. DOI: 10.1155/2012/436537. URL: <http://www.hindawi.com/journals/ijfr/2012/436537/>.
- Lu, D, Chen, Q, Wang, G, Liu, L, Li, G, and Moran, E (2014). "A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems." In: *International Journal of Digital Earth* December, pp. 1–43. ISSN: 1753-8947. DOI: 10.1080/17538947.2014.990526.
- MacArthur, RH and Horn, HS (1969). "Foliage Profile by Vertical Measurements." In: *Ecology* 50.5, pp. 802–804. DOI: 10.2307/1933693.
- MacDicken, KG, Sola, P, Hall, JE, Sabogal, C, Tadoum, M, and Wasseige, C de (2015). "Global progress toward sustainable forest management." In: *Forest Ecology and Management* 352, pp. 47–56. ISSN: 0378-1127. DOI: 10.1016/j.foreco.2015.02.005. URL: <http://dx.doi.org/10.1016/j.foreco.2015.02.005>.
- Magnussen, S, Næsset, E, Gobakken, T, and Frazer, G (2012). "A fine-scale model for area-based predictions of tree-size-related attributes derived from LiDAR canopy heights." In: *Scandinavian Journal of Forest Research* 27.3, pp. 312–322. ISSN: 02827581. DOI: 10.1080/02827581.2011.624116.
- Malhi, Y, Roberts, JT, Betts, RA, Killeen, TJ, Li, W, and Nobre, CA (2008). "Climate change, deforestation, and the fate of the Amazon." In: *Science* 319.5860, pp. 169–172. ISSN: 00368075. DOI: 10.1126/science.1146961.
- Martone, M, Rizzoli, P, Wecklich, C, González, C, Bueso-Bello, JL, Valdo, P, Schulze, D, Zink, M, Krieger, G, and Moreira, A (2018). "The global forest/non-forest map from TanDEM-X interferometric SAR data." In: *Remote Sensing of Environment* 205. August 2017, pp. 352–373. ISSN: 00344257. DOI: 10.1016/j.rse.2017.12.002. URL: <https://doi.org/10.1016/j.rse.2017.12.002>.
- Marvin, DC, Asner, GP, Knapp, DE, Anderson, CB, Martin, RE, Sinca, F, and Tupayachi, R (2014). "Amazonian landscapes and the bias in field studies of forest structure and biomass." In: *Proceedings of the National Academy of Sciences of the United States of America* 111.48, E5224–32. ISSN: 1091-6490. DOI: 10.1073/pnas.1412999111. arXiv: arXiv:1408.1149. URL: <http://www.ncbi.nlm.nih.gov/pubmed/25422434>.
- Mascaro, J, Asner, GP, Muller-Landau, HC, Van Breugel, M, Hall, J, and Dahlin, K (2011a). "Controls over aboveground forest carbon density on Barro Colorado Island, Panama." In: *Biogeosciences* 8.6, pp. 1615–1629. ISSN: 17264170. DOI: 10.5194/bg-8-1615-2011. URL: <http://www.biogeosciences.net/8/1615/2011/>.
- Mascaro, J, Detto, M, Asner, GP, and Muller-Landau, HC (2011b). "Evaluating uncertainty in mapping forest carbon with airborne LiDAR." In: *Remote Sensing of Environment* 115.12, pp. 3770–3774. ISSN: 00344257. DOI: 10.1016/j.rse.2011.07.019. URL: <http://dx.doi.org/10.1016/j.rse.2011.07.019>.

- Mascaro, J, Asner, GP, Dent, DH, DeWalt, SJ, and Denslow, JS (2012). "Scale-dependence of aboveground carbon accumulation in secondary forests of Panama: A test of the intermediate peak hypothesis." In: *Forest Ecology and Management* 276, pp. 62–70. ISSN: 03781127. DOI: 10.1016/j.foreco.2012.03.032. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0378112712001995>.
- Mascaro, J, Asner, GP, Davies, S, Dehgan, A, and Saatchi, S (2014). "These are the days of lasers in the jungle." In: *Carbon balance and management* 9.1, p. 7. ISSN: 1750-0680. DOI: 10.1186/s13021-014-0007-0. URL: <http://www.cbjournal.com/content/9/1/7>.
- McMahon, SM and Parker, GG (2015). "A general model of intra-annual tree growth using dendrometer bands." In: *Ecology and Evolution* 5.2, pp. 243–254. ISSN: 20457758. DOI: 10.1002/ece3.1117.
- Meyer, V, Saatchi, SS, Chave, J, Dalling, JW, Bohlman, S, Fricker, GA, Robinson, C, Neumann, M, and Hubbell, S (2013). "Detecting tropical forest biomass dynamics from repeated airborne lidar measurements." In: *Biogeosciences* 10.8, pp. 5421–5438. ISSN: 17264170. DOI: 10.5194/bg-10-5421-2013. URL: <http://www.biogeosciences.net/10/5421/2013/>.
- Meyer, V, Saatchi, S, Clark, DB, Keller, M, Vincent, G, Ferraz, A, Espírito-Santo, F, D’Oliveira, MVN, Kaki, D, and Chave, J (2018). "Canopy Area of Large Trees Explains Aboveground Biomass Variations across Nine Neotropical Forest Landscapes." In: *Biogeosciences Discussions*, pp. 1–38. ISSN: 1810-6285. DOI: 10.5194/bg-2017-547. URL: <https://www.biogeosciences-discuss.net/bg-2017-547/>.
- Miehe, G, Miehe, S, Vogel, J, Co, S, and La, D (2007). "Highest Treeline in the Northern Hemisphere Found in Southern Tibet." In: *Mountain Research and Development* 27.2, pp. 169–173. ISSN: 0276-4741. DOI: 10.1659/mrd.0792. URL: <http://www.bioone.org/doi/abs/10.1659/mrd.0792>.
- Mitchard, ET, Saatchi, SS, Baccini, A, Asner, GP, Goetz, SJ, Harris, NL, and Brown, S (2013). "Uncertainty in the spatial distribution of tropical forest biomass: a comparison of pan-tropical maps." In: *Carbon balance and management* 8.1, p. 10. ISSN: 1750-0680. DOI: 10.1186/1750-0680-8-10. URL: <http://www.ncbi.nlm.nih.gov/pubmed/24161143>.
- Moorcroft, APR, Hurtt, GC, and Pacala, SW (2001). "A Method for Scaling Vegetation Dynamics : The Ecosystem Demography Model (ED)." In: *Ecological Monographs* 71.4, pp. 557–585. DOI: 10.2307/3100036.
- Moreira, A (2015). *Lecture slides from the course Synthetic Aperture Radar (SAR): Basics, Theory and Future Developments*.
- Moreira, A, Prats-iraola, P, Younis, M, Krieger, G, Hajnsek, I, and Papathanassiou, KP (2013). "A Tutorial on Synthetic Aperture Radar." In: *IEEE Geoscience and Remote Sensing Magazine* March, pp. 6–43. DOI: 10.1109/MGRS.2013.2248301.
- Moreira, A, Krieger, G, Hajnsek, I, Papathanassiou, K, Younis, M, Lopez-Dekker, F, Huber, S, Villano, M, and Pardini, M (2015). "Tandem-L: A highly innovative bistatic SAR mission for global observation of dynamic processes on the earth’s surface." In: *IEEE Geoscience and Remote Sensing Magazine* June, pp. 8–23. DOI: 10.1109/MGRS.2015.2437353.
- Morsdorf, F, Frey, O, Meier, E, Itten, KI, and Allgöwer, B (2008). "Assessment of the influence of flying altitude and scan angle on biophysical vegetation products derived from airborne laser scanning." In: *International Journal of Remote Sensing* 29.5, pp. 1387–1406. ISSN: 13665901. DOI: 10.1080/01431160701736349.
- Moser Jr, JW (1980). "Historical chapters in the development of modern forest growth and yield theory." In: *Forecasting forest and stand dynamics: proceedings of the Workshop held at the School of Forestry, Lakehead University, Thunderbay, Ontario*, pp. 42–61.
- Murphy, M, Evans, JS, and Storfer, A (2010). "Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics." In: *Ecology* 91.1, pp. 252–61. ISSN: 0012-9658. DOI: 10.1890/08-0879.1. URL: <http://www.ncbi.nlm.nih.gov/pubmed/20380214>.
- Myers, N, Myers, N, Mittermeier, R, Mittermeier, R, Fonseca, GB, Fonseca, GB, Kent, J, and Kent, J (2000). "Biodiversity hotspots for conservation priorities." In: *Nature* 403.6772, pp. 853–8. ISSN: 0028-0836. DOI: 10.1038/35002501. arXiv: 0208024 [gr-qc]. URL: <http://www.ncbi.nlm.nih.gov/pubmed/10706275>.

- Myneni, RB, Nemani, RR, and Running, SW (1997). "Estimation of global leaf area index and absorbed par using radiative transfer models." In: *IEEE Transactions on Geoscience and Remote Sensing* 35.6, pp. 1380–1393. ISSN: 01962892. DOI: 10.1109/36.649788.
- Næsset, E (2002). "Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data." In: *Remote Sensing of Environment* 80.1, pp. 88–99. ISSN: 00344257. DOI: 10.1016/S0034-4257(01)00290-5. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425701002905>.
- Næsset, E (2009). "Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data." In: *Remote Sensing of Environment* 113.1, pp. 148–159. ISSN: 00344257. DOI: 10.1016/j.rse.2008.09.001. URL: <http://dx.doi.org/10.1016/j.rse.2008.09.001>.
- Næsset, E, Bollandsås, OM, Gobakken, T, Gregoire, TG, and Ståhl, G (2013). "Model-assisted estimation of change in forest biomass over an 11year period in a sample survey supported by airborne LiDAR: A case study with post-stratification to provide "activity data"." In: *Remote Sensing of Environment* 128, pp. 299–314. ISSN: 00344257. DOI: 10.1016/j.rse.2012.10.008. URL: <http://dx.doi.org/10.1016/j.rse.2012.10.008>.
- Nguyen, DT (2017). *Big Data and LiDAR: Developing a Framework to Process Country Wide LiDAR Point Cloud Data Sets Using Parallelization on a Computer Cluster*. Tech. rep. Leipzig: Helmholtz Centre for Environmental Research, p. 12. URL: <https://www.hnee.de/en/Programmes/Master-degree/Forest-Information-Technology/students-results/projects-results/summer-term-2017/Student-Research-Projects-2017-E9256.htm>.
- Ni-Meister, W, Jupp, DLB, and Dubayah, R (2001). "Modeling lidar waveforms in heterogeneous and discrete canopies." In: *IEEE Transactions on Geoscience and Remote Sensing* 39.9, pp. 1943–1958. ISSN: 01962892. DOI: 10.1109/36.951085. URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=951085>.
- Ninan, KN and Inoue, M (2013). "Valuing forest ecosystem services: What we know and what we don't." In: *Ecological Economics* 93, pp. 137–149. ISSN: 09218009. DOI: 10.1016/j.ecolecon.2013.05.005. arXiv: arXiv:1011.1669v3. URL: <http://dx.doi.org/10.1016/j.ecolecon.2013.05.005>.
- Olson, DM, Dinerstein, E, Wikramanayake, ED, Burgess, ND, Powell, GVN, Underwood, EC, D'amico, JA, Itoua, I, Strand, HE, Morrison, JC, Loucks, CJ, Allnut, TF, Ricketts, TH, Kura, Y, Lamoreux, JF, Wettengel, WW, Hedao, P, and Kassem, KR (2001). "Terrestrial Ecoregions of the World: A New Map of Life on Earth." In: *BioScience* 51.11, p. 933. ISSN: 0006-3568. DOI: 10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2. arXiv: j.1744-7429.2008.00471.x. URL: <https://academic.oup.com/bioscience/article/51/11/933-938/227116>.
- Ørka, HO and Hauglin, M (2016). "Use of remote sensing for mapping of non-native conifer species." In: *INA fagrapport* 33, p. 76. ISSN: 1891-2281. URL: [https://static02.nmbu.no/mina/publikasjoner/mina\\_fagrapport/pdf/mif33.pdf](https://static02.nmbu.no/mina/publikasjoner/mina_fagrapport/pdf/mif33.pdf).
- Palace, MW, Sullivan, FB, Ducey, MJ, Treuhaft, RN, Herrick, C, Shimbo, JZ, and Mota-E-Silva, J (2015). "Estimating forest structure in a tropical forest using field measurements, a synthetic model and discrete return lidar data." In: *Remote Sensing of Environment* 161, pp. 1–11. ISSN: 00344257. DOI: 10.1016/j.rse.2015.01.020. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425715000383>.
- Pan, Y, Birdsey, RA, Fang, J, Houghton, R, Kauppi, PE, Kurz, WA, Phillips, OL, Shvidenko, A, Lewis, SL, Canadell, JG, Ciais, P, Jackson, RB, Pacala, SW, McGuire, AD, Piao, S, Rautiainen, A, Sitch, S, and Hayes, D (2011). "A Large and Persistent Carbon Sink in the World's Forests." In: *Science* 333.6045, pp. 988–993. ISSN: 0036-8075. DOI: 10.1126/science.1201609. URL: <http://www.sciencemag.org/cgi/doi/10.1126/science.1201609>.
- Pan, Y, Birdsey, RA, Phillips, OL, and Jackson, RB (2013). "The Structure, Distribution, and Biomass of the World's Forests." In: *Annual Review of Ecology, Evolution, and Systematics* 44.1, pp. 593–622. ISSN: 1543-592X. DOI: 10.1146/annurev-ecolsys-110512-135914. URL: <http://www.annualreviews.org/doi/10.1146/annurev-ecolsys-110512-135914>.
- Patenaude, G, Hill, RA, Milne, R, Gaveau, DLA, Briggs, BBJ, and Dawson, TP (2004). "Quantifying forest above ground carbon content using LiDAR remote sensing." In: *Remote Sensing*

- of *Environment* 93.3, pp. 368–380. ISSN: 00344257. DOI: 10.1016/j.rse.2004.07.016. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425704002366>.
- Paulick, S, Dislich, C, Homeier, J, Fischer, R, and Huth, A (2017). “The carbon fluxes in different successional stages: modelling the dynamics of tropical montane forests in South Ecuador.” In: *Forest Ecosystems* 4.1, p. 5. ISSN: 2197-5620. DOI: 10.1186/s40663-017-0092-0. URL: <http://forestecosyst.springeropen.com/articles/10.1186/s40663-017-0092-0>.
- Pereira, HM et al. (2013). “Essential biodiversity variables.” In: *Science* 339.6117, pp. 277–278. ISSN: 10959203. DOI: 10.1126/science.1229931.
- Pettorelli, N, Laurance, WF, O’Brien, TG, Wegmann, M, Nagendra, H, and Turner, W (2014). “Satellite remote sensing for applied ecologists: Opportunities and challenges.” In: *Journal of Applied Ecology* 51.4, pp. 839–848. ISSN: 13652664. DOI: 10.1111/1365-2664.12261.
- Pettorelli, N et al. (2016). “Framing the concept of satellite remote sensing essential biodiversity variables: challenges and future directions.” In: *Remote Sensing in Ecology and Conservation* 2.3, pp. 122–131. ISSN: 20563485. DOI: 10.1002/rse2.15.
- Plummer, SE (2000). “Perspectives on combining ecological process models and remotely sensed data.” In: *Ecological Modelling* 129.2-3, pp. 169–186. ISSN: 03043800. DOI: 10.1016/S0304-3800(00)00233-7.
- Poorter, L et al. (2016). “Biomass resilience of Neotropical secondary forests.” In: *Nature* 530.7589, pp. 211–214. ISSN: 0028-0836. DOI: 10.1038/nature16512. URL: <http://www.nature.com/doi/10.1038/nature16512>.
- Pretzsch, H (2009). “Forest Dynamics, Growth and Yield.” In: *Forest Dynamics, Growth and Yield*. Springer. ISBN: 978-3-540-88306-7. DOI: 10.1007/978-3-540-88307-4. arXiv: arXiv:1011.1669v3.
- Proisy, C, Coutron, P, Pélissier, R, Barbier, N, and Engel, J (2007). “Monitoring canopy grain of tropical forest using Fourier-based textural ordination (FOTO) of very high resolution images.” In: *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 4324–4326. DOI: 10.1109/IGARSS.2007.4423808.
- Proisy, C, Barbier, N, Guérout, M, and Pélissier, R (2012). “Biomass Prediction in Tropical Forests : The Canopy Grain Approach.” In: *Remote Sensing of Biomass: Principles and applications*. InTech, pp. 59–76. ISBN: 9789533073156. DOI: 10.5772/696.
- Puliti, S, Solberg, S, Næsset, E, Gobakken, T, Zahabu, E, Mauya, E, and Malimbwi, RE (2017). “Modelling above ground biomass in Tanzanian miombo woodlands using TanDEM-X WorldDEM and field data.” In: *Remote Sensing* 9.10, pp. 1–13. ISSN: 20724292. DOI: 10.3390/rs9100984.
- Qi, W and Dubayah, RO (2016). “Combining Tandem-X InSAR and simulated GEDI lidar observations for forest structure mapping.” In: *Remote Sensing of Environment* 187.2016, pp. 253–266. ISSN: 0034-4257. DOI: <http://dx.doi.org/10.1016/j.rse.2016.10.018>. URL: <http://www.sciencedirect.com/science/article/pii/S003442571630390X>.
- R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. URL: <http://www.r-project.org/>.
- Rammig, A, Jupp, T, Thonicke, K, Tietjen, B, Heinke, J, Ostberg, S, Lucht, W, Cramer, W, and Cox, P (2010). “Estimating the risk of Amazonian forest dieback.” In: *New Phytologist* 187.3, pp. 694–706. ISSN: 0028646X. DOI: 10.1111/j.1469-8137.2010.03318.x.
- Ranson, KJ, Sun, G, Knox, RG, Levine, ER, Weishampel, JF, and Fifer, ST (2001). “Northern forest ecosystem dynamics using coupled models and remote sensing.” In: *Remote Sensing of Environment* 75.2, pp. 291–302. ISSN: 00344257. DOI: 10.1016/S0034-4257(00)00174-7.
- Raunonen, P, Kaasalainen, M, Åkerblom, M, Kaasalainen, S, Kaartinen, H, Vastaranta, M, Holopainen, M, Disney, M, and Lewis, P (2013). “Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data.” In: *Remote Sensing* 5.2, pp. 491–520. ISSN: 2072-4292. DOI: 10.3390/rs5020491. URL: <http://www.mdpi.com/2072-4292/5/2/491/htm>.
- Reigber, A and Moreira, A (2000). “First demonstration of airborne SAR tomography using multibaseline L-band data.” In: *IEEE Transactions on Geoscience and Remote Sensing* 38.5, pp. 2142–2152. DOI: 10.1109/36.868873.

- Reineke, LH (1933). "Perfecting a stand-density index for even-aged forests." In: *J. Agric. Res.* 46.7, pp. 627–638. ISSN: 00959758. URL: [https://www.fs.fed.us/psw/publications/cfres/cfres\\_1933\\_reineke001.pdf](https://www.fs.fed.us/psw/publications/cfres/cfres_1933_reineke001.pdf).
- Réjou-Méchain, M et al. (2014). "Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks." In: *Biogeosciences* 11.23, pp. 6827–6840. ISSN: 17264189. DOI: 10.5194/bg-11-6827-2014.
- Rödig, E, Huth, A, Bohn, F, Rebmann, C, and Cuntz, M (2017a). "Estimating the carbon fluxes of forests with an individual-based forest model." In: *Forest Ecosystems* 4.1, p. 4. ISSN: 2197-5620. DOI: 10.1186/s40663-017-0091-1. URL: <http://forestecosyst.springeropen.com/articles/10.1186/s40663-017-0091-1>.
- Rödig, E, Cuntz, M, Heinke, J, Rammig, A, and Huth, A (2017b). "Spatial heterogeneity of biomass and forest structure of the Amazon rain forest: Linking remote sensing, forest modelling and field inventory." In: *Global Ecology and Biogeography* 26.11, pp. 1292–1302. ISSN: 14668238. DOI: 10.1111/geb.12639.
- Rödig, E, Cuntz, M, Rammig, A, Fischer, R, Taubert, F, and Huth, A (2018). "The importance of forest structure for carbon flux estimates in the Amazon rainforest." In: *Environmental Research Letters* in press. DOI: <https://doi.org/10.1088/1748-9326/aabc61>.
- Rouse, JW, Haas, RH, Schell, JA, and Deering, DW (1974). *Monitoring vegetation systems in the Great Plains with ERTS*. URL: <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740022614.pdf>.
- Roy, DP et al. (2014). "Landsat-8: Science and product vision for terrestrial global change research." In: *Remote Sensing of Environment* 145, pp. 154–172. ISSN: 00344257. DOI: 10.1016/j.rse.2014.02.001. URL: <http://dx.doi.org/10.1016/j.rse.2014.02.001>.
- Running, SW and Nemani, RR (1988). "Relating seasonal patterns of the AVHRR vegetation index to simulated photosynthesis and transpiration of forests in different climates." In: *Remote Sensing of Environment* 24.2, pp. 347–367. ISSN: 00344257. DOI: 10.1016/0034-4257(88)90034-X.
- Saatchi, SS, Harris, NL, Brown, S, Lefsky, M, Mitchard, ETA, Salas, W, Zutta, BR, Buermann, W, Lewis, SL, Hagen, S, Petrova, S, White, L, Silman, M, and Morel, A (2011). "Benchmark map of forest carbon stocks in tropical regions across three continents." In: *Proceedings of the National Academy of Sciences of the United States of America* 108.24, pp. 9899–9904. ISSN: 1091-6490. DOI: 10.1073/pnas.1019576108. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3116381&tool=pmcentrez&rendertype=abstract>.
- Santoro, M and Cartus, O (2018). "Research pathways of forest above-ground biomass estimation based on SAR backscatter and interferometric SAR observations." In: *Remote Sensing* 10.4, p. 608. ISSN: 20724292. DOI: 10.3390/rs10040608. URL: <http://www.mdpi.com/2072-4292/10/4/608>.
- Sasaki, N and Putz, FE (2009). "Critical need for new definitions of "forest" and "forest degradation" in global climate change agreements." In: *Conservation Letters* 2.5, pp. 226–232. DOI: 10.1111/j.1755-263X.2009.00067.x.
- Schutz, BE, Zwally, HJ, Shuman, CA, Hancock, D, and DiMarzio, JP (2005). "Overview of the ICESat mission." In: *Geophysical Research Letters* 32.21, pp. 1–4. ISSN: 00948276. DOI: 10.1029/2005GL024009.
- Sexton, JO, Song, X-P, Feng, M, Noojipady, P, Anand, A, Huang, C, Kim, D-H, Collins, KM, Channan, S, Dimiceli, C, and Townshend, JR (2013). "Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS Vegetation Continuous Fields with lidar-based estimates of error." In: *International Journal of Digital Earth* 8947.July 2015, p. 130321031236007. ISSN: 1753-8947. DOI: 10.1080/17538947.2013.786146. URL: <http://dx.doi.org/10.1080/17538947.2013.786146>.
- Shugart Jr, HH and West, DC (1980). "Forest succession models." In: *BioScience* 30.5, pp. 308–313. DOI: 10.2307/1307854.
- Shugart, HH (1984). *A theory of forest dynamics: the ecological implications of forest succession models*. New York, USA: Springer, p. 278. URL: <http://www.osti.gov/scitech/biblio/5642300>.
- Shugart, HH (2003). *A theory of forest dynamics: The ecological implications of forest succession models*. Caldwell. New York, NJ: The Blackburn Press, p. 278. ISBN: 978-1930665750.

- Shugart, HH, Asner, GP, Fischer, R, Huth, A, Knapp, N, Le Toan, T, and Shuman, JK (2015). "Computer and remote-sensing infrastructure to enhance large-scale testing of individual-based forest models." In: *Frontiers in Ecology and the Environment* 13.9, pp. 503–511. ISSN: 1540-9295. DOI: 10.1890/140327. URL: <http://www.esajournals.org/doi/10.1890/140327>.
- Shugart, HH, Wang, B, Fischer, R, Ma, J, Fang, J, Yan, X, Huth, A, and Armstrong, AH (2018). "Gap models and their individual-based relatives in the assessment of the consequences of global change." In: *Environ. Res. Lett* 13. ISSN: 1748-9326. DOI: 10.1088/1748-9326/aaaacc. URL: <https://doi.org/10.1088/1748-9326/aaaacc>.
- Singh, M, Evans, D, Friess, D, Tan, B, and Nin, C (2015). "Mapping Above-Ground Biomass in a Tropical Forest in Cambodia Using Canopy Textures Derived from Google Earth." In: *Remote Sensing* 7.5, pp. 5057–5076. ISSN: 2072-4292. DOI: 10.3390/rs70505057. URL: <http://www.mdpi.com/2072-4292/7/5/5057/>.
- Sitch, S, Smith, B, Prentice, IC, Arneth, A, Bondeau, A, Cramer, W, Kaplan, JO, Levis, S, Lucht, W, Sykes, MT, and Others (2003). "Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model." In: *Global Change Biology* 9.2, pp. 161–185. DOI: 10.1046/j.1365-2486.2003.00569.x.
- Solberg, S, Næsset, E, Gobakken, T, and Bollandsås, O-M (2014). "Forest biomass change estimated from height change in interferometric SAR height models." In: *Carbon balance and management* 9.1, p. 5. ISSN: 1750-0680. DOI: 10.1186/s13021-014-0005-2. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4159577&tool=pmcentrez&rendertype=abstract>.
- Solberg, S, Gizachew, B, Næsset, E, Gobakken, T, Bollandsås, MO, Mauya, WE, Olsson, H, Malimbwi, R, and Zahabu, E (2015). "Monitoring forest carbon in a Tanzanian woodland using interferometric SAR: A novel methodology for REDD+." In: *Carbon Balance and Management* 10.1. ISSN: 17500680. DOI: 10.1186/s13021-015-0023-8. URL: <http://dx.doi.org/10.1186/s13021-015-0023-8>.
- Solberg, S, Hansen, EH, Gobakken, T, Næsset, E, and Zahabu, E (2017). "Biomass and InSAR height relationship in a dense tropical forest." In: *Remote Sensing of Environment* 192, pp. 166–175. ISSN: 00344257. DOI: 10.1016/j.rse.2017.02.010. URL: <http://dx.doi.org/10.1016/j.rse.2017.02.010>.
- Stark, SC, Leitold, V, Wu, JL, Hunter, MO, Castilho, CV de, Costa, FRC, McMahon, SM, Parker, GG, Shimabukuro, MT, Lefsky, MA, Keller, M, Alves, LF, Schiatti, J, Shimabukuro, YE, Brandão, DO, Woodcock, TK, Higuchi, N, Camargo, PB de, Oliveira, RC de, and Saleska, SR (2012). "Amazon forest carbon dynamics predicted by profiles of canopy leaf area and light environment." In: *Ecology Letters* 15.12, pp. 1406–1414. ISSN: 14610248. DOI: 10.1111/j.1461-0248.2012.01864.x. URL: <http://www.ncbi.nlm.nih.gov/pubmed/22994288>.
- Stark, SC, Enquist, BJ, Saleska, SR, Leitold, V, Schiatti, J, Longo, M, Alves, LF, Camargo, PB, and Oliveira, RC (2015). "Linking canopy leaf area and light environments with tree size distributions to explain Amazon forest demography." In: *Ecology Letters* 18.7, pp. 636–645. ISSN: 14610248. DOI: 10.1111/ele.12440.
- Stavros, EN, Schimel, D, Pavlick, R, Serbin, S, Swann, A, Duncanson, L, Fisher, JB, Fassnacht, F, Ustin, S, Dubayah, R, Schweiger, A, and Wennberg, P (2017). "ISS observations offer insights into plant function." In: *Nature Ecology and Evolution* 1.7, pp. 1–4. ISSN: 2397334X. DOI: 10.1038/s41559-017-0194. URL: <http://dx.doi.org/10.1038/s41559-017-0194>.
- Sun, G and Ranson, KJ (2000). "Modeling Lidar returns from vegetation canopies." In: *IEEE Transactions on Geoscience and Remote Sensing* 38.6, pp. 2617–2626. DOI: 10.1109/36.885208.
- Tang, H, Dubayah, R, Swatantran, A, Hofton, M, Sheldon, S, Clark, DB, and Blair, B (2012). "Retrieval of vertical LAI profiles over tropical rain forests using waveform lidar at La Selva, Costa Rica." In: *Remote Sensing of Environment* 124, pp. 242–250. ISSN: 00344257. DOI: 10.1016/j.rse.2012.05.005. URL: <http://dx.doi.org/10.1016/j.rse.2012.05.005>.
- Taubert, F, Frank, K, and Huth, A (2012). "A review of grassland models in the biofuel context." In: *Ecological Modelling* 245, pp. 84–93. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2012.04.007. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0304380012001731>.
- Taubert, F, Hartig, F, Dobner, H-J, and Huth, A (2013). "On the challenge of fitting tree size distributions in ecology." In: *PloS one* 8.2, e58036. ISSN: 1932-6203. DOI: 10.1371/journal.

- pone.0058036. URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3585190&tool=pmcentrez&rendertype=abstract>.
- Taubert, F, Fischer, R, Groeneveld, J, Lehmann, S, Müller, MS, Rödiger, E, Wiegand, T, and Huth, A (2018). "Global patterns of tropical forest fragmentation." In: *Nature* 554.7693, pp. 519–522. ISSN: 14764687. DOI: 10.1038/nature25508. URL: <http://dx.doi.org/10.1038/nature25508>.
- Tello, M, Cazcarra-Bes, V, Pardini, M, and Papathanassiou, K (2015). "Structural classification of forest by means of L-band tomographic SAR." In: *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 5288–5291. DOI: 10.1109/IGARSS.2015.7327028.
- Thomas, SC and Malczewski, G (2007). "Wood carbon content of tree species in Eastern China: Interspecific variability and the importance of the volatile fraction." In: *Journal of Environmental Management* 85.3, pp. 659–662. ISSN: 03014797. DOI: 10.1016/j.jenvman.2006.04.022.
- Thomas, SC and Baltzer, JL (2002). "Tropical forests." In: *Encyclopedia of life sciences*. URL: <http://homepages.wmich.edu/~kohlers/bios105/files/tropicalforests.pdf>.
- Thünen-Institut (2012). *Dritte Bundeswaldinventur - Ergebnisdatenbank*. URL: <https://bwi.info> (visited on 10/24/2018).
- Tietjen, B and Huth, A (2006). "Modelling dynamics of managed tropical rainforests-An aggregated approach." In: *Ecological Modelling* 199.4, pp. 421–432. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2005.11.045.
- Treuhaft, RN, Madsen, SN, Moghaddam, M, and Zyl, JJ (1996). "Vegetation characteristics and underlying topography from interferometric radar." In: *Radio Science* 31.6, pp. 1449–1485. DOI: 10.1029/96RS01763.
- Treuhaft, RN, Law, BE, and Gregory, P (2004). "Forest Attributes from Radar Interferometric Structure and Its Fusion with Optical Remote Sensing." In: *BioScience* 54.6, pp. 561–571. DOI: 10.1641/0006-3568(2004)054[0561:FAFRIS]2.0.CO;2.
- Treuhaft, R, Gonçalves, F, Roberto, J, Keller, M, Palace, M, Madsen, SN, Sullivan, F, and Graça, PMLA (2015). "Tropical-Forest Biomass Estimation at X-Band From the Spaceborne TanDEM-X Interferometer." In: *IEEE Geoscience and Remote Sensing Letters* 12.2, pp. 239–243. DOI: 10.1109/LGRS.2014.2334140.
- Treuhaft, R, Lei, Y, Gonçalves, F, Keller, M, Santos, JR dos, Neumann, M, and Almeida, A (2017). "Tropical-forest structure and biomass dynamics from TanDEM-X radar interferometry." In: *Forests* 8.8, pp. 1–28. ISSN: 19994907. DOI: 10.3390/f8080277.
- Tuanmu, MN and Jetz, W (2015). "A global, remote sensing-based characterization of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling." In: *Global Ecology and Biogeography* 24.11, pp. 1329–1339. ISSN: 14668238. DOI: 10.1111/geb.12365.
- Vastaranta, M, Wulder, MA, White, JC, Pekkarinen, A, Tuominen, S, Ginzler, C, Kankare, V, Holopainen, M, Hyypä, J, and Hyypä, H (2013). "Airborne laser scanning and digital stereo imagery measures of forest structure: Comparative results and implications to forest mapping and inventory update." In: *Canadian Journal of Remote Sensing* 39.5, pp. 382–395. ISSN: 07038992. DOI: 10.5589/m13-046.
- Verrelst, J, Malenovsky, Z, Tol, C van der, Camps-Valls, G, Gastellu-Etchegorry, JP, Lewis, P, North, P, and Moreno, J (2018). "Quantifying Vegetation Biophysical Variables from Imaging Spectroscopy Data: A Review on Retrieval Methods." In: *Surveys in Geophysics*, pp. 1–41. ISSN: 15730956. DOI: 10.1007/s10712-018-9478-y. URL: <https://doi.org/10.1007/s10712-018-9478-y>.
- Vincent, G, Sabatier, D, and Rutishauser, E (2014). "Revisiting a universal airborne light detection and ranging approach for tropical forest carbon mapping: Scaling-up from tree to stand to landscape." In: *Oecologia* 175.2, pp. 439–443. ISSN: 00298549. DOI: 10.1007/s00442-014-2913-y.
- Wang, M, White, N, Grimm, V, Hofman, H, Doley, D, Thorp, G, Cribb, B, Wherritt, E, Han, L, Wilkie, J, and Hanan, J (2018). "Pattern-oriented modelling as a novel way to verify and validate functional-structural plant models: A demonstration with the annual growth module of avocado." In: *Annals of Botany* 121.5, pp. 941–959. ISSN: 10958290. DOI: 10.1093/aob/mcx187.
- Whitmore, TC (1990). *An introduction to tropical rain forests*. Oxford, UK: Clarendon Press.

BIBLIOGRAPHY

- Whittaker, RH (1970). *Communities and ecosystems*. London, UK: Macmillan, p. 162. ISBN: 9780024273901.
- Wichmann, V, Bremer, M, Lindenberger, J, Rutzinger, M, Georges, C, and Petrini-Monteferrri, F (2015). "Evaluating the Potential of Multispectral Airborne Lidar for Topographic Mapping and Land Cover Classification." In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences II-3/W5.2012*, pp. 113–119. ISSN: 2194-9050. DOI: 10.5194/isprsannals-II-3-W5-113-2015. URL: <http://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/II-3-W5/113/2015/>.
- Wiegand, T, Jeltsch, F, Hanski, I, and Grimm, V (2003). "Using Pattern-Oriented Modeling for Revealing Hidden Information: A Key for Reconciling Ecological Theory and Application." In: *Oikos* 100.2, pp. 209–222. DOI: 10.1034/j.1600-0706.2003.12027.x.
- Wulder, MA, White, JC, Nelson, RF, Næsset, E, Ørka, HO, Coops, NC, Hilker, T, Bater, CW, and Gobakken, T (2012). "Lidar sampling for large-area forest characterization: A review." In: *Remote Sensing of Environment* 121, pp. 196–209. ISSN: 00344257. DOI: 10.1016/j.rse.2012.02.001. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425712000855>.
- Zanne, AE, Lopez-Gonzalez, G, Coomes, DA, Ilic, J, Jansen, S, Lewis, SL, Miller, RB, Swenson, NG, Wiemann, MC, and Chave, J (2009). *Data from: Towards a worldwide wood economics spectrum*. DOI: doi:10.5061/dryad.234. URL: <https://doi.org/10.5061/dryad.234>.
- Zhao, K, Suarez, JC, Garcia, M, Hu, T, Wang, C, and Londo, A (2018). "Utility of multitemporal lidar for forest and carbon monitoring: Tree growth, biomass dynamics, and carbon flux." In: *Remote Sensing of Environment* 204, pp. 883–897. ISSN: 00344257. DOI: 10.1016/j.rse.2017.09.007. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0034425717304194>.
- Zhao, M, Heinsch, FA, Nemani, RR, and Running, SW (2005). "Improvements of the MODIS terrestrial gross and net primary production global data set." In: *Remote Sensing of Environment* 95.2, pp. 164–176. ISSN: 00344257. DOI: 10.1016/j.rse.2004.12.011.
- Zolkos, SG, Goetz, SJ, and Dubayah, R (2013). "A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing." In: *Remote Sensing of Environment* 128, pp. 289–298. DOI: 10.1016/j.rse.2012.10.017.

## DANKSAGUNG

---

Im Folgenden möchte ich mich bei all den Menschen bedanken ohne deren Unterstützung die Erstellung dieser Dissertation nicht möglich gewesen wäre. Dies sind zuallererst meine beiden Betreuer Andreas Huth und Rico Fischer, die die Idee zu diesem Forschungsprojekt hatten und das Vertrauen in mich, dass ich der Richtige für dessen Umsetzung bin. Ich danke beiden dafür, dass ich das Wissenschaftlerhandwerk bei ihnen erlernen durfte, mit all seinen Facetten, die über die reine Forschungstätigkeit hinausgehen. Besonders dankbar bin ich für die vielen schönen gemeinsamen Konferenzreisen und die exzellenten Kontakte, die ich dabei knüpfen konnte. Andreas danke ich zudem für das breite Fachwissen er mir vermittelt und durch sein beharrliches Hinterfragen aller Details gefördert hat. Rico danke ich für sein stets überaus konstruktives Feedback und sein allzeit offenes Ohr für die Sorgen, Nöte und Bugs im Quellcode.

Allen Kollegen in der Abteilung Ökologische Systemanalyse (ÖSA) am UFZ sei herzlich dafür gedankt, dass sie so eine angenehme Arbeitsatmosphäre schaffen. Das ist vor allem auch das Verdienst der Leiterin Karin Frank und der unermüdlichen Hilfe durch Gaby Nagel und Heike Reichelt in administrativen und Michael Müller und Andreas Thiele in technischen Fragen. Besonders möchte ich mich bei Edna, Sara, Julia, Meike, Jessica und Eckart sowie den vielen Gästen bedanken mit denen ich das Büro 224 in den vergangenen vier Jahren teilen durfte. Danke für die schöne gemeinsame Zeit, die gegenseitige Unterstützung und die Beiträge aller, die 224 zu einem interdisziplinären, skaleninvarianten Kompetenzzentrum gemacht haben. Bei allen ÖSA-Doktoranden, besonders meiner Kohorte, Maria, Ulrike, Meike, Felix und Henning bedanke ich mich für die gegenseitige Ermunterung auf dem nicht immer einfachen Weg durch die Promotionszeit.

Nur durch gegenseitige Hilfe und den fortwährenden Austausch bei den gemeinsamen Programmier Tagen und Seminaren sind die vielfältigen Projekte der Waldmodelliergruppe möglich. Danke daher an die Forminder Franziska, Edna, Ulrike, Jessica, Julia, Rico, Friedrich, Mateus, Sebastian L., Sebastian P., Hans und Andreas. Insbesondere danke ich Franziska und Rico für die Koordination und Organisation vieler Gruppenaktivitäten. Weiterhin danke ich Stephan Getzin für die gute Zusammenarbeit als Lidar-Taskforce und viele wertvolle Ratschläge.

Ebenfalls möchte ich mich bei den Kollegen aus der Helmholtz-Allianz für Fernerkundung bedanken: für die Inspiration aus vielen Gesprächen und die Geduld beim Umgang und Austausch komplizierter Datensätze. Insbesondere gilt mein Dank Kostas Papathanassiou, Florian Kugler, Victor Cazcarra-Bes und Marivi Tello vom DLR und Peter Biber und Michael Heym von der TU München. Außerdem bedanke ich mich bei Ralph Dubayah von der NASA und seinen Mitarbeitern im GEDI Science Team für hilfreiche Diskussionen und den offenen Erkenntnisaustausch. Weiterhin bedanke ich mich bei allen Menschen, die an der Erhebung von Daten beteiligt waren, welche im Rahmen dieser Arbeit verwendet wurden. Angesichts des Aufwandes den die Durchführung von Feld- und Fernerkundungskampagnen erfordert kann man deren Leistung und die Bereitschaft die Datensätze zu teilen nicht hoch genug anrechnen. Bei den Dozenten der HNE Eberswalde, Jan-Peter Mund und Alfred Schultz, möchte ich mich für die Ermunterung zur Promotion und ihr Interesse an deren

Voranschreiten sowie die Möglichkeit Lehr- und Betreuungstätigkeiten an der HNE zu übernehmen bedanken. Ich danke den Forest Information Technology Studenten Jonathan Falke, Duc-Thang Nguyen und Angel Naya-Geiger, deren Projektarbeiten wertvolle Beiträge geliefert haben, um die langfristige Vision dieser Dissertation zu verwirklichen.

Ich danke meinen Freunden aus gemeinsamen Eberswalder Studienzeiten, insbesondere Valentin, Isaac, Monika und Fabian, für den andauernden Austausch und Wissenstransfer und auch das ein oder andere Gespräch jenseits von R und Fernerkundung. Meinem Freundeskreis aus gemeinsamen Würzburger Studienzeiten danke ich für die schönsten Ausbrüche aus dem Arbeitsalltag, aber auch die Faszination über die vielfältigen Tätigkeiten die wir ehemaligen Bio-Erstis nun alle ausüben und die Horizont-erweiternden Diskussionen darüber.

Meiner lieben Familie gilt besonderer Dank. Ich danke meinen Eltern Roswitha und Manfred dafür, dass sie mein Interesse für Naturwissenschaften und meine Studienentscheidungen, inklusive der vielen Auslandsaufenthalte, immer unterstützt und gefördert haben. Ich danke meinem Bruder Julian für die Einblicke in Bereiche jenseits der eigenen Filterblase. Meinen beiden Omas danke ich für ihr Interesse an meiner Arbeit und ihre Erinnerung daran, dass ich schon als Kind den Berufswunsch „Propellor“ geäußert haben soll, dem ich ja mit der Promotion einen Schritt näher komme. Meiner Lebensgefährtin Cristina danke ich, dafür dass sie mich antreibt, ob durch Ermunterung oder auch Kritik, für ihren Blick auf das große Ganze und für die Freude des Zusammenlebens, wofür sie große Opfer bringt. Unserer Hündin Maya gebührt Dank für den richtigen Ausgleich zwischen Waldfernerkundung und Waldnaherkundung, so dass ich dank ihr nie vor lauter simulierten Bäumen den Wald nicht mehr gesehen habe.

## ERKLÄRUNG

---

### **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.

Aufgrund der Zusammenarbeit mit Kollegen bei der Konzeptentwicklung und der Ausarbeitung der Kapitel dieser Arbeit als Publikationen wurde an vielen Stellen die 'Wir'-Form verwendet. Nachfolgend sind die Personen aufgeführt, die mir bei der Auswahl und Auswertung von Material in der jeweils beschriebenen Weise unentgeltlich geholfen haben.

**Kapitel 1:** Kommentare und Sprachliche Korrekturen: Andreas Huth, Rico Fischer.

**Kapitel 2:** Ko-Autoren des Manuskripts: Rico Fischer, Andreas Huth.

**Kapitel 3:** Ko-Autoren des Manuskripts: Andreas Huth, Florian Kugler, Konstantinos Papathanassiou, Richard Condit, Stephen Hubbell, Rico Fischer.

**Kapitel 4:** Ko-Autoren des Manuskripts: Rico Fischer, Andreas Huth.

**Kapitel 5:** Kommentare und Sprachliche Korrekturen: Andreas Huth, Rico Fischer.

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.

---

Ort, Datum

---

Unterschrift