

SEEING WITHIN THE CANOPY: MEASURING THREE-DIMENSIONAL FOREST TRAITS
AND PROCESSES ACROSS ECOSYSTEMS AND SPATIAL SCALES

By

Aaron Giusti Kamoske

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Geography—Doctor of Philosophy

2021

ABSTRACT

SEEING WITHIN THE CANOPY: MEASURING THREE-DIMENSIONAL FOREST TRAITS AND PROCESSES ACROSS ECOSYSTEMS AND SPATIAL SCALES

By

Aaron Giusti Kamoske

From the bottom of their roots to the tops of their canopies, forests provide benefits for all of Earth's inhabitants including cultural and spiritual significance, economic opportunities, clean air and water, habitat for flora and fauna, and recreation and aesthetic values. Yet these important ecosystems are being lost at an alarming rate due to resource extraction and urbanization. With forests' irreplaceable services to humans, flora, and fauna alike, and their central role in carbon mitigation strategies, forest loss could have severe impacts on Earth's biodiversity and humanity.

However, not all forests are the same. Instead, they consist of a diversity of species, ages, and structures which directly impact the processes that drive carbon sequestration. For example, light use efficiency, photosynthetic capacity, and trace gas exchange are affected by within-canopy radiation regimes and turbulence environments which are directly and indirectly regulated by the horizontal and vertical distribution of foliage within the canopy. Functional traits (e.g., leaf mass per area and foliar nitrogen content) and structural traits (e.g., leaf area density) drive these processes while showing significant variation between and within plant functional types and vertically through forest canopies. These plant functional types and forest traits also appear in different locations across the landscape due to soils, topography, climate, historic landscape conditions, and management activities which directly impacts forest biodiversity.

To improve our estimates of processes related to carbon cycling and biodiversity, a better understanding of the three-dimensional variation of forest canopy traits is needed. Airborne remote sensing platforms that make use of hyperspectral and lidar data have recently been operationalized,

which provide an opportunity to examine forest functional and structural traits across spatial extents not possible by field surveys alone. This dissertation utilizes these airborne platforms and explicit field testing to estimate three-dimensional forest traits across ecosystems while quantifying the effects of biodiversity, topography, and biogeography on the spatial variation and distribution of these traits.

Chapter 1 introduces the concepts and questions raised in this dissertation. Chapter 2 addresses the impacts of spatial scale, pulse density, and canopy penetration on forest structure estimates from two airborne lidar systems, while offering solutions to enhance the accuracy of these estimates by standardizing spatial grains, limiting understory inflation, and utilizing Beer-Lambert coefficients. Chapter 3 assesses the influence of lidar derived forest structure, abiotic gradients, and management regions on the spatial patterns of remotely sensed top-of-canopy and total canopy nitrogen showing that total canopy estimates correspond to different ecological processes and exhibit unique spatial patterns than traditional top-of-canopy nitrogen estimates. Chapter 4 examines how taxonomic, functional, and phylogenetic diversity vary across eastern US forests, while assessing to what degree remotely sensed metrics are correlated with in situ biodiversity measures concluding that canopy structure is a critical predictor of forest biodiversity when combined with forest functional and topographic metrics. Chapter 5 summarizes the results and charts a path forward for research on forest structure, function, and diversity.

Overall, this dissertation shows that it is critical to consider forest structural and functional traits together to accurately estimate the spatial distribution and variation of canopy processes and biodiversity, while helping to paint a clearer picture of how forests function in a time of rapid global change.

This thesis is dedicated to the trees.
“...so that the several hundred kinds of hawthorn will not have to laugh
at the single name they’re forced to share.”
– Richard Powers

ACKNOWLEDGMENTS

First, I want to acknowledge the National Science Foundation's Macrosystems Biology program that funded much of my degree. Next, I thank my advisor and friend Dr. Kyla Dahlin for all her support. I also thank my committee members and manuscript co-authors, Drs. Scott Stark, David Rothstein, Ashton Shortridge, Shawn Serbin, Phoebe Zartnetske, Quentin Read, and Sydne Record – for their guidance I am forever thankful. This thesis project involved three summers of field work across multiple states in the Eastern United States and I want to thank the many people who assisted me and made this research possible. Lastly – and most importantly – I thank my wife, Morgan, whose countless hours of encouragement, support, and laughter made the hard times pass and the good times better. And my parents who always knew I was able, even when I did not know myself.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER 1. INTRODUCTION	1
Research Context	1
Dissertation Focus and Organization	4
CHAPTER 2. LEAF AREA DENSITY FROM AIRBORNE LIDAR: COMPARING SENSORS AND RESOLUTIONS IN A TEMPERATE BROADLEAF FOREST ECOSYSTEM	8
Introduction.....	8
Materials and Methods.....	12
Study Site.....	12
Hemispherical Photography for LAI Estimation	13
Lidar Acquisition and Processing	14
Leaf Area Density from Airborne Lidar	17
LAD Profile Extraction.....	19
Comparing LAD Estimates.....	20
Comparing LAI and Total Leaf Area Estimates	20
Results.....	21
Lidar Penetration of Forest Canopies	21
Beer-Lambert Coefficients.....	23
LAD Profile Estimates	24
Total Leaf Area Estimates	27
Discussion	29
Measuring Leaf Area Density from Above	29
Lidar System Considerations	31
Ecological Implications	32
Looking Forward	33
Conclusion	33
Acknowledgments.....	34
Data Availability.....	35
CHAPTER 3. LEAF TRAITS AND CANOPY STRUCTURE TOGETHER EXPLAIN CANOPY FUNCTIONAL DIVERSITY: AN AIRBORNE REMOTE SENSING APPROACH	36
Introduction.....	36
Materials and Methods.....	39
Site Description.....	39
Airborne Remote Sensing Data	40
Field Data Collection and Lab Methodologies	41
Lidar Methods.....	43

Hyperspectral Imagery Methods.....	46
Remote Sensing Fusion: Total Canopy N.....	48
Raster Differences Across Scales	49
Environmental Driver Analysis	50
Results.....	51
Trait Prediction with PLSR: From Leaf to Canopy	51
Within Canopy Leaf Traits: Lidar and HSI	54
Top-of-Canopy and Total Canopy N: Differing Spatial Patterns	56
Regional Patterns and Environmental Drivers: Assessing Spatial Structure.....	58
Discussion.....	63
Scaling and Mapping Leaf and Canopy Traits	64
Measuring Ecosystem Function: Top-of-Canopy %N vs. Total Canopy N	66
Abiotic and Management Drivers of Foliar and Canopy N.....	67
Model Uncertainty and Data Concerns.....	68
Looking Forward	69
Conclusions.....	70
Acknowledgments.....	71
Data Availability.....	72
CHAPTER 4. MAPPING MULTIPLE DIMENSIONS OF FOREST BIODIVERSITY WITH AIRBORNE HYPERSPECTRAL AND LIDAR REMOTE SENSING.....	73
Introduction.....	73
Materials and Methods.....	76
Study Sites	76
Calculating Tree Diversity Metrics within NEON Field Plots	77
Remote Sensing Data.....	79
Forest Structural and Topographic Diversity from Lidar Remote Sensing	79
Hyperspectral Remote Sensing Reflectance Metrics.....	81
Influence of Remote Sensing Metrics on Biodiversity	82
Detecting Biodiversity and Remote Sensing Metric Variation.....	83
Results.....	84
Variation of Biodiversity and Remote Sensing Metrics	84
LME Models	86
Clustering Diversity and Remote Sensing Metrics.....	91
Discussion.....	94
Using Remote Sensing to Measure Biodiversity	94
Biodiversity Variation Across Eastern Temperate Forests.....	95
Remote Sensing Metrics and Biodiversity Heterogeneity	96
Model Uncertainty and Data Concerns.....	97
Looking Forward	98
Conclusions.....	98
Acknowledgments.....	99
Data Availability.....	100
CHAPTER 5. CONCLUSIONS	101
Summary of Results.....	101

Recommendations for Future Research	102
APPENDICES	104
APPENDIX A: Chapter 3 Supplementary Materials	105
APPENDIX B: Chapter 4 Supplementary Materials	115
BIBLIOGRAPHY	129

LIST OF TABLES

Table 1. SERC Beer-Lambert Coefficients	24
Table 2. SERC Whole Canopy Statistics	26
Table 3. SERC Reduced Canopy Statistics.....	26
Table 4. TALL PLSR Model Results	52
Table 5. TALL Model Coefficients	61
Table 6. LMER Model Results	88
Table A.2. Mean and SD for field samples (%N and LMA)	107
Table A.7. Abiotic and management variables.....	112
Table B.1. Field site information	116
Table B.2. Functional traits and phylogeny	117
Table B.3. Hyperspectral variables.....	118
Table B.4. Lidar derived variables.....	119
Table B.5. Topographic variables.....	120
Table B.6. Beer Lambert coefficients	121
Table B.7. Results from individual models	122
Table B.8. Results from PCA and Cluster Analysis	123
Table B.9. ANOVA results for dimensions of biodiversity.....	124
Table B.13. ANOVA results for remote sensing variables.....	128

LIST OF FIGURES

Figure 1. Traditional 2D Remote Sensing and 3D Remote Sensing.....	4
Figure 2. Map of Study Area at SERC	13
Figure 3. LAD Methodology	18
Figure 4. Lidar Spatial Resolutions	22
Figure 5. Lidar Within Canopy Returns	23
Figure 6. SERC LAD Profiles	25
Figure 7. SERC LAD Spatial Patterns.....	28
Figure 8. Map of Study Area at TALL	40
Figure 9. Total Canopy Modeling Methodology	45
Figure 10. Canopy Functional Trait Maps.....	53
Figure 11. Total Canopy N Map.....	55
Figure 12. Top-of-Canopy N and Total Canopy N Differences	57
Figure 13. Canopy N Variograms.....	58
Figure 14. Influence of Elevation on Canopy N.....	59
Figure 15. TALL Model Coefficients.....	60
Figure 16. Total Overstory Foliar N Comparison.....	65
Figure 17. NEON Field Sites Map.....	77
Figure 18. Biodiversity and Remote Sensing Metric Variation.....	85
Figure 19. Marginal and Conditional R^2 Values.....	87
Figure 20. Final LMER Model Results	90
Figure 21. Clustered Metric Variation	93

Figure A.1. Field data from TALL	106
Figure A.3. PLSR output from laboratory %N estimation	108
Figure A.4. PLSR output from HSI %N estimation	109
Figure A.5. PLSR output from HSI LMA estimation.....	110
Figure A.6. Output from within-canopy trait prediction model.....	111
Figure A.8. LAI plots.....	113
Figure B.10. Hyperspectral variables histograms	125
Figure B.11 Lidar variable histograms	126
Figure B.12. Topography variables histograms.....	127

CHAPTER 1. INTRODUCTION

Research Context

A fundamental goal in biogeography, developed over 250 years ago by Alexander von Humboldt, is to blend empirical field research with quantitative methods to understand how environmental and anthropogenic changes affect the distribution, abundance, and biodiversity of plant species across the globe (Nicolson 1987; Schaefer 1953). This objective has been further developed by community ecologists using advanced statistical methodologies, contemporary technologies, and empirical field research to not only understand – as von Humboldt set out to do – but also to predict the spatial distributions of species, traits, and biodiversity across ecosystems (Keddy 1992). This ability to predict spatial distributions of key ecosystem process, species, and traits is especially critical in forested ecosystems as they link the atmosphere and the vast majority of Earth’s terrestrial biomass (Ozanne et al. 2003), as well as provide key ecosystem functions including habitat for most of Earth’s biodiversity, clean water and air, and a plethora of uses related to medicine, spirituality, recreation, and local economies. Also, forested ecosystems store ~45% of terrestrial carbon (Bonan 2008; Bunker et al. 2005; Tilman et al. 2006) and contribute ~50% of terrestrial net primary production (Bonan 2008; Hooper et al. 2012; Isbell et al. 2015). Moreover, tree species, canopy traits, and forest biodiversity are not constant in space and instead show significant heterogeneity across landscapes (Chambers et al. 2007; Asner et al. 2014) and within the full vertical and horizontal extent of the canopy (the ‘canopy volume’; Ellsworth et al. 1993).

This heterogeneity can be measured by the variation of canopy structural and functional traits in three-dimensional space. These structural traits include leaf area density (LAD; the total leaf area per unit of volume) which explains the horizontal and vertical spatial variation of leaf area within a canopy (Weiss et al. 2004) and canopy clumping (e.g., a measure of foliage

aggregation relative to a random spatial distribution of leaf material within the canopy; Pisek et al. 2018). Important leaf functional traits such as leaf mass per area (LMA; the ratio between leaf dry mass and leaf area) and foliar nitrogen (foliar N; g/m_L^2 ; m_L = meter of leaf area) have been identified as key predictors of plant functional diversity (Diaz et al. 2016) and show strong relationships to photosynthetic carbon assimilation (Field and Mooney 1986), primary productivity (Ollinger and Smith 2005), and other photosynthetic parameters (Niinemets et al. 2015) within the canopy volume (Poorter et al. 2009, Niinemets, 2007). This relationship between three-dimensional forest structure and canopy function drives important processes such as net photosynthetic carbon assimilation (Niinemets 2007), resource use and efficiency (Hardiman et al. 2013), and woody growth (Stark et al. 2012). For example, trees with high LMA and low nitrogen content produce leaves that will likely survive longer, but require more resources during growth, while trees with low LMA and high nitrogen content will produce leaves that do not survive as long but are cheap to grow (the leaf economics spectrum; Wright et al. 2004). At the same time, leaves alter their orientation to maximize the amount of light they receive, which can change the local light environment by casting shade on neighboring leaves and other plants, leading to changes in LMA and nitrogen content to maximize photosynthesis in response to light conditions (e.g., lower LMA values in the shaded parts of canopy and higher LMA in sunlit areas; Niinemets 2007). However, this variation of traits within the canopy volume and across landscapes is usually simplified by traditional remote sensing products and Earth System Models (ESMs).

Spaceborne remote sensing has played a significant role in understanding the terrestrial carbon cycle for decades (Tucker and Sellers 1986; Schimel 1995). Insights from NASA satellite missions, such as the Advanced Very High Resolution Radiometer (AVHRR), Landsat, and the Moderate Resolution Imaging Spectroradiometer (MODIS) have shaped our understanding of how

ecosystems function, driving the development of global carbon models (DeFries et al. 1999; DeFries et al. 2002). Yet these satellite systems and models often assume a two-dimensional world (Bonan et al. 2014; Fisher et al. 2018) – ignoring critical information about the three-dimensional structure of forest canopies, which is vital to ecological processes related to carbon sequestration (Parker et al. 2004; Hardiman et al. 2011; Niinemets 2007; Bonan et al. 2008; Lawrence et al. 2007). For instance, the Community Earth System Model assigns the canopy two layers – sunlit and shaded – and applies the same model parameters to one of 15 non-crop plant functional types (PFTs; Hurrell et al. 2013) and then uses these models at global scales with very little consideration for differences within the canopy volume (Braghiere et al. 2019). However, functional and structural trait variation within PFTs has been shown to be critically important to ecosystem functioning by maximizing resource use efficiency through adaptation to local environmental conditions (Williams et al. 2017). With new spaceborne and airborne technologies, we have an opportunity to think about the terrestrial carbon cycle in three-dimensions allowing for more realistic estimations of the variation of forest structural and functional traits within PFTs, across landscapes, and throughout the canopy volume (Taylor et al. 2015; Stark et al. 2012). This will also help unlock important insights into how forests function in a time of rapid anthropogenic change (Antonarakis et al. 2014; Hardiman et al. 2013) while improving the performance of ESMs (Bonan et al. 2014; Bonan et al. 2012).

These new spaceborne and airborne remote sensing technologies can measure critical canopy functional and structural traits in three-dimensions – which are ignored by traditional 2D remote sensing platforms (Figure 1) – through the fusion of lidar and hyperspectral data (Asner and Martin 2009; Dahlin et al. 2013; Asner et al. 2010). Lidar has been used to measure three-dimensional forest structural traits such as LAD across landscapes at scales relevant to forest

managers and ecologists (e.g., at scales equal to or finer than an individual tree’s canopy; Kamoske et al. 2019). At the same time, hyperspectral data have been used extensively to measure plant functional traits such as foliar N and LMA across entire landscapes with high degrees of accuracy and precision (Martin et al. 2008; Asner et al. 2011). The fusion of these two data types can capture the diversity of traits and processes vital to canopy photosynthetic capacity, light use efficiency, exchange of water vapor and gases, and the within-canopy light regime – all of which play critical roles in carbon sequestration and are directly related to the three-dimensional structure of forest canopies (Asner et al. 2009; Dahlin et al. 2013).

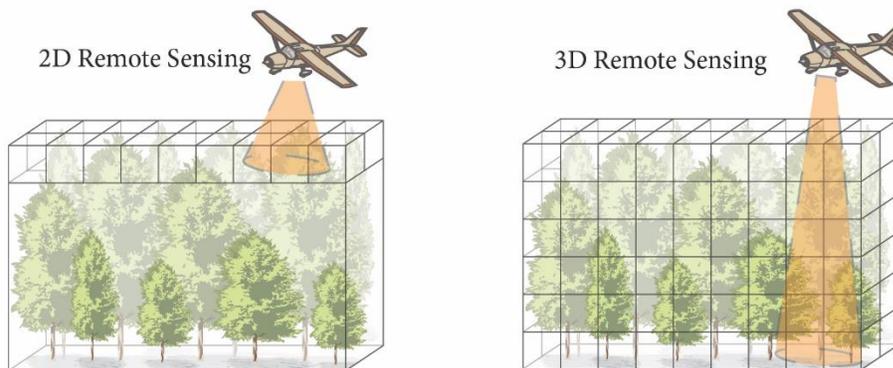


Figure 1. Traditional 2D Remote Sensing and 3D Remote Sensing. 3D remote sensing allows for the collection of information about the entire forest canopy volume that is ignored by traditional 2D remote sensing techniques.

This dissertation aims to determine how airborne remote sensing platforms can be used to measure forest functional and structural traits within the canopy and biodiversity across landscapes. This leads to a question in community and forest ecology that remains unanswered: Do landscape scale patterns of forest traits and biodiversity vary when whole plant structure is considered?

Dissertation Focus and Organization

The goal of this work is to address questions about the ability of airborne remote sensing platforms to accurately measure and map forest traits within the canopy volume and biodiversity across eastern US temperate forests. Chapters 2 through 4 are self-contained studies that address the following interrelated research questions:

- 1) Can new publicly available airborne platforms be used to accurately measure forest structural traits within the canopy volume?
- 2) Are leaf-level functional traits within the canopy volume able to be estimated using a combination of field data and remotely sensed airborne lidar and hyperspectral data?
- 3) Across eastern United States temperate forests, can taxonomic, phylogenetic, and functional diversity be predicted using topographic, spectral, and structural diversity data derived from airborne remote sensing platforms?

Chapter 2 examines the ability of two contemporary airborne lidar systems – the National Ecological Observatory Network’s Airborne Observation Platform (NEON AOP) and NASA Goddard’s Lidar, Hyperspectral, and Thermal Imager (NASA G-LiHT) – to estimate the three-dimensional structure of forest canopies in a temperate forest ecosystem. These two systems vary greatly in their survey and instrument specifications, collection goals, and laser pulse densities, leading to statistically significant differences in the lidar point clouds (e.g., high-density point clouds from NASA G-LiHT and low-density point clouds from the NEON AOP) that impact the accuracy of forest structural estimates. Based on these findings, a reproducible and open-source methodology to overcome these differences was developed to address the impacts of the spatial scale of analysis and differences in lidar pulse density on structure estimates. This standardized approach helps to bridge the gap between ecologists and forest managers who could use lidar data

in research and management plans and remote sensing scientists who use these data on a regular basis.

Chapter 3 presents remote sensing-based estimates of within-canopy functional traits in multiple forest ecosystem types. With a fusion of hyperspectral and lidar data from the NEON AOP and field-collected foliar trait data, the impacts of forest structure on spatial patterns of foliar nitrogen are assessed across a landscape consisting of a mosaic of open longleaf pine and dense broadleaf deciduous forests. In addition, the influence of abiotic gradients and management regimes on top-of-canopy percent nitrogen and total canopy nitrogen are examined. Resulting maps suggest that in contrast with top-of-canopy values which show high LMA and low N values in needleleaf species and low LMA and high N values in broadleaf species, total canopy nitrogen variation is dampened across this landscape resulting in relatively homogeneous spatial patterns due to broadleaf and needleleaf species having more similar total amounts of nitrogen within their canopies.

Chapter 4 addresses the persistent goal in community ecology to understand and predict the spatial distributions of species, traits, and biodiversity across ecosystems (Keddy 1992). This chapter assesses the ability of fine grain remotely sensed metrics related to vegetation health, forest structure, spectral diversity, and topography to predict biodiversity across eastern US temperate forests. Results show that these commonly used remotely sensed metrics do not completely capture the spatial patterns observed in the biodiversity variables, suggesting that new remote sensing metrics may need to be developed to better capture this variation. Moreover, different forest regions (e.g., Southern mixed (oak-pine), mesophytic (Appalachian oaks), Oak-hickory, Northern hardwoods; Dyer 2006) exhibit high and low diversity field plots, but both high and low diversity

plots may not be different between each forest region (e.g., a high diversity oak-hickory plot is not significantly different than a high diversity oak-pine plot).

Chapter 5 summarizes and discusses these data and findings from the three preceding chapters, offers suggestions for future research, and identifies contributions these works offer to the disciplines of community ecology and remote sensing.

CHAPTER 2. LEAF AREA DENSITY FROM AIRBORNE LIDAR: COMPARING SENSORS AND RESOLUTIONS IN A TEMPERATE BROADLEAF FOREST ECOSYSTEM.

Citation: Kamoske, A.G., K.M. Dahlin, S.C. Stark, and S.P. Serbin. 2019. Leaf area density from airborne Lidar: Comparing sensors and resolutions in a temperate broadleaf forest ecosystem. Forest Ecology and Management 433: 364-375.

Introduction

With terrestrial ecosystems storing around 11 gigatonnes of atmospheric carbon dioxide (CO₂) per year, approximately one third of anthropogenic emissions, forests are a critical component of the Earth's carbon cycle (Le Quere et al. 2015; Pan et al. 2011b). Forest processes that play an essential role in carbon sequestration are closely related to the three-dimensional structure of forest canopies (Parker et al. 2004; Hardiman et al. 2011). The horizontal and vertical distribution of foliage within a canopy directly and indirectly regulates the canopy-scale light use efficiency (LUE, Ellsworth and Reich 1993; Kitajima et al. 2005), photosynthetic capacity, and exchanges of water vapor, CO₂, and other trace gases (Baldocchi et al. 1988) in a number of important ways, including by defining the within-canopy radiation regime (Meir et al. 2002; Niinemets 2007) and turbulence environment. This variability, in turn, has significant impacts on forest productivity and thus carbon storage (Hardiman et al. 2013). However, these relationships are not static in 2- or 3-dimensional space; leaf physiological traits vary across landscapes (Serbin et al. 2014) and leaves at varying vertical positions within the canopy are physiologically unique due to differing light environments (Poorter et al. 2009). Given differences in leaf physiology and morphology, a better knowledge of how these properties vary vertically and horizontally within the canopy will provide a better estimate of carbon storage (Niinemets et al. 2015). Due to this within-canopy variation of light and foliar traits, inclusion of the three-dimensional structural diversity of a forest canopy is

critical to making improvements to carbon storage estimates by Earth system models (ESMs) (Bonan et al. 2012).

The structural diversity within a canopy and across a landscape is a critical component of ecological models that scale processes from leaf to landscape (Jarvis and McNaughton, 1986). Currently, many of these models, such as the Community Land Model, treat the canopy as only having two types of leaves – sunlit and shaded (Bonan et al. 2014). This lack of information about the three-dimensional canopy is one of a host of factors contributing to the uncertainty and disparity in predicting carbon uptake by terrestrial ecosystems (Fisher et al. 2017). By incorporating this vertical and horizontal structure, models can provide a better representation of forested landscapes, thus reducing model uncertainty and improving estimates of ecosystem productivity and landscape-scale functions (Bonan et al. 2014). However, the benefits of accurately measuring the three-dimensional structure of a forest canopy are not limited to ESMs.

Understanding the effects of disturbances on forested ecosystems is vital to long-term quantification of carbon storage (Goodale et al. 2002; Pan et al. 2011a). Defoliation from invasive insects and pathogens (Hummel and Agee 2003), stand replacement and thinning from fire (Collins et al. 2011), stress and mortality from drought (Anderegg et al. 2013), gap creation from wind (Hanson and Lorimer 2007), and a host of other disturbance impacts can affect the structure of a forest. Moreover, these changes in forest structure can significantly affect ecosystem processes related to carbon uptake (Gough et al. 2013). With forest managers facing increasingly complex disturbances, the ability to map and measure forest structure across landscapes is critical to developing forest management plans that consider the impacts of these disturbances on forest health, resilience, and function (Becknell et al. 2015).

While well-established field-based methods to measure the vertical and horizontal distribution of leaves within an individual tree's canopy exist, applying these methods at a landscape scale is challenging due to time, labor, and access constraints (Zheng and Moskal 2009). These measurements are made primarily by two methodologies (Hosoi and Omasa 2007); by lowering a probe through the canopy and recording the height and frequency of foliage contact with the probe (e.g., inclined point quadrat method; Wilson 1960) or by using a telephoto lens to measure the proportion of leaves in a given area at set height intervals looking up into the canopy (e.g., canopy closure method; MacArthur and Horn 1969). These methodologies are ultimately used to estimate leaf area index (LAI; the one-sided leaf area per unit of ground area; Chen and Black 1992) and leaf area density (LAD; the total leaf area per unit of volume; Weiss et al. 2004) which provide critical information about ecosystem processes and functions related to forest structure (Detto et al. 2015). The estimation of these variables is influenced by a variety of factors, including the assumption that leaves are distributed randomly throughout the canopy and the scale at which the measurements were taken. For example, the distribution of clusters of leaves, stems, and branches, the spatial structure of gaps in the forest, and the disturbance histories of a landscape (Silva et al. 2017) may influence the error and bias of forest structural estimates depending on the sampling scale used (e.g., an individual tree, a plot, or a forest stand; Roussel et al. 2017). However, emerging technologies present opportunities to evaluate factors influencing LAI and LAD estimates and to overcome prior limitations to extracting this critical information across landscapes, at varying temporal resolutions and with high accuracy.

Airborne light detection and ranging (Lidar) directly measures the distance between a sensor and an object using laser pulses. Lidar sensors provide a high repetition rate of these measurements (as high as 33,000 pulses per second) and when applied to forests these lidar pulses

act as a canopy probe, allowing for the estimation of the three-dimensional internal structure of a forest canopy (Lefsky et al. 2002). In contrast, traditional passive optical remote sensing systems (e.g., Landsat) produce two-dimensional images of sunlight reflected off the top of the canopy, which does not capture the complex vertical and horizontal structure of a forest canopy (Morsdorf et al. 2006). Lidar derived structural measurements have the potential to improve the accuracy and resolution of studies that have traditionally relied on two-dimensional remote sensing or field surveys, including estimates of defoliation from invasive pests and pathogens (Meng et al. 2018), predicting above ground carbon dynamics (Taylor et al. 2015; Stark et al. 2012; 2015), measuring forest stand successional stages (Falkowski et al. 2009), within-canopy habitat modeling (Smart et al. 2012), and ecosystem trait upscaling (Antonarakis et al. 2014). Additionally, with a host of applications lidar data are becoming more widely available at larger spatial and temporal scales.

In the United States, two airborne systems are acquiring publicly available lidar data at a wide-range of locations covering many ecoregions with increasing frequency. The National Ecological Observatory Network's Airborne Observation Platform (Kampe et al. 2010; NEON AOP) is collecting airborne data at more than 60 sites throughout the United States in 2018 (its first year of full operations), with survey areas ranging from 100 to 300 km² around each site. Data will be collected at NEON sites on a semi-annual basis for the next 30 years, offering an unprecedented opportunity to address long-term ecological questions (Hinckley et al. 2016). NASA Goddard's Lidar, Hyperspectral & Thermal Imager (NASA G-LiHT) is another airborne system, which currently flies site-specific missions for NASA-funded studies (Cook et al. 2013), with data publicly available in over 30 US states and territories and several countries dating back to 2011. With the launch of the NEON AOP and the continuing collection by NASA G-LiHT, lidar is becoming more readily available than ever before. This abundance of data will further grow with

the launch of NASA's Global Ecosystem Dynamics Investigation (GEDI) system on the International Space Station, which will provide waveform lidar coverage of temperate and tropical forests between 51 degrees North and South, beginning in late 2018 (Stavros et al. 2017). While lidar data are becoming more available at wider spatial extents and temporal scales, a critical gap remains between landscape and macrosystem ecologists who want to ask questions at broad spatial scales and remote sensing scientists who are more familiar with the opportunities and challenges of using these data (Turner et al. 2015; Pettoirelli et al. 2014; Mairota et al. 2015).

In this study, we help address this gap by describing a reproducible and open-source methodology for estimating LAD and LAI from airborne lidar. We compare LAD estimates derived from publicly available point cloud data produced by the NEON AOP and NASA G-LiHT lidar systems, which differ in survey and instrument specifications, collection goals, and laser pulse densities. Furthermore, we use hemispherical photographs as a means to calibrate our lidar derived LAD and LAI estimates. We also address the impacts of the spatial scale of analysis and differences in canopy penetration and pulse density on LAD and LAI estimates while offering potential solutions to enhance the accuracy of these estimates.

Materials and Methods

Study Site

Field measurements and lidar data were acquired at the Smithsonian Environmental Research Center (SERC), approximately 16 km south of Annapolis, Maryland, USA (Figure 2). SERC is a relocatable terrestrial NEON (neonscience.org) site and contains a mixed-species deciduous forest with American sweetgum (*Liquidambar styraciflua*) and tulip tree (*Liriodendron tulipifera*) dominating the overstory. Mockernut hickory (*Carya tomentosa*), white oak (*Quercus alba*), and

American beech (*Fagus grandifolia*) are also common, with ironwood (*Carpinus caroliniana*) and other small tree species forming a dense understory (Parker 1995). SERC contains approximately 11 km² ranging in elevation from zero to 40 meters above sea level, and with slopes ranging from zero to 34 degrees.

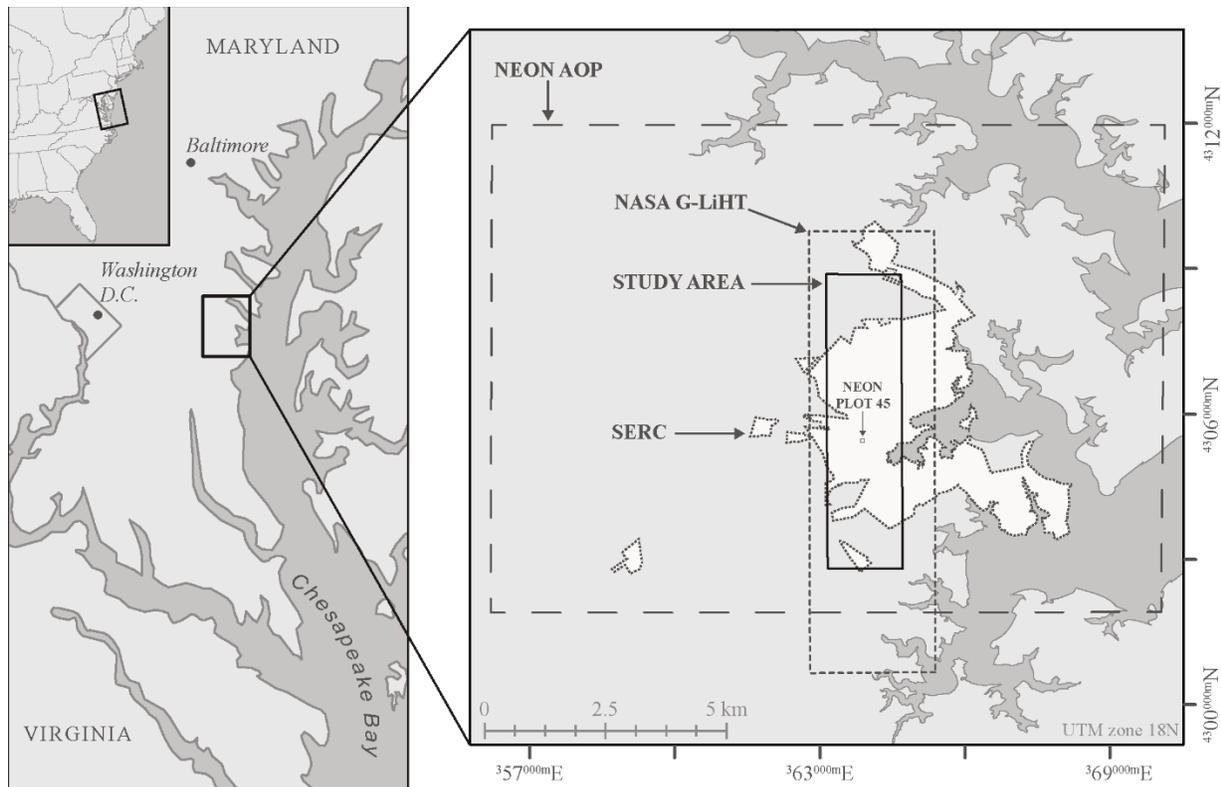


Figure 2. Map of Study Area at SERC. The study area (solid line) is an overlapping subset of the NEON AOP (long dashed line) and NASA G-LiHT (short dashed line) flight boundaries. The noncontiguous boundary of SERC is shown in white with a dotted outline, while NEON plot 45 is shown by a small dashed box within the SERC boundary.

Hemispherical Photography for LAI Estimation

We collected hemispherical photographs between July 23 and August 7, 2017 to coincide with G-LiHT and NEON flights, using a Canon EOS Rebel T6 camera with an 8 mm circular fish-eye lens (180-degree angle of view). We placed the camera on a leveled tripod one meter above the ground

to reduce the influence of ground vegetation and provide a vertical picture of the canopy. Each location was recorded using a Trimble GEO7x GPS. We took a single hemispherical photograph at 48 different locations along six transects and three plots within half a kilometer of the NEON flux tower (38.89° N, -76.56° W), with all photographs taken before sunrise or under uniformly cloudy conditions. Using the Digital Hemispherical Photography (DHP) software (Leblanc et al. 2005), we processed the hemispherical photographs for effective plant area index (PAI_e), which includes leaf and woody material in the gap fraction calculation (Miller 1967). Most studies use PAI_e as a proxy for LAI due to the difficulty of correcting for non-foliage elements in these photographs (Richardson et al. 2009), and hereafter we refer to PAI_e as LAI. Further, we set the zenith angle within the DHP software to match the scanning angle of each lidar sensor, to better relate the ground measured LAI to lidar derived LAI (Sabot et al. 2014; Richardson et al. 2009; Solberg et al. 2006; Korhonen et al. 2011).

Lidar Acquisition and Processing

NASA G-LiHT and NEON AOP collected lidar data on July 31, 2017 and between July 20 and August 10, 2017, respectively. NASA G-LiHT data were collected using a Riegl VQ-480i lidar sensor, operating at a wavelength of 1550 nm, with a scan angle of +/-30 degrees, a pulse repetition frequency of 300 kHz, a beam divergence of 0.3 mRad, and an average point density of 15.86 pts/m². NEON AOP data were collected using an ALTM Gemini lidar sensor, operation at a wavelength of 1064 nm, with a scan angle of +/-18 degrees, a pulse repetition frequency (PRF) of 100 kHz, a beam divergence of 0.8 mRad, and an average point density of 3.15 pts/m². Differences in the specifications of lidar systems can have substantial impacts on subsequent LAD and LAI estimations. Below we describe these parameters and how they relate to measuring forest canopies.

The scan angle is the range of angles at which the sensor scans the landscape (Figure 3A). By increasing the scan angle, the lidar pulses will cover a larger area and have a greater chance of encountering a gap in the canopy, thus having a higher probability of penetrating a dense forest canopy. The pulse repetition frequency (PRF; Figure 3C) is the number of pulses per second that a sensor produces, measured in cycles per second or kilohertz (kHz). A lower PRF results in fewer pulses produced per second, thus negatively affecting the density of the point cloud and the probability of penetrating a dense canopy. A related measure is beam divergence (Figure 3B), which is an angular measure describing how the laser beam widens as the distance between the sensor and the ground grows, measured in milliradians (mRad). A large mRad value will cause the laser's energy to be spread across a wider area, likely reducing its ability to penetrate a dense forest canopy and producing a lower signal-to-noise ratio (Gatziolis and Andersen 2008). Together, these parameters are pivotal to producing a high-quality lidar dataset with precise and accurate information about the internal structure of the forest canopy. Combined, these parameters determine point cloud density, forest canopy penetration, and the proportion of ground returns in the dataset, the latter of which is an essential measurement in the estimation of LAI and LAD.

We downloaded lidar point clouds as .las and .laz files through the NEON (National Ecological Observatory Network 2017) and NASA (Cook et al. 2013) data portals. Lidar pulses within the point cloud were classified as ground or not-ground by NEON and NASA prior to downloading. The code to reproduce these analyses is available as an R package on GitHub (canopyLazR; see Data Availability Statement). Using the R programming language (R Core Team, 2016), we loaded the lidar files into the workspace as individual datasets using the rlas library (Roussel, 2016). Due to the different spatial footprints of the NEON AOP and NASA G-LiHT flights, we took a subset of the overlapping data, thus returning two datasets (NEON AOP

and NASA G-LiHT) with matching spatial extents, each with an area of 7.2 km² (Figure 2). Next, we transformed the point clouds into voxelized arrays using the R libraries plyr (Wickham, 2011) and fields (Nychka et al. 2015). The first slice of the voxelized array contains the lowest ground height for each column of voxels. While using the lowest ground height might introduce some amount of uncertainty (Khosravipour et al. 2015), we choose the lowest ground height rather than the mean ground height so that we would not eliminate any understory vegetation that occurred below the mean ground elevation. The next slice contains the height of the canopy and each subsequent slice contains the number of pulses that occur in the given voxel. We set the voxel height to one meter but used multiple horizontal resolutions for this study: 1x1m, 2x2m, 5x5m, 10x10m, 20x20m (NEON vegetation plot resolution; National Ecological Observatory Network, 2017), and 30x30m (Landsat pixel resolution). If there is not a ground return present in the vertical column, we assigned it a NA value, and this column was not used in subsequent analyses. We removed these voxels because we wanted to account for the total canopy LAD and not just the upper canopy. In addition, the data were voxelized at each of these spatial resolutions independently, not by aggregating finer resolutions together. For instance, a 10x10 meter voxel contains all lidar returns that occur within the given spatial extent, whereas aggregating finer spatial resolutions together would result in the removal of many voxels that have upper canopy returns but no ground returns, which would result in a NA value for the entire vertical column. To better compare the voxelized data independent of changes in topography across the study site, we created a voxelized canopy height model from the lidar array so that each column of voxels was scaled to the distance from the ground, thus the ground has a height of zero. This eliminates the effects of topography on the dataset making comparisons between voxel columns easier (Lovell et al. 2003).

Leaf Area Density from Airborne Lidar

We estimated LAD from voxelized lidar data using an approach based on the method established by MacArthur and Horn (1969) and similar to other published methods (Stark et al. 2012; Zhao and Popescu, 2009; Solberg et al. 2006; Sumida et al. 2009; Bouvier et al. 2015) (Figure 3). With this methodology, we calculated LAD by counting the number of lidar pulses that enter and exit each voxel in a given vertical column. Within each voxel, LAD is estimated as:

$$LAD_{i-1,i} = \ln \left(\frac{S_e}{S_t} \right) \frac{1}{k\Delta z}$$

where for each vertical column of voxels, i is a voxel in a sequentially ordered vertical column of the canopy, S_e is the number of pulses entering the given voxel, S_t is the number of pulses exiting the same voxel, k is an extinction coefficient, and z represents the height of a voxel. The term k represents a Beer-Lambert Law extinction coefficient, which describes the attenuation of light by a medium or an object. When applied to forest canopies this derived value includes a correction for the non-random distribution and orientation of the foliage and the thickness of the leaf material and the forest canopy. Thus, as the canopy becomes denser and more leaves are encountered, the penetration of lidar pulses will diminish causing sample sizes for estimating LAD to decrease and error to increase.

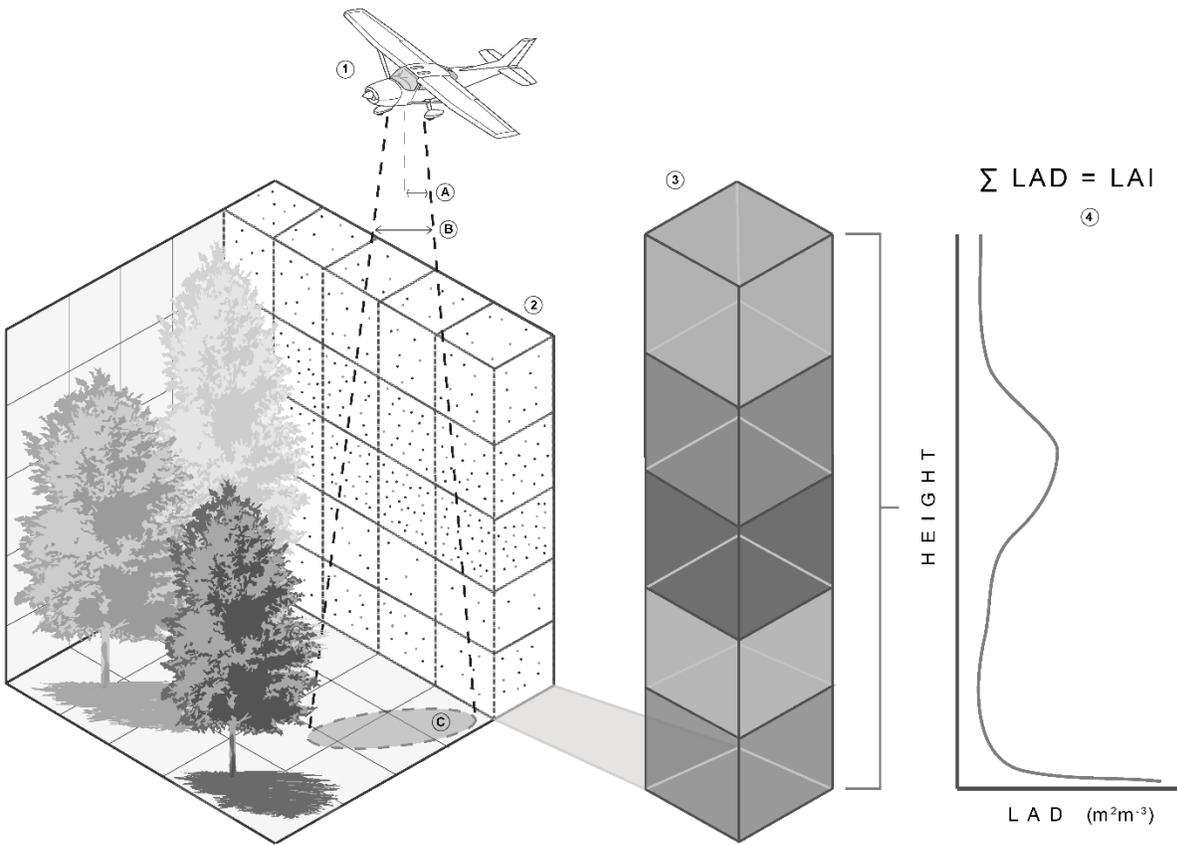


Figure 3. LAD Methodology. Lidar pulses from the airborne sensor penetrate the forest canopy and either bounce off leaf or woody material, or hit the ground, and return to the plane (1). These height measurements are then voxelized at the desired spatial resolution (2). The MacArthur and Horn method is then applied to voxelized columns of lidar returns (3) returning a LAD profile of the given area (4). The sum of LAD values in a column of voxels with a ground return is equal to the LAI of that vertical column. Further, the scan angle (A), beam divergence (B), and point density (C) sensor-survey parameters are highlighted.

An extinction coefficient can be used to better relate field measured LAI with lidar estimated LAD or LAI. To estimate the extinction coefficient, we first estimated LAD from the lidar data with the extinction coefficient set to one. We removed vertical columns of voxels without at least one ground return from further analysis since this indicates that some of the canopy column was not sampled by the lidar sensor, thus preventing estimation of LAD in unsampled voxels and

LAI which relies on the sum of all column voxels. We then extracted the lidar estimated LAI values at the same coordinates the hemispherical photographs were taken, using the field recorded GPS locations. Next, we plotted each extracted lidar estimated LAI value against the same LAI estimate from a hemispherical photograph. The slope of the linear model fit without an intercept, which is used because the Beer-Lambert law assumes that there is a true zero intercept, estimates the extinction coefficient (Klingberg et al. 2017). We then estimated LAD from the lidar array again, this time including the extinction coefficient in the above equation, resulting in a LAD estimate for each voxel in a given vertical column that is adjusted to more closely resemble the hemispherical photography approach. We repeated this process for both sensors at each of the six spatial resolutions, resulting in 12 voxelized arrays containing LAD estimates.

LAD Profile Extraction

To extract LAD vertical profiles for individual point locations, we converted the adjusted LAD arrays to raster stacks with each raster layer representing a 1-meter interval of the forest canopy using the raster package (Hijmans 2016) in R. We then generated 50 random 20x20m plots across the study area to compare LAD estimates from each of the airborne systems. At each plot, the mean LAD estimate of all raster cells, at all heights within the canopy, that were either completely or partially within the plot was extracted. We repeated this for each of the 12 raster stacks previously generated (each sensor and six spatial resolutions). To better visualize this information, we extracted this same data from a permanent NEON forest plot (NEON plot 45; see Fig 1), using coordinates from the NEON data portal.

Comparing LAD Estimates

To compare the LAD estimates from NEON AOP and NASA G-LiHT, we used linear regression to model the relationship between the estimates from each sensor and each spatial resolution. We split the data into two categories, ground to the top of the canopy (TOC) (all data) and 10 meters above the ground to TOC (removing understory data), due to inflated LAD values that may occur in the understory of the canopy at coarser spatial resolutions (Stark et al. 2012). We speculate that this happens due to topographic effects within a given cell, where the range of ground elevations is greater than the voxel height. Point returns from the understory vegetation at these higher elevations will be counted in a voxel higher in the canopy than where they actually occur, due to a cell containing only a single ground elevation. Moreover, the vertical distribution of leaf area can change with the age of a forest stand, causing higher LAD values in the understory of younger and older stands, with LAD values peaking in the upper canopy in middle-aged stands (Brown and Parker 1994). While 10 meters might not be the best height to address LAD estimate uncertainty for a single spatial resolution, we chose this value to easily compare results across multiple spatial resolutions; studies conducted with a single spatial resolution should choose a height cutoff based on the data present. Next, we calculated R^2 , 95% confidence intervals, lines of best fit, slopes, and Spearman's ρ , along with standard deviations for LAD estimations for each of the 6 spatial resolutions using the R programming language.

Comparing LAI and Total Leaf Area Estimates

To compare LAD estimates across the landscape, we calculated the mean LAI and total leaf area across the study area (TLA, km^2) by taking the sum of all LAI estimates across the study area. We did this for only the 10x10m resolution, because it was the finest resolution that produced the most

stable results between sensors. Due to the NEON AOP data having a large number of pixels with no ground returns, and thus a larger number of NA values, our estimates are slightly biased when compared to NASA G-LiHT. We have left these NA values in the analysis to better represent the results one could achieve if only a single data set was available. To better understand the differences between sensor estimates, we calculated LAI and TLA for three height subsets: ground to TOC, 10 to 20 meters above the ground, and 20 meters above the ground to TOC. We calculated the mean difference in LAI across the study area and between each sensor by calculating the mean of absolute values at each pixel of NASA G-LiHT LAI minus NEON AOP LAI. We also calculated the difference in TLA between each sensor by taking the difference in TLA estimates from NASA G-LiHT and NEON AOP. We calculated the percent difference by subtracting the NEON AOP values from the G-LiHT values and dividing by G-LiHT values. While mean LAI and TLA yield similar results, particularly if LAI values are normally distributed, here we present both, as mean LAI represents the average expected value at the local scale, while TLA represents the aggregated leaf area across the landscape. While TLA across a landscape may seem overly coarse for studies of fine scale variation in forest structure, TLA measurements will connect to scales relevant to global land surface models and eddy covariance towers.

Results

Lidar Penetration of Forest Canopies

Lidar ground return counts increase with the coarsening of spatial resolution for both the NASA G-LiHT and NEON AOP platforms (Figure 4). However, there were notable differences between the two systems at finer spatial resolutions. At a 1x1 meter resolution within the study area, NASA G-LiHT had ground returns in 75.53% of the raster cells, while NEON AOP had 37.52%. While

differences persist, both sensors were within 10% of each other at a 10x10 meter resolution where NASA G-LiHT had 99.17% ground returns and NEON AOP had 90.98% ground returns.

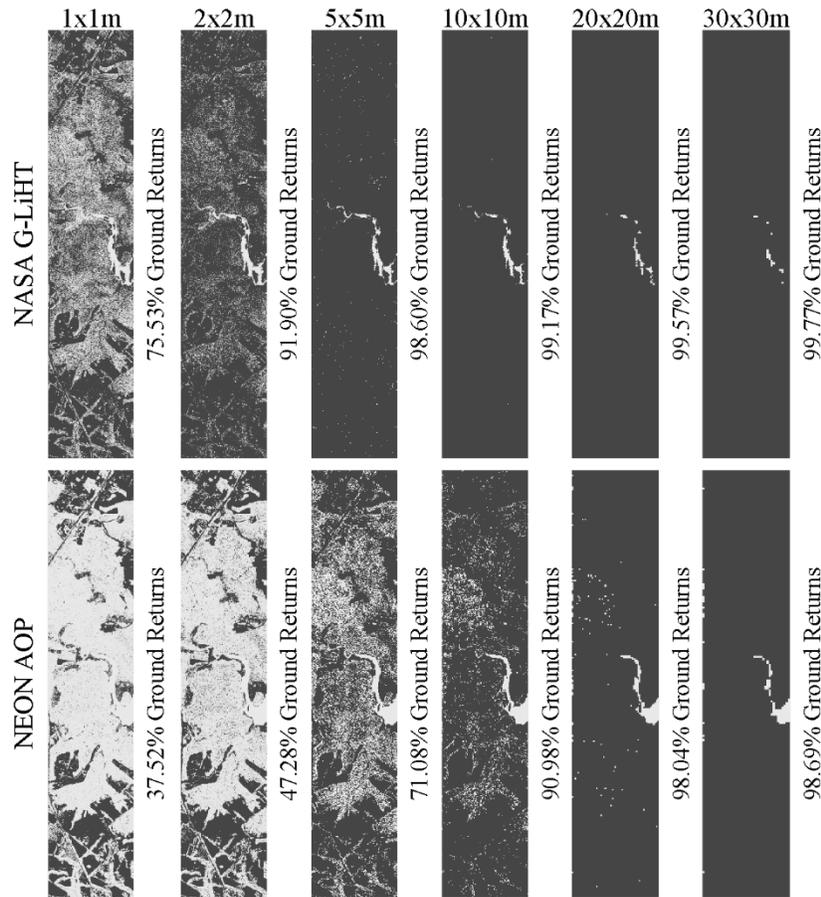


Figure 4. Lidar Spatial Resolutions. Six different spatial resolutions were used for this study, which are shown here. The percentage of ground returns for each spatial resolution are shown, which increase with the coarsening of spatial resolution. Black pixels are locations of ground returns and light gray pixels do not have ground returns.

These two airborne systems also exhibit differences in the depth of canopy penetration (Figure 5). Within NEON plot 45, a 20x20 meter NEON vegetation plot, NASA G-LiHT had 20,475 lidar returns while NEON AOP had 1,299, or 94% less returns than NASA G-LiHT. When we binned these returns by height, the 25th, 50th, 75th, and 90th percentiles occurred at similar heights above the ground. However, the 10th percentile of returns for NASA G-LiHT occurred

much deeper in the canopy at 11.5 meters above the ground, while the 10th percentile of returns for NEON AOP occurred at 20.3 meters above the ground. These findings highlight the differences in canopy penetration between the two sensors, as the distribution of lidar returns is skewed more deeply into the canopy in the G-LiHT data.

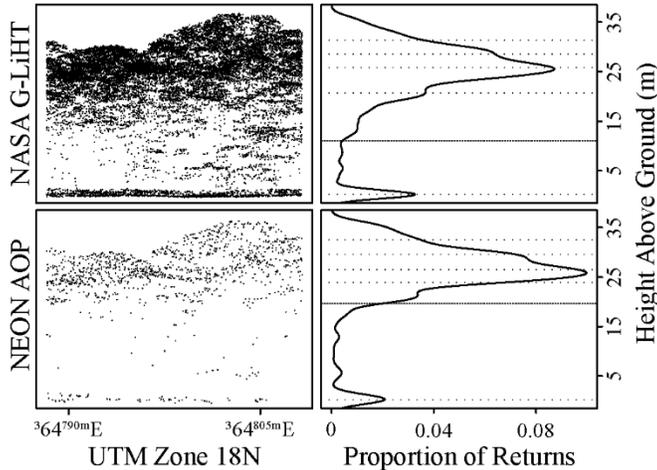


Figure 5. Lidar Within Canopy Returns. We used a permanent 20x20 meter NEON vegetation plot (NEON plot 45) to extract lidar point return information. The left plots show all lidar pulses within the plot, with NASA G-LiHT having 20,475 returns and the NEON AOP having 1,299 returns within the same plot area. The right plots show the proportion of returns at each meter above the ground. The dotted lines show the ground (0 meters), 10th percentile, 25th percentile, 50th percentile, 75th percentile, and 90th percentile. The 10th percentile is shown as a slightly darker dotted line to highlight the differences between sensors deep in the canopy.

Beer-Lambert Coefficients

Using the field-measured LAI together with the lidar LAI estimates, we calculated a broad range of Beer-Lambert coefficients at different spatial resolutions (Table 1). We found that Beer-Lambert coefficients decrease with the coarsening of spatial resolutions regardless of airborne system. These coefficients approach 0.5, which is commonly used in closed-canopy forest

ecosystems (Stark et al. 2012; Vose et al. 1995; Burton et al. 1991), at these coarser spatial resolutions.

Table 1. SERC Beer-Lambert Coefficients. Beer-Lambert coefficients for each lidar sensor at each spatial resolution based on relationships with hemispherical photographs.

	Beer-Lambert Coefficients					
	1x1	2x2	5x5	10x10	20x20	30x30
NASA G-LiHT	1.01	0.87	0.83	0.67	0.58	0.48
NEON AOP	1.34	1.12	0.85	0.68	0.63	0.56

LAD Profile Estimates

We observed a general increase in the agreement between NASA G-LiHT and the NEON AOP LAD values with coarsening spatial resolution in both sets of data (ground to TOC and 10 meters above the ground to TOC) (Figure 6). When the entire dataset is considered (row A, Figure 6), R^2 values increase and the 95% confidence interval becomes narrower as the spatial resolution becomes coarser while the line of best fit changes drastically based on the position of the spuriously large values in the lower canopy, as discussed in Section 2.6. However, while row B in Figure 6 (10 meters above ground to TOC) shows a similar loosely correlated relationship at finer spatial resolutions and R^2 values increasing as the spatial resolution becomes coarser, the line of best fit approaches the 1:1 line as the spatial resolution coarsens, signifying a stronger relationship between these two datasets. The tightening of this relationship begins to occur at a 10-meter spatial resolution ($R^2 = 0.87$), which is also the spatial resolution where over 90% of cells have a ground return value in both datasets.

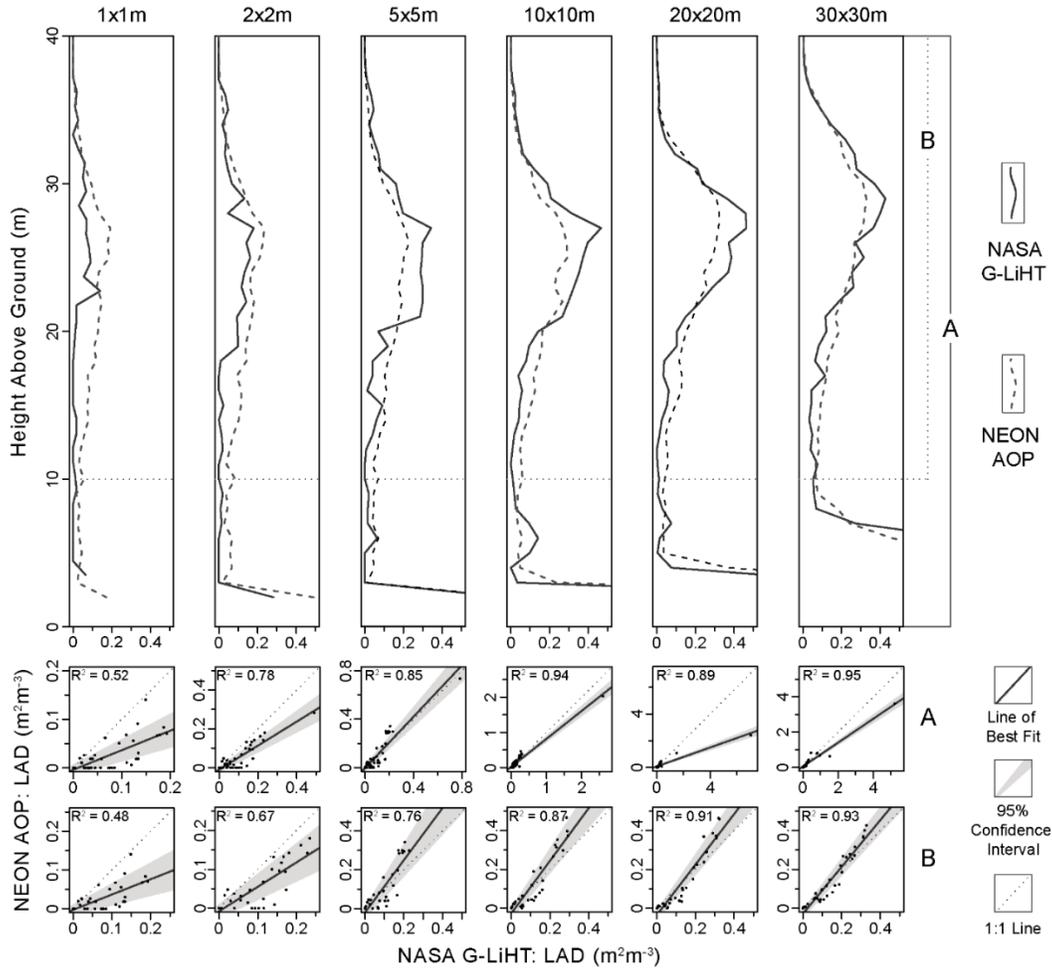


Figure 6. SERC LAD Profiles. All spatial resolutions considered are shown for NEON Plot 45 with NASA G-LiHT and NEON AOP LAD profiles in the first row of plots. LAD values were cut off at 0.5 for visualization, so that differences in the upper canopy can be seen. Plots in row A show the relationship between NASA G-LiHT and NEON AOP LAD data at each spatial resolution from the ground to TOC. Plots in row B, show the same relationship, but only including data from 10 meters above ground to TOC. All R^2 values are significant at $p < 0.001$.

To consider the broader landscape variation and patterns, we generated 50 random 20x20 plots and extracted the same data as in the case of NEON Plot 45. At these 50 randomly located plots, we observed the same general relationships described above when all the data from the ground to TOC were employed (Table 2) and when only the data from 10 meters above the ground to the TOC were employed (Table 3). These findings show that NEON Plot 45 is not an anomaly

and instead is representative of the relationships between these two lidar datasets across the landscape.

Table 2. SERC Whole Canopy Statistics. LAD profiles (for all voxels) for 50 random 20x20m plots from each spatial resolution and for each sensor (NASA and NEON) were extracted. For every plot at each spatial resolution and for each sensor, R^2 , RMSE, and slope were calculated from a linear regression. Spearman's ρ was also calculated. SD = standard deviation.

LAD Estimates: Ground to Top of Canopy												
	1x1		2x2		5x5		10x10		20x20		30x30	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
R2	0.66	0.27	0.76	0.19	0.85	0.17	0.89	0.15	0.95	0.07	0.93	0.10
RMSE	0.01	0.01	0.02	0.01	0.03	0.02	0.05	0.04	0.05	0.05	0.08	0.07
Slope	0.36	0.23	0.53	0.22	0.91	0.38	1.04	0.52	1.22	1.26	0.91	0.46
Spearman's ρ	0.83	0.13	0.90	0.08	0.95	0.04	0.97	0.05	0.97	0.05	0.98	0.03

Table 3. SERC Reduced Canopy Statistics. LAD profiles (for voxels 10 meters above the ground to TOC) for 50 random 20x20m plots from each spatial resolution and for each sensor (NASA and NEON) were extracted. For every plot at each spatial resolution and for each sensor, R^2 , RMSE, and slope were calculated from a linear regression. Spearman's ρ was also calculated. SD = standard deviation.

LAD Estimates: 10 Meters Above Ground to Top of Canopy												
	1x1		2x2		5x5		10x10		20x20		30x30	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
R2	0.55	0.22	0.67	0.17	0.83	0.13	0.90	0.05	0.93	0.05	0.91	0.08
RMSE	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.01
Slope	0.35	0.19	0.46	0.22	1.01	0.29	1.08	0.31	1.05	0.22	0.92	0.21
Spearman's ρ	0.76	0.18	0.80	0.18	0.93	0.05	0.94	0.11	0.98	0.02	0.96	0.06

Total Leaf Area Estimates

Across the entire study area, results reflect our findings at the plot level (Figure 7). From 20 meters to TOC, NEON AOP has slightly higher LAI and TLA estimates than NASA G-LiHT due to the point cloud being skewed towards the top of the canopy. Even with these differences, there was less than a 5% difference in TLA between the two sensors at this height interval. Between 10 and 20 meters above the ground, there was a 10% difference between in TLA between the sensors, with NASA G-LiHT having slightly higher estimates due to the point cloud being skewed lower in the canopy than NEON AOP. When 10 meters above the ground to TOC is considered, there is a 2% difference between TLA estimates between the two sensors. However, when the ground to top of canopy is considered there is a much larger difference between TLA estimates at 17%.

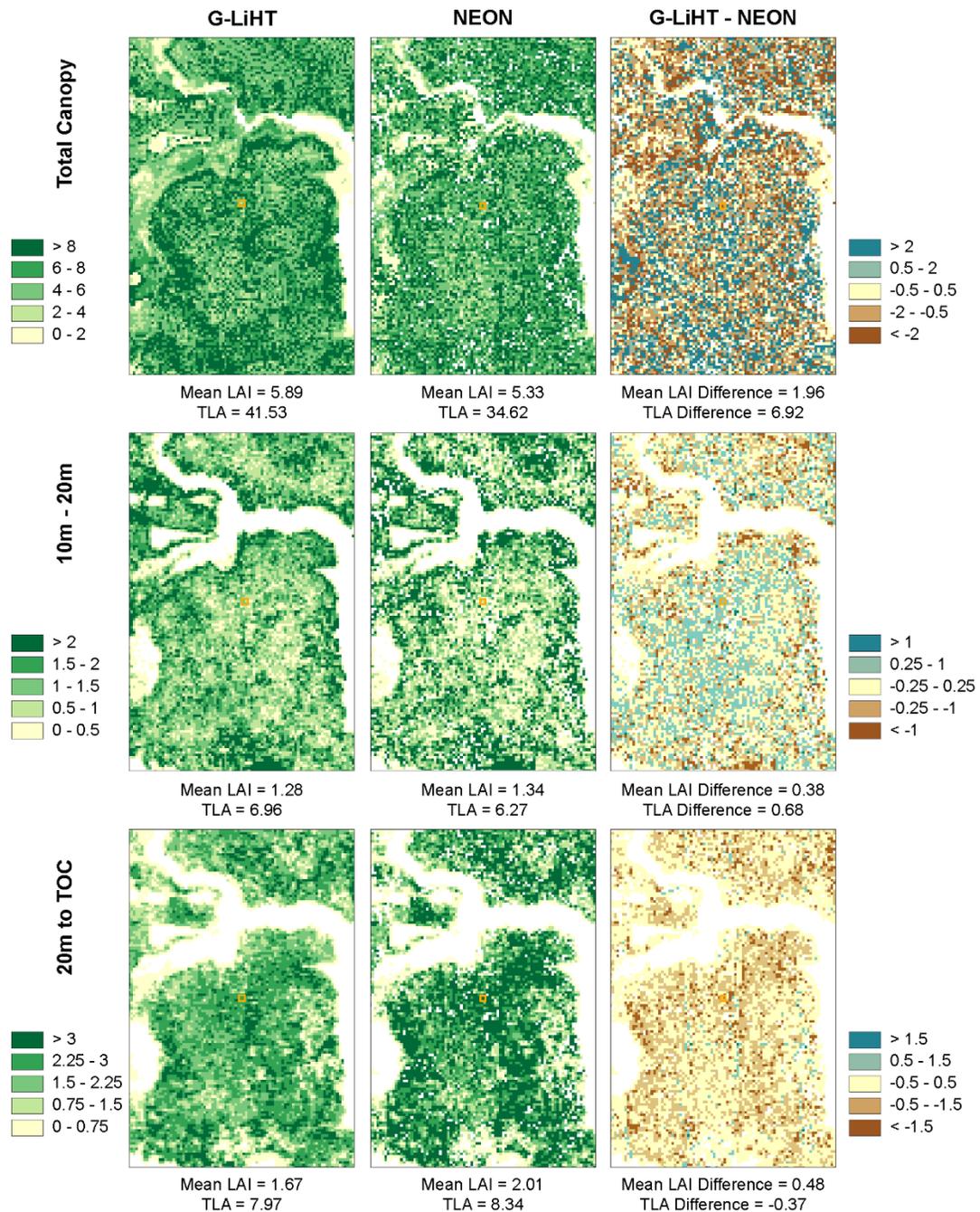


Figure 7. SERC LAD Spatial Patterns. Total leaf area (TLA; km²) was estimated at three different canopy height intervals (ground to TOC, 10m-20m, 20m to TOC, at the 10x10 meter resolution across the entire study area. A subset of the study area is shown here for visualization purposes, but LAI and TLA values are calculated from the entire study area. Differences between the G-

LiHT estimates and NEON estimates were also calculated. Mean LAI differences were calculated as the mean of the absolute values of the differences. NEON Plot 45 is shown as an orange square for reference. TLA is the total km² of leaf for the study area (7.2 km², solid line in Fig 1).

Discussion

Measuring Leaf Area Density from Above

While there are differences between these two airborne systems, our analysis can serve as a case study on how to estimate LAD at the appropriate spatial resolution for a given airborne lidar data set. Since these total canopy calculations are limited by the need to have a ground return in a given raster cell, this is often the most significant limiting factor in the estimation of LAD. The lack of ground returns can severely limit the spatial coverage of LAD estimates across the landscape, which can result in missing data within the study area, thus finding a balance between spatial resolution and spatial coverage is often the first step in these calculations. While it would be ideal to compare these datasets for other research sites, this is often not possible because only one of the datasets is available for a given area, thus the need to have comparable LAD estimates between sensors. By excluding voxels that lack ground returns we ensure that the entire canopy is accounted for in our LAD estimates, while providing forest structural estimates at an ecologically relevant scale (10x10 meters is similar to a canopy dominant tree crown) to help answer landscape- and macro-scale questions.

We have shown that even low pulse density NEON AOP lidar data can be used to successfully estimate LAD, but these measurements come at the expense of spatial resolution. LAD estimates from between sensors began to stabilize at around a 10-meter resolution; however, there is still the need to remove a portion of the understory for close agreement, likely due to errors related to topographic changes inflating LAD estimates near the ground. Such understory inflation

is particularly evident with NASA G-LiHT due to its high pulse density and canopy penetration. On the other hand, because of this high pulse density and canopy penetration LAD estimates from NASA G-LiHT can be calculated at finer spatial resolutions, allowing for the consideration of the full LAD profile. However, there is a tradeoff between high data density (NASA G-LiHT) and spatial extent and temporal coverage (NEON AOP). With NEON AOP flying the same sites on a semiannual basis, the ability to have yearly, landscape scale analyses might outweigh the need for finer resolutions of LAD estimates. While NASA G-LiHT has a higher pulse density and greater canopy penetration, which allows for finer resolution analysis, traditionally the flights have covered smaller areas, are not conducted on a yearly basis, and are project driven. That said, the G-LiHT archive is extensive with data in over 30 US states and territories and several countries at the time of this study (<https://gliht.gsfc.nasa.gov/>), highlighting the potential to generate broad spatial estimates across a range of vegetation types at fine spatial scales.

We show that the low-density point cloud from the NEON AOP can be used to estimate LAD within the forest canopy, with minimal differences (around 2%), as long as the inflated understory estimates are removed. While removing the understory of the canopy from the dataset is not ideal, the temporal and spatial coverage of the NEON AOP provide a unique opportunity to monitor forest ecosystems in ways that were not previously possible. Additionally, our LAD estimates per voxel from 10 meters above the ground to the top of canopy are within the ranges found by other research conducted at SERC (LAI of 4 to 7; LAD of 0.1 to 0.5) using field-based techniques (Parker and Tibbs 2004; Brown and Parker 1994). This offers additional evidence that we can obtain accurate LAD estimates from airborne lidar systems with differing parameters and from point clouds with varying degrees of density and canopy penetration. While we show that

these measurements are accurate in the dense forests of SERC, more research is needed in different biomes to further test the abilities of airborne lidar to estimate LAD across landscapes.

Lidar System Considerations

All lidar collections are not the same. Lidar sensor specifications (beam divergence, scan angle, etc.) have large impacts on the density and quality of the data and are tuned to the specific data collection goal. These specifications are important to consider before processing the data and can help the researcher determine how to best use the data. Specifications such as scan angle determine how LAI estimates from hemispherical photographs need to be constrained during processing (see Section 2.2), while beam divergence and pulse repetition frequency can help determine the quality of the point cloud and canopy penetration. Likewise, differences in the wavelength the lidar sensor operates at can affect how pulses are reflected within the canopy. For instance, leaves typically have a higher reflectance at 1064 nm (NEON AOP) than at 1550 nm (NASA G-LiHT), while bark has a higher reflectance at 1550 nm; this could lead to bark and branches having a slightly higher impact on returned pulses for NASA G-LiHT and leaves having a slightly higher impact for the NEON AOP. While the extent of the impact due to these differences would be difficult to quantify without ray tracing and a well-defined architecture, this would most likely lead to slightly different point clouds if all other variables were held constant. Examining the underlying metadata and understanding what the goals of the data collection mission are can help determine how the resulting lidar data can be used for a specific research project. For example, we have shown here that NEON AOP lidar data are adequate for measuring LAD in the forest canopy at a 10-meter resolution, but not for detecting variations in the understory if the understory is dense.

Ecological Implications

While it would be ideal to use a spatial resolution that mimics the fine scale variation found within a forest canopy (e.g., 1x1 meters), this may not always be possible due to data availability.

Since the lidar data available for ecological studies differs from site to site, it is challenging to compare studies and combine analyses over such heterogeneous collections. By developing standardized approaches for LAD and LAI estimation that are accurate and consistent regardless of the sensor used, analyses can more easily compare multiple studies while encompassing varied data sources, resulting in an opportunity for robust quantitative comparison and hypothesis testing. We have shown that LAD and LAI estimates at 10x10 meters, that are fine-tuned with hemispherical photographs, are in line with field-based measurements across two very different airborne lidar systems. Thus, we propose a resolution of 10x10 meters to estimate LAD and LAI, with inflated understory LAD estimates removed, as a viable standard resolution for landscape to macro-scale studies that use lidar data collected with lower pulse densities (e.g., less than 20 pulses per m²). While a 10x10 meter resolution will not be fine enough to investigate leaf level processes or the structural components of individual trees, airborne lidar with moderate to low pulse densities is still well situated for the investigation of landscape to macro-scale trends. When higher pulse densities are available from airborne, ground, and drone based lidar systems, there is the potential to model biophysical processes occurring at the leaf level (Wu et al. 2018), to investigate the role of fine-scale heterogeneity on canopy function (Atkins et al. 2018), and to consider the structural components of individual trees (Hosoi and Omasa 2006). As these types of high-density lidar data become more readily available, additional detailed analyses will be needed to quantify structural and functional processes at these finer scales.

Looking Forward

With lidar data becoming more readily available, it is important to consider the end user and their needs. Airborne and spaceborne platforms like NASA G-LiHT, NEON AOP, and GEDI are collecting and will continue to collect a large catalog of lidar data across a variety of ecoregions, allowing researchers the opportunity to ask and answer new questions about forest structure at large spatial scales. To support these new lines of research, we present a reproducible workflow and encourage other researchers to do the same, so that the scientific community as a whole can use these data in a consistent and standardized manner. While there are many other approaches to estimating LAD from airborne lidar (e.g., McNeil et al. 2016; Detto et al. 2015), we have shown that our methodology produces accurate estimates that are based on well-established field-based methodologies. With this large influx of data, we have a unique opportunity to not only use lidar data in new ways, but also to incorporate the resulting products into research projects that may have never considered using lidar data previously.

Conclusion

Lidar has become a common data type in the remote sensing community and with this large influx of data, there are many unique opportunities to incorporate it into different ecological studies. Here we have presented a reproducible methodology to produce LAD and LAI estimates from airborne lidar with R code available for other researchers to use. We also highlight the importance of airborne lidar survey parameters that dictate pulse return density and ultimately determine the coverage of LAD and LAI estimates within survey areas, while providing ideas on how this data can be used in ecological and forest studies. Furthermore, we show that a spatial resolution of 10x10 meters can successfully estimate LAD with either of these two moderate to low pulse

density airborne sensors. While lidar data has been used to inform management and conservation decisions related to the estimation of aboveground biomass and productivity (Hughes et al. 2018; Socha et al. 2017), the response of forests to large-scale disturbances (Hoffman et al. 2018), the impacts of drought on forest health (Paz-Kagan et al. 2018), and the conservation of biodiversity (Garabedian et al. 2017; Mao et al. 2018), these methodologies can be difficult to reproduce. To help bridge this gap between ecologists and forest managers who could use lidar data in research and management plans and remote sensing scientists who use this data on a regular basis, we provide an open source, reproducible, and standardized workflow to calculate LAD and LAI from airborne lidar data.

Acknowledgments

Thanks to Logan Brissette for field assistance and to the Smithsonian Environmental Research Center, especially Geoffrey Parker, Patrick Megonigal, and Sean McMahon, for providing site access and space. This work was supported in part by the NSF Macrosystem Biology Program award #1702379. The National Ecological Observatory Network is a program sponsored by the National Science Foundation and operated under cooperative agreement by Battelle Memorial Institute. NASA's G-LiHT is a program sponsored by the NASA Goddard Space Flight Center's Internal Research and Development program and NASA's Terrestrial Ecology, Carbon Cycle and Carbon Monitoring System programs. This material is based in part upon work supported by the National Science Foundation through the NEON Program. Shawn P. Serbin was partially supported by the United States Department of Energy contract No. DE-SC0012704 to Brookhaven National Laboratory.

Data Availability

Lidar point clouds are available at: <http://data.neonscience.org> and <https://glihtdata.gsfc.nasa.gov>.

R package to estimate LAD and LAI from airborne lidar data is provided through GitHub at: <https://github.com/akamoske/canopyLazR>. Hemispherical photographs and shapefile with locations can be found on figshare at: <https://doi.org/10.6084/m9.figshare.6955142.v1>.

CHAPTER 3. LEAF TRAITS AND CANOPY STRUCTURE TOGETHER EXPLAIN CANOPY FUNCTIONAL DIVERSITY: AN AIRBORNE REMOTE SENSING APPROACH

Citation: Kamoske, A.G., K.M. Dahlin, S.P. Serbin, and S.C. Stark. 2020. Leaf Traits and Canopy Structure Together Explain Canopy Functional Diversity: An Airborne Remote Sensing Approach. Ecological Applications e2230.

Introduction

The relationship between forest structure and function is a major focus of ecosystem ecology; however, most studies have focused on measurements within traditional forest plots (Ellsworth and Reich 1993; Parker et al. 2004; Gough et al. 2019; Atkins et al. 2018; Fahey et al. 2015; Pedro et al. 2017). These studies have shown that the integral relationship between structure and function drives important canopy processes such as net photosynthetic carbon assimilation (Niinemets 2007), resource use and efficiency (Hardiman et al. 2013), and woody growth (Stark et al. 2012), as well as critical ecosystem processes such as net primary production (Scheuermann et al. 2018; Hardiman et al. 2011). Since the individual traits that drive this structure-function relationship are not constant in space and instead show significant heterogeneity across landscapes (Chambers et al. 2007; Asner et al. 2014), a core question in ecosystem ecology is: Do landscape scale patterns of forest functional traits change when whole plant structure is considered?

In addition to this significant spatial variation, plant functional and structural traits also vary in three-dimensional space due to a host of different long-term abiotic growth conditions, crown position within the canopy and competition for light, as well as within-canopy fluctuating light environments across the full vertical and horizontal extent of the canopy (the ‘canopy volume’; Ellsworth et al. 1993). These differing light and growth environments drive variation of important leaf functional traits including leaf mass per area (LMA; the ratio between leaf dry mass

and leaf area) and foliar nitrogen (foliar N; g/m_G^2 ; m_G = meter of ground) within the canopy volume (Poorter et al. 2009, Niinemets, 2007). Moreover, horizontal and vertical patterns of these traits in growth environments create heterogeneous distributions of leaves in three dimensions causing significant variation in canopy-scale carbon assimilation across plant functional types (Niinemets 2015). This variation can be attributed to differing light environments related to the effects of multiple scattering, within-canopy shading, and the density of plant material above and around a given leaf (Stark et al. 2012; Harding et al. 2001). Resulting tradeoffs between light interception, photosynthetic capacity, and construction costs (e.g., the leaf economics spectrum) leads to broadly predictable variation in photosynthetic strategies across the plant species comprising global terrestrial biomes (Reich et al. 1997; Wright et al. 2004).

At the canopy scale, forest structural traits can be used to describe the architectural properties that define a leaf's growth environment. These traits include leaf area density (LAD; the total leaf area per unit of volume) which characterizes the horizontal and vertical spatial variation of leaf area within a canopy (Weiss et al. 2004) and canopy clumping (a measure of foliage aggregation relative to a random spatial distribution of leaf material within the canopy; Pisek et al. 2018). When combined with information on plant function, structural diversity yields important insights into vegetation growth and carbon cycling (Niinemets 2012), however both can be challenging to quantify at scales larger than vegetation plot without advanced remote sensing technologies (Asner and Martin 2009).

Remote sensing has played a significant role in understanding the global terrestrial carbon cycle for decades (Tucker and Sellers 1986; Schimel 1995; Running et al. 2004; Schimel et al. 2015), with a more recent focus on the use of hyperspectral imagery and lidar to measure forest function and structure. By utilizing hundreds of narrow spectral bands, airborne passive optical

hyperspectral imagery (HSI; also known as imaging spectroscopy) provides detailed two-dimensional (2D) information on the spectral and functional properties of leaves at the top of the canopy (Ollinger et al. 2002; Townsend et al. 2003; Asner et al. 2015; Singh et al. 2015; Dahlin et al. 2013). Lidar is an active remote sensing system that utilizes laser pulses to measure distance, which can then be used to accurately estimate the three dimensional (3D) and internal structure of forest canopies across a range of plants in different biomes (Stark et al. 2012; Kamoske et al. 2019; Shao et al. 2019; Smith et al. 2019). While passive optical data can also be used to estimate variables related to forest structure, including clumping index (Pisek et al. 2018) and 3D point clouds through structure-from-motion methods (Dandois et al. 2013; Iglhaut et al. 2019), the results are not as robust as active methods like lidar for generating 3D plant information. Compared to HSI data, lidar can yield detailed insights into plant architecture but does not provide the information necessary to map leaf functional traits across space and time, a combination of these data sources is required to provide a complete picture of vegetation structural and functional diversity. However, few publicly available opportunities and platforms exist for the simultaneous collections of these two complementary technologies (Cook et al. 2013, Kampe et al. 2010), limiting our ability to combine landscape-scale information about forest structural and functional traits that play critical roles in whole-canopy processes like carbon assimilation.

In this study, we take steps towards addressing the question of how leaf traits and structural heterogeneity determine whole canopy function by considering how spatial patterns of top-of-canopy and total canopy traits vary across a heterogeneous landscape. We detail a reproducible methodology for estimating functional and structural diversity within the canopy volume from airborne lidar and hyperspectral data from the National Ecological Observatory Network's Airborne Observation Platform (NEON AOP; Kampe et al. 2010). We compare the spatial patterns

of 3D whole canopy traits derived from our fusion of lidar and hyperspectral data with traditional 2D remote sensing derived top-of-canopy traits. In addition, we examine the influence of topography, geology, and management regimes on these two measurements of functional diversity at a NEON site consisting of patches of open longleaf pine and dense broadleaf deciduous forests, located in Alabama, USA. These insights could lead to a better understanding of how we scale fine-resolution ecological processes to landscape, continental, and global models (Schimel et al. 2019).

Materials and Methods

Site Description

Field measurements and remote sensing data were acquired in Talladega National Forest – Oakmulgee Ranger District (TALL) in west-central Alabama, USA (Figure 8). TALL is a core NEON site covering 5,300 hectares with a mean annual temperature of 17° C and a mean annual precipitation of 1350 mm. TALL consists of a mosaic of forest types, with higher elevation areas containing an overstory of longleaf pine (*Pinus palustris*) and loblolly pine (*Pinus taeda*), while white oak (*Quercus alba*), Southern red oak (*Quercus falcata*), chestnut oak (*Quercus montana*), blackjack oak (*Quercus marilandica*), mockernut hickory (*Carya tomentosa*), pignut hickory (*Carya glabra*), sweetgum (*Liquidambar styraciflua*), and tulip tree (*Liriodendron tulipifera*) are present in lower elevation bottomlands. TALL is an actively managed site with ongoing logging, restoration, and prescribed burning projects (USDA Forest Service 2005).

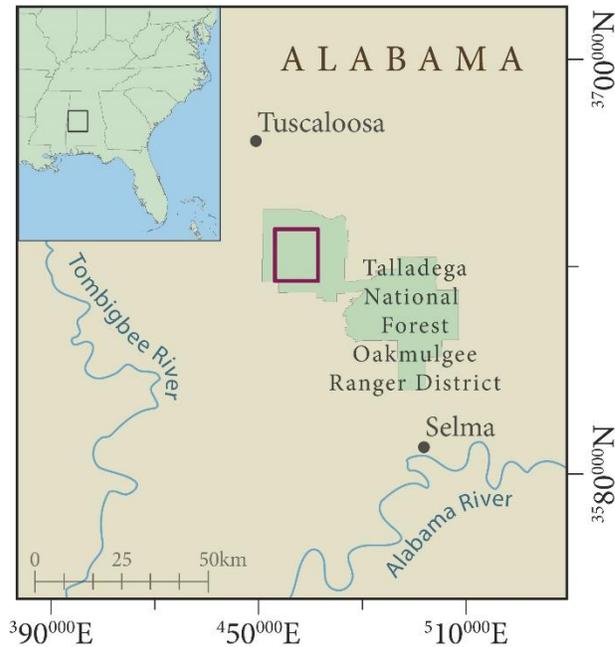


Figure 8. Map of Study Area at TALL. Location of field site. Purple rectangle represents the extent of the aerial data collection of the NEON AOP. Inset map shows the extent of the larger map view within the southeastern United States.

Airborne Remote Sensing Data

The NEON AOP collected remotely sensed data from April 27 to April 29, 2018 at TALL. The NEON AOP employs a full-range hyperspectral sensor (380 to 2500 nm; 5 nm bands), a high-resolution RGB camera, and a lidar system (Kampe et al. 2010). Flights occurred at an altitude of 1000m, resulting in hyperspectral measurements at a one-meter resolution. The lidar system for this collection was a Riegl Q780 Laser Measurement System operated at a scan angle of +/- 18 degrees, and a beam divergence of 0.8 mRad, resulting in an average point density of 9.48 pts/m².

Field Data Collection and Lab Methodologies

In May 2018, shortly after the AOP collection, we collected leaves from throughout the canopy volume, targeting the dominant species at TALL (10 species total; listed in section 2.1). Foliar samples were collected using a Big Shot line launcher (SherrillTree, Greensboro, NC) and a pole pruner, with each sample's height estimated using a laser range finder and meter marks on the set line. We collected sample locations using a Trimble GEO7x GPS (Trimble, Sunnyvale, CA), which were later differentially corrected with Trimble's GPS Pathfinder Office software. As we collected samples from the canopy, they were wrapped in a damp paper towel, sealed in a plastic bag, and placed in a cooler with ice packs. In total we collected 156 foliar samples from the canopy dominant species (Appendix A.1 and Appendix A.2). In addition to leaf samples, we took 120 hemispherical photographs across the site, following the protocol described in Kamoske et al. (2019).

Leaf samples were processed the same day in our mobile laboratory. For each sample (a small branch with multiple leaves) we took three reflectance measurements from different leaves with a SVC HR-1024i Spectroradiometer with an attached LC-RP-Pro leaf clip foreoptic (Spectra Vista Corporation, Poughkeepsie, NY), which collects data from 340 to 2500 nm with a bandwidth of approximately 2 nm. Leaves from broadleaf samples were placed directly into the leaf clip, while we created mats from needleleaf samples by laying the needles vertically next to one another while taping the ends together. For needleleaf samples, only the needles and not the taped ends were placed into the leaf clip. After each sample, the instrument was recalibrated using a white Spectralon panel. We then collected a minimum of 500 mg of leaf material from the sample using a pair of scissors that were sterilized between each sample. These pieces of leaf material were imaged on a flatbed scanner and processed for area using imageJ software (Schneider et al. 2012).

We placed the leaf material in a paper coin envelope and dried the samples at 70° C for at least 48 hours. After drying, we weighed the leaf samples and calculated leaf mass per leaf area (LMA; g/m_L^{-2} ; m_L = meter of leaf material). A subset of these samples ($n = 40$, ~4 per species) were re-dried, ground to a fine powder using a ball mill (2000 Geno Grinder; Spex Sample Prep, Cridersville, OH, USA), with 1.50-2.50 mg weighed in 0.1-mil tin foil vials (AX26DR; Mettler Toledo, Columbus, OH, USA), and used to determine the C:N ratio and elemental N content (g N/g leaf, \%) employing a CHNS/O elemental analyzer operated in CHN mode, according to the manufacturer's instructions (2400 Series II CHNS/O Analyzer; Perkin Elmer, Waltham, MA, USA) at Brookhaven National Laboratory (Upton, NY).

To build a leaf-scale model of %N to apply to the remaining samples in lieu of determining foliar N in the lab, we used the laboratory calculated %N values and the associated mean reflectance values for each wavelength, to train a partial least squares regression model (PLSR; Serbin et al. 2014; Singh et al. 2015). We withheld 20% of the samples using a weighted random approach, based on the %N values, as validation data ($n = 8$) that wasn't used to develop the model and used the remaining samples ($n = 32$) as model training data. Using a jackknife approach that randomly withholds 20% of the training data through 50 iterations, we calculated a PRESS statistic (up to 15 components) for each iteration. We then selected the number of components for our final model using the lowest PRESS statistic that balanced predictive accuracy between the training and validation datasets. We applied these equations to the validation data to assess model accuracy. We then applied the final PLSR coefficients to the reflectance measurements of all 156 leaf samples to determine PLSR derived %N values. We used the PLSR predicted values in subsequent analysis. This methodology follows the process and code described in Serbin et al. (2014), with all analysis performed in R using the pls package (Mevik and Wehrens 2015).

Lidar Methods

Lidar data was processed for LAD (m_L^2/m_G^3 ; m_G = meter of ground) at a 10x10 meter spatial resolution using the canopyLazR package on GitHub (Kamoske et al. 2019). The canopyLazR package uses the methods described by MacArthur and Horn (1969) and is similar to other published methods (Stark et al. 2012; Zhao and Popescu 2009; Solberg et al. 2006; Sumida et al. 2009). By normalizing the point cloud to height above ground, LAD is calculated by counting the number of lidar pulses that enter and exit each voxel in each vertical column of data that has at least one ground return. After removing the bottom 10 meters of the canopy due to noise caused by topographic variation (Kamoske et al. 2019), a stack of rasters containing LAD estimates for each 1-meter slice of the canopy above this threshold is returned (mean canopy height at TALL is 25 meters). LAI is then calculated by taking the sum of LAD values within a given column of voxels within the canopy. While the TALL lidar data set has a considerably higher point density than the NEON lidar data used in Kamoske et al (2019), here we elected to keep this relatively conservative approach to aggregating and filtering these data as these lidar point clouds were processed as part of a larger study where we wanted to maintain data uniformity across sites. Moreover, topographic issues have been shown to be common when using lidar data for DEM generation (Bater and Coops 2009), which are further amplified when using low-density lidar data. To calibrate the lidar derived LAI estimates to field collected data, we processed field-collected hemispherical photographs for LAI using the DHP software (Leblanc et al. 2005). We then calculated the slope of a regression equation between these measurements and the lidar derived LAI estimates (Appendix A.8; Sabol et al. 2014; Richardson et al. 2009). This slope is used as an extinction coefficient in the Beer-Lambert portion of the LAD equation described in Kamoske et al. (2019) and in Appendix A.8. For TALL we used an extinction coefficient of 0.4982. Here we

opted to use a single extinction coefficient for the entire site, rather than separate coefficients for broadleaf, needleleaf, and mixed species pixels due to difficulties in detecting species differences with lidar data.

Based on our previous work in Kamoske et al. (2019), we then applied a canopy height and LAI mask to each processed LAD raster to minimize noise in the lidar dataset. Using Tukey's outlier test ($k = 1.5$), we removed all outliers from the upper end of the dataset, which resulted in all pixels with a canopy height greater than 44 meters being removed as well as all pixels with a LAI value greater than 6 (0.002% of pixels). While a LAI value of 6 is a statistical output, it is also greater than our highest field-collected plot-scale LAI value of 4.35. We also removed all pixels with a LAI value equal to 0. Using these masked LAD tiles, we calculated 26 lidar derived forest structural attributes in raster format at a 10x10 meter resolution. These include filled canopy volume, canopy porosity, and canopy distribution metrics described in Hardiman et al. (2013), top-of-canopy rugosity, and canopy euphotic, oligophotic, and empty zone metrics described in Lefsky et al. (1999), canopy height metrics described in Shi et al. (2018), and within canopy rugosity described in Hardiman et al. (2011). All code to calculate these metrics is provided in the canopyLazR package on our GitHub page (<https://github.com/akamoske/canopyLazR>; <http://doi.org/10.5281/zenodo.3987340>). An overall diagram of our workflow is shown in Figure 9.

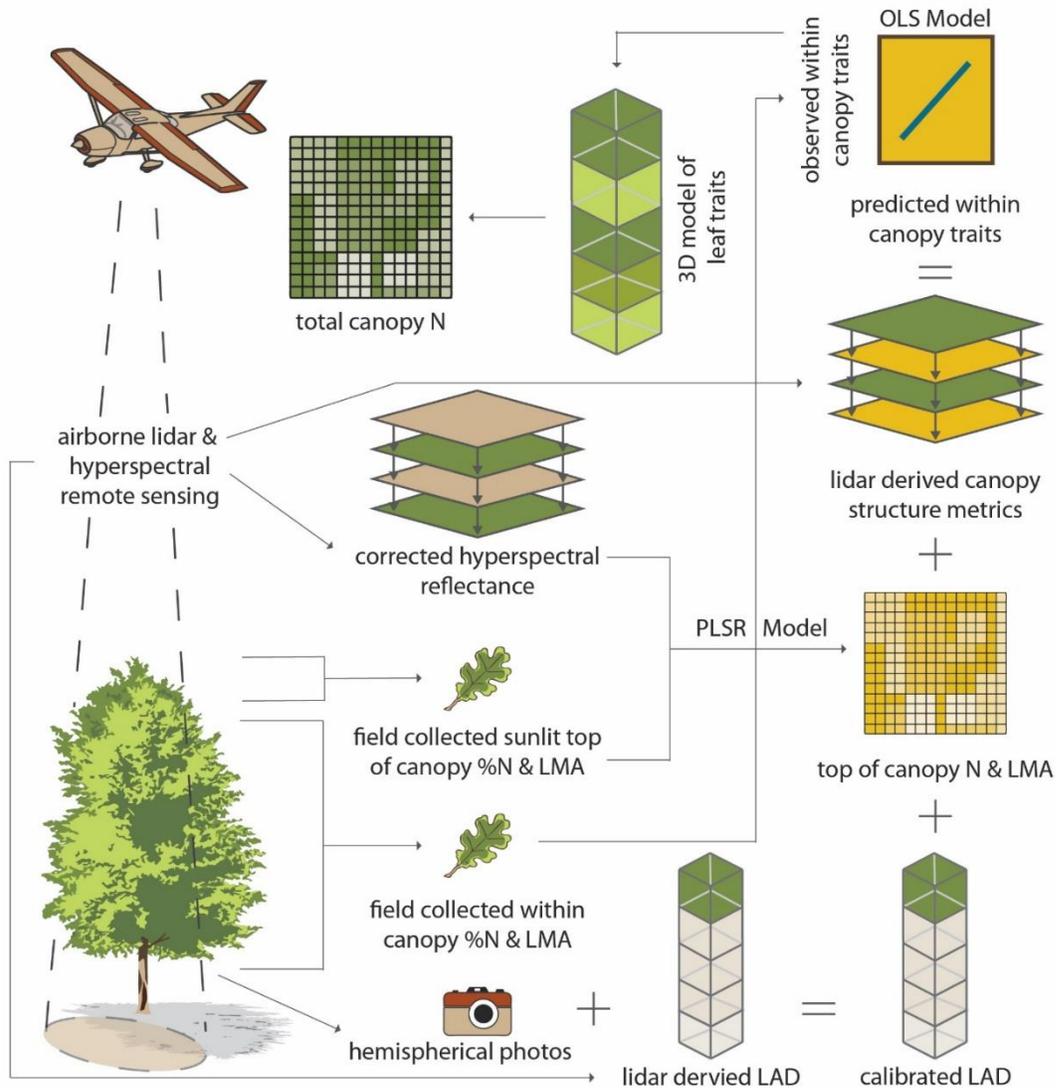


Figure 9. Total Canopy Modeling Methodology. Workflow diagram showing our methodology for within canopy trait modeling. LAD = leaf area density (m_L^2/m_G^3), LMA = leaf mass per area (g/m_L^2), N = foliar nitrogen content (g N/g leaf %), total canopy N = total canopy nitrogen content (g/m^2). Field collected sunlit top-of-canopy %N & LMA refers to leaf samples that were collected at the top of the canopy, were constantly sunlit, and had no leaves above (i.e., no sun impediment). Field collected within canopy %N & LMA refer to leaf samples that were collected within the canopy (i.e., not constantly sunlit, shaded, and with other leaves surrounding them).

Hyperspectral Imagery Methods

We processed the atmospherically corrected, HSI reflectance data before analysis. First, we removed all flight lines from April 27 due to cloudiness, as well as the horizontal (east-west) flight lines from April 29 and April 30. The remaining north-south flight lines covered the entire TALL site (April 29 and April 30 flights covered the same area as the April 27 flights). Next, we visually identified noisy bands in the dataset and removed all bands that were below 500 nm, between 1350 and 1450 nm, between 1800 and 2000 nm, and all bands above 2400 nm. We then calculated a narrowband NDVI mask (red = 674 nm; NIR = 830 nm; $NDVI > 0.5$) to remove all non-vegetated pixels from further analysis (Dahlin et al. 2014). We used this relatively high NDVI value of 0.5 in order to leave only healthy green vegetated pixels during the subsequent corrections and analysis. We also calculated a brightness mask to remove all shaded pixels using Tukey's outlier test ($k = 1.5$), where all pixels that have a reflectance below this cutoff at 800 nm are considered outliers and removed. This is a modified version of the methodologies presented by Clark et al. (2005) and Gougeon (1995), which removes all pixels that are less than the mean reflectance value at 800nm. Following this, we applied a topographic correction to reduce the effects of terrain, view, and illumination on the reflectance data by normalizing the sunlit area within a pixel without changing the sun and sensor positions or the orientation, geometry, and structure of the canopy while also accounting for diffuse radiation (Soenen et al. 2005). Lastly, we applied a bidirectional reflectance distribution function effects correction (BRDF) with a thick Ross kernel and a dense Li kernel to remove the anisotropic scattering properties of vegetation that result in flight line artifacts (Colgan et al. 2012; Collings et al. 2010; Schlapfer et al. 2015; Wanner et al. 1995; Weyermann et al. 2015). Annotated R code to apply these corrections is available on our GitHub page as the `hypRspec` package

(<https://github.com/akamoske/hypRspec>; <https://zenodo.org/record/3987336>).

From the resulting images, we extracted reflectance data for all top of canopy field samples. Due to potential image orthorectification errors, GPS uncertainty, and field challenges, we visually assessed GPS point locations and, when necessary, moved the GPS locations, by hand, 1-2 meters to the most appropriate pixel based on a canopy height model and pixel brightness. Due to flight line overlap, many samples had multiple reflectance values. In these cases, we kept the reflectance data from whichever image produced the brightest total reflectance across all bands. We choose to take the brightest reflectance value rather than the median here, in order to filter pixels that were possibly affected by collection issues related to adverse weather conditions that would not be resolved during the topographic and BRDF correction process.

Once reflectance spectra for all top of canopy samples ($n = 52$) were extracted, we developed PLSR models for top-of-canopy %N and LMA (Ollinger et al. 2002; Townsend et al. 2003; Singh et al. 2015) using the same methodology and code described for the laboratory data. For the LMA model, we removed all lab measured LMA values that were greater than 259 g/m^2 based on the results from a Tukey's extreme outlier test ($k = 3$). This outlier test removed 6 samples from the dataset. We removed these outliers from the dataset prior to fitting our models, due to PLSR being sensitive to outliers during the calibration and validation process (Martens and Martens 2000). Once PLSR coefficients were calculated for top-of-canopy LMA and %N, we applied them to the corrected HSI data, resulting in a 1x1 meter raster for each trait (%N and LMA). We then filtered the trait maps to remove all extreme outlier pixels ($k = 3$) and values less than 0 from each 1x1 meter raster that result from the errors associated with reflectance values collected during image collection. This resulted in 0.09% of the pixels being removed from the final raster. Next, we resampled the mosaicked image to a 10x10 meter spatial resolution using the

mean value within a given kernel, to match the spatial resolution of the lidar derived rasters. Following this, we mosaicked the flight line rasters together with the mean of overlapping pixels used in the final raster. All analysis was performed in the R programming language and is available on our GitHub page as the hypRspec package (<https://github.com/akamoske/hypRspec> <https://zenodo.org/record/3987336>).

Remote Sensing Fusion: Total Canopy N

To model within canopy LMA, we extracted data from the 26 previously calculated lidar structural attribute rasters, and top-of-canopy %N and LMA rasters, for all 156-field sample locations. We also included the height and depth (e.g., distance from the top of canopy) for each of the samples in the model. We then removed all top-of-canopy samples ($n = 52$) since these were used in previous steps and were predicted using the HSI data and PLSR. We then tested the correlation (Pearson's R) between each variable and within canopy LMA. To avoid multicollinearity, variables with correlations greater than 0.5 to each other were considered too correlated and the predictor most correlated with LMA was kept for further analysis. We then split the dataset into validation data (20%; $n = 20$) and training data (80%; $n = 84$) using a weighted approach based on species sample counts. Using the previously determined variables we developed an ordinary least squares (OLS) regression model from the training data. To determine the best combination of variables for our final model predicting within canopy LMA, we used backwards stepwise AIC model selection (Burnham et al. 2011; Mascaro et al. 2011). We then applied the resulting coefficients to the validation dataset to examine the overall predictive accuracy of our model. Because we did not see a substantial variation of within canopy %N in our data (Appendix A.1)

or in the literature (Serbin et al. 2014; Bachofen et al. 2020), we used top-of-canopy %N values for our within canopy %N values in lieu of creating another predictive model.

We then applied the final model coefficients to the raster data to create a three-dimensional model of within canopy LMA (g/mL^2), with any value less than zero set to NA (due to predictive inaccuracy and noise in the raster data). Lastly, we used these three-dimensional models to calculate within canopy N per meter of ground area (g/mG^2 ; mG = meter of ground) using the following equation:

$$N_{tot} = \sum_{i=10}^h N_{TOC} * LMA_i * LAD_i$$

where N_{tot} is the total canopy N (g/mG^2) for each 10x10 meter pixel, i refers to each 1 m layer of the canopy, starting at 10 m (layers below 10 m were not considered in this analysis), h is the maximum height of each column of voxels, N_{TOC} is the top-of-canopy N (%), LMA_i is the LMA at each voxel i (g/mL^2) and LAD_i is the LAD at each voxel i . This resulted in a two-dimensional raster for the entire AOP collection area that summarizes functional and structural traits within the canopy volume. We also calculated foliar biomass using the same equation described above but withholding the N_{TOC} values. Lastly, we removed all extreme outliers from the raster images using Tukey's outlier test ($k = 3$). All analysis was performed in the R programming language.

Raster Differences Across Scales

To test whether the distinction between leaf-level and canopy traits was scale dependent, we tested the differences between the top-of-canopy and total canopy N rasters at multiple spatial grains. First, we scaled the original 10x10 meter data to 30x30 and 250x250 meter resolutions to match Landsat and MODIS pixels using the raster package in R (Hijmans 2019). Next, we randomly

extracted 10,000 points from the 10x10 m and 30x30 m rasters and 1,000 points from the 250x250 m raster. We then used a linear regression to test the correlations between the two rasters at each spatial resolution. To compare the spatial patterns of the two rasters, we scaled and centered the rasters using the scale function in the raster package and then subtracted the normalized total canopy N raster from the normalized top-of-canopy %N raster.

To compare the overall spatial patterns of the two maps, we extracted 10,000 random points from the top-of-canopy and total canopy rasters at the 10x10 m resolution and fit variograms to these samples. We compared estimates of spatial autocorrelation as well as differences in the nugget, sill, and range of the variograms.

Environmental Driver Analysis

To understand the influence of abiotic gradients and management practices on the spatial patterns of top-of-canopy %N and total canopy N (g/mG^2), we assessed and analyzed the spatial patterns of the data, using multiple regression and Moran's I to test these relationships.

To quantify the abiotic gradients and management practices, we calculated 26 topographic, geologic, and management variables using ArcGIS, QGIS, and R (Appendix A.7). Topographic variables were calculated from the 10x10 meter lidar data, geologic variables were downloaded from the USGS (Horton 2017), and management variables were downloaded from the US Forest Service (<https://data.fs.usda.gov/geodata/edw/datasets.php>). All variables were transformed into rasters for subsequent analysis.

We performed a Monte Carlo test with 1,000 simulations to calculate a distribution of model coefficients, Moran's I of the residuals, and R^2 . During each simulation, we extracted 10,000 random points from the rasters. We then standardized all non-binary variables (Gelman

2007; mean = 0, standard deviation = 0.5) to allow direct comparison between model coefficients. We developed two regression models, one for top-of-canopy %N and one for total canopy N (g/mG^2). For each simulation and for each regression model we used the following methodology. First, we tested the correlation between each variable (Pearson's R) to avoid multicollinearity, with correlations greater than 0.5 considered to be too correlated and the predictor most correlated with N kept for further analysis. Using the remaining variables, we developed an OLS regression equation. With these results, we used backwards stepwise AIC model selection to determine the best combination of variables for each of our final models. Any remaining variables with non-significant coefficients ($p\text{-value} > 0.05$) were then removed. We then used these variables in a final OLS regression. To test for spatial autocorrelation, we calculated Moran's I on the model residuals. All analysis was performed with the R programming language.

Results

Trait Prediction with PLSR: From Leaf to Canopy

To predict leaf level %N, we used a PLSR model with five components to produce the best results between training and validation data (Table 4; Appendix A.3). This model had an R^2 of 0.90 for the training data, an R^2 of 0.78 for the validation data, and an R^2 of 0.87 when applied to all the data. All models had a $p\text{-value} < 0.001$. Across the lab-measured %N samples, values ranged from 0.55 to 2.64% and PLSR-predicted values ranged from 0.40 to 2.64%. For subsequent steps, we used PLSR-predicted values.

Table 4. TALL PLSR Model Results. PLSR model results (R^2). All models have p -values < 0.001 .

	Training Data	Validation Data	All Data
Lab %N PLSR	0.9	0.78	0.87
HSI %N PLSR	0.61	0.57	0.56
HSI LMA PLSR	0.72	0.77	0.73

To predict the top-of-canopy %N from the HSI data, we used a PLSR model with five components. This model had an R^2 of 0.61 for the training data, an R^2 of 0.57 for the validation data, and an R^2 of 0.56 when applied to all the data (Table 4; Appendix A.4). All models had a p -value < 0.001 . After applying the PLSR coefficients across the images and removing extreme outliers using a Tukey's outlier test ($k = 3$), %N values ranged from 0.004 to 3.048% (Figure 10a), which is comparable to the ranges of %N found in Eastern US temperate forests by Serbin et al. (2014).

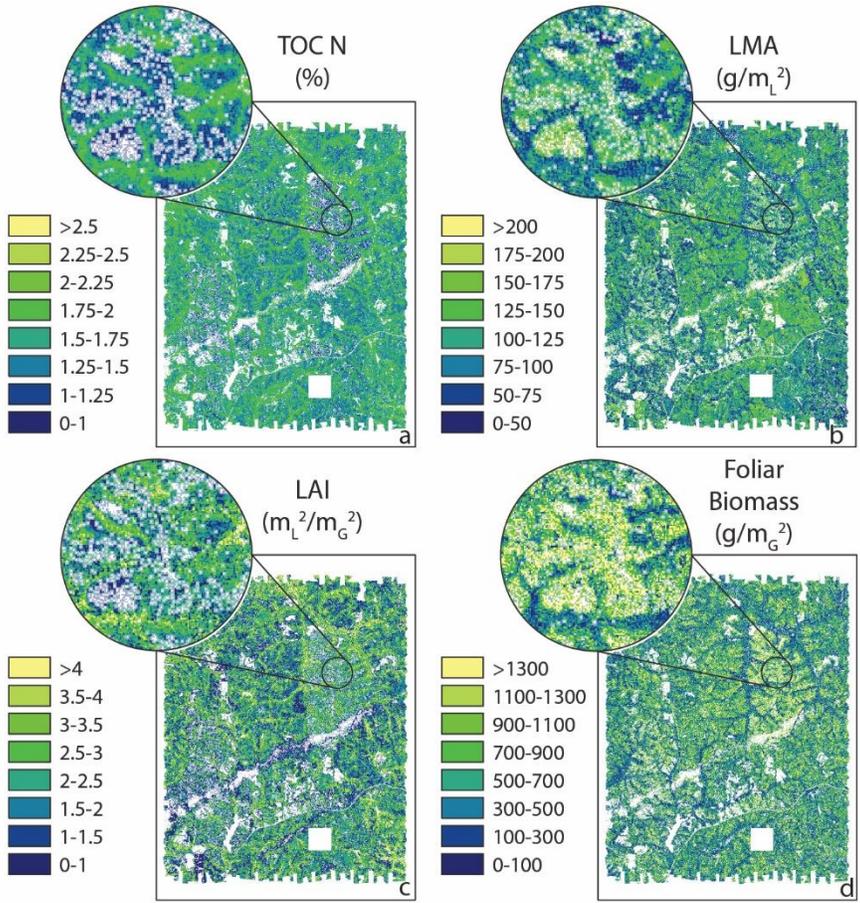


Figure 10. Canopy Functional Trait Maps. Maps of functional and structural traits derived from NEON AOP HSI and lidar data. TOC = top-of-canopy; m_L^2 refers to square meters of leaf material, while m_G^2 refers to square meters of ground. Call out circle is a 1km radius around the NEON flux tower at this site, shown as a star.

To predict LMA from the HSI data, we used a PLSR model with eight components. This model had an R^2 of 0.72 for the training data, an R^2 of 0.77 for the validation data, and an R^2 of 0.73 when applied to all the data (Table 4; Appendix A.5). All models had a p-value < 0.001 . Across the field measured samples, LMA values ranged from 20.72 to 326.02 g/m_L^2 . After applying the PLSR coefficients to the images and removing extreme outliers using a Tukey's outlier test ($k = 3$), LMA values ranged from 0.041 to 356.7 g/m_L^2 (Figure 10b). While these values

are extrapolated outside of the range of values used in our PLSR model, they are comparable to LMA ranges found globally by Poorter et al. (2009).

Within Canopy Leaf Traits: Lidar and HSI

To predict within canopy LMA, our final model consists of four lidar-derived metrics. These metrics included top-of-canopy %N, sample height, euphotic zone depth, and standard deviation of LAD within a column of voxels. Our final model for within canopy LMA had an R^2 of 0.51 for the training data and an R^2 of 0.50 for our validation data (Appendix A.6). Both models had a p-value < 0.001 .

After summing all within canopy values we calculated the total amount of N (g/m_G^2 ; Figure 11), foliar biomass (g/m_G^2 ; Figure 10d), and LAI ($\text{m}_L^2/\text{m}_G^2$; Figure 10c) for each pixel. We then removed extreme outliers using Tukey's outlier test ($k = 3$). Values greater than 15 g/m_G^2 were removed from the total canopy N raster (0.03% of raster pixels), values greater than 2465 g/m_G^2 were removed from the foliar biomass raster (0.46% of raster pixels), and values greater than $7 \text{ m}_L^2/\text{m}_G^2$ were removed from the LAI raster (0.03% of raster pixels).

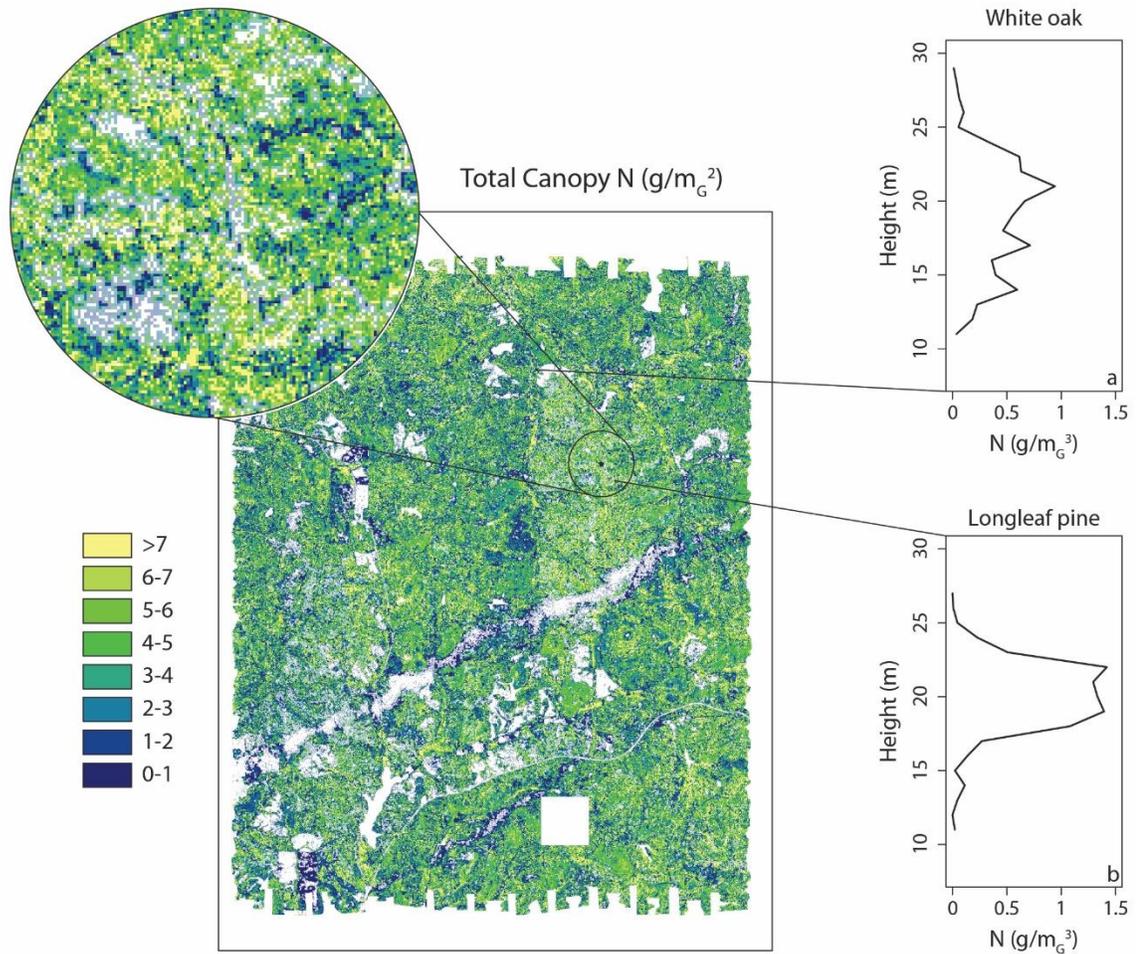


Figure 11. Total Canopy N Map. Map of total canopy N (g/m_G^2) and within canopy N (g/m_G^3) profiles from white oak (total foliar N = $6.99 \text{ g}/\text{m}_G^2$) and longleaf pine (total foliar N = $7.93 \text{ g}/\text{m}_G^2$). Locations were extracted based on the GPS positions of field samples. Call out circle is a 1km radius around the NEON flux tower at this site, show as a star.

To illustrate the differences in canopy profiles of within canopy N (g/m_G^3) we extracted data from the total canopy rasters using the GPS locations of a white oak (Figure 11a) and longleaf pine (Figure 11b) sample from our field data. The total amount of N in the white oak sample was $6.99 \text{ g}/\text{m}_G^2$ while there was $7.93 \text{ g}/\text{m}_G^2$ in the canopy of the longleaf pine sample. Moreover, the profiles of each sample illustrate differing within canopy allocation strategies for the two species.

Top-of-Canopy and Total Canopy N: Differing Spatial Patterns

After normalizing (mean = 0, SD = 1) the top-of-canopy %N and total canopy N (g/m_G^2) rasters for equal comparison, there was no relationship between the two variables at any of the spatial resolutions, showing that these differences are not scale dependent (Figure 12, panels a, b, & c). Prior to normalization, we used linear regression to test the relationship between the two variables at each spatial resolution (Figure 12, panels d, e, & f). All linear regressions were significant (p-value < 0.05), but the largest R^2 value was 0.02 showing a very weak relationship between top-of-canopy and total canopy N across spatial resolutions. This lack of relationship shows that as data is aggregated together at coarser spatial resolutions, resulting in pixels containing multiple PFTs rather than single species, there are still distinct differences between top-of-canopy and total canopy N.

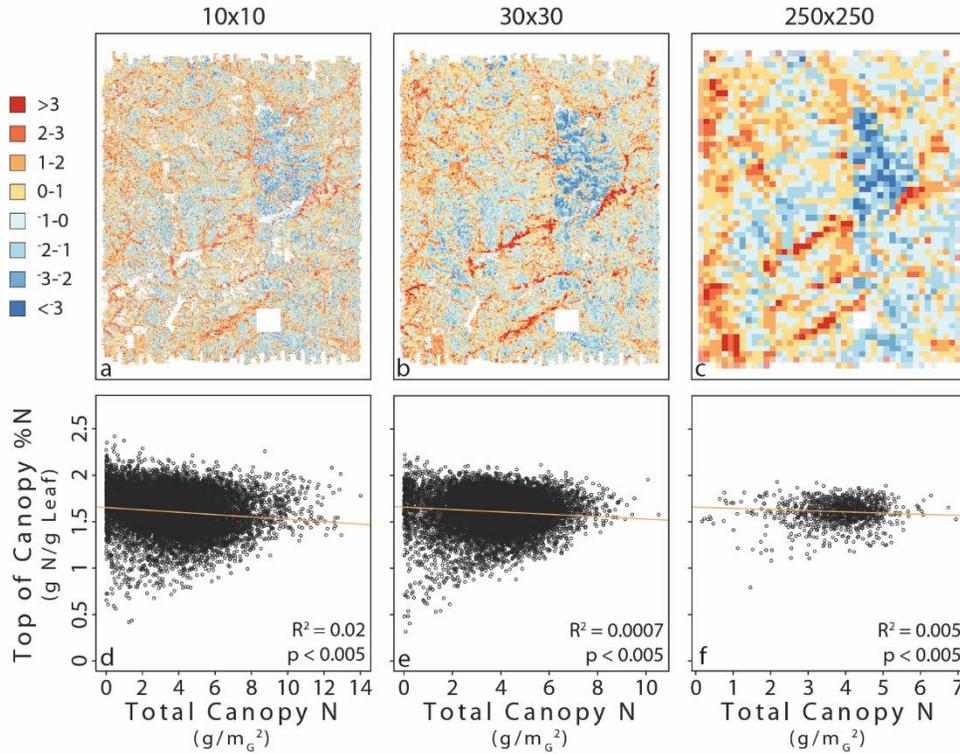


Figure 12. Top-of-Canopy N and Total Canopy N Differences. Maps of the scaled and centered differences between top-of-canopy %N and total canopy N (g/m^2) at three different spatial resolutions: 10x10 m (NEON AOP lidar), 30x30 m (Landsat), 250x250 m (MODIS). Regression results showing no relationship between the two measurements.

To assess differences in spatial patterns across the landscape, we calculated variograms for the top-of-canopy %N and total canopy N (g/m^2) datasets (Figure 13). A comparison of Moran's I values for the two normalized (mean = 0, SD = 1) datasets showed that the top-of-canopy %N map was substantially more spatially autocorrelated (Moran's I = 0.026) than the total canopy N map (Moran's I = 0.014). For the non-normalized datasets, top-of-canopy %N samples exhibit spatial autocorrelation up to a distance of 1200 meters, while total canopy N (g/m^2) samples are spatially autocorrelated up to a distance of 700 meters. Partial sill measurements also differ

substantially, showing differences in variability between pairs of points, with top-of-canopy %N having a value of 0.23 and total canopy N (g/m_G^2) having a value of 0.09. The shapes of the variograms indicate that top-of-canopy %N is grouped into clusters of similar values (lower nugget, longer range), while the total canopy N values are more evenly distributed (higher nugget, shorter range).

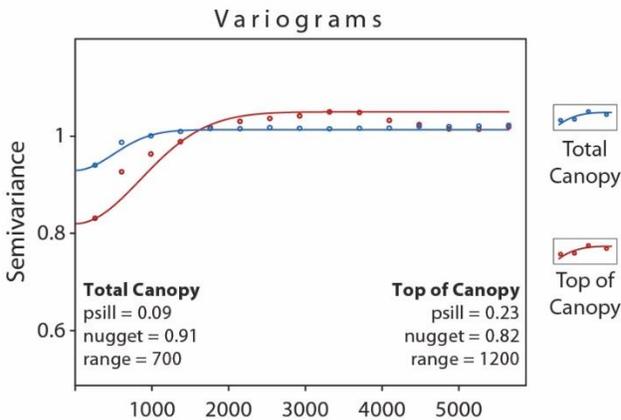


Figure 13. Canopy N Variograms. Variograms for normalized (mean = 0, SD = 1) Top-of-Canopy %N and Total Canopy N (g/m_G^2). 10,000 random samples were extracted from both datasets.

Regional Patterns and Environmental Drivers: Assessing Spatial Structure

Elevation visually appeared to be a strong driver of leaf trait spatial distributions in our maps (Figure 10). To quantify this relationship, we looked at the influence of elevation on top-of-canopy %N, total canopy N (g/m_G^2), and the normalized difference between these two datasets (Figure 14). Top-of-canopy %N was related to elevation ($R^2 = 0.13$), while total canopy N (g/m_G^2) was not related to elevation ($p > 0.05$). Therefore, the correlation between the normalized difference of these two estimates and elevation ($R^2 = 0.06$) is mostly due to the stronger correlation between elevation and top-of-canopy %N.

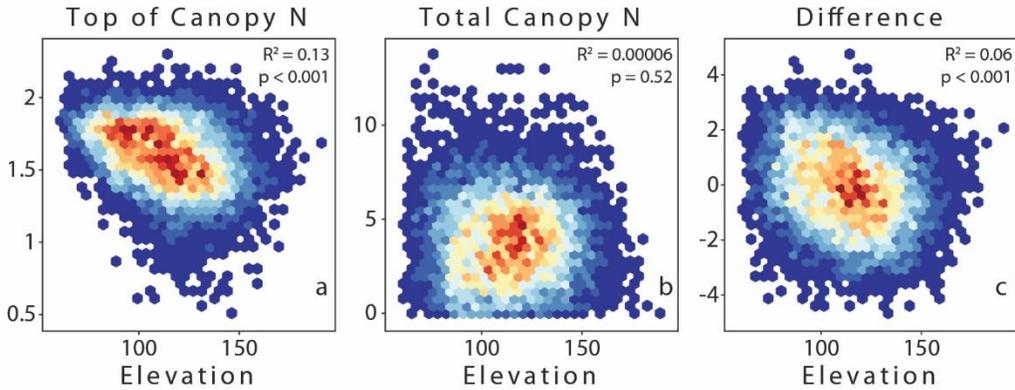


Figure 14. Influence of Elevation on Canopy N. Heatmaps showing the relationship between top-of-canopy %N, total canopy N (g/mG^2), the normalized difference between these two measurements, and elevation. Y-axis units for each plot is given at the top of the plot.

To more broadly understand the effects of abiotic gradients and management regimes on leaf and canopy functional traits, we performed a Monte Carlo simulation on the abiotic and management rasters to compile a distribution of results. Models predicting top-of-canopy %N had a mean R^2 of 0.24 with a standard deviation of 0.009. Eleven of the predictors appeared in over 20% of the models (Figure 15), seven variables appeared in no models, and 7 variables appeared in all the models (Table 5).

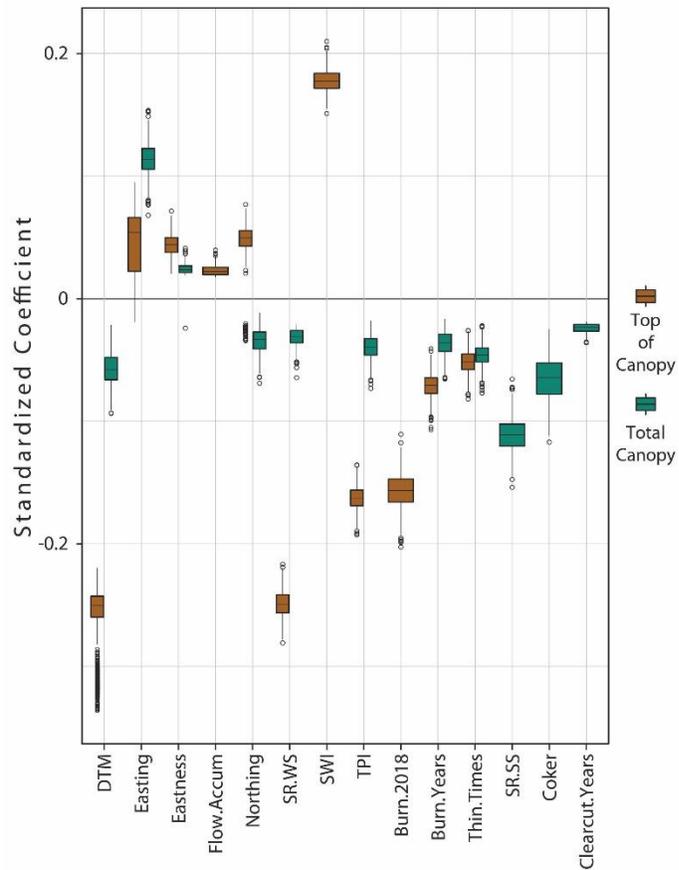


Figure 15. TALL Model Coefficients. Coefficients from standardized variables (mean = 0, SD = 0.5) from Monte Carlo simulations with variables that appeared in at least 20% of the regressions. All coefficients have a p-value of <0.05.

Table 5. TALL Model Coefficients. Mean standardized coefficients (mean = 0, SD = 0.5), standard deviation of coefficients, and percent of models each variable was present from Monte Carlo simulations. All coefficients have a p-value of <0.05.

		Top-of-Canopy			Total Canopy		
		Mean Coefficient	Standard Deviation	Models Present (%)	Mean Coefficient	Standard Deviation	Models Present (%)
TOPOGRAPHIC	DTM	-0.259	0.027	100.0	-0.057	0.013	99.8
	Easting	0.047	0.023	30.9	0.114	0.013	100.0
	Eastness	0.044	0.009	99.8	0.024	0.005	20.9
	Flow Accumulation	0.023	0.004	35.3	-0.017	0.018	11.2
	Northing	0.047	0.017	85.1	-0.034	0.009	70.0
	Northness	NA	NA	0.0	0.027	0.006	4.7
	Surface Roughness	NA	NA	0.0	NA	NA	0.0
	Slope	NA	NA	0.0	NA	NA	0.0
	Solar Radiation – Summer Solstice	NA	NA	0.0	-0.111	0.013	100.0
	Solar Radiation – Winter Solstice	-0.249	0.011	100.0	-0.032	0.007	44.3
	Soil Wetness Index	0.178	0.009	100.0	NA	NA	0.0
	Topographic Position Index	-0.162	0.009	100.0	-0.040	0.009	97.1
	Topographic Ruggedness Index	NA	NA	0.0	NA	NA	0.0
GEOLOGIC	Alluvial	-0.091	0.151	1.3	0.153	0.158	8.3
	Coker	-0.144	0.015	18.1	-0.065	0.017	86.6
	Eutaw	0.146	0.013	18.1	0.021	0.036	5.7
	Gordo	-0.023	0.008	7.9	0.038	0.009	7.6

Table 5. continued

MANAGEMENT	Prescribed Burn 2018	-0.157	0.014	100.0	0.034	0.020	3.0
	Times Burned	NA	NA	0.0	0.025	0.010	7.8
	Years Since Last Burn	-0.071	0.010	100.0	-0.036	0.010	90.9
	Times Chemically Treated	-0.009	0.020	4.8	-0.017	0.017	4.3
	Years Since Last Chemical Treatment	-0.021	0.003	2.1	-0.002	0.023	2.5
	Times Clearcut	-0.013	0.020	7.5	-0.015	0.019	2.4
	Years Since Last Clearcut	-0.023	0.008	17.3	-0.024	0.004	24.6
	Times Thinned	-0.052	0.009	100.0	-0.046	0.009	98.2
	Years Since Last Thinning	NA	NA	0.0	-0.041	0.006	1.8

The only major topographic predictor (coefficient > 0.1) with a positive coefficient was soil wetness index (SWI), while major topographic predictors with a negative coefficient included elevation (DTM), solar radiation at the winter solstice (SR.WS), and TPI (topographic position index). The only major geologic predictor (coefficient > 0.1) with a negative coefficient was Coker substrate, while Eutaw substrate had a positive coefficient and was a major geologic predictor. The only major management variable (coefficient > 0.1) was areas burned in 2018 and it had a negative coefficient.

Total canopy N (g/mG^2) models had a mean R^2 of 0.03 with a standard deviation of 0.003. Eleven of the predictors appeared in over 20% of the models (Figure 15), 4 variables appeared in no models and 2 variables appeared in all the models (Table 5). Solar radiation at the summer solstice was the only major topographic predictor (coefficient > 0.1) with a negative coefficient,

while the only major topographic predictor with a positive coefficient was distance from western collection boundary (easting). Alluvial substrate was the only major geologic predictor (coefficient > 0.1) and it had a positive coefficient. There were no major management (coefficient > 0.1) predictors in the total canopy regressions.

For both regression models many of the management variables appeared in only a small percentage of the total models. This is because these management practices were only completed across a small fraction of the entire landscape, and these areas were not randomly sampled in each iteration of the Monte Carlo simulation.

The residuals of both regression models exhibited some spatial autocorrelation with top-of-canopy %N having a mean Moran's I of 0.03 with a standard deviation of 0.001 and total canopy N (g/mG^2) having a mean Moran's I of 0.008 with a standard deviation of 0.0006. While this spatial autocorrelation of the residuals would indicate that there is a trend present that we are not capturing, the aim of these regression was not predictive, but instead to compare the influence of these abiotic and management variables between the two functional traits estimates.

Discussion

We used airborne remote sensing and field-collected trait data to show that when three-dimensional forest structure is considered, different patterns of N appear across this landscape than are produced by two-dimensional top-of-canopy functional trait estimates. This analysis demonstrates that canopy functional diversity is not equivalent to leaf functional diversity, which illustrates the dampened variation in total canopy N between PFTs and across this landscape when compared to the heterogeneous spatial patterns produced by leaf functional diversity. This suggests that these two measurements correspond to different ecological processes and that relationships

between plant carbon assimilation and leaf functional traits must be considered in the context of canopy vertical structural heterogeneity.

Scaling and Mapping Leaf and Canopy Traits

Many studies have used HSI data to estimate plant functional traits and lidar data to measure forest structure, with much success across a wide variety of ecoregions (Dahlin et al. 2013; Asner et al. 2015; Stark et al. 2015; Smith et al. 2019). By combining 3D structural traits from lidar and 2D functional traits from HSI, we show that a fusion of these two data types can be used to model traits within the canopy volume. Moreover, our findings are within the ranges reported in field-based studies for LAD (Parker and Tibbs, 2003; Brown and Parker, 2004), %N (Serbin et al. 2014), LMA (g/mL^2 ; Poorter et al. 2009), and total canopy N (g/mG^2 ; Cole and Rapp 1981; Figure 16).

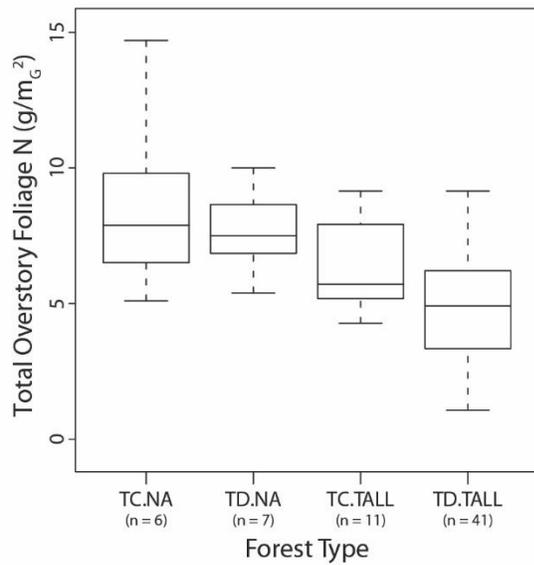


Figure 16. Total Overstory Foliar N Comparison. North American Foliar N values vs. TALL Foliar N values. TC = Temperate Coniferous, TD = Temperate Deciduous. North American N (NA) values come from Cole and Rapp (1981). Because our values calculated at TALL do not include the lowest 10 meters of the canopy, ANOVA results ($p < 0.001$) show a significant difference between NA and TALL values but not between Forest Types (TC and TD).

Our study focuses on an ecoregion consisting of closed-canopy broadleaf stands and sparser needleleaf forests, with our within-canopy trait estimates being reliable across these two plant functional types (PFTs). In addition, our within-canopy model utilizes variables related to the differences in PFTs (top-of-canopy %N), the local light environment (standard deviation of LAD within a column of voxels), and light capture (euphotic zone depth). These variables have been shown to be critical to canopy level processes (Field and Mooney 1986; Hardiman et al. 2001; Lefsky et al. 1999).

While our results show that we can accurately model foliar functional traits within the canopy volume in this ecosystem, more research is needed in different biomes to test the ability of HSI and lidar to accurately estimate within-canopy traits.

Measuring Ecosystem Function: Top-of-Canopy %N vs. Total Canopy N

While both foliar N and LMA have been identified as key drivers of plant functional diversity (Díaz et al. 2016) and have shown strong correlations with leaf photosynthesis in temperate ecosystems (Field and Mooney 1986; Evans 1989), we show that the spatial patterns of leaf-level top-of-canopy %N are not equivalent to those of total canopy N (g/m_G^2). Top-of-canopy leaf-level traits reflect key differences between PFTs, with needleleaf species exhibiting low %N and high LMA, while broadleaf species have higher %N and lower LMA (Appendix A.1). These fundamental differences in functional and structural traits between PFTs produce distinct dendritic patterns across this landscape corresponding to topographic features including drainages, which are dominated by broadleaf species, and slopes and ridges, which are dominated by pines (Figure 10a, b, c. & d). However, when three-dimensional canopy structure is considered (i.e., total canopy N), these distinct landscape patterns are dampened (Figure 11).

Figure 14 further shows that these distinct spatial patterns related to elevation are not reflected in our estimates of total canopy N (g/m_G^2). This may suggest that canopy architectural differences between PFTs are causing unique distributions of N within the canopies of individual trees (Figure 11a & b), and that these differences represent trade-offs since different PFTs exhibit similar total quantities of N (g/m_G^2) in their canopies (Figure 16). In this case, differences over a leaf function-structural architecture trade-off produces the dampened spatial patterns we see in this landscape (Figure 11).

Given the importance of N for photosynthesis, these dampened spatial patterns may not be surprising. By varying LMA, individual trees will distribute N (g/m_L^2) throughout their canopies in ways to maximize their nitrogen use efficiency, utilizing as much of the available N (g/m_G^2) as possible. Lower total N (g/m_G^2) within the canopy volume would imply lower production, a disadvantage that would be hard to reconcile between PFTs in the same ecosystem. While N-fixing trees could change these patterns, we observed no N-fixing trees in this landscape and overall, this area appears to have low N-fixing tree abundance (Staccone et al. 2020).

Abiotic and Management Drivers of Foliar and Canopy N

Following community assembly theory (Keddy 1992), abiotic drivers have been shown to predict species and leaf trait distributions within landscapes with both remote sensing and field observations (Dahlin et al. 2012; Kraft et al. 2008). We show that these same types of drivers can be used to predict top-of-canopy %N in this system, but not total canopy N (g/m_G^2).

Top-of-canopy %N patterns have consistently strong topographic, substrate, and management predictors, with many of these predictors being related to the distribution of PFTs across this landscape. For example, higher elevation areas that receive more solar radiation during the winter months and that were treated with a prescribed burn in 2018 prior to NEON AOP flights had consistently lower top-of-canopy %N values. This describes the spatial distribution of needleleaf species in this ecosystem. Conversely, lower elevation areas with a high soil water content had consistently higher top-of-canopy %N values, describing the distribution of broadleaf species in this environment. These relationships suggest that the spatial patterns of top-of-canopy %N are closely related to the spatial distribution of species within this ecosystem.

In contrast, variables related primarily to forest structural changes and water availability were the main drivers of total canopy N (g/m_G^2), even though these relationships were considerably weaker, though still significant. For instance, areas that had been clear-cut, thinned, or burned had lower total canopy N (g/m_G^2) estimates than areas that did not have a documented management history. This relationship is most likely due to management activities resulting in significant structural changes to forest stands and the removal of foliar biomass during these activities. Furthermore, areas that received high solar radiation in the summer months also had lower estimates of total canopy N (g/m_G^2). This could be due to microclimatic effects. Water stress in these sunnier, drier areas may cause a reduction in growth and, therefore, total canopy N (g/m_G^2), as light availability is not likely to be a limiting factor in this system.

Model Uncertainty and Data Concerns

There are many possible sources of error and uncertainty to consider when scaling traits from leaf to landscape, including those related to field and GPS collections, laboratory equipment, remote sensing sensors, and statistical methodologies. While we did not conduct a formal assessment of uncertainty as it propagates through this study, our findings are within the ranges reported in many field-based studies (see section 4.1). Our final PLSR models did show a systematic bias of slightly underestimating N and LMA in needleleaf species (Appendix A.3 and Appendix A.4), which could partially explain the differing landscape-scale relationships between total canopy and top-of-canopy N. This could possibly be improved by the inclusion of forest structure metrics such as LAI in the PLSR models. However, due to the low density lidar data we are forced to estimate structural traits at a coarser spatial resolution (10x10 meters) than the HSI data (1x1 meter).

Because some field samples are closer than 10 meters to one another, and thus exist within the same pixel, the inclusion of structural traits did not correct this bias.

While understory shade tolerant plants play an important role in ecosystem functioning (Valladares et al. 2016), we ignored the lowest 10 meters of the forest canopy where many of these species occur due to limitations with the lidar data from the NEON AOP (Kamoske et al. 2019). As current lidar sensors within the NEON AOP are upgraded, we will be able to ask important questions about the role of the understory in ecosystem functioning.

In this study we only considered healthy green forest vegetation, which may partially explain the weaker relationships between environmental variables and canopy functional and structural traits. More research is needed into how HSI and PLSR perform in stressed terrestrial environments and across more heterogeneous landscapes.

The development of a universal model to predict leaf- and canopy-level traits was beyond the scope of this project; however, as more within-canopy foliar traits are collected across a diversity of ecosystems, PFTs, and tree species, these models will become more robust and can be applied to other regions.

Looking Forward

With airborne and spaceborne platforms like the NEON AOP, NASA Goddard's Lidar, Hyperspectral, & Thermal Imager (G-LiHT; Cook et al. 2013), the Global Ecosystem Dynamics Investigation (GEDI; Stavros et al. 2017), and the proposed Surface Biology and Geology Mission (SBG; National Academies of Sciences, Engineering, and Medicine 2018) collecting HSI and lidar data across a variety of ecoregions, there is a unique opportunity for researchers to ask and answer

questions related to how forest canopies function across landscapes and continents, rather than just the leaves at the top of the canopy.

In support of these new questions about ecosystem function, we present a reproducible methodology to model foliar traits throughout the entire canopy volume. We also show that the spatial patterns produced by traditional top-of-canopy measurements of %N are dramatically different than those produced when three-dimensional forest structure is considered. While more research is needed to test these relationships in different ecoregions and across latitudinal gradients, this ever-increasing availability of HSI and lidar data will provide new and exciting opportunities.

These opportunities may raise several questions about the drivers of canopy function. For example: A) What is the role of soil nutrient availability and heterogeneity in canopy function? and B) How are these relationships affected by latitudinal gradients and climate regimes? Further research is needed into these questions to better understand the drivers behind ecosystem functioning in horizontal and vertical space as well as through time.

Conclusions

Forest structural and functional diversity drive critical canopy processes related to carbon sequestration; however, structure and function are rarely considered in unison at ecosystem scales. Here we show that when forest structure is considered, the patterns produced by the total amount of N (g/m^2) within the canopy volume are substantially different from the patterns produced by top-of-canopy %N. Furthermore, since total canopy N variation is dampened relative to leaf-level variation over a landscape characterized by variable PFT dominance, we find evidence of canopy architecture and leaf function tradeoffs. Patterns of total N are driven by different abiotic gradients

and management regimes, further showing the differences between these two estimates of ecosystem function.

These differing spatial patterns, as well as differing abiotic and management drivers, show that canopy functional diversity is not equivalent to leaf functional diversity. By not considering structure and function together, there could be impacts on how we scale fine-resolution ecological processes to landscape, continental, and global models. However, with new space- and airborne remote sensing platforms collecting HSI and lidar data across a variety of ecoregions, we have an opportunity to think about the terrestrial carbon cycle in three dimensions. This new approach will potentially unlock important insights into how forests function in a time of rapid anthropogenic and environmental change.

Acknowledgments

Thanks to the Talladega National Forest - Oakmulgee Ranger District for providing site access, to NEON staff for providing technical support, and to O. Jain for providing assistance in the field. This work was supported in part by the NSF Macrosystem Biology Program award #1702379. The NEON is a program sponsored by the National Science Foundation and operated under cooperative agreement by Battelle Memorial Institute. This material is based in part upon work supported by the National Science Foundation through the NEON Program. Shawn P. Serbin was partially supported by the United States Department of Energy contract No. DE-SC0012704 to Brookhaven National Laboratory.

Data Availability

Lidar and HSI data are available at: <http://data.neonscience.org>. R package to estimate structural traits from airborne lidar data is provided through our GitHub at:

<https://github.com/akamoske/canopyLazR> and as a stable DOI at

<http://doi.org/10.5281/zenodo.3987340>. R package to pre-process HSI data, extract reflectance data, and apply PLSR coefficients is provided through our GitHub at:

<https://github.com/akamoske/hypRspec> and as a stable DOI at

<https://zenodo.org/record/3987336>. Reflectance spectra and trait data are available through the

ECOSIS database at: <https://data.ecosis.org/dataset/2018-talladega-national-forest--leaf-level-reflectance-spectra-and-foliar-traits>. Laboratory measured trait data are available through the

TRY database (dataset ID = 714) at: www.try-db.org.

CHAPTER 4. MAPPING MULTIPLE DIMENSIONS OF BIODIVERSITY WITH AIRBORNE HYPERSPECTRAL AND LIDAR REMOTE SENSING

The following co-authors contributed to this study: K.M. Dahlin, Q.D. Read, S. Record, S.P. Serbin, S.C. Stark, and P.L. Zarnetske.

Introduction

A fundamental goal in community ecology is to understand and predict the spatial distributions of species, traits, and biodiversity across ecosystems (Keddy 1992). However, there are many ways of measuring distinct dimensions of biodiversity, including taxonomic, functional phylogenetic. Each dimension of biodiversity may be determined by different abiotic and biotic drivers and ecological processes, while following unique spatial and temporal patterns (Gaston 2000; Lomolino et al. 2010). For example, taxonomic diversity, the relative abundance of species, is affected by environmental change due to management practices and disturbance regimes (Baiser et al. 2012; Olden and Rooney 2006, Li et al. 2020) and has been linked to carbon storage in forest biomes (Cavanaugh et al. 2014). Functional diversity, the community-wide variation in structure or functional traits that affect how species interact with their environment (e.g., leaf nutritional or physiological properties, shade tolerance, canopy height, leaf area index, etc.), is critical for determining biodiversity-ecosystem function relationships (Baiser and Lockwood 2011; Flynn et al. 2011). Phylogenetic diversity, the overall relatedness of species in a community (Srivastava et al. 2012) is influenced by the spatial clustering of closely related species that occupy similar environments (Cavender-Bares et al. 2009).

The impacts of these different dimensions of biodiversity on the observable properties of forest canopies is not well known, but is critical because forest canopies link the atmosphere and the vast majority of Earth's terrestrial biomass (Ozanne et al. 2003; Bonan 2008), provide key

ecosystem services such as carbon sequestration (Bunker et al. 2005; Hooper et al. 2012; Isbell et al. 2015), and are severely impacted by rapid global change (Parmesan and Yohe 2003; Cardinale et al. 2012; Hooper et al. 2012; Brook et al. 2008; Urban 2015; Stocker et al 2013; Smith et al. 2015). By considering these different dimensions of biodiversity in forest ecosystems, we may be able to advance our understanding of the impacts of global change on the spatial and temporal patterns of biodiversity.

Much of the current understanding of the spatial distribution of these dimensions of biodiversity at continental scales has come from satellite remote sensing products (Turner et al. 2003; Pettorelli et al. 2014; Bush et al. 2017; Duro et al. 2007). Remotely sensed products have helped clarify the scale-dependence of topography and biogeography as drivers of patterns of biodiversity (Zarnetske et al. 2019; Read et al. 2020; Record et al. 2020). Even though these remote sensing products have provided critical insights and standardized measurements over a range of spatial scales for decades (e.g., from individual forest stands to continents; He et al. 2015), there are important differences between these measurements and ground-based biodiversity observations due to differences in scale (Tews et al. 2004). Coarse-scale ecological observations (e.g., MODIS data at 250x250 meters) can cause dominant landscape features to homogenize measurements (Boyce 2006; Cooper et al. 2019), leading to the omission of fine-scale (e.g., an individual field plot or tree canopy) heterogeneity. This has been shown to have critical impacts on ecosystem functioning due to maximizing resource use efficiency. However, high-resolution airborne remote sensing platforms may help resolve these issues related to the spatial scale of observation.

The National Ecological Observatory Network's Airborne Observation Platform (NEON AOP) provides a unique opportunity to determine the role of remotely sensed metrics on the

prediction of biodiversity by collecting airborne lidar and hyperspectral data (e.g., 380-2500nm; 5 nm bands) across a variety of ecosystems at fine spatial grains (i.e., pixel resolution of 1 meter). NEON also collects ground data at individual field plots within the footprint of these landscape-scale airborne observations across a network of 81 systematically sampled sites across the US (Kampe et al. 2010; Thorpe et al. 2016; Barnett et al. 2019). Airborne lidar has been used to measure metrics critical to mapping biodiversity across landscapes, such as the structural heterogeneity of forests (Stark et al. 2015; Kamoske et al. 2019; Shao et al. 2019; Cosovic et al. 2020) and topographic diversity (Dahlin et al. 2012). Hyperspectral imagery has recently been used to measure the spectral diversity of ecosystems at fine spatial scales, which can be related to plant biodiversity (Gholizadeh et al. 2019; Dahlin 2016; Laliberté et al. 2020; Cavender-Bares et al. 2017; Wang and Gamon 2019).

While enthusiasm for the application of hyperspectral and lidar remote sensing to map biodiversity has grown in recent years (Jetz et al. 2019; Stavros et al. 2017) most studies have focused on within-site diversity mapping in a single biome (Dahlin 2016; Gholizadeh et al. 2018; Gholizadeh et al. 2019; Wang et al. 2018). Yet, for a biodiversity mapping program to be operationalized (e.g., the Group on Earth Observations Biodiversity Observation Networks Essential Biodiversity Variables; Jetz et al. 2019), methods must work across multiple sites and biomes and address multiple dimensions of biodiversity. Resolving relationships between canopy observations and biodiversity within and among ecosystems is essential to advance our understanding of the patterns of and changes in biodiversity as well as the nature of diversity-function relationships due to the rapid acquisition at global scales of remotely sensed data that is not reproducible with field methods alone (LaRue et al. 2019).

In this study, we ask how well commonly used remotely sensed spectral and forest structural heterogeneity metrics can explain different dimensions of alpha diversity at multiple sites across a wide (i.e., 10°) latitudinal gradient within one biome – temperate broadleaf forests – consisting of multiple forest regions (e.g., Dyer 2006; Southern mixed (oak-pine), mesophytic (Appalachian oaks), Oak-hickory, Northern hardwoods, etc.). We also consider whether topographic variability and biogeographic differences more broadly influence these relationships between remote sensing metrics and biodiversity. We aim to address three questions critical to understanding forest biodiversity in this temperate forest biome: (1) Which remotely sensed metric, or combination of metrics, best predicts alpha taxonomic, functional, and phylogenetic diversity across a latitudinal gradient of temperate forest regions, (2) how do these different dimensions of biodiversity vary within and among these forest regions, and (3) do remote sensing metrics capture this variation? We also detail a replicable methodology for estimating the structural, spectral, and topographic heterogeneity of temperate forests from airborne lidar and hyperspectral data.

Materials and Methods

Study Sites

Field measurements and remote sensing data were acquired from five climatically and ecologically diverse NEON sites located along a latitudinal gradient of eastern US temperate forest regions (Figure 17). These sites include, from south to north and followed by their four-letter NEON site abbreviations, Talladega National Forest (TALL), Oak Ridge National Laboratory (ORNL), Mountain Lake Biological Station (MLBS), the Smithsonian Environmental Research Center (SERC), and Harvard Forest (HARV). Across all sites, mean annual precipitation ranges

from 967-1350 mm, mean annual temperature ranges from 8-17 °C, mean canopy height ranges from 18-38 meters, elevation ranges from 15-1126 meters, and airborne imagery collections range from 110-355 km² (individual site information in Appendix B.1).

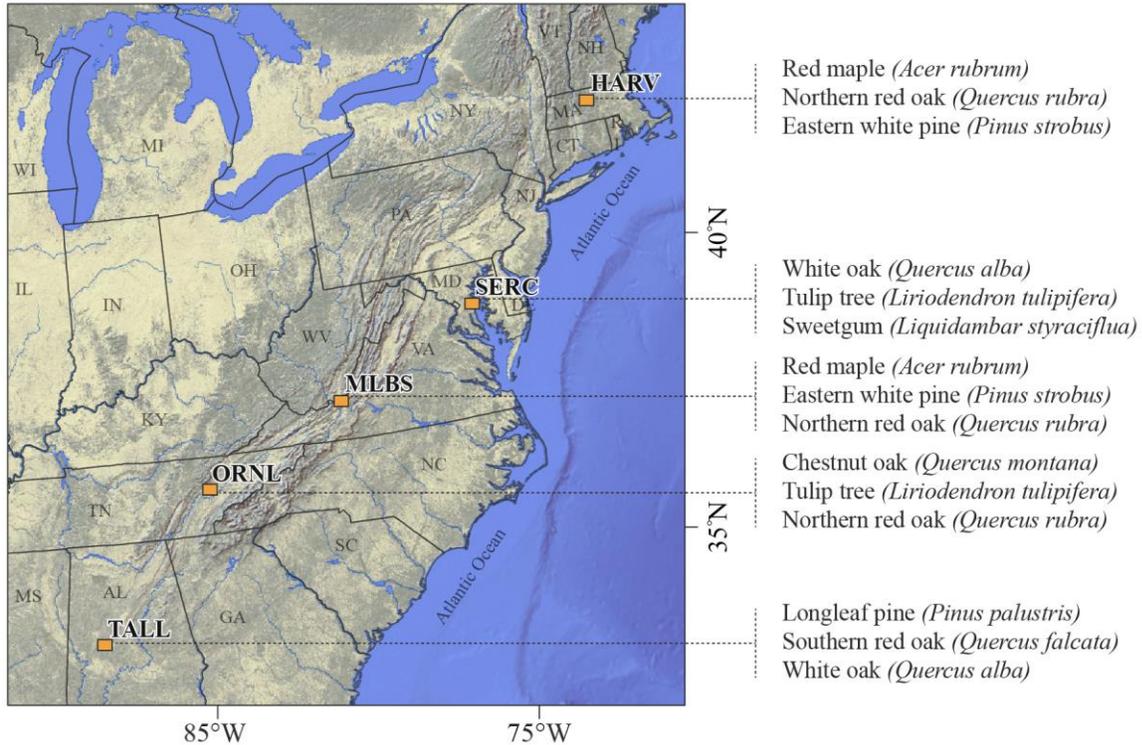


Figure 17. NEON Field Sites Map. Map showing NEON field sites used in this study (orange squares) and three of the most common tree species for each site based on field observations (individual site information in Appendix B.1).

Calculating Tree Diversity Metrics within NEON Field Plots

To quantify tree taxonomic, functional, and phylogenetic diversity at the NEON plot scale (i.e., 40x40m), we downloaded woody-plant species data from the NEON data portal (National Ecological Observatory Network 2020) for the same year that the NEON AOP flights and our field work were conducted (i.e., 2018 for all sites, except 2017 for SERC) and filtered it to retain only

living tree growth forms. We then used the stem diameters of each individual tree to calculate the relative abundance of each species per plot by summing the total basal area of each species and dividing it by the total basal area of all species in each plot (Auclair and Cottam 1971; Whitehead 1978). To quantify alpha taxonomic diversity within each field plot we used these abundance weighted species richness measurements to calculate Shannon's Diversity Index (Spellerberg and Fedor 2003). To calculate phylogenetic and functional diversity, we compiled widely available functional trait data from Kattge et al. (2020) and Stevens et al. (2020) and tree species phylogeny from Potter and Woodall (2012) and Potter and Koch (2014) for all species present in the field plots described above using the basal area abundance weights (Appendix B.2) and resolved any discrepancies in species names using the most recent taxonomy listed by IUCN (iucnredlist.org). Since most species had at least one missing functional trait value, we used the phylogeny to impute the missing values with the *Rphylopars* R package and then created a Gower distance matrix of normalized functional traits using the imputed trait dataset and phylogeny for all the species (Read et al. 2020). We then created a cophenetic distance matrix based on the tree species phylogeny data. Next, we calculated a community-level mean pairwise distance (MPD) metric for both functional and phylogenetic datasets. While there are many different methods to calculate alpha taxonomic, functional, and phylogenetic diversity (Jost 2006; Jost 2007), we used the above metrics because they are widely used and easily interpreted due to averaging all pairwise distances so that very distantly related species are more heavily weighted (Read et al. 2020).

Since several of the plots only had one species present, we could not calculate these functional and phylogenetic diversity metrics as they cannot be defined with only one data point (Read et al. 2020). After removing these plots from the dataset, there were 19 plots for TALL, 14 plots for ORNL, 32 plots for MLBS, 14 plots for SERC, and 17 plots for HARV.

Remote Sensing Data

To better understand the role of canopy observations on different dimensions of biodiversity we processed airborne lidar and hyperspectral data (e.g., 1m spatial resolution) from the NEON AOP into 43 metrics related to spectral diversity, vegetation health, canopy structure, and topography (Appendix B.3, B.4, & B.5). We used remotely sensed data from 2018 for all sites except for SERC (2017), all of which was collected during peak greenness as defined by MODIS NDVI (Kampe et al. 2010). Two different lidar systems operated at the same specifications were used for these collections (Appendix B.1; Kamoske et al. 2019). We calculated each metric at its nominal resolution, and then aggregated the results to produce a single value for each NEON plot, calculating the mean, minimum, maximum, range, and standard deviation of each metric that did not already produce a single value (e.g., convex hull volume).

Forest Structural and Topographic Diversity from Lidar Remote Sensing

To calibrate lidar structural diversity estimates with Beer-Lambert extinction coefficients, we collected hemispherical photographs across each site at locations representing the diversity of tree species and stand structures in conjunction with NEON AOP flights following the methodology outlined in Kamoske et al. (2019). We then calculated plant area index (hereafter referred to as LAI; Miller 1967), which is widely used as a proxy for LAI due to the difficulty of correcting for non-foliage elements (Richardson et al. 2009), using the Digital Hemispherical Photography software (DHP; Leblanc et al. 2005) and setting the zenith angle to match the scanning angle of each lidar sensor (Appendix B.1; Sabol et al. 2014; Richardson et al. 2009; Solberg et al. 2006; Korhonen et al. 2011).

To estimate three-dimensional canopy structural diversity, we processed the lidar data for leaf area density (LAD; the total leaf area per unit of volume) at a 10x10m spatial resolution using our *canopyLazR* R package (Kamoske et al. 2019; github.com/akamoske/canopyLazR), which is similar to other published methods (MacArthur and Horn 1969; Stark et al. 2012; Zhao and Popescu 2009; Solberg et al. 2006; Sumida et al. 2009). First, we normalized the point cloud to height above the ground and then calculated LAD by counting the number of lidar pulses that enter and exit each voxel in each vertical column of data that has at least one ground return. Due to this relatively coarse lidar data exhibiting noise caused by topographic variation in the LAD results and to have an easily comparable dataset, we removed the LAD estimates from the bottom 5m of the canopy (Kamoske et al. 2019). We then calibrated the LAD estimates for each individual site using a Beer-Lambert extinction coefficient derived by calculating the slope of a regression equation between hemispherical photograph derived LAI and lidar estimated LAI (e.g., Appendix B.6; Sabol et al. 2014; Richardson et al. 2009). To remove non-forest pixels, we applied a canopy height and LAI mask to the upper end of to each LAD dataset using Tukey's outlier test ($k=1.5$) and then removed all pixels where LAI equals zero (Kamoske et al. 2019). With these masked LAD rasters, we calculated 21 forest structural metrics at a 10x10m resolution for each field-plot (Appendix B.4).

To quantify topographic diversity at each site, we calculated nine variables using QGIS and the 10x10m lidar derived digital terrain model (Appendix B.5).

Hyperspectral Remote Sensing Reflectance Metrics

We processed the atmospherically corrected 1m resolution hyperspectral imagery from the NEON AOP before analysis using our *hypRspec* R package on GitHub (Kamoske et al. 2020; github.com/akamoske/hypRspec). After removing all flight lines re-flown due to cloudiness and keeping the less cloudy ones, we visually identified noisy bands in the data and removed all wavelengths that were below 500nm, between 1350 and 1450nm, between 1800 and 2000nm, and above 2400nm. We then calculated a narrowband NDVI mask (red = 674nm; NIR = 830nm; NDVI > 0.5) to remove all unlikely-to-be-vegetated pixels from further analysis (Dahlin et al. 2014). To remove all shaded pixels, we used Tukey's outlier test ($k = 1.5$), where all pixels that have a reflectance below the lower threshold are considered outliers and removed (Kamoske et al. 2020). We then applied a topographic correction to reduce the effects of terrain, view, and illumination (Soenen et al. 2005) and a bidirectional reflectance distribution function effects correction (BRDF) with a thick Ross kernel and a dense Li kernel to remove the anisotropic scattering properties of vegetation that result in flight line artifacts (Colgan et al. 2012; Collings et al. 2010; Schlapfer et al. 2015; Wanner et al. 1995; Weyermann et al. 2015; Wang et al. 2020).

Using this corrected hyperspectral data, we calculated 13 hyperspectral reflectance and principal component analysis (PCA) derived metrics (Appendix B.3) by extracting reflectance spectra from all pixels within each field plot and using the mean value if a pixel occurred in multiple flight lines. To calculate the PCA based metrics, we used the extracted data from all sites as a single dataset and PCA to reduce the dimensionality of this data (Venables and Ripley 2002). We used the first two principal components (PCs), which captured 97.6% of the overall variation in the dataset, in subsequent analyses.

Influence of Remote Sensing Metrics on Biodiversity

To quantify the relative importance of our 43 remote sensing derived metrics related to the structural, spectral, and topographic heterogeneity of eastern temperate forests on different dimensions of alpha diversity we used a combination of linear mixed effect modeling (LME; Pinheiro and Bates 2000; Gotelli and Ellison 2013) and stepwise AIC model selection (Burnham et al. 2010; Mascaro et al. 2011). To allow for direct comparison between model coefficients, we standardized all 43 metrics and the three diversity variables (Gelman 2008; mean = 0, standard deviation = 0.5). For each dimension of biodiversity, we calculated a single model for each predictor type individually (i.e., hyperspectral, lidar, and topography) and a single model with all predictors combined.

First, to avoid multicollinearity, defined here as variables with a Pearson's R greater than 0.5, we tested the correlation between each pair of predictor variables and kept the variable most correlated with each dimension of biodiversity for further analysis. Using the remaining variables, we developed a LME model using each of these variables as a fixed effect and site (i.e., TALL, ORNL, etc.) as a random effect to allow for inferences to extend to differences between sites in general rather than between the five sites for which we had data (Pinheiro and Bates 2000; Gotelli and Ellison 2013). We included these site level differences to help account for critical large-scale biogeographical and management differences between sites (Dupouey et al. 2002; Bengtsson et al. 2000; Reich et al. 2001; Dambrine et al. 2007).

We then used stepwise AIC model selection to determine the best combination of predictor variables for each of our models (Safken et al. 2018). We tested backward, forward, and both direction stepwise variable selection, all of which resulted in the same predictor variables for each model. We then removed any remaining variables with non-significant coefficients (p-value >

0.05) and developed final LME models with site as a random effect. Lastly, we performed LME partial regression analysis using the final metrics in each model grouped by type (i.e., hyperspectral, lidar, and topography) to determine the percentage of the final model's R^2 value that each group of metrics represents (i.e., metric R^2 / final model R^2).

Detecting Biodiversity and Remote Sensing Metric Variation

The preceding analyses indicated that plot-scale biodiversity within sites might vary as much or more than biodiversity among sites. To better understand and quantify this biodiversity variation and to determine if remotely sensed metrics follow the same general patterns, we used a combination of PCA, k-means clustering (Ding and He 2004; Wagstaff et al. 2001), and analysis of variance (ANOVA; Girden 1992; Tabachnick and Fidell 2007).

First, we used PCA on the three dimensions of biodiversity to reduce the dimensionality of the field data and then used k-means to cluster the observations together by calculating Euclidean distances between all the observations. Because it can be difficult to determine how many clusters are needed to meaningfully group a dataset, we computed the within-group sum of squares for each set of clusters ranging from 2 to 18. We then generated a scree plot from these results and determined the appropriate number of clusters to be between two and five; however, using a set of two clusters was easily interpreted based on having a relatively large cluster mean within each cluster. Using a k-means with 10 random starting points, with the assignment with the lowest within-cluster variation used for further analysis, and two centers, we clustered the PC values while avoiding clumping the values together in space. We then assigned a cluster value to each field plot based on these two groups which represent high and low diversity plots (i.e., a high diversity

plot/cluster has high taxonomic, functional, and phylogenetic diversity) and used ANOVA and Tukey's post hoc tests to examine the inter- and intra-site differences between the dimensions of biodiversity and remotely sensed metrics.

Results

Variation of Biodiversity and Remote Sensing Metrics

In sites with spatially distinct broadleaf and needleleaf stands (i.e., TALL; Kamoske et al. 2020) there were a wider range of values across spectral and structural diversity metrics compared to sites dominated by broadleaf species (i.e., SERC, ORNL, and MLBS; Figure 18). However, topographic variables did not follow these same patterns, showing variation independent of broadleaf or needleleaf dominated stands. Functional diversity variation was uniform between sites, but phylogenetic diversity showed a wider range of values within sites consisting of a significant number of both broadleaf and needleleaf individuals (i.e., TALL and HARV) than the other broadleaf dominated sites (Figure 18). Moreover, the three dimensions of biodiversity exhibit wide ranging values within the same site.

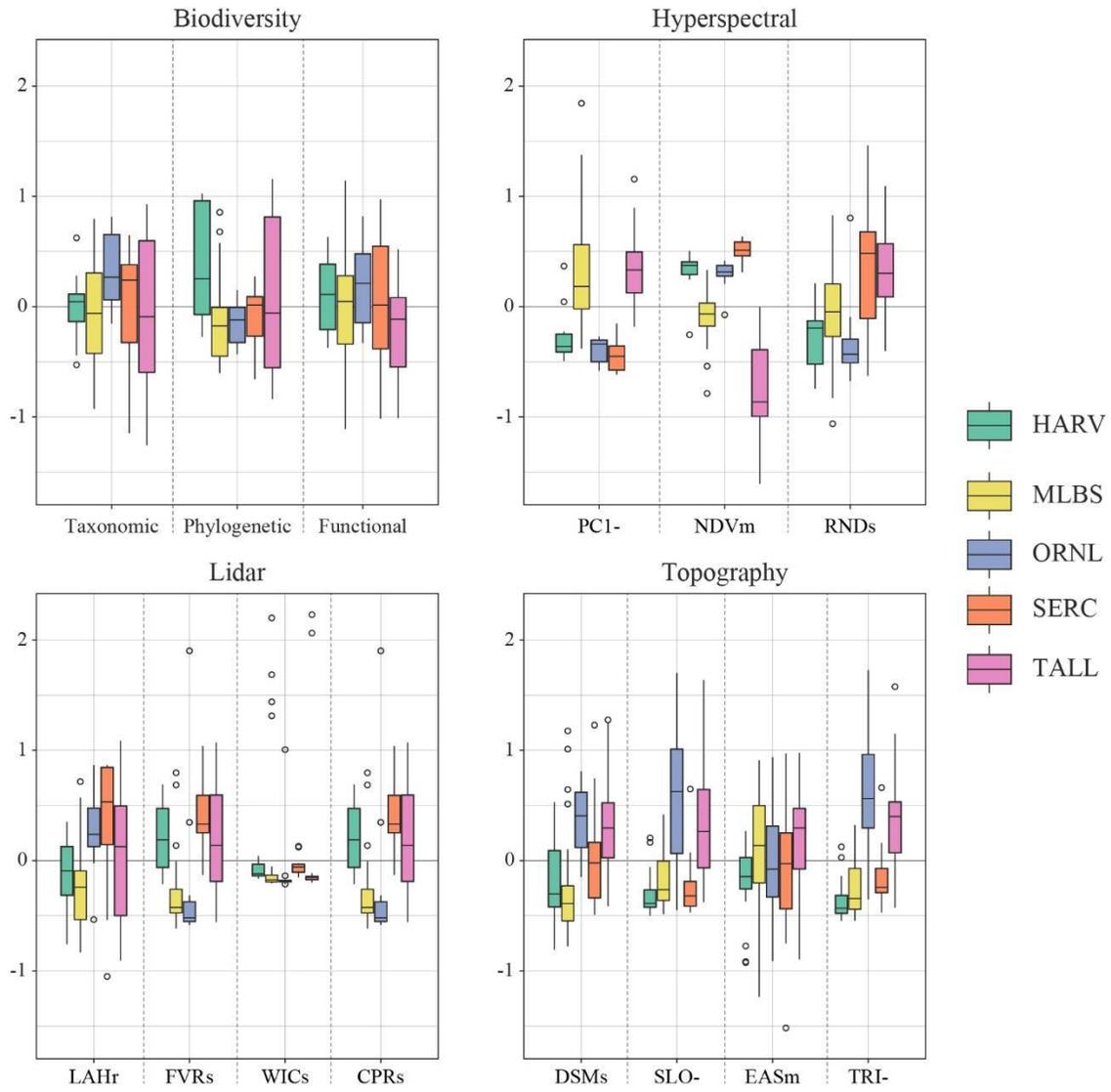


Figure 18. Biodiversity and Remote Sensing Metric Variation. Boxplots showing variation of normalized metrics used in final models. Metric abbreviations found in Table 6, with the last symbol signifying mean (*m*), minimum (*-*), maximum (*+*), range (*r*), or standard deviation (*s*).

LME Models

Models that included metrics derived from all three predictor types (i.e., hyperspectral, lidar, and topography) performed better than each individual predictor type (Figure 19; Table 6). In the best performing models, fixed effects explained all the variation in the taxonomic and functional diversity models, whereas random effects (i.e., site) had the largest influence on the phylogenetic model with the model R^2 value increasing from 0.33 to 0.70 with the inclusion of site (Table 6). The three best models included significant hyperspectral, lidar, and topographic metrics (Table 6); however, of the individual models, lidar explained the most taxonomic and functional diversity variation (Appendix B.7), showing the importance of forest structure to different dimensions of biodiversity. Given the improved model performance using all three predictor types in a single model, we used these to further examine the ability to predict different dimensions of biodiversity.

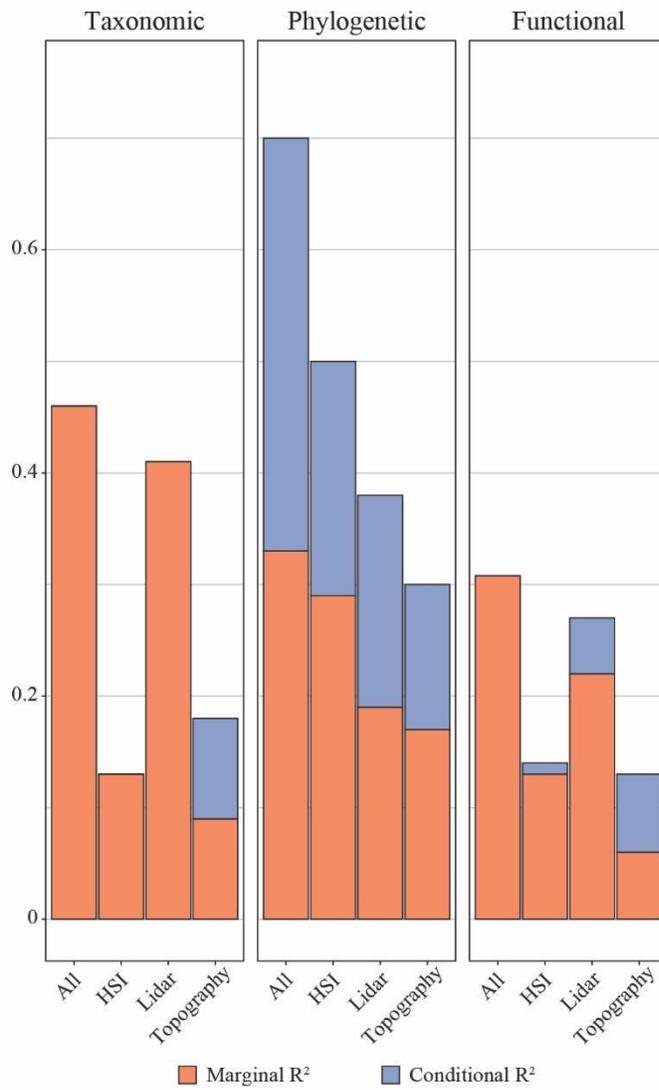


Figure 19. Marginal and Conditional R^2 Values. Marginal (only fixed effects) and Conditional (fixed and random effects) R^2 values for each model representing each sensor individually and all sensors combined.

Table 6. LMER Model Results. Final LMER Full Model results, showing marginal R^2 , conditional R^2 , and RMSE. Airborne remote sensing derived predictor variables used in final models. Metrics may include range, minimum, maximum, mean, or standard deviation and are signified as such in subsequent tables and figures. All metrics calculated in this study with references can be found in Appendix B.3, B.4, & B.5.

	Marginal R^2	Conditional R^2	RMSE	Type	Variable	Symbol	Description
Taxonomic Diversity	0.46	0.46	0.30	Lidar	Maximum Leaf Area Density Height	LAH	Height (m) of the largest leaf area density measurement within a column of the canopy
					Within Canopy Rugosity	WIC	The standard deviation of the central cell and the 8 surrounding pixels based on their LAD _{SD} measurement
					Canopy Filled Voxel Ratio	FVR	Ratio (%) of voxels within a column of the canopy that contain plant material
				Topography	Eastness	EAS	How eastward a pixel's slope is facing: derived from $\sin(\text{aspect})$
					Digital Surface Model	DSM	Elevation (m) of all objects on Earth's Surface
					Slope	SLO	Slope of pixel (degrees)
Hyperspectral	Principal Component 1	PC1	Value of the first principal component				
Phylogenetic Diversity	0.33	0.70	75.18	Lidar	Maximum Leaf Area Density Height	LAH	Height (m) of the largest leaf area density measurement within a column of the canopy
				Topographic	Topographic Roughness Index	TRI	The mean difference between the central cell and the 8 surrounding cells
				Hyperspectral	Normalized Difference Vegetation Index	NDV	Vegetation health remote sensing index
					Red-Edge Normalized Difference Vegetation Index	RND	Vegetation health remote sensing index
Functional Diversity	0.31	0.31	0.03	Lidar	Maximum Leaf Area Density Height	LAH	Height (m) of the largest leaf area density measurement within a column of the canopy
					Canopy Porosity Ratio	CPR	Ratio (%) of voxels within a column of the canopy that do not contain plant material
				Topographic	Slope	SLO	Slope of pixel (degrees)
				Hyperspectral	Principal Component 1	PC1	Value of the first principal component

To further understand model performance we examined the residuals, normalized coefficient values, and the performance of each group of metrics in the final models (Figure 20). All models showed randomly dispersed residuals that were not clustered by functional group (i.e., broadleaf, needleleaf, or mixed forest) or by site (Figure 20A; Appendix B.7). All models included the range of maximum LAD heights metric, and several of the models included similar metrics (i.e., minimum slope and the minimum first PC). Moreover, metrics representing an individual

sensor were not universally positive or negative, instead showing a wide range of influence on each of the final models (Figure 20B; Appendix B.7). For example, the phylogenetic diversity model's hyperspectral metrics had positive coefficients, while both taxonomic and functional diversity model's hyperspectral metrics had negative coefficients. Within the final taxonomic and functional diversity models, lidar metrics had the largest influence representing 65% and 52% of the total model R^2 respectively, while site had the largest influence on the phylogenetic model representing 60% of the total model R^2 (Figure 20C).

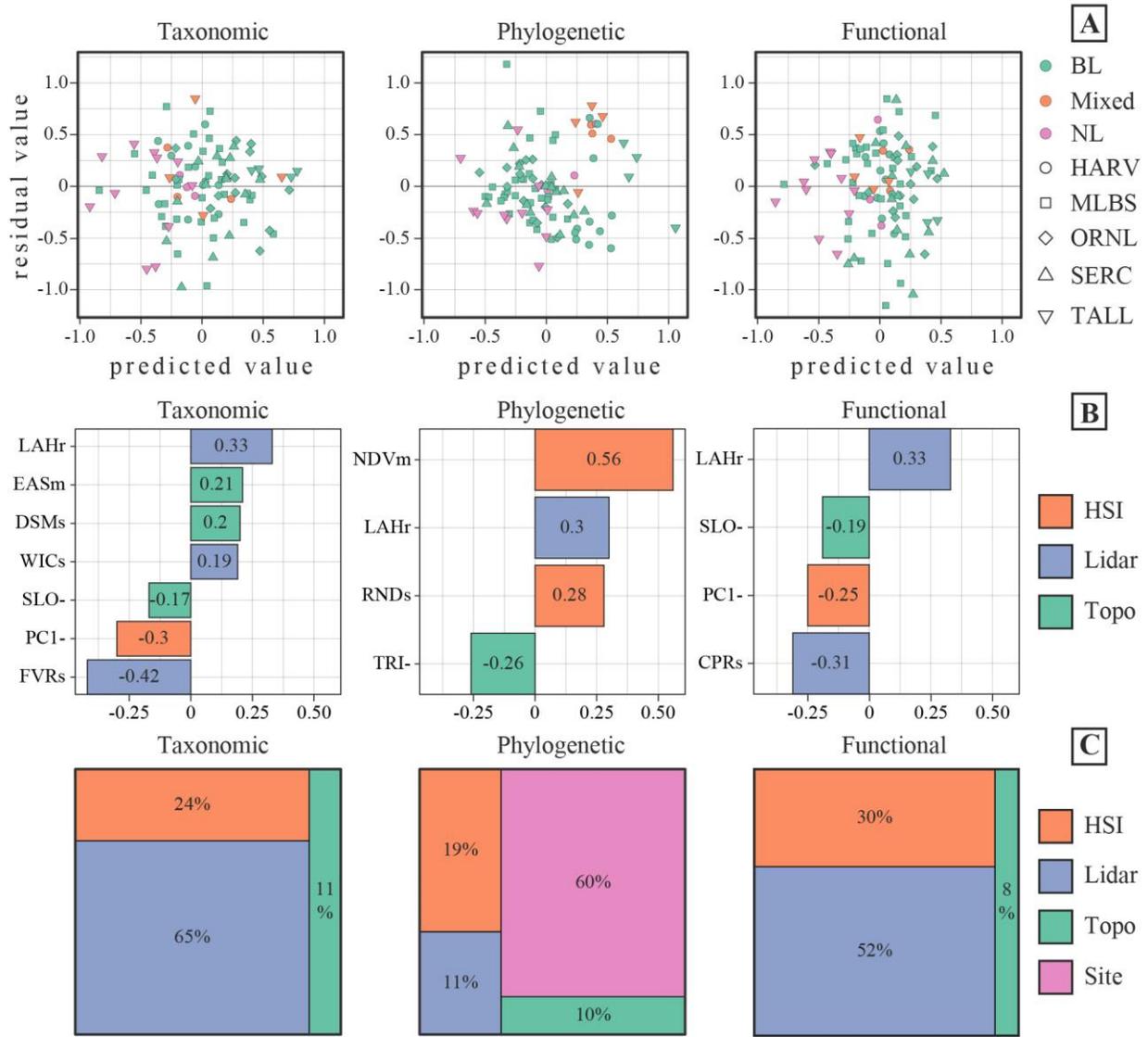


Figure 20. Final LMER Model Results. Results of final LMER Models showing (A) residuals, (B) model coefficients, and (C) model percentage of R^2 of each sensor. BL = > 66% broadleaf species, NL = > 66% needleleaf species, Mixed = mixed broadleaf and needleleaf species (i.e., between 33 and 66% broadleaf). Row B is labeled by metric abbreviations found in Table 6, with the last symbol signifying mean (m), minimum (-), maximum (+), range (r), or standard deviation (s).

Clustering Diversity and Remote Sensing Metrics

To better understand the variation within the biodiversity data we performed a PCA on the biodiversity metrics, cluster analysis, and ANOVA. PCA results showed that the first PC explained 75.83% of the variation in the field-derived biodiversity data and consisted of taxonomic, phylogenetic, and functional diversity values representing low diversity field plots (Appendix B.8). While we used only the first PC in our analysis, the inclusion of additional PCs did not change the clustering results. Moreover, we opted to use the first PC rather than the individual biodiversity values, because we were interested in biodiversity patterns in general and not each dimension of biodiversity individually for this analysis. K-means clustering resulted in two clusters, which for the first PC represented low-diversity plots (cluster 1) and high-diversity plots (cluster 2) and showed that there were high and low diversity field plots within each individual site (Appendix B.8).

To better visualize the inter- and intra-site variation of the three dimensions of biodiversity, we plotted histograms classified by cluster (Figure 21). Visual interpretation showed distinct differences between low and high diversity field plots within each site and across all sites together. Statistical analysis using ANOVA and Tukey's post hoc test confirmed that within each site, and across all sites together, these high and low diversity clusters were significantly different from one another (Appendix B.9). However, when field plot differences between individual sites were examined, there were no significant differences between taxonomic and functional diversity, and very few significant differences between phylogenetic diversity (Appendix B.9).

These relationships were not as clear when examining the variation of remote sensing variables in the high and low diversity field plots. For example, visual interpretation of Figure 21

showed almost no differences between three representative remote sensing metrics in low and high diversity field plots within each site and between all sites (additional metrics can be found in Appendix B.10, B.11, & B.12). Within each site, and across all sites together, statistical analysis confirmed that remote sensing metrics in high and low diversity field plots were not significantly different from one another with minor exceptions (e.g., metrics at TALL, one lidar variable across all sites, and several hyperspectral variables at certain sites; bolded values in Appendix B.13). Moreover, when differences between individual sites were examined a similar pattern emerges, with most metrics not showing significant differences between high and low diversity field plots (Appendix B.13).

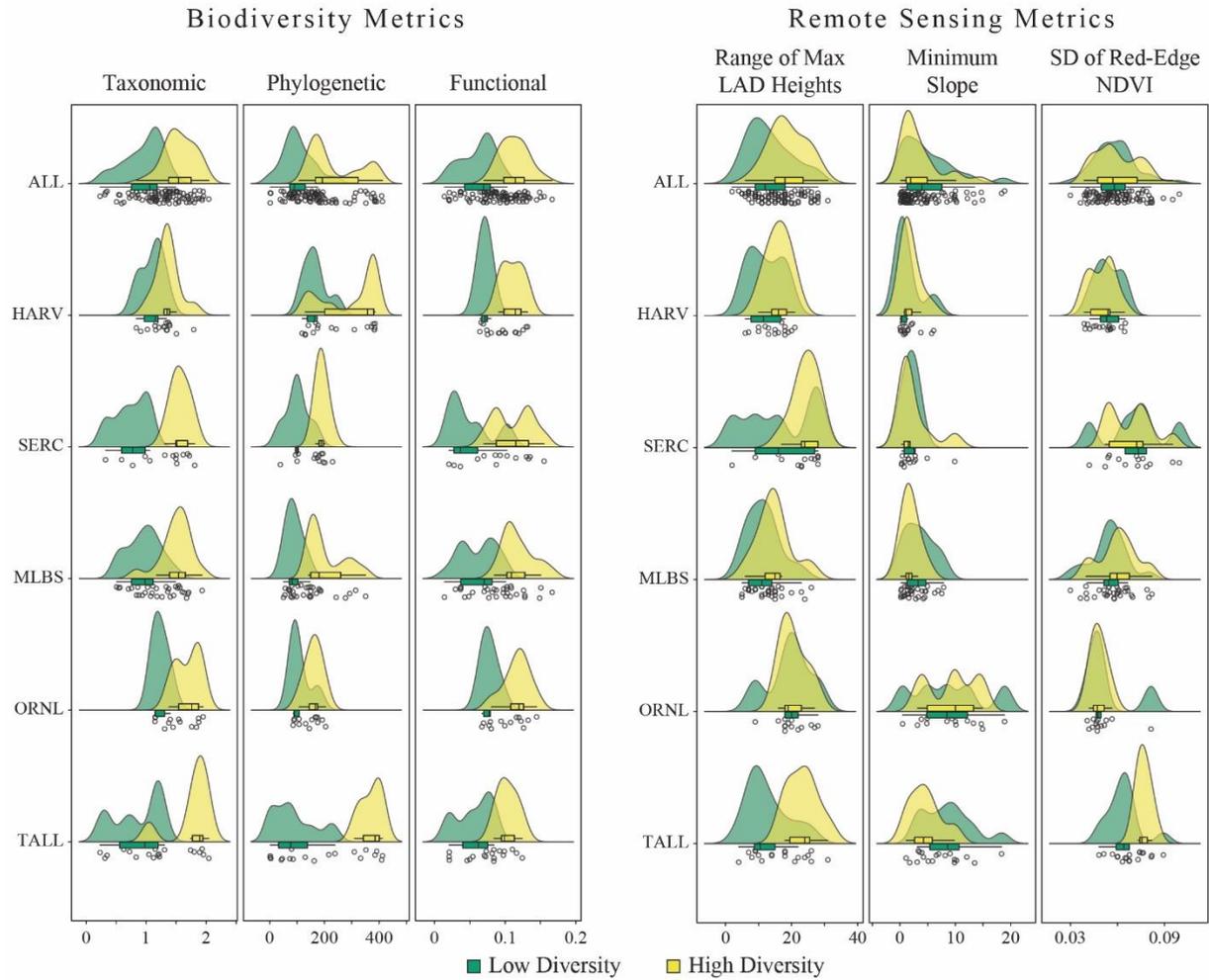


Figure 21. Clustered Metric Variation. Histograms of each dimension of biodiversity classified by each cluster and histograms of three remote sensing variables (one for each sensor) classified by each cluster. Histograms for all remote sensing variables found in the final models can be found in Appendix B.10, B.11, and B.12.

Discussion

We used airborne remote sensing to measure different dimensions of biodiversity across Eastern US temperate forest ecosystems, showing that spectral diversity, canopy structural heterogeneity, and topography together predict the spatial distribution of biodiversity metrics. This analysis demonstrates that, within the temperate forest biome explored here, while there are high and low diversity field plots within each site representing a given forest region, there are not significant differences in a high or low diversity field plot between each forest regions.

Using Remote Sensing to Measure Biodiversity

Many studies have used hyperspectral (Cavender-Bares et al. 2016; Asner and Martin 2016; Feret and Asner 2014) or lidar (Bergen et al. 2009; Cosovic et al. 2020; Simonson et al. 2012) data to measure biodiversity in a range of ecosystems; however, far fewer have relied on the fusion of these data (Leutner et al. 2012; Zhao et al. 2018). While most of these studies have focused on taxonomic diversity or leaf functional traits, we show that an integration of these two data types can be used to explain variation in different dimensions of alpha diversity.

Unsurprisingly, our models show that a combination of lidar, hyperspectral, and topographic metrics better explain biodiversity patterns than each predictor type alone (Fig. 2; Appendix B.7). However, the influence of metrics derived from each individual predictor type show that lidar derived three-dimensional forest structure has the largest influence on the estimates of taxonomic and functional biodiversity in this biome (Fig. 3C). Site-level differences had a large influence on our phylogenetic model suggesting that there are important biogeographic factors

were not possible to include in this analysis, such as species distribution maps or fine-scale soils data.

For example, while our results show that we can explain a large fraction of field plot biodiversity across this eastern temperate forest biome (Table 6), the inclusion of metrics related to soils, forest age, disturbance history, and climate could provide greater context for the results but are either not currently available across all sites or are measured at coarse spatial grains that are unsuitable for this fine-grain study. For instance, there are known differences in current and historic land use across NEON sites, which influence vegetation (e.g., prescribed burns at TALL and historic land-use at HARV (Foster 1992)).

Biodiversity Variation Across Eastern Temperate Forests

While there are significant differences between high and low biodiversity clusters within an individual site or forest region, there are very limited significant differences between individual sites (Appendix B.8 & B.9). Moreover, our PCA and clustering analysis shows that there are both high and low diversity field plots located within each site or forest region (Figure 21; Appendix B.8). This suggests that at this spatial scale and within eastern US temperate forests, biodiversity is dependent on intra-site differences related to the spatial distribution of plant functional types (i.e., broadleaf, mesic, and needleleaf dominated forest stands) within a given forest region and not differences between these same forest regions. This means that even though these forest regions consist of different combinations of tree species and traits, a given forest region is not any more or less diverse than another region and instead distributions of species and traits within a forest region drive biodiversity differences.

Remote Sensing Metrics and Biodiversity Heterogeneity

The spatial variability of remotely sensed metrics shows a complicated picture in their ability to detect differences between high and low diversity field plots (Fig. 5; Appendix B.13; Appendix B.10, B.11, & B.12). This may partially explain why our predictive models did not do a better job at detecting biodiversity within eastern US temperate forests. For example, the lidar derived variable that explains the range of maximum LAD heights is significant in all three biodiversity models. However, when considering the variation between field plots there are only significant differences at TALL, which exhibits distinct spatial patterns related to the distribution of needleleaf and broadleaf species across the landscape (Kamoske et al. 2020), and across all sites together. This suggests that in more broadleaf-dominated closed canopy forests, there are not significant differences between the distribution of leaf material within the canopy in high and low diversity plots.

Similar patterns of variation, as described above, are found across all other lidar, hyperspectral, and topographic metrics used in our study (Appendix B.13; Appendix B.10, B.11, & B.12). These differing relationships may be driven by a combination of the remotely sensed data being collected during peak greenness at each site, making subtle differences in lidar and hyperspectral derived metrics more difficult to detect due to each site being collected at different months, and by our study only focusing on a single biome consisting of a similar mix of tree species between field sites. Further research into the impacts of including other unique and diverse biomes and forest regions, using remotely sensed data collected at different times of the year to capture phenological differences, and the development of additional remote sensing-based metrics is needed to better explain the relationship between canopy observations and dimensions of biodiversity.

Model Uncertainty and Data Concerns

Remote sensing studies can contain many possible sources of error and uncertainty related to scale, sensors, and statistical methodologies and the remote sensing data used in this study was collected over multiple years using two different lidar sensors. Using lower pulse density lidar data also requires a coarser spatial resolution (i.e., 10x10m; Kamoske et al 2019) than the hyperspectral data (i.e., 1x1m). While these data are derived at these nominal resolutions, they are ultimately aggregated to match the spatial grain of the field plots (i.e., 40x40m) and may be representing processes occurring at different scales.

Due to these limitations with the lidar data we ignored the lowest 5m of the forest canopy (Kamoske et al. 2019), which is where many important understory shade tolerant plant species occur (Valladares et al. 2016); however, our LAD estimates are within the ranges reported in field-based studies (Parker and Tibbs, 2003; Brown and Parker, 2004). As current lidar sensors within the NEON AOP are upgraded, we will be able to ask important questions about the role of the understory in these different dimensions of biodiversity.

In this study we filtered all field data for only living trees to match our remote sensing variables which only considered healthy green forest vegetation. This may partially explain the weaker statistical relationships between our field derived biodiversity metrics and our remote sensing variables due to this reduced heterogeneity. More research will be needed into how unhealthy vegetation and stressed and/or disturbed environments impact these relationships.

Looking Forward

There is a unique opportunity for researchers to ask and answer questions related to the spatial distribution of different dimensions of biodiversity not only within a given biome, but also across continents with airborne and spaceborne hyperspectral and lidar platforms like the NEON AOP, NASA Goddard's Lidar, Hyperspectral, & Thermal Imager (G-LiHT; Cook et al. 2013), the Global Ecosystem Dynamics Investigation (GEDI; Stavros et al. 2017), and the proposed Surface Biology and Geology Mission (SBG; National Academies of Sciences, Engineering, and Medicine 2018).

To support these ongoing questions about remote sensing of biodiversity, we present a reproducible methodology to calculate lidar, hyperspectral, and topographic derived metrics that we show are related to different dimensions of alpha diversity within this temperate forest biome representing different forest regions. We show that a fusion of metrics derived from these different remote sensing sensors perform better at measuring biodiversity than each predictor type alone and that forest structure plays a significant role in all models. Moreover, our results suggest that while there are significant intra-site differences between our biodiversity variables due to differing local forest stand types (i.e., broadleaf, mesic, and needleleaf), there are also very few significant differences across individual sites or forest regions. While more research is needed to test these relationships across different ecoregions and at continental scales, the ever-increasing availability of hyperspectral and lidar data will provide new and exciting opportunities.

Conclusions

Lidar derived metrics related to three-dimensional forest structure, hyperspectral derived spectral diversity metrics related to foliar chemistry and health, and lidar derived topographic variables can

be used to estimate different dimensions of alpha diversity within this eastern US temperate forest biome, with forest structure having a large influence on all models. Further examination of the variation of remote sensing metrics within high and low diversity field plots shows a complicated relationship that does not mimic the patterns found in the biodiversity metrics. This suggests that these commonly used remotely sensed metrics do not completely explain the effects of each dimension of biodiversity on observable canopy properties.

While this study focuses on a single biome representing multiple forest regions at the fine spatial grain of individual field plots, these findings can be applied to studies focused on continental scales. With an abundance of hyperspectral and lidar data being collected across a variety of biomes with new space- and airborne remote sensing platforms, we have an opportunity to expand these methodologies to unlock important insights into how different dimensions of biodiversity vary and respond to global change.

Acknowledgments

Thank you to NEON, HARV, MLBS, ORNL, SERC, and TALL and their respective staff members for providing data and site access and to L. Brissette, O. Jain, S. Igwe, and R. Nagelkirk for helping in the field. This work was supported in part by the NSF DEB awards #1702379, #1926567, and #1926568. The National Ecological Observatory Network is a program sponsored by the National Science Foundation and operated under cooperative agreement by Battelle Memorial Institute. This material is based in part upon work supported by the National Science Foundation through the NEON Program. SR was partially supported by the Bryn Mawr College K.G. Research Fund. SPS was partially supported by the United States Department of Energy contract No. DE-SC0012704 to Brookhaven National Laboratory. QDR was supported by the

National Socio-Environmental Synthesis Center (SESYNC) under funding received from NSF DBI-1639145.

Data Availability

Lidar and HSI data are available at: <http://data.neonscience.org>. R package to estimate structural traits from airborne lidar data is provided through our GitHub at:

<https://github.com/akamoske/canopyLazR>. R package to pre-process HSI data is provided

through our GitHub at: <https://github.com/akamoske/hypRSpec>. R code to calculate

hyperspectral metrics is provided through our GitHub at: <https://>

github.com/akamoske/SpectralDiversity.

CHAPTER 5. CONCLUSIONS

Summary of Results

The preceding chapters present several new findings which utilize contemporary technologies and novel methodologies. These findings support the use of airborne hyperspectral and lidar remote sensing for monitoring and measuring critical forest processes and biodiversity. At the same time, these findings support the field of Geography by building upon the biogeographical framework laid out by Alexander von Humboldt, developing new geospatial methodologies that can be utilized by GIScientists, and acknowledging and quantifying the role of human development in temperate forest ecosystems.

Chapter 2 indicates that moderate and low pulse density point clouds derived from airborne lidar can be used to successfully estimate three-dimensional forest structure in closed canopy forest ecosystems (Kamoske et al. 2019). Chapter 3 shows that within-canopy functional traits can be predicted from airborne remote sensing and that, in contrast with traditional measurements of top-of-canopy N values, total canopy N variation is dampened across the landscape resulting in relatively homogenous spatial patterns (Kamoske et al. 2020). Chapter 4 suggests that commonly used remotely sensed metrics do not completely explain the effects of biodiversity on observable canopy properties. Moreover, there are high and low diversity field plots within each site representing a given forest region, but no significant differences in measures of taxonomic, phylogenetic, or functional diversity between these field plots across forest regions.

Together these findings from Chapters 2, 3, and 4 address a fundamental goal in community ecology – to understand and predict the spatial distributions of species, traits, and biodiversity across ecosystems (Keddy 1992) – by showing that airborne lidar and hyperspectral remote sensing data can be used to estimate the variation of forest functional and structural traits

across landscapes and within forest canopies, while examining the influence of biogeographic and management regimes on these traits. Moreover, forest structural and functional diversity drive critical canopy processes related to carbon sequestration (Parker et al. 2004; Hardiman et al. 2001; Ellsworth and Reich 1993; Baldocchi et al. 1998; Hardiman et al. 2013) and are directly impacted by different dimensions of biodiversity (Baiser et al. 2012; Olden and Rooney 2006; Cavanaugh et al. 2014; Baiser and Lockwood 2011; Flynn et al. 2011; Cavender-Bares et al. 2009). In previous work, structure and function have rarely been considered together at ecosystem scales (Kamoske et al. 2020). By not considering forest structural and functional traits together, landscape, continental, and global models may be misrepresenting these fine resolution ecological processes (Bonan et al. 2014; Bonan et al. 2012). New space- and airborne remote sensing platforms collecting hyperspectral and lidar data across the world offer an exciting opportunity to expand on this dissertation's findings to think about the terrestrial carbon cycle in three dimensions and unlock important insights into how forests function in a time of rapid anthropogenic and environmental change (Cook et al. 2013; Kampe et al. 2010; Stavros et al. 2017; National Academies of Sciences, Engineering, and Medicine 2018; Jetz et al. 2019).

This dissertation offers a field and remote sensing based assessment of the role of three-dimensional observable canopy traits on forest processes and biodiversity, the results from which indicate that airborne remote sensing can be used in a variety of novel methodologies to better understand the spatial distribution of forest processes and dimensions of biodiversity. Suggestions for future research outlined in the next section could expand upon and solidify these findings.

Recommendations for Future Research

While more research is needed to test the findings described throughout this dissertation in different biomes and across larger latitudinal gradients, the ever-increasing availability of

hyperspectral and lidar data will provide new and exciting opportunities. These opportunities will raise several questions related to the drivers of canopy functioning and different dimensions of biodiversity. For example: A) What is the role of soil nutrient availability, unhealthy vegetation, disturbed environments, and changing climates on critical canopy processes driven by biodiversity and biogeography? And B) Do these relationships hold when other unique biomes are considered such as boreal forests, grasslands, or savanna ecosystems? By addressing these questions in future research, we can continue to assess the importance of considering forest structural and functional diversity together to depict canopy and ecosystem processes with more detail, accuracy, and precision.

APPENDICES

APPENDIX A

Chapter 3 Supplementary Materials

Figure A.1. Field data from TALL. This shows within canopy variation of LMA and %N. Canopy positions (Bottom, Middle, Top) were designated via visual assessment in the field.

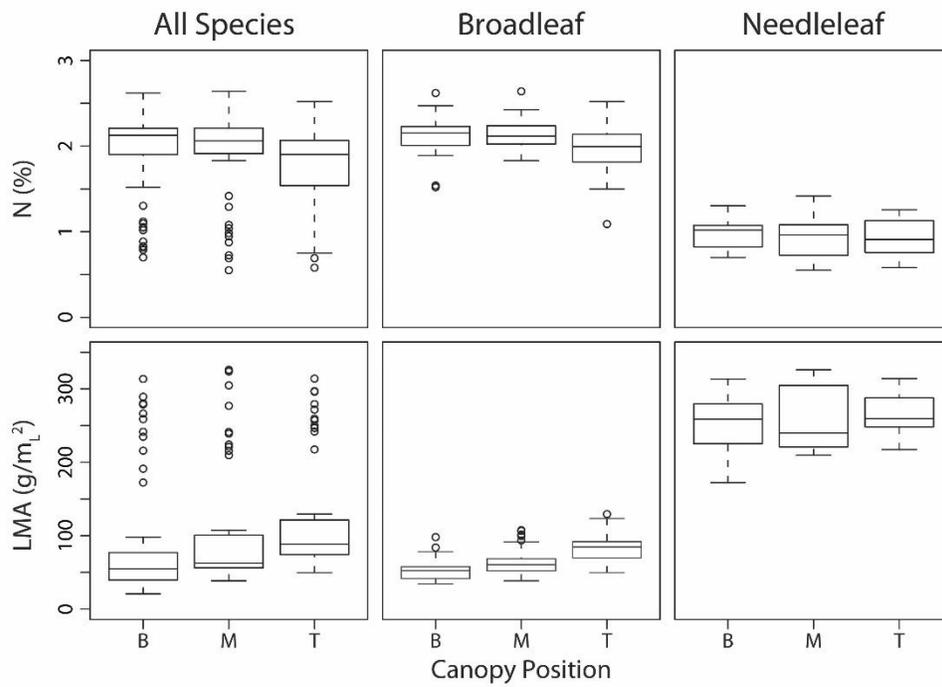


Table A.2. Mean and SD for field samples (%N and LMA). These are categorized by general position in the canopy which was determined by visual assessment.

	Bottom		Middle		Top		All	
	%N	LMA	%N	LMA	%N	LMA	%N	LMA
<i>Carya glabra</i> - pignut hickory	1.71 ± 0.43	47.7 ± 10.69	1.98 ± 0.06	48.28 ± 5.24	1.91 ± 0.2	90.29 ± 18.93	1.94 ± 0.18	53.59 ± 16.53
<i>Carya tomentosa</i> - mockernut hickory	1.95 ± 0.15	54.83 ± 13.88	1.96 ± 0.11	55.8 ± 10.11	1.86 ± 0.18	72.5 ± 17.1	1.93 ± 0.15	60.72 ± 15.2
<i>Liquidambar styraciflua</i> - sweetgum	2.04 ± 0.22	42.41 ± 9.11	2.12 ± 0.15	52.98 ± 10.75	1.85 ± 0.42	78.61 ± 8.41	2 ± 0.29	58 ± 18.01
<i>Liriodendron tulipifera</i> - tulip tree	2.48 ± 0.13	45.77 ± 8.15	2.49 ± 0.09	52.78 ± 8.7	2.25 ± 0.12	64.42 ± 13.37	2.41 ± 0.16	54.33 ± 12.45
<i>Pinus palustris</i> - longleaf pine	0.99 ± 0.29	260.81 ± 27.96	0.83 ± 0.34	291.18 ± 42.26	0.72 ± 0.26	281.82 ± 25.47	0.86 ± 0.3	276.86 ± 33.01
<i>Pinus taeda</i> - loblolly pine	1.06 ± 0.24	235.75 ± 56.97	0.96 ± 0.29	225.36 ± 14.15	1.1 ± 0.18	253.72 ± 25.74	1.04 ± 0.23	239.24 ± 35.94
<i>Quercus alba</i> - white oak	1.92 ± 0.1	57.13 ± 2.31	1.87 ± 0.15	66.02 ± 9.03	1.78 ± 0.12	82.15 ± 9.24	1.86 ± 0.13	68.43 ± 12.81
<i>Quercus falcata</i> - Southern red oak	1.81 ± 0.28	53.11 ± 6.47	2 ± 0.13	78.54 ± 14.33	1.88 ± 0.2	90.89 ± 7.2	1.9 ± 0.21	74.18 ± 18.71
<i>Quercus marilandica</i> - blackjack oak	1.83 ± 0.12	76.52 ± 19.18	1.87 ± 0.11	94.46 ± 21.19	1.61 ± 0.17	111.82 ± 17.51	1.77 ± 0.17	94.26 ± 23.31
<i>Quercus montana</i> - chestnut oak	1.97 ± 0.12	47.95 ± 6.68	2.06 ± 0.1	54.89 ± 15.37	1.99 ± 0.19	77.23 ± 14.1	2.01 ± 0.14	60.02 ± 17.44

Figure A.3. PLSR output from laboratory %N estimation.

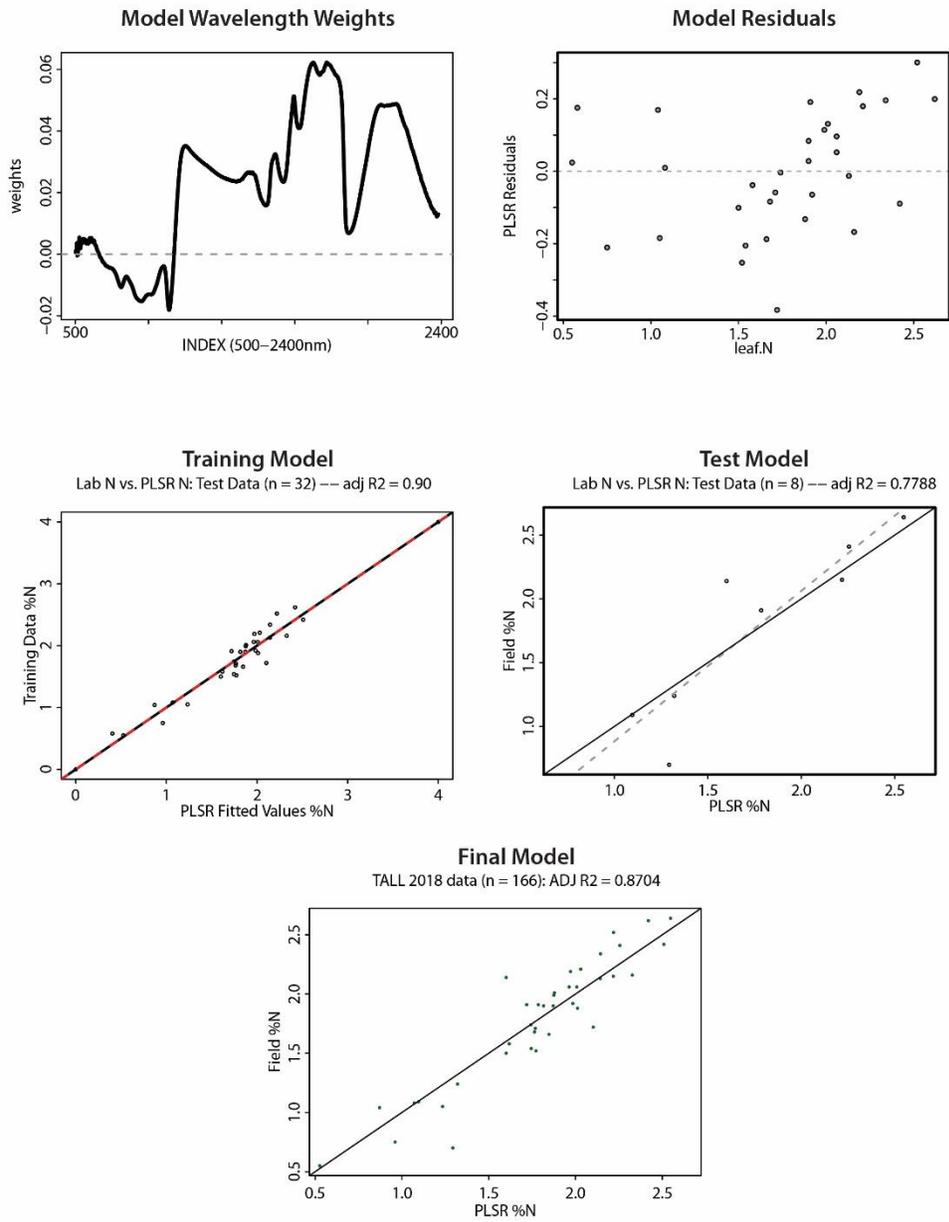


Figure A.4. PLSR output from HSI %N estimation.

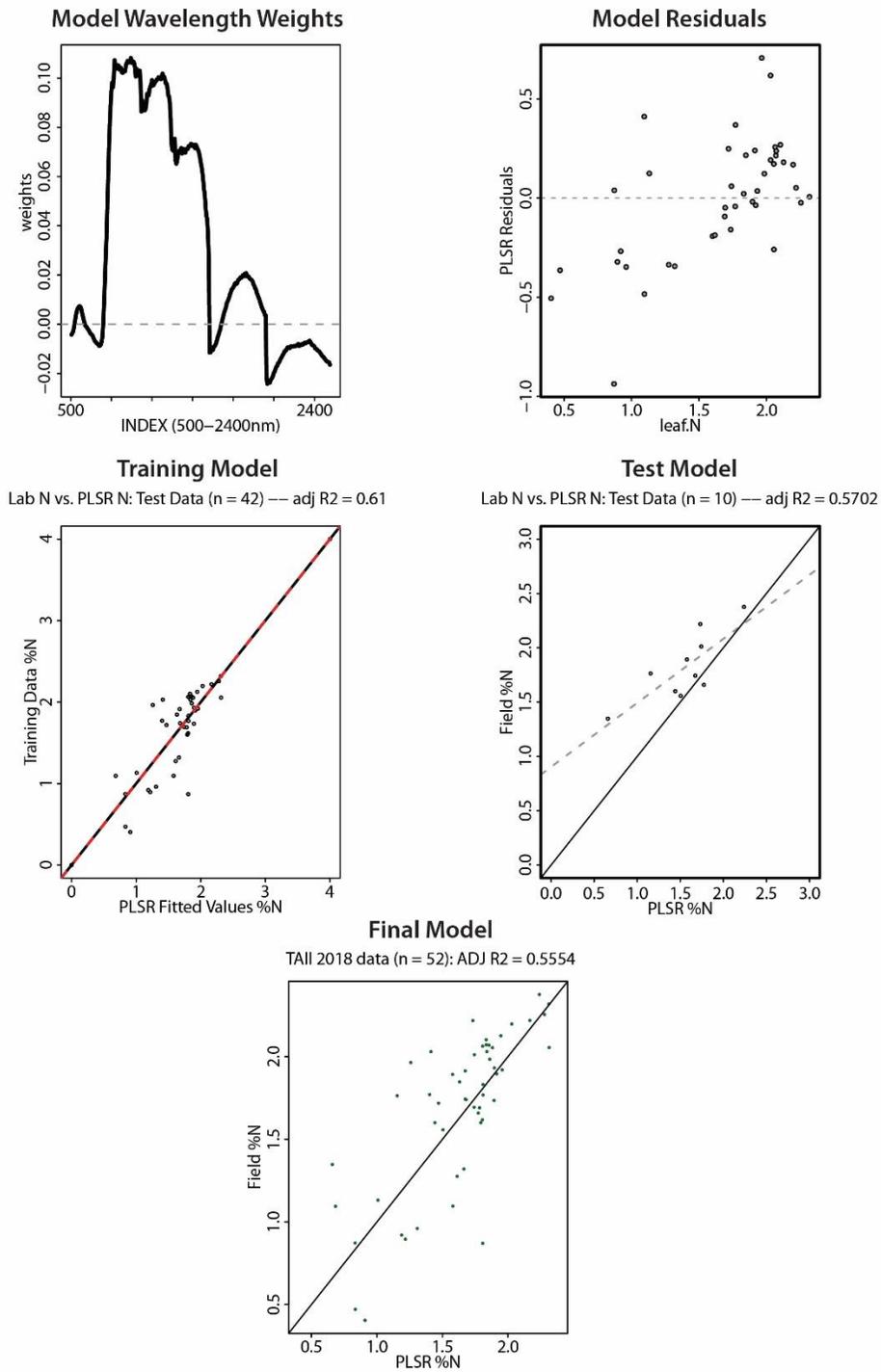


Figure A.5. PLSR output from HSI LMA estimation.

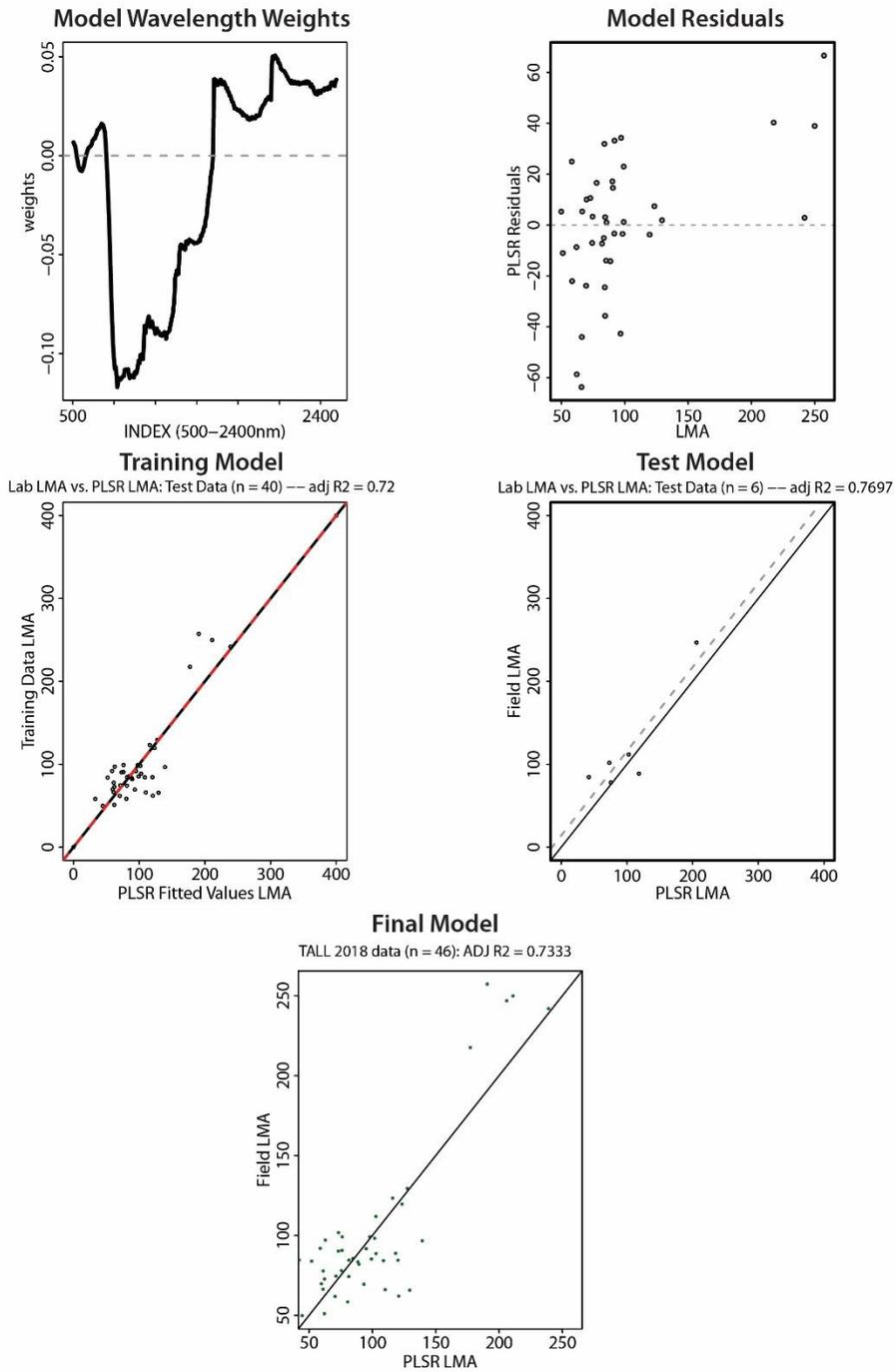


Table A.7. Abiotic and management variables. Names and references.

	Variable	Source
Topographic Variables	Digital Terrain Model	NEON AOP Lidar; R programming language
	Eastness - Aspect	NEON AOP Lidar; QGIS programming language
	Flow Accumulation	NEON AOP Lidar; ArcGIS programming language
	Meters from Northern Collection Boundary	NEON AOP Lidar; R programming language
	Meters from Western Collection Boundary	NEON AOP Lidar; R programming language
	Northness - Aspect	NEON AOP Lidar; QGIS programming language
	Slope	NEON AOP Lidar; QGIS programming language
	Soil Wetness Index	NEON AOP Lidar; QGIS programming language
	Solar Radiation - Summer Solstice	NEON AOP Lidar; ArcGIS programming language
	Solar Radiation - Winter Solstice	NEON AOP Lidar; ArcGIS programming language
	Surface Roughness	NEON AOP Lidar; QGIS programming language
	Topographic Position Index	NEON AOP Lidar; R programming language
	Topographic Roughness Index	NEON AOP Lidar; R programming language
Geologic Variables	Alluvial Substrate	Horton 2017; https://doi.org/10.5066/F7WH2N65
	Coker Substrate	Horton 2017; https://doi.org/10.5066/F7WH2N65
	Eutaw Substrate	Horton 2017; https://doi.org/10.5066/F7WH2N65
	Gordo Substrate	Horton 2017; https://doi.org/10.5066/F7WH2N65
Management Variables	Area burned in 2018 before NEON AOP flights	https://data.fs.usda.gov/geodata/edw/datasets.php
	Times burned since 2007 (first year of data)	https://data.fs.usda.gov/geodata/edw/datasets.php
	Times chemically treated since 2011	https://data.fs.usda.gov/geodata/edw/datasets.php
	Times clear cut since 1991 (first year of data)	https://data.fs.usda.gov/geodata/edw/datasets.php
	Times thinned since 1993 (first year of data)	https://data.fs.usda.gov/geodata/edw/datasets.php
	Years since last chemical treatment	https://data.fs.usda.gov/geodata/edw/datasets.php
	Years since last clear cut	https://data.fs.usda.gov/geodata/edw/datasets.php
	Years since last forest thinning treatment	https://data.fs.usda.gov/geodata/edw/datasets.php
	Years since last prescribed burn	https://data.fs.usda.gov/geodata/edw/datasets.php

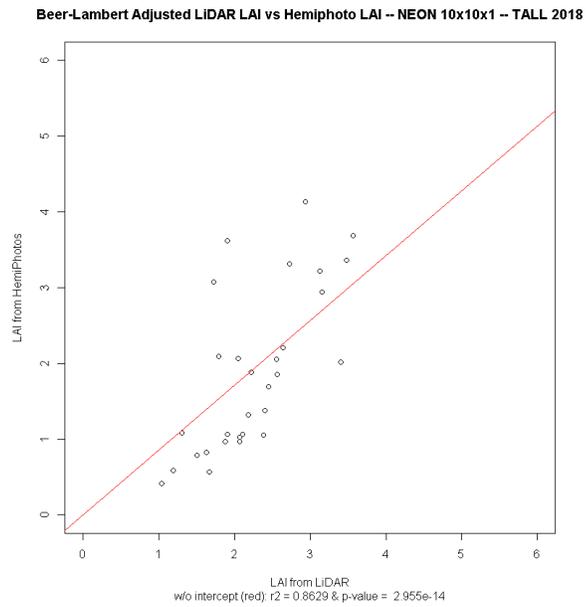
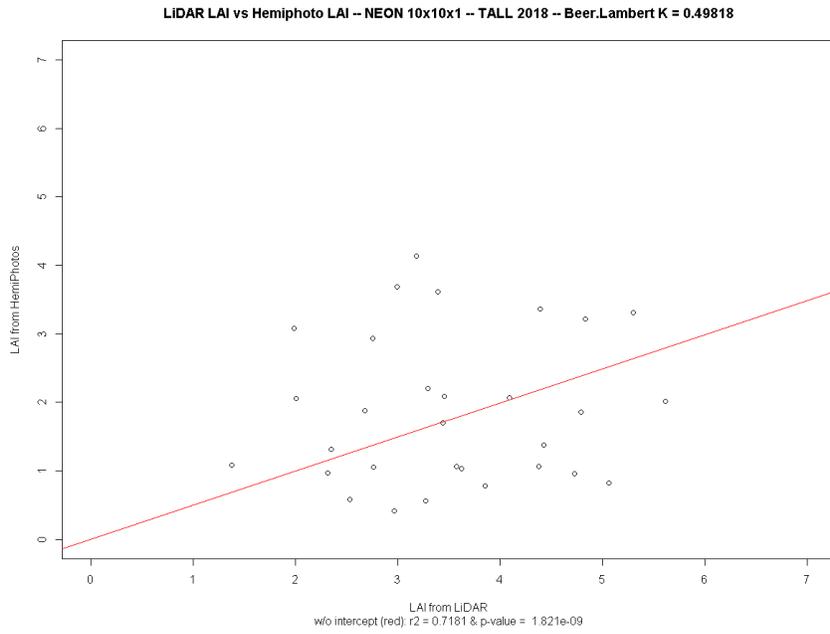
Figure A.8. LAI plots. Top Figure shows the relationship between LAI derived from hemispherical photographs and raw LAI derived from lidar, used to calculate a Beer Lambert extinction coefficient ($R^2 = 0.7181$). Bottom figure shows the relationship between Beer-Lambert adjusted LAI derived from lidar and LAI derived from hemispherical photographs ($R^2 = 0.8629$). LAD is calculated as the following:

Within each voxel, LAD is estimated as:

$$LAD_{i-1,i} = \ln \left(\frac{S_e}{S_t} \right) \frac{1}{k\Delta z}$$

where for each vertical column of voxels, i is a voxel in a sequentially ordered vertical column of the canopy, S_e is the number of pulses entering the given voxel, S_t is the number of pulses exiting the same voxel, k is an extinction coefficient, and z represents the height of a voxel. Together, the term $1/k\Delta z$ represents a Beer-Lambert Law extinction coefficient, which relates reflectance and absorbance of light to the thickness and angle of a surface. Thus, as the canopy becomes denser and more leaves are encountered, the penetration of lidar pulses will diminish causing sample sizes for estimating LAD to decrease and error to increase.

Appendix A.8. Continued.



APPENDIX B

Chapter 4 Supplementary Materials

Table B.1. Field site information. Names and locations, abbreviations, and attributes, listed from south to north. MAP = mean annual precipitation (mm); MAT = mean annual temperature (°C); MCH = mean canopy height (m); Area = total area of AOP collection (km²); Collection dates refer to when the NEON AOP collected airborne remote sensing data; Lidar System refers to the lidar sensor brand.

Name and State	Abbr.	Latitude (Degrees)	Elev. (m)	MAP (mm)	MAT (c)	MCH (m)	Area (km ²)	Collection Dates	Lidar System
Harvard Forest, Massachusetts	HARV	42.54	351	967	8	26	340	Aug. 28 - Sept. 05, 2018	Gemini
Smithsonian Environmental Research Center, Maryland	SERC	38.89	15	1107	14	38	110	July 31, 2017	Gemini
Mountain Lake Biological Station, Virginia	MLBS	37.38	1126	1030	13	18	170	June 15, 2018	Riegl
Oak Ridge National Laboratory, Tennessee	ORNL	35.96	334	1222	15	28	355	May 11 - May 13, 2018	Riegl
Talladega National Forest, Alabama	TALL	32.95	135	1350	17	25	150	April 27 - April 30, 2018	Riegl

Table B.2. Functional traits and phylogeny.

Trait	Unit	Reference
Bark Thickness	cm	Kattge et al. 2020; Stevens et al. 2020
Specific Leaf Area	$\text{mm}^2 \text{mg}^{-1}$	Kattge et al. 2020; Stevens et al. 2020
Leaf N Content by Mass	mg g^{-1}	Kattge et al. 2020; Stevens et al. 2020
Leaf P Content by Mass	mg g^{-1}	Kattge et al. 2020; Stevens et al. 2020
Leaf Thickness	$\text{mm}^2 \text{mg}^{-1}$	Kattge et al. 2020; Stevens et al. 2020
Stomatal Conductance per unit Leaf Area	$\text{mmol m}^{-2} \text{s}^{-1}$	Kattge et al. 2020; Stevens et al. 2020
Photosynthetic Rate per unit Leaf Area	$\text{umol m}^{-1} \text{s}^{-1}$	Kattge et al. 2020; Stevens et al. 2020
Specific Stem Density (Wood Density)	mg mm^{-3}	Kattge et al. 2020; Stevens et al. 2020
Seed Dry Mass	mg	Kattge et al. 2020; Stevens et al. 2020
Rooting Depth	m	Kattge et al. 2020; Stevens et al. 2020
Maximum Lifespace	years	Kattge et al. 2020; Stevens et al. 2020
Tree Species Phylogeny	NA	Potter and Koch 2014; Potter and Woodall 2012

Table B.3. Hyperspectral variables.

	Variable	Symbol	Description	References
Hyperspectral Indices	Convex Hull Volume	CHV	<i>Range based spectral diversity metric generated from the smallest possible convex hull volume within a plot</i>	Dahlin et al. 2016; Cornwell et al. 2006; Habel et al. 2015; Diaz et al. 2016; Smith et al. 2013
	Sum of Squares	SS	<i>Variance based spectral diversity metric generated from the the sum of the total spectral variance of the principal components</i>	Laliberté et al. 2020
	Total Variance	TV	<i>Variance based spectral diversity metric describing the total spectral variance of the principal components</i>	Dahlin et al. 2016
	Principal Component 1	PC1	<i>Value of the first principal component</i>	Oldeland et al. 2010
	Principal Component 2	PC2	<i>Value of the second principal component</i>	Oldeland et al. 2010
	Coefficient of Variation	CV	<i>Variance based spectral diversity metric based on the relative variability of reflectance measurements within a plot</i>	Gholizadeh et al. 2018; Gholizadeh et al. 2019
	Normalized Difference Vegetation Index	NDV	<i>Vegetation health remote sensing index</i>	Rouse et al. 1974
	Photochemical Reflectance Index	PRI	<i>Light use efficiency remote sensing index</i>	Gamon et al. 1992
	Red-Edge Vegetation Stress Index	RVS	<i>Plant stress remote sensing index</i>	Merton and Huntington 1999
	Red-Edge Normalized Difference Vegetation Index	RND	<i>Vegetation health remote sensing index</i>	Gitelson and Merzlyak 1994
	Shortwave Infrared 1 Mean Reflectance	SW1	<i>Mean reflectance of wavelength from 1500 to 1800nm</i>	Ollinger 2011
	Shortwave Infrared 2 Mean Reflectance	SW2	<i>Mean reflectance of wavelengths from 2000 to 2400nm</i>	Ollinger 2011
	Near Infrared Mean Reflectance	NIR	<i>Mean reflectance of wavelengths from 800 to 1350nm</i>	Ollinger 2011

Table B.4. Lidar derived variables.

	Variable	Symbol	Description	References
Lidar Metrics	Canopy Height	CH	<i>Height (m) of canopy within a given pixel</i>	Lefsky et al 1999
	Canopy Empty Volume	EV	<i>Volume of the space within the canopy that does not contain plant material</i>	Lefsky et al 1999
	Canopy Euphotic Depth	ED	<i>Depth (m) of the uppermost 65% of plant material within the canopy</i>	Lefsky et al 1999
	Canopy Euphotic Leaf Area	ELA	<i>Total amount of plant material within the uppermost 65% of plant material within the canopy</i>	Lefsky et al 1999
	Canopy Euphotic Volume	CEV	<i>Volume of the uppermost 65% of plant material within the canopy</i>	Lefsky et al 1999
	Canopy Filled Voxel Ratio	FVR	<i>Ratio (%) of voxels within a column of the canopy that contain plant material</i>	Hardiman et al 2013
	Leaf Area Index	LAI	<i>The one-sided leaf area per unit of ground area</i>	Chen and Black 1992
	Maximum Leaf Area Density	MLA	<i>Largest leaf area density measurement within a column of the canopy</i>	Hardiman et al 2013
	Maximum Leaf Area Density Height	LAH	<i>Height (m) of the largest leaf area density measurement within a column of the canopy</i>	Hardiman et al 2013
	Mean Leaf Area Density Height	LAM	<i>Mean height (m) of the the total amount of leaf material within a column of the canopy</i>	Hardiman et al 2013
	Canopy Oligophotic Leaf Area	OLA	<i>Total amount of plant material within the bottommost 35% of plant material within the canopy</i>	Lefsky et al 1999
	Canopy Oligophotic Volume	COV	<i>Volume of the bottommost 35% of plant material within the canopy</i>	Lefsky et al 1999
	Canopy Porosity Ratio	CPR	<i>Ratio (%) of voxels within a column of the canopy that do not contain plant material</i>	Hardiman et al 2013
	Leaf Area Density 10 th Quantile	L10	<i>Height (m) of the 10th quantile of leaf material within a column of the canopy</i>	Shi et al 2018
	Leaf Area Density 25 th Quantile	L25	<i>Height (m) of the 25th quantile of leaf material within a column of the canopy</i>	Shi et al 2018
	Leaf Area Density 50 th Quantile	L50	<i>Height (m) of the 50th quantile of leaf material within a column of the canopy</i>	Shi et al 2018
	Leaf Area Density 75 th Quantile	L75	<i>Height (m) of the 75th quantile of leaf material within a column of the canopy</i>	Shi et al 2018
	Leaf Area Density 90 th Quantile	L90	<i>Height (m) of the 90th quantile of leaf material within a column of the canopy</i>	Shi et al 2018
	Standard Deviation of Leaf Area Density	LSD	<i>The standard deviation of the leaf area density measurement within a column of the canopy</i>	Hardiman et al 2011
	Top of Canopy Rugosity	CR	<i>Sum of the intercell difference between the central cell and their CH_{lidar} measurements converted to a volume</i>	Lefsky et al 1999
Within Canopy Rugosity	WIC	<i>The standard deviation of the central cell and the 8 surrounding pixels based on their LAD_{sp} measurement</i>	Hardiman et al 2011	

Table B.5. Topographic variables.

	Variable	Symbol	Description	References
Topographic Variables	Digital Surface Model	DSM	<i>Elevation (m) of all objects on Earth's Surface</i>	QGIS
	Digital Terrain Model	DTM	<i>Elevation (m) of the Earth's Surface</i>	QGIS
	Eastness	EAS	<i>How eastward a pixel's slope is facing: derived from sin(aspect)</i>	QGIS
	Latitude	LAT	<i>Latitude (degrees) of the pixel centroid</i>	QGIS
	Longitude	LON	<i>Longitude (degrees) of the pixel centroid</i>	QGIS
	Northness	NOR	<i>How northward a pixel is: derived from cos(aspect)</i>	QGIS
	Slope	SLO	<i>Slope of pixel (degrees)</i>	QGIS
	Topographic Position Index	TPI	<i>The intercell difference between the central cell and the mean of the 8 surrounding cells</i>	QGIS
	Topographic Roughness Index	TRI	<i>The mean difference between the central cell and the 8 surrounding cells</i>	QGIS

Table B.6. Beer Lambert coefficients.

Site	Year	Beer Lambert Coefficient
TALL	2018	0.4982
ORNL	2018	0.8354
MLBS	2018	0.8776
SERC	2017	0.6784
HARV	2018	0.7796

Table B.7. Results from individual models.

	Sensor	Marginal R ²	Conditional R ²	RMSE	AIC	Variable	Coefficient
Taxonomic Diversity	All	0.46	0.46	0.36	122.7	max_lad_height_range	0.3308
						pc1_min	-0.2982
						filled_voxel_ratio_sd	-0.4171
						dsm_sd	0.1972
						within_canopy_rugosity_sd	0.1874
						slope_min	-0.1724
						eastness_mean	0.2081
	HSI	0.13	0.13	0.46	144.9	NIR_sd	0.3268
						pc1_sd	-0.2681
	Lidar	0.41	0.41	0.38	126.3	max_lad_height_range	0.3144
						filled_voxel_ratio_sd	-0.4086
						euphotic_tla_min	0.2446
						quantile_90_sd	0.2686
within_canopy_rugosity_min						-0.2167	
filled_voxel_ratio_max	-0.1804						
Topography	0.09	0.18	0.46	148.6	dsm_sd	0.2609	
					slope_min	-0.2990	
Phylogenetic Diversity	All	0.33	0.70	0.36	123.7	max_lad_height_range	0.3004
						tri_min	-0.2610
						NDVI_mean	0.5557
						reNDVI_sd	0.2823
	HSI	0.29	0.50	0.39	130.3	pc2_min	-0.3275
						NDVI_mean	0.7163
						reNDVI_sd	0.3737
						PRI_range	-0.3438
	Lidar	0.19	0.38	0.41	133.5	euphotic_tla_min	0.2785
						max_lad_height_range	0.2936
	Topography	0.17	0.30	0.42	136.1	tpi_mean	-0.2649
						tri_min	-0.3200
Functional Diversity	All	0.31	0.31	0.41	131.9	max_lad_height_range	0.3336
						porosity_ratio_sd	-0.2451
						pc1_min	-0.3058
						slope_min	-0.1884
	HSI	0.13	0.14	0.46	140.5	pc1_min	-0.3632
	Lidar	0.22	0.27	0.43	133.5	max_lad_height_range	0.4103
						porosity_ratio_sd	-0.3147
Topography	0.06	0.13	0.47	147.8	slope_min	-0.2483	

Table B.8. Results from PCA and Cluster Analysis

	PCA Results			
	PC1	PC2	PC3	
Taxonomic Diversity	-0.5948	0.4331	0.6773	Results from PCA showing normalized loading coefficients
Phylogenetic Diveristy	-0.5291	-0.8452	0.0758	
Functional Diversity	-0.6052	0.3133	-0.7318	
	PC1 = 75.83% of variation			
	PC2 = 16.89% of variation			
	PC3 = 07.29% of variation			
	Cluster Analysis Results			
	PC1			Results from Cluster Analysis showing loading values for each PC
Cluster 1	1.300			
Cluster 2	-1.147			
	Plots in Cluster 1	Plots in Cluster 2		Results from Cluster Analysis showing the number of plots in each cluster at each site
HARV	6	11		
MLBS	17	15		
ORNL	5	9		
SERC	5	9		
TALL	12	7		

Table B.9. ANOVA results for dimensions of biodiversity. Bolded values represent p-values

≤ 0.05 meaning that these sites are significantly different from one another.

ANOVA Results			
	Taxonomic	Phylogenetic	Functional
ALL	0	0	0
HARV	0.019	0.012	0
SERC	0	0	0.003
MLBS	0	0	0
ORNL	0.001	0.017	0.001
TALL	0	0	0

Table shows p-values from ANOVA examining plots within each site and within all sites.

Taxonomic Diversity				
	HARV	SERC	MLBS	ORNL
SERC	1			
MLBS	0.99	0.99		
ORNL	0.4	0.48	0.12	
TALL	0.98	0.98	1	0.15

Table shows p-values from Post Hoc Test examining between site differences.

Functional Diversity				
	HARV	SERC	MLBS	ORNL
SERC	0.98			
MLBS	0.92	1		
ORNL	0.99	0.84	0.67	
TALL	0.25	0.67	0.56	0.11

Table shows p-values from Post Hoc Test examining between site differences.

Phylogenetic Diversity				
	HARV	SERC	MLBS	ORNL
SERC	0.08			
MLBS	0	0.99		
ORNL	0.02	0.99	1	
TALL	0.43	0.84	0.4	0.57

Table shows p-values from Post Hoc Test examining between site differences.

Figure B.10. Hyperspectral variables histograms. Histograms of all hyperspectral remote sensing variables that appear in final models.

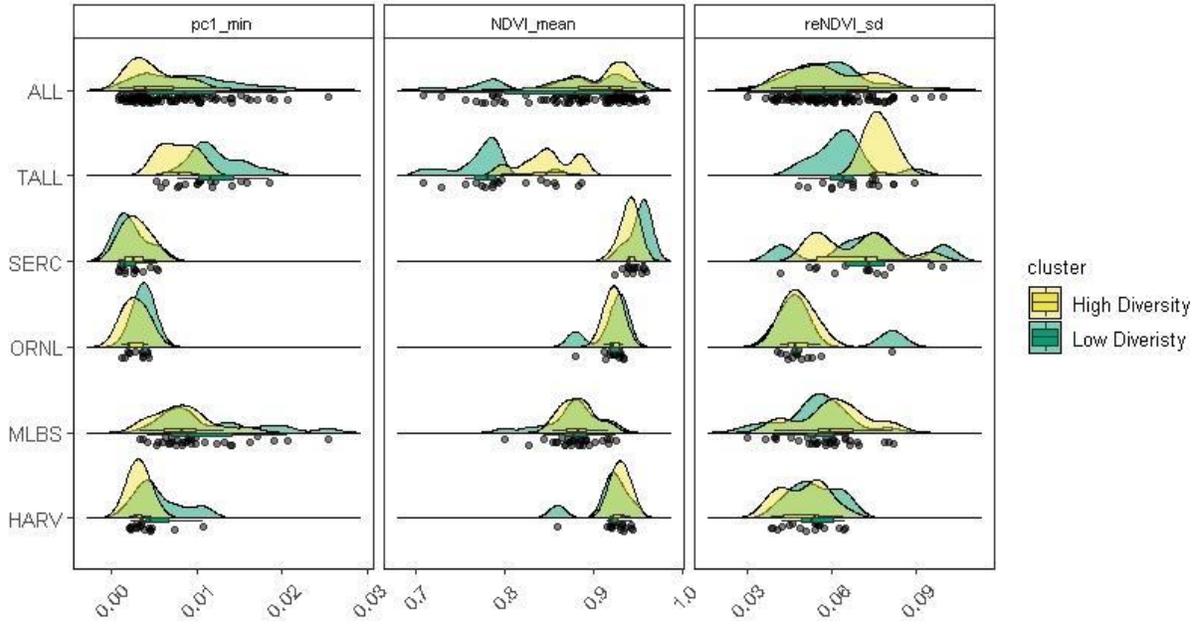


Figure B.11. Lidar variable histograms. Histograms of all lidar remote sensing variables that appear in final models.

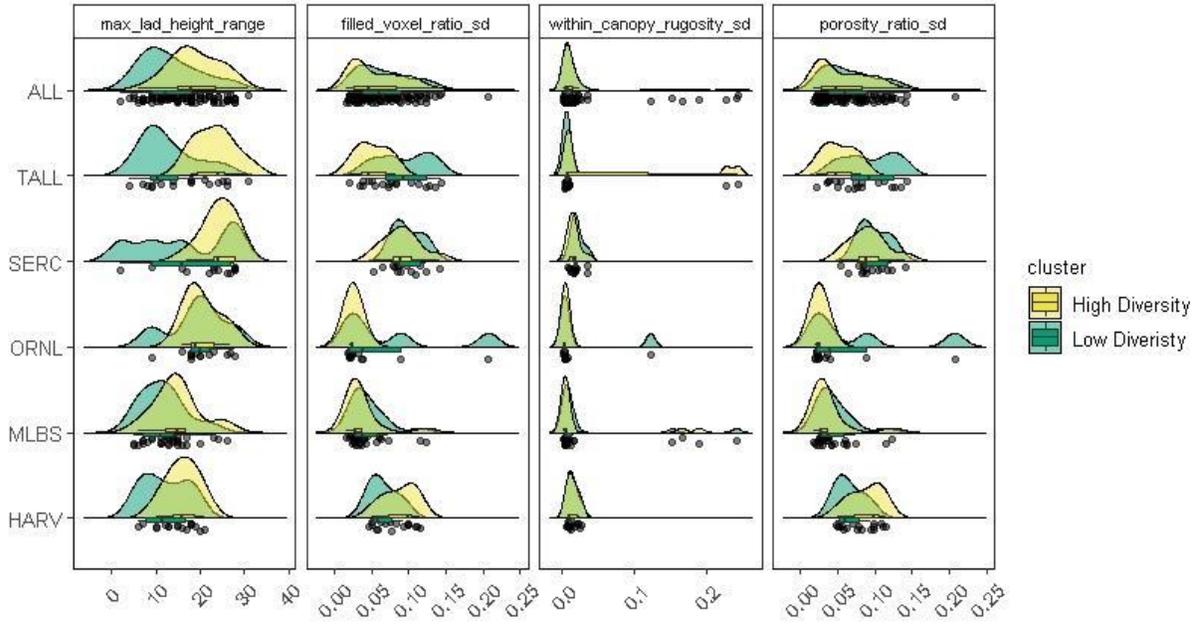


Figure B.12. Topography variables histograms. Histograms of all topographic remote sensing variables that appear in final models.

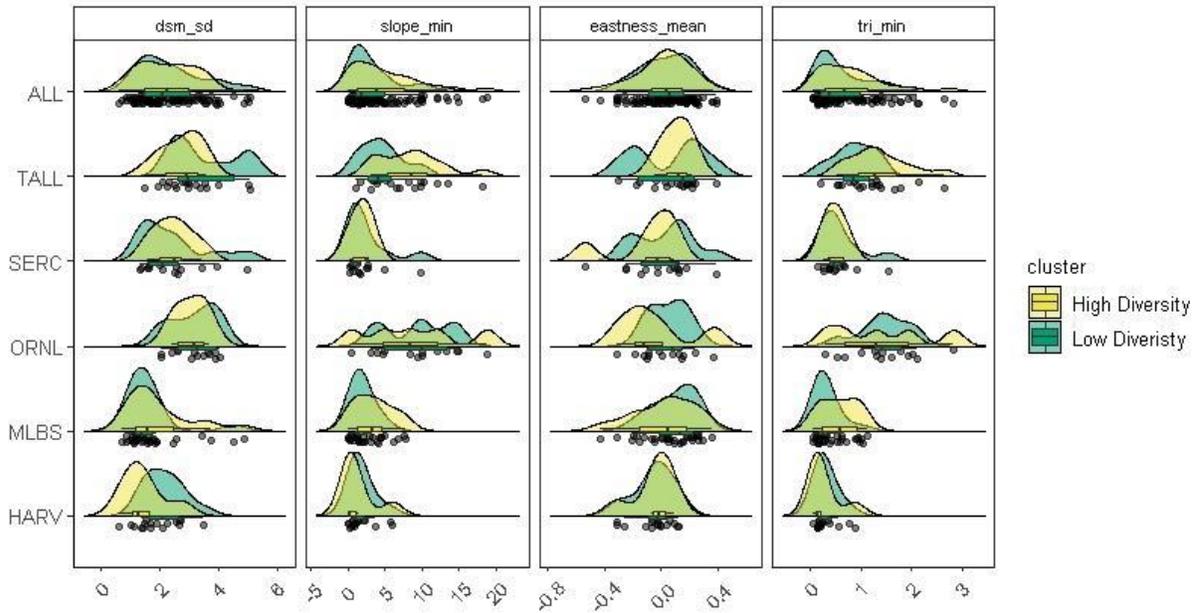


Table B.13. ANOVA results for remote sensing variables. Bolded values represent p-values

≤ 0.05

ANOVA Results											
	Lidar				Hyperspectral			Topography			
	Max LAD Height Range	Filled Voxel Ratio SD	Within Canopy Rugosity SD	Porosity Ratio SD	PC1 Min	NDVI Mean	reNDVI SD	DSM SD	Slope Min	Easterness Mean	TRI Min
ALL	0.000	0.103	0.635	0.103	0.000	0.000	0.705	0.527	0.131	0.529	0.130
HARV	0.076	0.059	0.923	0.059	0.026	0.125	0.363	0.053	0.631	0.940	0.738
SERC	0.070	0.589	0.354	0.589	0.627	0.038	0.767	0.969	0.660	0.323	0.633
MLBS	0.080	0.353	0.922	0.353	0.075	0.601	0.186	0.335	0.046	0.291	0.025
ORNL	0.691	0.074	0.206	0.074	0.110	0.640	0.391	0.803	0.900	0.350	0.895
TALL	0.002	0.013	0.047	0.013	0.003	0.000	0.005	0.061	0.048	0.523	0.074

Max LAD Height Range				Within Canopy Rugosity SD					
	HARV	SERC	MLBS	ORNL		HARV	SERC	MLBS	ORNL
SERC	0.040				SERC	1.000			
MLBS	0.964	0.001			MLBS	0.930	0.984		
ORNL	0.169	0.994	0.009		ORNL	1.000	1.000	0.919	
TALL	0.941	0.262	0.400	0.623	TALL	0.920	0.976	1.000	0.907

Filled Voxel Ratio SD				Porosity Ratio SD					
	HARV	SERC	MLBS	ORNL		HARV	SERC	MLBS	ORNL
SERC	0.872				SERC	0.872			
MLBS	0.004	0.000			MLBS	0.004	0.000		
ORNL	0.048	0.002	1.000		ORNL	0.048	0.002	1.000	
TALL	1.000	0.787	0.005	0.059	TALL	1.000	0.787	0.005	0.059

PC1 Min				reNDVI SD					
	HARV	SERC	MLBS	ORNL		HARV	SERC	MLBS	ORNL
SERC	0.948				SERC	0.001			
MLBS	0.000	0.000			MLBS	0.657	0.031		
ORNL	0.991	1.000	0.000		ORNL	0.999	0.001	0.471	
TALL	0.000	0.000	0.997	0.000	TALL	0.002	1.000	0.037	0.001

NDVI Mean				TRI Min					
	HARV	SERC	MLBS	ORNL		HARV	SERC	MLBS	ORNL
SERC	0.827				SERC	0.795			
MLBS	0.011	0.000			MLBS	0.914	0.996		
ORNL	1.000	0.737	0.044		ORNL	0.000	0.002	0.000	
TALL	0.000	0.000	0.000	0.000	TALL	0.000	0.010	0.000	0.741

DSM SD				Slope Min					
	HARV	SERC	MLBS	ORNL		HARV	SERC	MLBS	ORNL
SERC	0.640				SERC	0.999			
MLBS	0.999	0.294			MLBS	0.943	0.999		
ORNL	0.010	0.469	0.001		ORNL	0.000	0.000	0.000	
TALL	0.012	0.588	0.000	1.000	TALL	0.000	0.005	0.001	0.662

reNDVI SD				Easterness Mean					
	HARV	SERC	MLBS	ORNL		HARV	SERC	MLBS	ORNL
SERC	0.872				SERC	0.997			
MLBS	0.004	0.000			MLBS	0.397	0.803		
ORNL	0.048	0.002	1.000		ORNL	0.917	0.996	0.986	
TALL	1.000	0.787	0.005	0.059	TALL	0.229	0.577	0.992	0.886

Table shows p-values from ANOVA examining remote sensing variables at each plot within each site and within all sites.

Tables showing p-values from Post Hoc Test for all lidar variables between each pair of sites.

Tables showing p-values from Post Hoc Test for all hyperspectral variables between each pair of sites.

Tables showing p-values from Post Hoc Test for all topography variables between each pair of sites.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Anderegg, W.R.L., J.M. Kane, and L.D.L Anderegg. 2013. Consequences of widespread tree mortality triggered by drought and temperature stress. *Nature Climate Change* 3, 30–36.
- Antonarakis, A.S., J.W. Munger, and P.R. Moorcroft. 2014. Imaging spectroscopy- and lidar-derived estimates of canopy composition and structure to improve predictions of forest carbon fluxes and ecosystem dynamics. *Geophysical Research Letters* 41, 2535–2542.
- Asner, G.P., and R.E. Martin. 2009. Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. *Frontiers in Ecology and Environment* 7(5): 269–276.
- Asner, G.P., Powell, G.V.N., Mascaro, J., Knapp, D.E., Clark, J.K., Jacobson, J., KennedyBowdoin, T., Balaji, A., Paez-Acosta, G., Victoria, E., Secada, L., Valqui, M., Hughes, R.F., 2010. High-resolution forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of Sciences of the USA* 107(38): 16738-16742.
- Asner, G.P., Martin, R.E., Tupayachi, R., Emerson, R., Martinez, P., Sinca, F., Powell, G.V.N., Wright, S.J., Lugo, A.E., 2011. Taxonomy and remote sensing of leaf mass per area (LMA) in humid tropical forests. *Ecological Applications* 21(1): 85-98.
- Asner, G.P., C. Anderson, R.E. Martin, D.E. Knapp, R. Tupayachi, and T. Kennedy-Bowdoin. 2014. Landscape-scale changes in forest structure and functional traits along an Andes-to-Amazon elevation gradient. *Biogeosciences* 11: 843–856.
- Asner, G. P., R. E. Martin, C. B. Anderson, and D. E. Knapp. 2015. Quantifying forest canopy traits: Imaging spectroscopy versus field survey. *Remote Sensing of Environment* 158:15–27.
- Asner, G.P., and R.E. Martin. 2016. Spectranomics: Emerging science and conservation opportunities at the interface of biodiversity and remote sensing. *Global Ecology and Conservation* 8: 212-219.
- Atkins, J. W., R. T. Fahey, B. H. Hardiman, and C. M. Gough. 2018. Forest Canopy Structural Complexity and Light Absorption Relationships at the Subcontinental Scale. *Journal of Geophysical Research: Biogeosciences* 123:1387–1405.
- Auclair, A.N., and G. Cottam. 1971. Dynamics of Black Cherry (*Prunus serotina* Ehrh.) in Southern Wisconsin Oak Forests. *Ecological Monographs* 41(2): 153-177.

- Bachofen, C., P. D'Odorico, and N. Buchmann. 2020. Light and VPD gradients drive foliar nitrogen partitioning and photosynthesis in the canopy of European beech and silver fir. *Oecologia* 192(2): 323-339.
- Baiser, B., and J. L. Lockwood. 2011. The relationship between functional and taxonomic homogenization. *Global Ecology and Biogeography* 20:134–144.
- Baiser, B., J. D. Olden, S. Record, J. L. Lockwood, and M. L. McKinney. 2012. Pattern and process of biotic homogenization in the New Pangaea. *Proceedings of the Royal Society of London B: Biological Sciences*: rspb20121651.
- Baldocchi, D.D., B.B. Hincks, and T.P. Meyers. 1988. Measuring Biosphere-Atmosphere Exchanges of Biologically Related Gases with Micrometeorological Methods. *Ecology* 69, 1331–1340.
- Barnett, D.T., P.A. Duffy, D.S. Schimel, R.E. Krauss, K.M. Irvine, F.W. Davis, J.E. Gross, et al. 2019. The terrestrial organism and biogeochemistry spatial sampling design for the National Ecological Observatory Network. *Ecosphere* 10(2): e02540.
- Bater, C.W, and N.C. Coops. 2009. Evaluation error associated with lidar-derived DEM interpolation. *Computers & Geosciences* 35(2): 289-300.
- Becknell, J.M., A.R. Desai, M.C. Dietze, C.A. Schultz, G. Starr, P.A. Duffy, J.F. Franklin, A. Pourmokhtarian, J. Jall, P.C. Stoy, M.W. Binford, L.R. Boring, and C.L. Staudhammer. 2015. Assessing interactions among changing climate, management, and disturbance in forests: A macrosystems approach. *BioScience* 65, 263-274.
- Bengtsson, J., S.G. Nilsson, A. Franc, and P. Menozzi. 2000. Biodiversity, disturbances, ecosystem function and management of European Forests. *Forest Ecology and Management* 132(1): 39-50.
- Bergen, K.M., S.J. Goetz, R.O. Dubayah, G.M. Henebry, C.T. Hunsaker, M.L. Imhoff, R.F. Nelson, G.G. Parker, and V.C. Radeloff. 2009. Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *Journal of Geophysical Research* 114: 1-13.
- Bonan, G. B. 2008. Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science* 320:1444–1449.
- Bonan, G.B., K.W. Oleson, R.A. Fisher, G. Lasslop, and M. Reichstein. 2012. Reconciling leaf physiological traits and canopy flux data: Use of the TRY and FLUXNET databases in the Community Land Model version 4. *Journal of Geophysical Research: Biogeosciences* 117, n/a-n/a.

- Bonan, G.B., M. Williams, R.A. Fisher, and K.W. Oleson. 2014. Modeling stomatal conductance in the earth system: linking leaf water-use efficiency and water transport along the soil–plant–atmosphere continuum. *Geoscientific Model Development* 7, 2193–2222.
- Bouvier, M., S. Durrieu, R.A. Fournier, and J.P. Renaud. 2015. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sensing of Environment* 156, 322–334.
- Boyce M.S. 2006. Scale for resource selection functions. *Diversity and Distributions* 12: 269–276.
- Braghiere, R. K., T. Quaife, E. Black, L. He, and J. M. Chen. 2019. Underestimation of Global Photosynthesis in Earth System Models Due to Representation of Vegetation Structure. *Global Biogeochemical Cycles* 33:1358–1369.
- Brook, B. W., N. S. Sodhi, and C. J. A. Bradshaw. 2008. Synergies among extinction drivers under global change. *Trends in Ecology & Evolution* 23:453–460.
- Brown, M. J., and G. G. Parker. 1994. Canopy light transmittance in a chronosequence of mixed-species deciduous forests. *Canadian Journal of Forest Research* 24:1694–1703.
- Bunker, D.E., F. DeClerck, J.C. Bradford, R.K. Colwell, I. Perfecto, O.L. Phillips, M. Sankaran, and S. Naeem. 2005. Species Loss and Aboveground Carbon Storage in a Tropical Forest. *Science* 310:1029–1031.
- Burnham, K. P., D. R. Anderson, and K. P. Huyvaert. 2011. AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons. *Behavioral Ecology and Sociobiology* 65:23–35.
- Burton, A.J., K.S. Pregitzer, and D.D. Reed. 1991. Leaf area and foliar biomass relationships in northern hardwood forests located along an 800 km acid deposition gradient. *Forest Science* 37, 1041–1059.
- Bush, A., R. Sollmann, A. Wilting, K. Bohmann, B. Cole, H. Balzter, et al. 2017. Connecting Earth Observation to high-throughput biodiversity data. *Nature Ecology & Evolution* 1: 1–9.
- Cardinale, B. J., J. E. Duffy, A. Gonzalez, D. U. Hooper, C. Perrings, P. Venail, A. Narwani, G. M. Mace, D. Tilman, D.A. Wardle, A. P. Kinzig, G. C. Daily, M. Loreau, J. B. Grace, A. Larigauderie, D. S. Srivastava, and S. Naeem. 2012. Biodiversity loss and its impact on humanity. *Nature* 486:59–67.
- Cavanaugh, K.C., J.S. Gosnell, S.L. Davis, J. Ahumada, P. Boundja, D.B. Clark, et al. 2014. Carbon storage in tropical forests correlates with taxonomic diversity and functional dominance on a global scale. *Global Ecology and Biogeography* 23(5): 563–573.

- Cavender-Bares, J., K.H. Kozak, P.V. Fine, and S.W. Kembel. 2009. The merging of community ecology and phylogenetic biology. *Ecology Letters* 12: 693-715.
- Cavender-Bares, J., J.E. Meireles, J.J. Couture, M.A. Kaproth, C.C. Kingdon, A. Singh, S.P. Serbin, A. Center, E. Zuniga, G. Pilza, and P.A. Townsend. 2016. Associations of Leaf Spectra with Genetic and Phylogenetic Variation in Oaks: Prospects for Remote Detection of Biodiversity. *Remote Sensing* 8(221): 1-17.
- Cavender-Bares, J., J.A. Gamon, S.E. Hobbie, et al. 2017. Harnessing plant spectra to integrate the biodiversity sciences across biological and spatial scales. *American Journal of Botany* 104(7): 966-969.
- Chambers, J.Q., G.P. Asner, D.C. Morton, L.O. Anderson, S.S. Saatchi, F.D.B. Espirito-Santo, M. Palace, and C. Souza Jr. 2007. Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends in Ecology & Evolution* 22(8): 414-423.
- Chan, K.M., M.R. Shaw, D.R. Cameron, E.C. Underwood, and G.C. Daily. 2006. Conservation planning for ecosystem services. *PLOS Biology*, 4(11): e379.
- Chapin, F. S., B. H. Walker, R. J. Hobbs, D. U. Hooper, J. H. Lawton, O. E. Sala, and D. Tilman. 1997. Biotic Control over the Functioning of Ecosystems. *Science* 277:500–504.
- Chen, J.M., and T.A. Black. 1992. Defining leaf area index for non-flat leaves. *Plant, Cell and Environment* 15, 421–429.
- Clark, M. L., D. A. Roberts, and D. B. Clark. 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment* 96:375–398.
- Cole, D. W., and M. Rapp. 1981. Elemental cycling in forest ecosystems. Pages 341–409 in D. E. Reichle, editor. *Dynamic Properties of Forest Ecosystems*. Cambridge University Press, Cambridge.
- Colgan, M. S., C. A. Baldeck, J. baptiste Féret, and G. P. Asner. 2012. Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data. *Remote Sensing* 4:3462–3480.
- Collins, B.M., R.G. Everett, and S.L. Stephens. 2011. Impacts of fire exclusion and recent managed fire on forest structure in old growth Sierra Nevada mixed-conifer forests. *Ecosphere*, art51.
- Collings, S., P. Caccetta, N. Campbell, and X. Wu. 2010. Techniques for BRDF correction of hyperspectral mosaics. *IEEE Transactions on Geoscience and Remote Sensing* 48:3733–3746.

- Cook, B. D., L. A. Corp, R. F. Nelson, E. M. Middleton, D. C. Morton, J. T. McCorkel, J. G. Masek, K. J. Ranson, V. Ly, and P. M. Montesano. 2013. NASA Goddard's LiDAR, hyperspectral and thermal (G-LiHT) airborne imager. *Remote Sensing* 5:4045–4066.
- Cooper, W.J., W.J. McShea, D.A. Luther, and T. Forrester. 2019. Incorporating local habitat heterogeneity and productivity measures when modelling vertebrate richness. *Environmental Conservation*: 1:8.
- Cornwell, W. K., L. D. W. Schwilk, and D. D. Ackerly. 2006. A trait-based test for habitat filtering: convex hull volume. *Ecology* 87:1465–1471.
- Cosovic, M., M.N. Bugalho, D. Thom, and J.G. Borges. 2020. Stand Structural Characteristics Are the Most Practical Biodiversity Indicators for Forest Management Planning Europe. *Forests* 11 (343): 1-24.
- Dahlin, K. M., G. P. Asner, and C. B. Field. 2012. Environmental filtering and land-use history drive patterns in biomass accumulation in a mediterranean-type landscape. *Ecological Applications* 22:104–118.
- Dahlin, K. M., G. P. Asner, and C. B. Field. 2013. Environmental and community controls on plant canopy chemistry in a Mediterranean-type ecosystem. *Proceedings of the National Academy of Sciences* 110:6895–6900.
- Dahlin, K. M., G. P. Asner, and C. B. Field. 2014. Linking vegetation patterns to environmental gradients and human impacts in a mediterranean-type island ecosystem. *Landscape Ecology* 29:1571–1585.
- Dahlin, K.M. 2016. Spectral diversity area relationships for assessing biodiversity in a wildland-agriculture matrix. *Ecological Applications* 26(8): 2758-2768.
- Dandois, J.P., and E.C. Ellis. 2013. High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. *Remote Sensing of Environment* 136: 259-276.
- Dambrine, E., J.L. Dupouey, L. Laut, L. Humbert, M. Thion, T. Beauvils, and H. Richard. 2007. Present Forest Biodiversity Patterns in France Related to Former Roman Agriculture. *Ecology* 88(6): 1430-1439.
- DeFries, R.S., Field, C.B., Fung, I., Collatz, G.J., Bounoua, C.L., 1999. Combining satellite data and biogeochemical models to estimate global effects of human-induced land cover change on carbon emissions and primary productivity. *Global Biogeochemical Cycles* 13(3): 803-815.
- DeFries, R.S., Houghton, R.A., Hansen, M.C., Field, C.B., Skole, D., Townshend, J., 2002. Carbon emissions from tropical deforestation and regrowth based on satellite observations

for the 1980s and 1990s. *Proceedings of the National Academy of Sciences of the USA* 99(22): 14256-14261.

- Detto, M., G.P. Asner, H.C. Muller-Landau, O. Sonnentag. 2015. Spatial variability in tropical forest leaf area density from multireturn lidar and modeling: Multireturn LiDAR and Tropical Forest. *Journal of Geophysical Research: Biogeosciences* 120, 294–309.
- Díaz, S., J. Kattge, J. H. C. Cornelissen, I. J. Wright, S. Lavorel, S. Dray, B. Reu, M. Kleyer, C. Wirth, I. Colin Prentice, E. Garnier, G. Bönisch, M. Westoby, H. Poorter, P. B. Reich, A. T. Moles, J. Dickie, A. N. Gillison, A. E. Zanne, J. Chave, S. Joseph Wright, S. N. Sheremet Ev, H. Jactel, C. Baraloto, B. Cerabolini, S. Pierce, B. Shipley, D. Kirkup, F. Casanoves, J. S. Joswig, A. Günther, V. Falczuk, N. Rüger, M. D. Mahecha, and L. D. Gorné. 2016. The global spectrum of plant form and function. *Nature* 529:167–171.
- Ding, C., and X. He. 2004. K-means clustering via principal component analysis. *Proceedings of the twenty-first international conference on Machine Learning*.
- Dupouey, J.L., E. Dambrine, J.D. Laffite, and C. Moares. 2002. Irreversible Impact of Past Land Use on Forest Soils and Biodiversity. *Ecology* 83(11): 2978-2984.
- Duro, D.C., N.C. Coops, M.A. Wulder, and T. Han. 2007. Development of a large area biodiversity monitoring system driven by remote sensing. *Progress in Physical Geography* 31(3): 235-260.
- Dyer, J. 2006. Revisiting the Deciduous Forests of Eastern North America. *BioScience* 56(4): 341-352.
- Ellsworth, D. S., and P. B. Reich. 1993. Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest. *Oecologia* 96:169–178.
- Evans, J. R. 1989. Photosynthesis and nitrogen relationship in leaves of C3 plants. *Oecologia* 78:9–19.
- Fahey, R. T., A. T. Fotis, and K. D. Woods. 2015. Quantifying canopy complexity and effects on productivity and resilience in late-successional hemlock-hardwood forests. *Ecological Applications* 25:834–847.
- Falkowski, M.J., J.S. Evans, S. Martinuzzi, P.E. Gessler, and A.T. Hudak. 2009. Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. *Remote Sensing of Environment* 113, 946–956.
- Feret, J.B., and G.P. Asner. 2014. Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecological Applications* 24(6): 1289-1296.

- Field, C. B., and H. A. Mooney. 1986. The photosynthesis-nitrogen relationship in wild plants. Pages 25–56 in T.J. Givnish, editor. *On the Economy of Plant Form and Function*. Cambridge University Press, Cambridge.
- Fisher, R.A., C.D. Koven, W.R.L. Anderegg, B.O. Christoffersen, M.C. Dietze, C.E. Farrior, J.A. Holm, G.C. Hurtt, R.G. Knox, P.J. Lawrence, J.W. Lichstein, M. Longo, A.M. Matheny, D. Medvigy, H.C. Muller-Landau, T.L. Powell, S.P. Serbin, H. Sato, J.K. Shuman, B. Smith, A.T. Trugman, T. Viskari, H. Verbeeck, E. Weng, C. Xu, X. Xu, T. Zhang, and P.R. Moorcroft. 2018. Vegetation demographics in Earth System Models: A review of progress and priorities. *Global Change Biology* 24, 35–54.
- Flynn, D. F. B., N. Mirotnick, M. Jain, M. I. Palmer, and S. Naeem. 2011. Functional and phylogenetic diversity as predictors of biodiversity–ecosystem-function relationships. *Ecology* 92:1573–1581
- Foster, D.R. 1992. Land-use history (1730-1990) and vegetation dynamics in Central New England, USA. *Ecology* 80(4): 753-771.
- Gamon, J.A., J. Penuelas, and C.B. Field. 1992. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment* 41: 35-44.
- Garabedian, J.E., C.E. Moorman, M.N. Peterson, and J.C. Kilgo. 2017. Use of LiDAR to define habitat thresholds for forest bird conservation. *Forest Ecology and Management* 399, 24-36.
- Gaston, K. J. 2000. Global patterns in biodiversity. *Nature* 405:220–227
- Gatziolis, D., and H.E. Andersen. 2008. A guide to LIDAR data acquisition and processing for the forests of the Pacific Northwest. (No. PNW-GTR-768). *U.S. Department of Agriculture, Forest Service*, Pacific Northwest Research Station, Portland, OR.
- Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*. 27:2865–2873.
- Gholizadeh, H., J.A. Gamon, A.I. Zyguelbaum, R. Wang, A.K. Schweiger, and J. Cavender-Bares. 2018. Remote sensing of biodiversity: Soil correction and data dimension reduction methods improve assessment of α -diversity (species richness) in prairie ecosystems. *Remote Sensing of Environment* 206: 240–253.
- Gholizadeh, H., J.A. Gamon, P.A. Townsend, et al. 2019. Detecting prairie biodiversity with airborne remote sensing. *Remote Sensing of Environment* 221: 38-49.
- Girden, E.R. 1992. ANOVA: Repeated measures. *Sage University Press*, London, UK.
- Gitelson, A., and M.N. Merzlyak. 1994. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology, B: Biology* 22: 247-252.

- Goodale, C.L., M.J. Apps, R.A. Birdsey, C.B. Field, L.S. Heath, R.A. Houghton, J.C. Jenkins, G.H. Kohlmaier, W. Kurz, S. Liu, G.J. Nabuurs, S. Nilsson, and A.Z. Shvidenko. 2002. Forest carbon sinks in the Northern Hemisphere. *Ecological Applications* 12(3): 891-899.
- Gelman, A. 2007. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine* 27:2865–2873.
- Gotelli, N., and A.M. Ellison. 2013. *A Primer of Ecological Statistics*. Sinauer Associates, Inc, Sunderland, MA.
- Gougeon, F. A. 1995. Comparison of possible multispectral classification schemes for tree crowns individually delineated on high spatial resolution multispectral images. *Canadian Journal of Remote Sensing* 21:1–9.
- Gough, C.M., B.S. Hardiman, L.E. Nave, G. Bohrer, K.D. Maurer, C.S. Vogel, K.J. Nadelhoffer, and P.S. Curtis. 2013. Sustained carbon uptake and storage following moderate disturbance in a Great Lakes forest. *Ecological Applications* 23, 1202–1215.
- Gough, C. M., J. W. Atkins, R. T. Fahey, and B. S. Hardiman. 2019. High rates of primary production in structurally complex forests. *Ecology* 100.
- Habel, K., R. Grasman, R. B. Gramacy, A. Stahel, and D. C. Sterratt. 2015. *geometry: Mesh Generation and Surface Tessellation*. R package version 0.3-6. <https://CRAN.R-project.org/package=geometry>
- Hansen, A.J., R.P. Neilson, V.H. Dale, C.H. Flather, L.R. Iverson, D.J. Currie, S. Shafer, R. Cook, and P.J. Bartlein. 2001. Global Change in Forests: Responses of Species, Communities, and Biomes Interactions between climate change and land use are projected to cause large shifts in biodiversity. *BioScience* 51:765–779
- Hanson, J.J., and C.G. Lorimer. 2007. Forest structure and light regimes following moderate windstorms: Implications for multi-cohort management. *Ecological Applications* 17, 1325–1340.
- Harding, D.J., M.A. Lefsky, G.G. Parker, and J.B. Blair. 2001. Laser altimeter canopy profiles methods and validation for closed-canopy, broadleaf forests. *Remote Sensing of Environment* 76: 283-297.
- Hardiman, B. S., C. M. Gough, A. Halperin, K. L. Hofmeister, L. E. Nave, G. Bohrer, and P. S. Curtis. 2013. Maintaining high rates of carbon storage in old forests: A mechanism linking canopy structure to forest function. *Forest Ecology and Management* 298:111–119.
- Hardiman, B. S., G. Bohrer, C. M. Gough, C. S. Vogel, P. S. Curtis, S. Vogel, S. Curtis, and S. Hardiman. 2011. The role of canopy structural complexity in wood net primary production of a maturing northern deciduous forest. *Ecology* 92: 1818-1827.

- He, K.S., B.A. Bradley, A.F. Cord, D. Rocchini, M.N. Tuanmu, and S. Schmidtlein. 2015. Will remote sensing shape the next generation of species distribution models? *Remote Sensing in Ecology and Conservation* 1: 4–18
- Hijmans, R. J. 2019. raster: Geographic Data Analysis and Modeling. *R package* version 2.9-5.
- Hinckley, E.L.S., G.B. Bonan, G.J. Bowen, C.P. Colman, P.A. Duffy, C.L. Goodale, B.Z. Houlton, E. Marín-Spiotta, K. Ogle, S.V. Ollinger, E.A. Paul, P.M. Vitousek, K.C. Weathers, and D.G. Williams. 2016. The soil and plant biogeochemistry sampling design for The National Ecological Observatory Network. *Ecosphere*, e01234.
- Hoffman K.M., A.J. Trant, W. Nijland, and B.M. Starzomski. 2018. Ecological legacies of fire detected using plot-level measurements and LiDAR in an old growth coastal temperate rainforest. *Forest Ecology and Management* 424, 11-20.
- Holling, C.S. 1973. Resilience and Stability of Ecological Systems. *Annual Review of Ecology and Systematics* 4:1–23
- Hooper, D. U., E. C. Adair, B. J. Cardinale, J. E. K. Byrnes, B. A. Hungate, K. L. Matulich, A. Gonzalez, J. E. Duffy, L. Gamfeldt, and M. I. O’Connor. 2012. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature* 486:105–108.
- Horton, J. D. 2017. The State Geologic Map Compilation (SGMC) geodatabase of the conterminous United States (ver. 1.1, August 2017). *U.S. Geological Survey data release*.
- Hosoi, F., Omasa, K., 2006. Voxel-based 3-D modeling of individual trees for estimating leaf area density using high resolution portable scanning lidar. *IEEE Transactions on Geoscience and Remote Sensing* 44, 3610–3618.
- Hosoi, F., and K. Omasa. 2007. Factors contributing to accuracy in the estimation of the woody canopy leaf area density profile using 3D portable lidar imaging. *Journal of Experimental Botany* 58, 3463–3473.
- Hughes, R.F., G.P. Asner, J.A. Baldwin, J. Mascaró, L.K.K. Bufile, and D.E. Knapp. 2018. Estimating aboveground carbon density across forest landscapes of Hawaii: Combining FIA plot-derived estimates and airborne LiDAR. *Forest Ecology and Management* 424, 323-337.
- Hummel, S., and J.K. Agee. 2003. Western spruce budworm defoliation effects of forest structure and potential fire behavior. *Northwest Science* 15, 1–49.
- Hurrell, J. W., M. M. Holland, P. R. Gent, S. Ghan, J. E. Kay, P. J. Kushner, J.-F. Lamarque, W. G. Large, D. Lawrence, K. Lindsay, W. H. Lipscomb, M. C. Long, N. Mahowald, D. R. Marsh, R. B. Neale, P. Rasch, S. Vavrus, M. Vertenstein, D. Bader, W. D. Collins, J. J. Hack, J. Kiehl, and S. Marshall. 2013. The Community Earth System Model: A Framework for Collaborative Research. *Bulletin of the American Meteorological Society* 94:1339–1360.

- Iglhaut, J., C. Cabo, S. Puliti, L. Piermattei, J. O'Connor, and J. Rosette. 2019. Structure from Motion Photogrammetry in Forestry: a Review. *Current Forestry Reports* 5(3): 155-168.
- Isbell, F., D. Craven, J. Connolly, M. Loreau, B. Schmid, C. Beierkuhnlein, T. M. Bezemer, C. Bonin, H. Bruelheide, E. De Luca, A. Ebeling, J. N. Griffin, Q. Guo, Y. Hautier, A. Hector, A. Jentsch, J. Kreyling, V. Lanta, P. Manning, S. T. Meyer, A. S. Mori, S. Naeem, P. A. Niklaus, H. W. Polley, P. B. Reich, C. Roscher, E. W. Seabloom, M. D. Smith, M. P. Thakur, D. Tilman, B. F. Tracy, W. H. Van Der Putten, J. Van Ruijven, A. Weigelt, W. W. Weisser, B. Wilsey, and N. Eisenhauer. 2015. Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature* 526:574–577.
- Jarvis, P.G., and K.G. McNaughton. 1986. Stomatal Control of Transpiration: Scaling Up from Leaf to Region. *Advances in Ecological Research* 15, 1-49.
- Jetz, W., M.A. McGeoch, R. Guralnick, S. Ferrier, J. Beck, M.J. Costello, M. Fernandex, et al. 2019. Essential biodiversity variables for mapping and monitoring species populations. *Nature Ecology & Evolution* 3: 539-551.
- Jost, L. 2006. Entropy and diversity. *Oikos* 113(2): 363-375.
- Jost, L. 2007. Partitioning Diversity into Independent Alpha and Beta Components. *Ecology* 88(10): 2427-2439.
- Kamoske, A. G., K. M. Dahlin, S. C. Stark, and S. P. Serbin. 2019. Leaf area density from airborne LiDAR: Comparing sensors and resolutions in a temperate broadleaf forest ecosystem. *Forest Ecology and Management* 433:364–375.
- Kamoske, A.G., K.M. Dahlin, S.P. Serbin, and S.C. Stark. 2020. Leaf traits and canopy structure together explain canopy functional diversity: an airborne remote sensing approach. *Ecological Applications* e2230.
- Kampe, T. U., B. R. Johnson, M. Kuester, and M. Keller. 2010. NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *Journal of Applied Remote Sensing* 4:043510.
- Kattge, J., G. Boenisch, S. Diaz, et al. 2020. TRY plant trait database - enhance coverage and open access. *Global Change Biology* 26: 119-188.
- Keddy, P.A. 1992. Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetation Science* 3(2): 157-164.
- Kitajima, K., 2004. Variation in Crown Light Utilization Characteristics among Tropical Canopy Trees. *Annals of Botany* 95, 535–547.

- Khosravipour, A., A.K. Skidmore, T. Wang, M. Isenburg, and K. Khoshelham. 2015. Effect of slope on treetop detection using a LiDAR Canopy Height Model. *ISPRS Journal of Photogrammetry and Remote Sensing* 104, 44-52.
- Klingberg, J., J. Konarska, F. Lindberg, L. Johansson, and S. Thorsson. 2017. Mapping leaf area of urban greenery using aerial LiDAR and ground-based measurements in Gothenburg, Sweden. *Urban Forestry & Urban Greening* 26, 31–40.
- Korhonen, L., I. Korpela, J. Heiskanen, and M. Maltamo. 2011. Airborne discrete-return LIDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index. *Remote Sensing of Environment* 115, 1065-1080.
- Kraft, N. J. B., R. Valencia, and D. D. Ackerly. 2008. Functional traits and niche-based tree community assembly in an Amazonian forest. *Science* 322:580–582.
- Laliberté, E., A.K. Schweiger, and P. Legendre. 2020. Partitioning plant spectral diversity into alpha and beta components. *Ecology Letters* 23: 370-380.
- LaRue, E.A., B.S. Hardiman, J.M. Elliot, and S. Fei. 2019. Structural diversity as a predictor of ecosystem function. *Environmental Research Letters* 14(11): 114011.
- Lawrence, P. J., and T. N. Chase. 2007. Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0). *Journal of Geophysical Research: Biogeosciences* 112.
- Le Quéré, C., R. Moriarty, R.M. Andrew, G.P. Peters, P. Ciais, P. Friedlingstein, et al. 2015. Global carbon budget 2014. *Earth System Science Data* 7, 47–85.
- Leblanc, S. G., J. M. Chen, R. Fernandes, D. W. Deering, and A. Conley. 2005. Methodology comparison for canopy structure parameters extraction from digital hemispherical photography in boreal forests. *Agricultural and Forest Meteorology* 129:187–207.
- Lefsky, M. A., W. B. Cohen, S. A. Acker, G. G. Parker, T. A. Spies, and D. Harding. 1999. Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir Western Hemlock Forests. *Remote Sensing of Environment* 70:339–361.
- Lefsky, M.A., W.B. Cohen, G.G. Parker, and D.J. Harding. 2002. Lidar Remote Sensing for Ecosystem Studies. *BioScience* 52, 19.
- Leutner, B.F., B. Reineking, J. Muller, M. Bachmann, C. Beierkuhnlein, S. Dech, and M. Wegmann. 2012. Modelling Forest α -Diversity and Floristic Composition – On the Added Value of LiDAR plus Hyperspectral Remote Sensing. *Remote Sensing* 4: 2818-2845.
- Levin, S.A. 1992. The problem of pattern and scale in ecology. *Ecology* 73(6): 1953-1967.

- Li, D., J.D. Olden, J.L. Lockwood, S. Record, M.L. McKinney, and B. Baiser. 2020. Changes in taxonomic and phylogenetic diversity in the Anthropocene. *Proc. R. Soc. B* 287: 20200777
- Lomolino, M. V., B. R. Riddle, R. J. Whittaker, and J. H. Brown. 2010. *Biogeography*. 4th edition. Sinauer.
- Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, M. A. Huston, D. Raffaelli, B. Schmid, D. Tilman, and D. A. Wardle. 2001. Biodiversity and Ecosystem Functioning: Current Knowledge and Future Challenges. *Science* 294:804–808.
- Lovell, J.L., D.L.B. Jupp, D.S. Culvenor, and N.C. Coops. 2003. Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. *Canadian Journal of Remote Sensing* 29, 607–622.
- MacArthur, R. H., and H. S. Horn. 1969. Foliage Profile by Vertical Measurements. *Ecology* 50:802–804.
- Mairota, P., B. Cafarelli, R.K. Didham, F.P. Lovergine, R.M. Lucas, H. Nagendra, D. Rocchini, and C. Tarantino. 2015. Challenges and opportunities in harnessing satellite remote sensing for biodiversity monitoring. *Ecological Informatics* 30, 207–214.
- Mao, L., J. Dennett, C.W. Bater, P. Tompalski, N.C. Coops, D. Farr, M. Kohler, B. White, J.J. Stadt, and S.E. Nielson. 2018. Using airborne laser scanning to predict plant species richness and assess conservation threats in the oil sands region of Alberta’s boreal forest. *Forest Ecology and Management* 409, 29-37.
- Martens, H., and M. Martens. 2000. Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Quality and Preference* 5(16): 1-16.
- Martin, M.E., Plourde, L.C., Ollinger, S.V., Smith, M.L., McNeil, B.E., 2008. A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sensing of Environment* 112: 3511-3519.
- Mascaro, J., G. P. Asner, H. C. Muller-Landau, M. Van Breugel, J. Hall, and K. Dahlin. 2011. Controls over aboveground forest carbon density on Barro Colorado Island, Panama. *Biogeosciences* 8:1615–1629.
- McNeil, B.E., J. Pisek, H. Lepisk, and E.A. Flamenco. 2016. Measuring leaf angle distribution in broadleaf canopies using UAVs. *Agriculture and Forest Meteorology* 218, 204-208.
- Meir, P., B. Kruijt, M. Broadmeadow, E. Barbosa, O. Kull, F. Carswell, A. Nobre, and P.G. Jarvis. 2002. Acclimation of photosynthetic capacity to irradiance in tree canopies in relation to leaf nitrogen concentration and leaf mass per unit area. *Plant, Cell and Environment* 25, 343–357.

- Meng, R., P.E. Dennison, F. Zhao, I. Shendryk, A. Rickert, R.P. Hanavan, B.D. Cook, and S.P. Serbin. 2018. Mapping canopy defoliation by herbivorous insects at the individual tree level using bi-temporal airborne imaging spectroscopy and LiDAR measurements. *Remote Sensing of Environment* 215, 170–183.
- Merton, R.N., and J.F. Huntington. 1999. Early simulation results of the ARIES-1 satellite sensor for multi-temporal vegetation research derived from AVIRIS. *Proceedings of the Eighth Annual JPL Airborne Earth Science Workshop*. NASA, Jet Propulsion Laboratory, Pasadena, California, USA. 8-14 February 1999.
- Mevik, B., and R. Wehrens. 2007. The pls Package: Principal Component and Partial Least Squares Regression in R. *Journal of Statistical Software* 18:1–23.
- Miller, J., 1967. A formula for average foliage density. *Australian Journal of Botany* 15, 141.
- Morsdorf, F., B. Kötz, E. Meier, K.I. Itten, and B. Allgöwer. 2006. Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. *Remote Sensing of Environment* 104, 50–61.
- National Academies of Sciences, Engineering, and Medicine. 2018. Thriving on Our Changing Planet: A Decadal Strategy 650 for Earth Observation from Space. *The National Academies Press*, Washington D.C.
- National Ecological Observatory Network, 2017. Data Products NEON.DP1.10058.001, NEON.DP1.30003.001. *Provisional data downloaded* from <https://data.neonscience.org> on 6 Nov 2017. Battelle, Boulder, CO, USA.
- National Ecological Observatory Network. 2020. Data Products NEON.DP1.10098.001. *Provisional data downloaded* from <http://data.neonscience.org> on 04 April 2020. Battelle, Boulder, CO, USA
- Nicolson, M. 1987. Alexander von Humboldt, Humboldtian science and the origins of the study of vegetation. *History of Science* (25): 167-193.
- Niinemets, Ü. 2012. Optimization of foliage photosynthetic capacity in tree canopies: Towards identifying missing constraints. *Tree Physiology* 32: 505-509.
- Niinemets, Ü. 2007. Photosynthesis and resource distribution through plant canopies. *Plant, Cell and Environment* 30:1052–1071.
- Niinemets, Ü., T. F. Keenan, and L. Hallik. 2015. A worldwide analysis of within-canopy variations in leaf structural, chemical, and physiological traits across plant functional types. *New Phytologist* 205:973–993.
- Nychka, D., Furrer, R., Paige, J., Sain, S., 2015. Fields: Tools for spatial data. *R package* version 8.4-1. www.image.ucar.edu/fields

- Olden, J. D., and T.P. Rooney. 2006. On defining and quantifying biotic homogenization. *Global Ecology and Biogeography* 15:113–120.
- Ollinger, S.V. 2011. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist* 189(2): 375-394.
- Ollinger, S. V, M. L. Smith, M. E. Martin, R. A. Hallett, C. L. Goodale, and J. D. Aber. 2002. Regional Variation in Foliar Chemistry and N Cycling among Forests of Diverse History and Composition. *Ecology* 83:339–355.
- Ollinger, S.V. and M.L. Smith. 2005. Net Primary Production and Canopy Nitrogen in a Temperate Forest Landscape: An Analysis Using Imaging Spectroscopy, Modeling, and Field Data. *Ecosystems* 8: 760-778.
- Ozanne, C.M.P., D. Anhug, S.L. Boulter, M. Keller, R.L. Kitching, C. Korner, F.C. Meinzer, A.W. Mitchell, T. Nakashizuka, P.L. Silva Dias, N.E. Stork, S.J. Wright, and M. Yoshimura. 2003. Biodiversity Meets the Atmosphere: A Global View of Forest Canopies. *Science* 301: 183-186.
- Pan, Y., R.A. Birdsey, J. Fang, R. Houghton, P.E. Kauppi, W.A. Kurz, O.L. Phillips, et al. 2011a. A Large and Persistent Carbon Sink in the World’s Forests. *Science* 333, 988–993.
- Pan, Y., J.M. Chen, R. Birdsey, K. McCullough, L. He, and F. Deng. 2011b. Age structure and disturbance legacy of North American forests. *Biogeosciences* 8, 715–732.
- Parker, G.G., 1995. Structure and microclimate of forest canopies. Pages 73-106 in Lowman M., Nadkarni, N., eds. *Forest Canopies: A review of research on a biological frontier*. San Diego: Academic Press.
- Parker, G. G., M.E. Harmon, M.A. Lefsky, J. Chen, R. Van Pelt, S.B. Weiss, S.C. Thomas, W.E. Winner, D.C. Shaw, and J.F. Franklin. 2004. Three-dimensional Structure of an Old-growth *Pseudotsuga-tsuga* Canopy and its Implications for Radiation Balance, Microclimate, and Gas Exchange. *Ecosystems* 7: 440-453.
- Parker, G. G., and D. J. Tibbs. 2004. Structural Phenology of the Leaf Community in the Canopy of a *Liriodendron tulipifera* L. Forest in Maryland, USA. *Forest Science* 50:387–397.
- Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421:37–42.
- Paz-Kagan, T., N.R. Vaughn, R.E. Martin, P.G. Brodrick, N.L. Stephenson, A.J. Das, K.R. Nydick, and G.P. Asner. 2018. Landscape-scale variation in canopy water content of giant sequoias during drought. *Forest Ecology and Management* 419-420, 291-304.
- Pedro, M. S., W. Rammer, and R. Seidl. 2017. Disentangling the effects of compositional and structural diversity on forest productivity. *Journal of Vegetation Science* 28:649–658.

- Pettorelli, N., W.F. Laurance, T.G. O'Brien, M. Wegmann, H. Nagendra, and W. Turner. 2014. Satellite remote sensing for applied ecologists: opportunities and challenges. *Journal of Applied Ecology* 51, 839–848.
- Pinheiro, J. C., and D.M. Bates. 2000. *Mixed-Effects Models in S and S-PLUS*. Springer, New York.
- Pisek, J., H. Buddenbaum, F. Camacho, J. Hill, J.L.R. Jensen, et al. 2018. Data synergy between leaf area index and clumping index Earth Observation production using photon recollision probability theory. *Remote Sensing of Environment* 215(15): 1-6.
- Poorter, H., Ü. Niinemets, L. Poorter, I. J. Wright, and R. Villar. 2009. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. *New Phytologist* 182:565–588.
- Potter, K.M., and F.H. Koch. 2014. Patterns of forest phylogenetic community structure across the United States and their possible forest health implications. *Forest Science* 60: 851-861.
- Potter, K.M., and C.W. Woodall. 2012. Trends over time in tree and seedling phylogenetic diversity indicate regional differences in forest biodiversity change. *Ecological Applications* 22: 517-531.
- R Core Team, 2016. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna Austria. <https://www.R-project.org>
- Read, Q.D., P.L. Zarnetske, S. Record, K.M. Dahlin, J.K. Costanza, A.O. Finley, K.D. Gaddis, J.M. Grady, M.L. Hobi, A.M. Latimer, S.L. Malone, S.V. Ollinger, S. Pau, and A.M. Wilson. 2019. Beyond counts and averages: Relating geodiversity to dimensions of biodiversity. *Global Ecology and Biogeography* 00: 1-15.
- Record, S. K.M. Dahlin, P.L. Zarnetske, Q.D. Read, S.L. Malone, K.D. Gaddis, et al. 2020. Remote Sensing of Geodiversity as a Link to Biodiversity. Editors: Cavender-Bares, J., J.A. Gamon, and P.A. Townsend in *Remote Sensing of Plant Biodiversity*. Springer Open press, Switzerland.
- Reich, P. B., M. B. Walters, and D. S. Ellsworth. 1997. From tropics to tundra: Global convergence in plant functioning. *Proceedings of the National Academy of Sciences of the United States of America* 94:13730–13734.
- Reich, P.B., P. Bakken, D. Carlson, L.E. Frelich, S.K. Friedman, and D.F. Grigal. 2001. Influence of Logging, Fire, and Forest Type on Biodiversity and Productivity in Southern Boreal Forests. *Ecology* 82(10): 2731-2748.
- Richardson, J. J., L. M. Moskal, and S. H. Kim. 2009. Modeling approaches to estimate effective leaf area index from aerial discrete-return LIDAR. *Agricultural and Forest Meteorology* 149:1152–1160.

- Rocchini, D., N. Balkenhol, G.A. Carter, G.M. Food, T.W. Gillespie, and K.S. He. 2010. Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecological Informatics* 5: 318–329.
- Rohde, K. 1992. Latitudinal gradients in species diversity: the search for the primary cause. *Oikos* 65: 514-527.
- Rouse, J.W., R.H. Haas, J.A. Schell, and D.W. Deering. 1974. Monitoring vegetation systems in the Great Plains with ERTAS, In: S.C. Freden, E.P. Mercanti, and M. Becker (eds.) *Third Earth Resources Technology Satellite-1 Symposium*. Volume 1: Technical Presentations, NASA SP-351, NASA, Washington, D.C., pp. 309-317.
- Roussel, J.R., 2016. Rlas: read and write ‘las’ and ‘laz’ binary file formats used for remote sensing data. *R package version 1.0.3*. <https://CRAN.R-project.org/package=rflas>
- Roussel, J.R., J. Caspersen, M. Beland, S. Thomas, and A. Achim. 2017. Removing bias from LiDAR-based estimates of canopy height: Accounting for the effects of pulse density and footprint size. *Remote Sensing of Environment* 198, 1-16.
- Running, S. W., R. R. Nemani, F. A. Heinsch, M. Zhao, M. Reeves, and H. Hashimoto. 2004. A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production. *BioScience* 54:547.
- Sabol, J., Z. Patočka, and T. Mikita. 2014. Usage of Lidar Data for Leaf Area Index Estimation. *GeoScience Engineering* 60:10–18.
- Safken, B., D. Rugamer, T. Kneib, and S. Geven. 2018. Conditional Model Selection in Mixed-Effects Models with cAIC4. *rXiv1803.05664*: 1–31
- Schaefer, F.K. 1953. Exceptionalism in Geography: A Methodological Examination. *Annals of the AAG* (43)3: 226-249.
- Scheuermann, C.M., L.E. Nave, R.T. Fahey, K.J. Nadelhoffer, and C.M. Gough. 2018. Effects of canopy structure and species diversity on primary production in upper Great Lakes forests. *Oecologia* 188: 405-415.
- Schimel, D. S. 1995. Terrestrial biogeochemical cycles: Global estimates with remote sensing. *Remote Sensing of Environment* 51:49–56.
- Schimel, D., R. Pavlick, J. B. Fisher, G. P. Asner, S. Saatchi, P. Townsend, C. Miller, C. Frankenberg, K. Hibbard, and P. Cox. 2015. Observing terrestrial ecosystems and the carbon cycle from space. *Global Change Biology* 21:1762–1776.
- Schimel, D., F. D. Schneider, and J. C. and E. Participants. 2019. Flux towers in the sky: global ecology from space. *New Phytologist* 224(2): 570-584.

- Schlapfer, D., R. Richter, and T. Feingersh. 2015. Operational BRDF effects correction for wide-field-of-view optical scanners (BREFCOR). *IEEE Transactions on Geoscience and Remote Sensing* 53:1855–1864.
- Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of Image Analysis. *Nature Methods* 9:671–675.
- Serbin, S. P., A. Singh, B. E. McNeil, C. C. Kingdon, and P. A. Townsend. 2014. Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecological Applications* 24:1651–1669.
- Shao, G., S. C. Stark, D. R. A. de Almeida, and M. N. Smith. 2019. Towards high throughput assessment of canopy dynamics: The estimation of leaf area structure in Amazonian forests with multitemporal multi-sensor airborne lidar. *Remote Sensing of Environment* 221:1–13.
- Shi, Y., T. Wang, A. K. Skidmore, and M. Heurich. 2018. Important LiDAR metrics for discriminating forest tree species in Central Europe. *ISPRS Journal of Photogrammetry and Remote Sensing* 137:163–174.
- Silva, C.A., A.T. Hudak, L.A. Vierling, C. Klauber, M. Garcia, A. Ferraz, M. Keller, J. Eitel, and S. Saatchi. 2017. Impacts of airborne lidar pulse density on estimating biomass stocks and changes in a selectively logged tropical forest. *Remote Sensing* 9, 1-19.
- Singh, A., S. P. Serbin, B. E. McNeil, C. C. Kingdon, and P. A. Townsend. 2015. Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. *Ecological Applications* 25:2180–2197.
- Simonson, W.D., H.D. Allen, and D.A. Coomes. 2012. Use of an Airborne Lidar System to Model Plant Species Composition and Diversity of Mediterranean Oak Forests. *Conservation Biology* 26(5): 840-850.
- Slot, M., G.H. Krause, B. Krause, G.G. Hernández, and K. Winter, 2018. Photosynthetic heat tolerance of shade and sun leaves of three tropical tree species. *Photosynthesis Research* 114: 119-130.
- Smart, L.S., J.J. Swenson, N.L. Christensen, and J.O. Sexton. 2012. Three-dimensional characterization of pine forest type and red-cockaded woodpecker habitat by small-footprint, discrete-return lidar. *Forest Ecology and Management* 281, 100–110.
- Smith, A. B., B. Sandel, N. J. B. Kraft, and S. Carey. 2013. Characterizing scale-dependent community assembly using the functional-diversity–area relationship. *Ecology* 94:2392–2402
- Smith, S. J., J. Edmonds, C. A. Hartin, A. Mundra, and K. Calvin. 2015. Near-term acceleration in the rate of temperature change. *Nature Climate Change* 5:333–336.

- Smith, M. N., S. C. Stark, T. C. Taylor, M. L. Ferreira, E. de Oliveira, N. Restrepo-Coupe, S. Chen, T. Woodcock, D. B. dos Santos, L. F. Alves, M. Figueira, P. B. de Camargo, R. C. de Oliveira, L. E. O. C. Aragão, D. A. Falk, S. M. McMahon, T. E. Huxman, and S. R. Saleska. 2019. Seasonal and drought-related changes in leaf area profiles depend on height and light environment in an Amazon forest. *New Phytologist* 222:1284–1297.
- Socha J., M. Pierzchalski, R. Balazy, and M. Ciesielski. 2017. Modelling top height growth and site index using repeated laser scanning data. *Forest Ecology and Management* 406, 307-317.
- Soenen, S. A., D. R. Peddle, and C. A. Coburn. 2005. SCS+C: a modified Sun-canopy-sensor topographic correction in forested terrain. *IEEE Transactions on Geoscience and Remote Sensing* 43:2148–2159.
- Solberg, S., E. Næsset, K. H. Hanssen, and E. Christiansen. 2006. Mapping defoliation during a severe insect attack on Scots pine using airborne laser scanning. *Remote Sensing of Environment* 102:364–376.
- Spellerberg, I.F., and P.J. Fedor. 2003. A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the ‘Shannon–Wiener’ Index. *Global Ecology and Biogeography* 12(3): 177-179.
- Srivastava, D. S., M. W. Cadotte, A. A. M. MacDonald, R. G. Marushia, and N. Mirotchnick. 2012. Phylogenetic diversity and the functioning of ecosystems. *Ecology Letters* 15:637–648.
- Staccone, A., W. Liao, S. Perakis, J. Compton, C. Clark, and D. Menge. A spatially explicit, empirical estimate of tree-based biological nitrogen fixation in forests of the United States. *Global Biogeochemical Cycles* 34(2): 1-18.
- Stark, S. C., B. J. Enquist, S. R. Saleska, V. Leitold, J. Schietti, M. Longo, L. F. Alves, P. B. Camargo, and R. C. Oliveira. 2015. Linking canopy leaf area and light environments with tree size distributions to explain Amazon forest demography. *Ecology Letters* 18:636–645.
- Stark, S. C., V. Leitold, J. L. Wu, M. O. Hunter, C. V. de Castilho, F. R. C. Costa, S. M. McMahon, G. G. Parker, M. T. Shimabukuro, M. A. Lefsky, M. Keller, L. F. Alves, J. Schietti, Y. E. Shimabukuro, D. O. Brandão, T. K. Woodcock, N. Higuchi, P. B. de Camargo, R. C. de Oliveira, and S. R. Saleska. 2012. Amazon forest carbon dynamics predicted by profiles of canopy leaf area and light environment. *Ecology Letters* 15:1406–1414.
- Stavros, E.N., D. Schimel, R. Pavlick, S. Serbin, A. Swann, L. Duncanson, J.B. Fisher, F. Fasnacht, S. Ustin, R. Dubayah, A. Schwigher, and P. Wennberg. 2017. ISS observations offer insights into plant function. *Nature Ecology and Evolution* 1: 1-4.

- Stevens, J.T., M.W. Kling, D.W. Schwilk, J.M. Varner, and J.M. Kane. 2020. Biogeography of fire regimes in western U.S. conifer forests: A trait-based approach. *Global Ecology and Biogeography* 29(5): 944-955.
- Stocker, T., Q. Dahe, and G.-K. Plattner. 2013. *Working Group I Contribution to the IPCC Fifth Assessment Report Climate Change 2013: The Physical Science Basis*.
- Sumida, A., T. Nakai, M. Yamada, K. Ono, S. Uemura, and T. Hara. 2009. Ground-based estimation of leaf area index and vertical distribution of leaf area density in a betula ermanii forest. *Silva Fennica* 43:799–816.
- Tabachnick, B.G., and L.S. Fidell. 2007. *Experimental Designs using ANOVA*. Duxbury, Belmont, CA.
- Taylor, P., G. Asner, K. Dahlin, C. Anderson, D. Knapp, R. Martin, J. Mascaro, R. Chazdon, R. Cole, W. Wanek, F. Hofhansl, E. Malavassi, B. Vilchez-Alvarado, and A. Townsend. 2015. Landscape-Scale Controls on Aboveground Forest Carbon Stocks on the Osa Peninsula, Costa Rica. *PLOS ONE* 10, e0126748.
- Tews, J., U. Brose, V. Grimm, K. Tielbörger, M.C. Wichmann, M. Schwager, and F. Jeltsch F. 2004. Animal species diversity driven by habitat heterogeneity/ diversity: the importance of keystone structures. *Journal of Biogeography* 31: 79–92.
- Thorpe, A.S., D.T. Barnett, S.C. Elmendorf, E.S. Hinckley, D. Hoekman, K.D. Jones, K.E. LeVan, C.L. Meier, L.F. Stanish, and K.M. Thibault. 2016. Introduction to the sampling designs of the National Ecological Observatory Network Terrestrial Observation System. *Ecosphere* 7(12): e01627.
- Tilman, D., J. Hill, and C. Lehman. 2006. Carbon-Negative Biofuels from Low-Input High-Diversity Grassland Biomass. *Science* 314:1598–1600
- Townsend, P. A., J. R. Foster, R. A. Chastian Jr, and W. S. Currie. 2003. Canopy nitrogen in the forests of the Central Appalachian Mountains using Hyperion and AVIRIS. *IEEE Transactions on Geoscience and Remote Sensing* 41:1347–1354.
- Tucker, C. J., and P. J. Sellers. 1986. Satellite remote sensing of primary production. *International Journal of Remote Sensing* 7:1395–1416.
- Turner, W., C. Rondinini, N. Pettorelli, B. Mora, A.K. Leidner, Z. Szantoi, G. Buchanan, et al. 2015. Free and open-access to satellite data are key to biodiversity conservation. *Biological Conservation*, 182: 173-176.
- Turner, W., S. Spector, N. Gardiner, M. Fladeland, E. Sterling, and M. Steininger. 2003. Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution* 18:306–314.

- Urban, M. C. 2015. Accelerating extinction risk from climate change. *Science* 348:571–573
- USDA Forest Service. 2005. Longleaf ecosystem-restoration project: Final environmental impact statement, National Forests in Alabama. *Talladega National Forest, Oakmulgee District*, USDA Forest Service, Brent, AL.
- Ustin S.L., and J.A. Gamon. 2010. Remote sensing of plant functional types. *New Phytologist* 186: 795–816.
- Valladares, F., L. Laanisto, Ü. Niinemets, and M. A. Zavala. 2016. Shedding light on shade: ecological perspectives of understorey plant life. *Plant Ecology and Diversity* 9:237–251.
- Venables, W. N., and B. D. Ripley. 2002. Modern applied statistics with S. *Stanford University, Stanford, California, USA*.
- Vose, J.M., N.H. Sullivan, B.D. Clinton, and P.V. Bolstad. 1995. Vertical leaf area distribution, light transmittance, and application of the Beer-Lambert Law in four mature hardwood stands in the southern Appalachians. *Canadian Journal of Forest Research* 25, 1036-1043.
- Wagstaff, K., C. Cardie, S. Rogers, and S. Schroedl. 2001. Constrained K-means Clustering with Background Knowledge. *Proceedings of the Eighteenth International Conference on Machine Learning*: 577-584.
- Wang, R., J.A. Gamon, J. Cavender-Bares, P.A. Townsend, and A.I. Zyguelbaum. 2018. The spatial sensitivity of the spectral diversity-biodiversity relationship: an experimental test in a prairie grassland. *Ecological Applications* 28: 541-556.
- Wang, R. and J.A. Gamon. 2019. Remote sensing of terrestrial plant biodiversity. *Remote Sensing of Environment* 231: 211-218.
- Wang, Z., A. Chlus, R. Geygan, Z. Ye, T. Zheng, A. Singh, J.J. Couture, J. Caender-Bares, E.L. Kruger, and P.A. Townsend. 2020. Foliar functional traits from imaging spectroscopy across biomes in eastern North America. *New Phytologist*: 1-18.
- Wanner, W., X. Li, and A. H. Strahler. 1995. On the derivation of kernels for kernel-driven models of bidirectional reflectance. *Journal of Geophysical Research* 100:21077.
- Weiss, M., F. Baret, G.J. Smith, I. Jonckheere, and P. Coppin. 2004. Review of methods for in situ leaf area index (LAI) determination Part II. Estimation of LAI, errors and sampling. *Agricultural and Forest Meteorology* 121: 37-53.
- Weyermann, J., M. Kneubühler, D. Schläpfer, and M. E. Schaepman. 2015. Minimizing reflectance anisotropy effects in airborne spectroscopy data using Ross-Li model inversion with continuous field land cover stratification. *IEEE Transactions on Geoscience and Remote Sensing* 53:5814–5823.

- Whitehead, D. 1978. The Estimation of Foliage Area from Sapwood Basal Area in Scots Pine. *Forestry* 51(2): 137-149.
- Wickham, H., 2011. The split-apply-combine strategy for data analysis. *Journal of statistical software* 40, 1-29.
- Williams, L. J., A. Paquette, J. Cavender-Bares, C. Messier, and P. B. Reich. 2017. Spatial complementarity in tree crowns explains overyielding in species mixtures. *Nature Ecology and Evolution* 1.
- Wilson, J.W., 1960. Inclined point quadrats. *New Phytologist* 59, 1–7.
- Wright, I. J., P. B. Reich, M. Westoby, D. D. Ackerly, Z. Baruch, F. Bongers, J. Cavender-Bares, T. Chapin, J. H. C. Cornelissen, M. Diemer, J. Flexas, E. Garnier, P. K. Groom, J. Gulias, K. Hikosaka, B. B. Lamont, T. Lee, W. Lee, C. Lusk, J. J. Midgley, M.-L. Navas, Ü. Niinemets, J. Oleksyn, N. Osada, H. Poorter, P. Poot, L. Prior, V. I. Pyankov, C. Roumet, S. C. Thomas, M. G. Tjoelker, E. J. Veneklaas, and R. Villar. 2004. The worldwide leaf economics spectrum. *Nature* 428:821–827.
- Wu, J., H. Kobayashi, S.C. Stark, R. Meng, K. Guan, N.N. Tran, S. Gao, W. Yang, N. Restrepo-Coupe, T. Miura, and R.C. Oliveira. 2018. Biological processes dominate seasonality of remotely sensed canopy greenness in an Amazon evergreen forest. *New Phytologist*, 217(4): 1507-1520.
- Yachi, S., and M. Loreau. 1999. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 96:1463–1468
- Zarnetske, P.L., Q.D. Read, S. Record, K.D. Gaddis, S. Pau, M.L. Hobi, S.L. Malone, et al. 2019. Towards connecting biodiversity and geodiversity across scales with satellite remote sensing. *Global Ecology and Biogeography* 28: 548-556.
- Zhao, K., and S. Popescu. 2009. Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA. *Remote Sensing of Environment* 113: 1628–1645.
- Zhao, Y., Y. Zeng, Z. Zheng, W. Dong, D. Zhao, B. Wu, and Q. Zhao. 2018. Forest species diversity mapping using airborne LiDAR and hyperspectral data in a subtropical forest in China. *Remote Sensing of Environment* 213: 104-114.