Title

Leaf traits and canopy structure together explain canopy functional diversity: an airborne remote sensing approach

Author names and affiliations

Aaron G. Kamoske[a] (corresponding author), Kyla M. Dahlin[a,b], Shawn P. Serbin[c], Scott C. Stark[d]

[a] *Michigan State University, Department of Geography, Environment, & Spatial Sciences*

673 Auditorium Rd. #116, East Lansing, MI 48824

[b] *Michigan State University, Program in Ecology, Evolutionary Biology, & Behavior*

103 Giltner Hall, 293 Farm Lane #103, East Lansing, MI 48824

[c] *Brookhaven National Laboratory, Environmental and Climate Sciences Department*

98 Rochester St, Upton, NY 11973

[d] *Michigan State University, Department of Forestry*

480 Wilson Rd #126., East Lansing, MI 48824

Author email addresses

Aaron G. Kamoske: *kamoskea@msu.edu*

Kyla M. Dahlin: *kdahlin@msu.edu*

Scott C. Stark: *scstark@msu.edu*

Shawn P. Serbin: *sserbin@bnl.gov*

Abstract

Plant functional diversity is strongly connected to photosynthetic carbon assimilation in terrestrial ecosystems. However, many of the plant functional traits that regulate photosynthetic capacity, including foliar nitrogen concentration and leaf mass per area, vary significantly between and within plant functional types and vertically through forest canopies, resulting in considerable landscape-scale heterogeneity in three dimensions. Hyperspectral imagery has been used extensively to quantify functional traits across a range of ecosystems but is generally limited to providing information for top-of-canopy leaves only. On the other hand, lidar data can be used to retrieve the vertical structure of forest canopies. Because these data are rarely collected at the same time, there are unanswered questions about the effect of forest structure on the 3-D spatial patterns of functional traits across ecosystems. In the United States, the National Ecological Observatory Network's Airborne Observation Platform (NEON AOP) provides an opportunity to address this structure-function relationship by collecting lidar and hyperspectral data together across a variety of ecoregions. With a fusion of hyperspectral and lidar data from the NEON AOP and field-collected foliar trait data, we assessed the impacts of forest structure on spatial patterns of N. In addition, we examine the influence of abiotic gradients and management regimes on top of canopy %N and total canopy N (i.e. the total amount of N (g/m^2) within a forest canopy) at a NEON site consisting of a mosaic of open longleaf pine and dense broadleaf deciduous forests. Our resulting maps suggest that in contrast with top-of-canopy values, total canopy N variation is dampened across this landscape resulting in relatively homogeneous spatial patterns. At the same time, we found that leaf functional diversity and canopy structural diversity showed distinct dendritic patterns related to the spatial distribution of plant functional types.

Keywords

Hyperspectral imagery; airborne lidar; functional traits; forest canopies; forest structure; landscape ecology

1. 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).

2. MATERIALS AND METHODS

2.1 Site Description

Field measurements and remote sensing data were acquired in Talladega National Forest – Oakmulgee Ranger District (TALL) in west-central Alabama, USA (Figure 1). 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).

2.2 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/m^2.

2.3 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 S1: Fig. S1 and Fig. S2). 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

185 the sample using a pair of scissors that were sterilized between each sample. These pieces of leaf
186 material were imaged on a flatbed scanner and processed for area using imageJ software
187 (Schneider et al. 2012). We placed the leaf material in a paper coin envelope and dried the
188 samples at 70° C for at least 48 hours. After drying, we weighed the leaf samples and calculated
189 leaf mass per leaf area (LMA; g/m_L^{-2}; m_L = meter of leaf material). A subset of these samples (n
190 = 40, ~4 per species) were re-dried, ground to a fine powder using a ball mill (2000 Geno
191 Grinder; Spex Sample Prep, Cridersville, OH, USA), with 1.50-2.50 mg weighed in 0.1-mil tin
192 foil vials (AX26DR; Mettler Toledo, Columbus, OH, USA), and used to determine the C:N ratio
193 and elemental N content (g N/g leaf, %) employing a CHNS/O elemental analyzer operated in
194 CHN mode, according to the manufacturer's instructions (2400 Series II CHNS/0 Analyzer;
195 Perkin Elmer, Waltham, MA, USA) at Brookhaven National Laboratory (Upton, NY).
196 To build a leaf-scale model of %N to apply to the remaining samples in lieu of
197 determining foliar N in the lab, we used the laboratory calculated %N values and the associated
198 mean reflectance values for each wavelength, to train a partial least squares regression model
199 (PLSR; Serbin et al. 2014; Singh et al. 2015). We withheld 20% of the samples using a weighted
200 random approach, based on the %N values, as validation data (n = 8) that wasn't used to develop
201 the model and used the remaining samples (n = 32) as model training data. Using a jackknife
202 approach that randomly withholds 20% of the training data through 50 iterations, we calculated a
203 PRESS statistic (up to 15 components) for each iteration. We then selected the number of
204 components for our final model using the lowest PRESS statistic that balanced predictive
205 accuracy between the training and validation datasets. We applied these equations to the
206 validation data to assess model accuracy. We then applied the final PLSR coefficients to the
207 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).

2.4 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 S1: Fig.S8; 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 S1: Fig. S8. 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 2.

2.5 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).

2.6 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 S1: Fig.S1) 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/m_L^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/m_G^2; m_G = 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/m_G^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/m_L^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.

2.7 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.

2.8 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/m_G^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 S1: Fig.S7). 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/m_G^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-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.

3. RESULTS

3.1 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 1; Appendix S1: Fig.S3). 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-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.

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 1; Appendix S1: Fig.S4). 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 3a), which is comparable to the ranges of %N found in Eastern US temperate forests by Serbin et al. (2014).

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 1; Appendix S1: Fig.S5). 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 3b). 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).

3.2 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 S1: Fig.S6). 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 4), foliar biomass (g/m_G^2; Figure 3d), and LAI (m_L^2/m_G^2; Figure 3c) 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 that 7 m_L^2/m_G^2 were removed from the LAI raster (0.03% of raster pixels).

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 4a) and longleaf pine (Figure 4b) sample from our field data. The total amount of N in the white oak

sample was 6.99 g/m_G^2 while there was 7.93 g/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.

3.3 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 5, panels a, b, & c). Prior to normalization, we used linear regression to test the relationship between the two variables at each spatial resolution (Figure 5, 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.

To assess differences in spatial patterns across the landscape, we calculated variograms for the top of canopy %N and total canopy N (g/m_G^2) datasets (Figure 6). 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_G^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).

3.4 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 3). 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 7). 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.

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 (Fig. 8), seven variables appeared in no models, and 7 variables appeared in all the models (Table 2). 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/m_G^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 (Fig. 8), 4 variables appeared in no models and 2 variables appeared in all the models (Table 2). 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/m_G^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.

4. 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.

4.1 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/m_L^2; Poorter et al. 2009), and total canopy N (g/m_G^2; Cole and Rapp 1981; Fig. 9).

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.

4.2 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 S1: Fig.S1). 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 (Fig. 3a, b. c. & d). However, when three-dimensional canopy structure is considered (i.e. total canopy N), these distinct landscape patterns are dampened (Fig. 4).

Fig. 7 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 (Fig. 4a & b), and that these differences represent trade-offs since different PFTs exhibit similar total quantities of N (g/m_G^2) in their canopies (Fig. 9). In this case, differences over a leaf function-structural architecture trade-off produces the dampened spatial patterns we see in this landscape (Fig. 4).

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).

4.3 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.

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

558 *4.4 Model Uncertainty and Data Concerns*

559 There are many possible sources of error and uncertainty to consider when scaling traits from
560 leaf to landscape, including those related to field and GPS collections, laboratory equipment,
561 remote sensing sensors, and statistical methodologies. While we did not conduct a formal
562 assessment of uncertainty as it propagates through this study, our findings are within the ranges
563 reported in many field-based studies (see section 4.1). Our final PLSR models did show a
564 systematic bias of slightly underestimating N and LMA in needleleaf species (Appendix S1: Fig.
565 S3 and Fig. S4), which could partially explain the differing landscape-scale relationships
566 between total canopy and top of canopy N. This could possibly be improved by the inclusion of
567 forest structure metrics such as LAI in the PLSR models. However, due to the low density lidar
568 data we are forced to estimate structural traits at a coarser spatial resolution (10x10 meters) than
569 the HSI data (1x1 meter). Because some field samples are closer than 10 meters to one another,
570 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.

4.5 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.

5. 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.

6. 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.

7. 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.

8. LITERATURE CITED

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., 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.

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.

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.

Bater, C.W, and N.C. Coops. 2009. Evaluation error associated with lidar-derived DEM interpolation. Computers & Geosciences 35(2): 289-300.

660 Brown, M. J., and G. G. Parker. 1994. Canopy light transmittance in a chronosequence of mixed-
661 species deciduous forests. Canadian Journal of Forest Research 24:1694–1703.

662 Burnham, K. P., D. R. Anderson, and K. P. Huyvaert. 2011. AIC model selection and
663 multimodel inference in behavioral ecology: Some background, observations, and
664 comparisons. Behavioral Ecology and Sociobiology 65:23–35.

665 Chambers, J.Q., G.P. Asner, D.C. Morton, L.O. Anderson, S.S. Saatchi, F.D.B. Espirito-Santo,
666 M. Palace, and C. Souza Jr. 2007. Regional ecosystem structure and function: ecological
667 insights from remote sensing of tropical forests. Trends in Ecology & Evolution 22(8): 414-
668 423.

669 Clark, M. L., D. A. Roberts, and D. B. Clark. 2005. Hyperspectral discrimination of tropical rain
670 forest tree species at leaf to crown scales. Remote Sensing of Environment 96:375–398.

671 Cole, D. W., and M. Rapp. 1981. Elemental cycling in forest ecosystems. Pages 341–409 *in* D.
672 E. Reichle, editor. Dynamic Properties of Forest Ecosystems. Cambridge University Press,
673 Cambridge.

674 Colgan, M. S., C. A. Baldeck, J. baptiste Féret, and G. P. Asner. 2012. Mapping savanna tree
675 species at ecosystem scales using support vector machine classification and BRDF
676 correction on airborne hyperspectral and LiDAR data. Remote Sensing 4:3462–3480.

677 Collings, S., P. Caccetta, N. Campbell, and X. Wu. 2010. Techniques for BRDF correction of
678 hyperspectral mosaics. IEEE Transactions on Geoscience and Remote Sensing 48:3733–
679 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.

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. 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. 2014. Linking vegetation patterns to environmental gradients and human impacts in a mediterranean-type island ecosystem. Landscape Ecology 29:1571–1585.

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.

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.

701 Ellsworth, D. S., and P. B. Reich. 1993. Canopy structure and vertical patterns of photosynthesis
702 and related leaf traits in a deciduous forest. Oecologia 96:169–178.

703 Evans, J. R. 1989. Photosynthesis and nitrogen relationship in leaves of C3 plants. Oecologia
704 78:9–19.

705 Fahey, R. T., A. T. Fotis, and K. D. Woods. 2015. Quantifying canopy complexity and effects on
706 productivity and resilience in late-successional hemlock-hardwood forests. Ecological
707 Applications 25:834–847.

708 Field, C. B., and H. A. Mooney. 1986. The photosynthesis-nitrogen relationship in wild plants.
709 Pages 25–56 in T.J. Givnish, editor. On the Economy of Plant Form and Function.
710 Cambridge University Press, Cambridge.

711 Gelman, A. 2007. Scaling regression inputs by dividing by two standard deviations. Statistics in
712 Medicine 27:2865–2873.

713 Gougeon, F. A. 1995. Comparison of possible multispectral classification schemes for tree
714 crowns individually delineatedon high spatial resolution meis images. Canadian Journal of
715 Remote Sensing 21:1–9.

716 Gough, C. M., J. W. Atkins, R. T. Fahey, and B. S. Hardiman. 2019. High rates of primary
717 production in structurally complex forests. Ecology 100.

718 Harding, D.J., M.A. Lefsky, G.G. Parker, and J.B. Blair. 2001. Laser altimeter canopy profiles
719 methods and validation for closed-canopy, broadleaf forests. Remote Sensing of
720 Environment 76: 283-297.

721 Hardiman, B. S., C. M. Gough, A. Halperin, K. L. Hofmeister, L. E. Nave, G. Bohrer, and P. S.
722 Curtis. 2013. Maintaining high rates of carbon storage in old forests: A mechanism linking
723 canopy structure to forest function. Forest Ecology and Management 298:111–119.

724 Hardiman, B. S., G. Bohrer, C. M. Gough, C. S. Vogel, P. S. Curtis, S. Vogel, S. Curtis, and S.
725 Hardiman. 2011. The role of canopy structural complexity in wood net primary production
726 of a maturing northern deciduous forest. Ecology 92: 1818-1827.

727 Hijmans, R. J. 2019. raster: Geographic Data Analysis and Modeling. R package version 2.9-5.

728 Horton, J. D. 2017. The State Geologic Map Compilation (SGMC) geodatabase of the
729 conterminous Unite States (ver. 1.1, August 2017. U.S. Geological Survey data release.

730 Iglhaut, J., C. Cabo, S. Puliti, L. Piermattei, J. O'Connor, and J. Rosette. 2019. Structure from
731 Motion Photogrammetry in Forestry: a Review. Current Forestry Reports 5(3): 155-168.

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

735 Kampe, T. U., B. R. Johnson, M. Kuester, and M. Keller. 2010. NEON: the first continental-
736 scale ecological observatory with airborne remote sensing of vegetation canopy
737 biochemistry and structure. Journal of Applied Remote Sensing 4:043510.

738 Keddy, P.A. 1992. Assembly and response rules: two goals for predictive community ecology.
739 Journal of Vegetation Science 3(2): 157-164.

740 Kraft, N. J. B., R. Valencia, and D. D. Ackerly. 2008. Functional traits and niche-based tree
741 community assembly in an Amazonian forest. Science 322:580–582.

742 Leblanc, S. G., J. M. Chen, R. Fernandes, D. W. Deering, and A. Conley. 2005. Methodology
743 comparison for canopy structure parameters extraction from digital hemispherical
744 photography in boreal forests. Agricultural and Forest Meteorology 129:187–207.

745 Lefsky, M. A., W. B. Cohen, S. A. Acker, G. G. Parker, T. A. Spies, and D. Harding. 1999. Lidar
746 Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir
747 Western Hemlock Forests. Remote Sensing of Environment 70:339–361.

748 MacArthur, R. H., and H. S. Horn. 1969. Foliage Profile by Vertical Measurements. Ecology
749 50:802–804.

750 Martens, H., and M. Martens. 2000. Modifiend Jack-knife estimation of parameter uncertainty in
751 bilinear modelling by partial least squares regression (PLSR). Food Quality and Preference
752 5(16): 1-16.

753 Mascaro, J., G. P. Asner, H. C. Muller-Landau, M. Van Breugel, J. Hall, and K. Dahlin. 2011.
754 Controls over aboveground forest carbon density on Barro Colorado Island, Panama.
755 Biogeosciences 8:1615–1629.

756 Mevik, B., and R. Wehrens. 2007. The pls Package: Principal Component and Partial Least
757 Squares Regression in R. Journal of Statistical Software 18:1–23.

National Academics 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.

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.

Ollinger, A. 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.

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.

777 Pedro, M. S., W. Rammer, and R. Seidl. 2017. Disentangling the effects of compositional and
778 structural diversity on forest productivity. Journal of Vegetation Science 28:649–658.

779 Pisek, J., H. Buddenbaum, F. Camacho, J. Hill, J.L.R. Jensen, et al. 2018. Data synergy between
780 leaf area index and clumping index Earth Observation production using photon recollision
781 probability theory. Remote Sensing of Environment 215(15): 1-6.

782 Poorter, H., Ü. Niinemets, L. Poorter, I. J. Wright, and R. Villar. 2009. Causes and consequences
783 of variation in leaf mass per area (LMA): a meta-analysis. New Phytologist 182:565–588.

784 Reich, P. B., M. B. Walters, and D. S. Ellsworth. 1997. From tropics to tundra: Global
785 convergence in plant functioning. Proceedings of the National Academy of Sciences of the
786 United States of America 94:13730–13734.

787 Richardson, J. J., L. M. Moskal, and S. H. Kim. 2009. Modeling approaches to estimate effective
788 leaf area index from aerial discrete-return LIDAR. Agricultural and Forest Meteorology
789 149:1152–1160.

790 Running, S. W., R. R. Nemani, F. A. Heinsch, M. Zhao, M. Reeves, and H. Hashimoto. 2004. A
791 Continuous Satellite-Derived Measure of Global Terrestrial Primary Production.
792 BioScience 54:547.

793 Sabol, J., Z. Patočka, and T. Mikita. 2014. Usage of Lidar Data for Leaf Area Index Estimation.
794 GeoScience Engineering 60:10–18.

795 Scheuermann, C.M., L.E. Nave, R.T. Fahey, K.J. Nadelhoffer, and C.M. Gough. 2018. Effects of
796 canopy structure and species diversity on primary production in upper Great Lakes forests.
797 Oecologia 188: 405-415.

798 Schimel, D. S. 1995. Terrestrial biogeochemical cycles: Global estimates with remote sensing.
799 Remote Sensing of Environment 51:49–56.

800 Schimel, D., R. Pavlick, J. B. Fisher, G. P. Asner, S. Saatchi, P. Townsend, C. Miller, C.
801 Frankenberg, K. Hibbard, and P. Cox. 2015. Observing terrestrial ecosystems and the
802 carbon cycle from space. Global Change Biology 21:1762–1776.

803 Schimel, D., F. D. Schneider, and J. C. and E. Participants. 2019. Flux towers in the sky: global
804 ecology from space.

805 Schlapfer, D., R. Richter, and T. Feingersh. 2015. Operational BRDF effects correction for wide-
806 field-of-view optical scanners (BREFCOR). IEEE Transactions on Geoscience and Remote
807 Sensing 53:1855–1864.

808 Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of
809 Image Analysis. Nature Methods 9:671–675.

810 Serbin, S. P., A. Singh, B. E. McNeil, C. C. Kingdon, and P. A. Townsend. 2014. Spectroscopic
811 determination of leaf morphological and biochemical traits for northern temperate and
812 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.

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.

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.

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.

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. Fassnacht, S. Ustin, R. Dubayah, A. Schwigher, and P. Wennberg. 2017. ISS observations offer insights into plant function. Nature Ecology and Evolution 1: 1-4.

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.

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.

854 Tucker, C. J., and P. J. Sellers. 1986. Satellite remote sensing of primary production.
855 International Journal of Remote Sensing 7:1395–1416.

856 USDA Forest Service. 2005. Longleaf ecosystem-restoration project: Final environmental impact
857 statement, National Forests in Alabama. Talladega National Forest, Oakmulgee District,
858 USDA Forest Service, Brent, AL.

859 Valladares, F., L. Laanisto, Ü. Niinemets, and M. A. Zavala. 2016. Shedding light on shade:
860 ecological perspectives of understorey plant life. Plant Ecology and Diversity 9:237–251.

861 Wanner, W., X. Li, and A. H. Strahler. 1995. On the derivation of kernels for kernel-driven
862 models of bidirectional reflectance. Journal of Geophysical Research 100:21077.

863 Weiss, M., F. Baret, G.J. Smith, I. Jonckheere, and P. Coppin. 2004. Review of methods for in
864 situ leaf area index (LAI) determination Part II. Estimation of LAI, errors and sampling.
865 Agricultural and Forest Meteorology 121: 37-53.

866 Weyermann, J., M. Kneubühler, D. Schläpfer, and M. E. Schaepman. 2015. Minimizing
867 reflectance anisotropy effects in airborne spectroscopy data using Ross-Li model inversion
868 with continuous field land cover stratifiAcation. IEEE Transactions on Geoscience and
869 Remote Sensing 53:5814–5823.

870 Wright, I. J., P. B. Reich, M. Westoby, D. D. Ackerly, Z. Baruch, F. Bongers, J. Cavender-Bares,
871 T. Chapin, J. H. C. Cornelissen, M. Diemer, J. Flexas, E. Garnier, P. K. Groom, J. Gulias,
872 K. Hikosaka, B. B. Lamont, T. Lee, W. Lee, C. Lusk, J. J. Midgley, M.-L. Navas, Ü.
873 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.

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.

9. Tables

	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

Table 1. PLSR model results (R^2). All models have p-values <0.001.

917

		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
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

Table 2. 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.

10. Figures Captions

Fig. 1. 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.

Fig. 2. 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).

Fig. 3. 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.

Fig. 4. Map of total canopy N (g/m_G^2) and within canopy N (g/mG3) profiles from white oak (total foliar N = 6.99 g/m_G^2) and longleaf pine (total foliar N = 7.93 g/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.

Fig. 5. Maps of the scaled and centered differences between top of canopy %N and total canopy N (g/m_G^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.

Fig. 6. 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.

Fig. 7. Heatmaps showing the relationship between top of canopy %N, total canopy N (g/m_G^2), the normalized difference between these two measurements, and elevation. Y-axis units for each plot is given at the top of the plot.

Fig. 8. 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.

Fig. 9. 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).

11. Figures

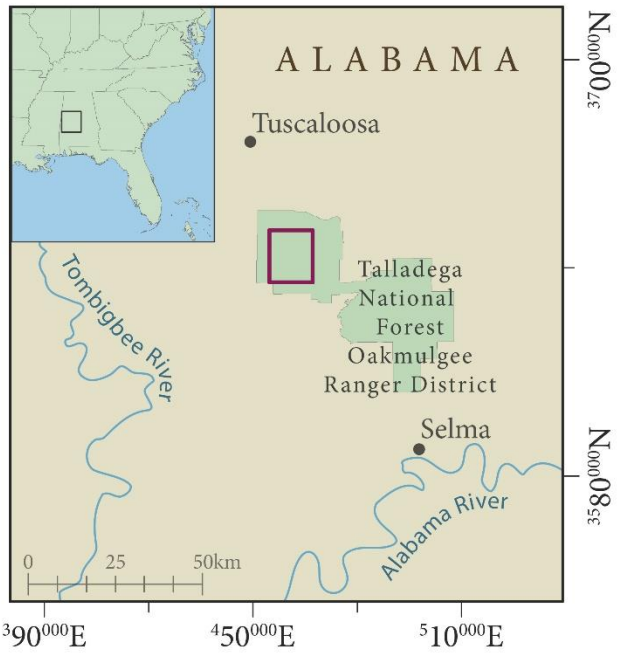


Figure. 1.

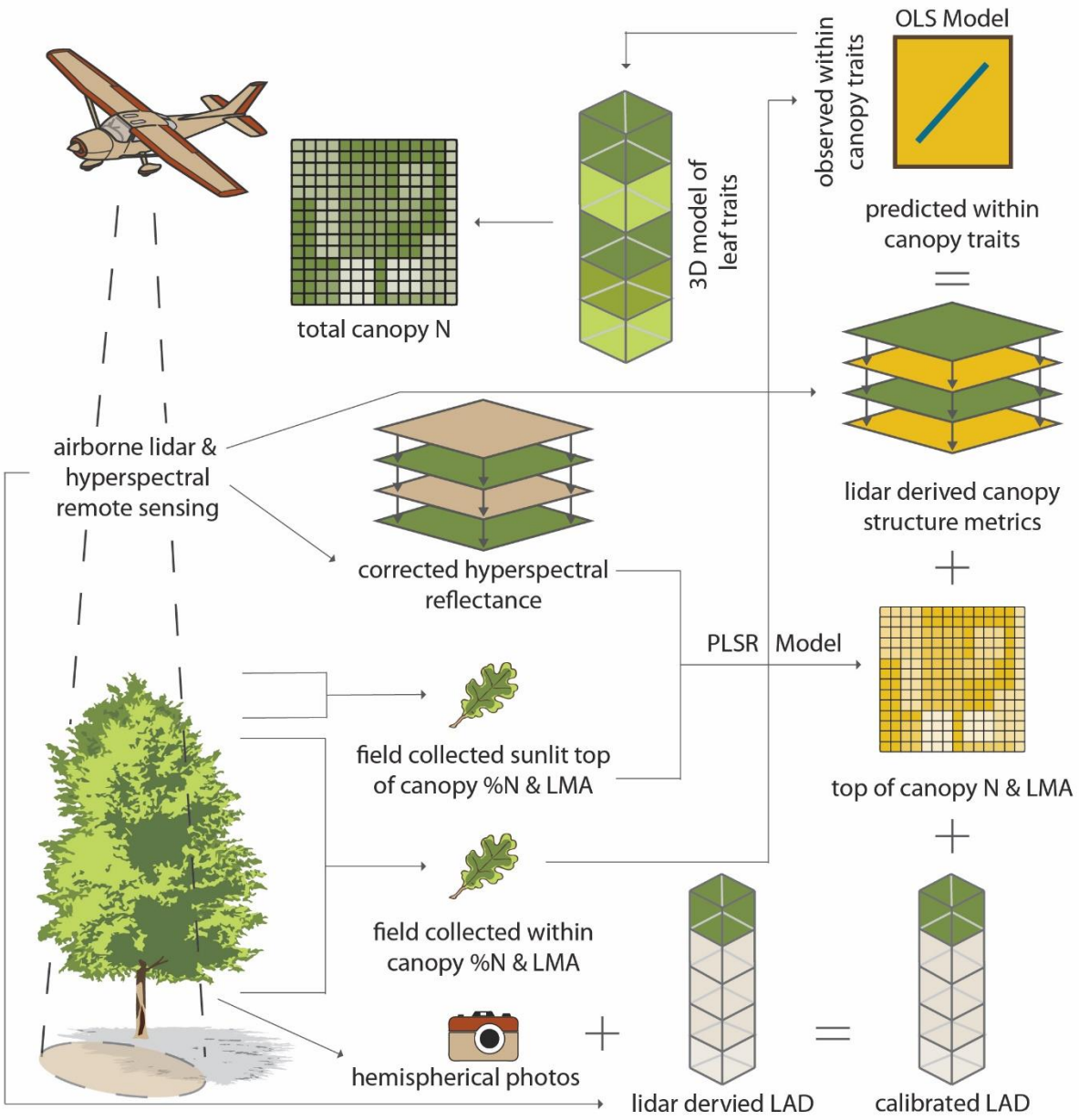

OLS Model
observed within canopy traits
predicted within canopy traits
3D model of leaf traits
total canopy N
airborne lidar & hyperspectral remote sensing
lidar derived canopy structure metrics
corrected hyperspectral reflectance
PLSR Model
field collected sunlit top of canopy %N & LMA
top of canopy N & LMA
field collected within canopy %N & LMA
hemispherical photos
lidar dervied LAD
calibrated LAD

Fig. 2

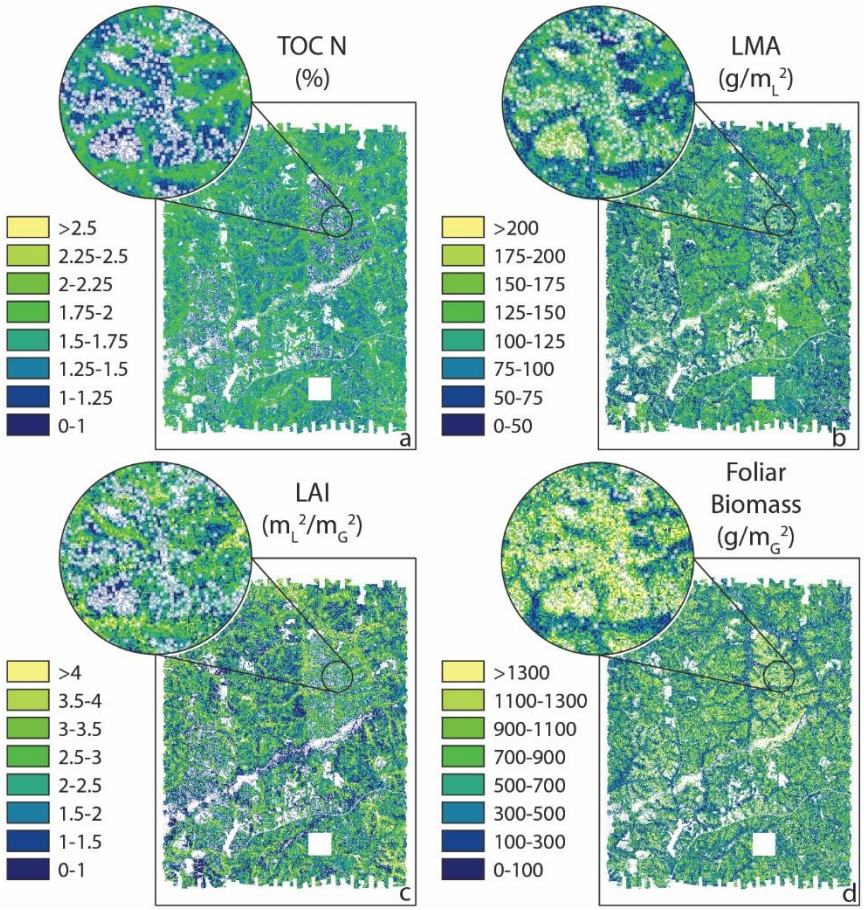

TOC N
(%)
>2.5
2.25-2.5
2-2.25
1.75-2
1.5-1.75
1.25-1.5
1-1.25
0-1
a
LMA
(g/m$_L^2$)
>200
175-200
150-175
125-150
100-125
75-100
50-75
0-50
b
LAI
(m$_L^2$/m$_G^2$)
>4
3.5-4
3-3.5
2.5-3
2-2.5
1.5-2
1-1.5
0-1
c
Foliar
Biomass
(g/m$_G^2$)
>1300
1100-1300
900-1100
700-900
500-700
300-500
100-300
0-100
d

Fig. 3

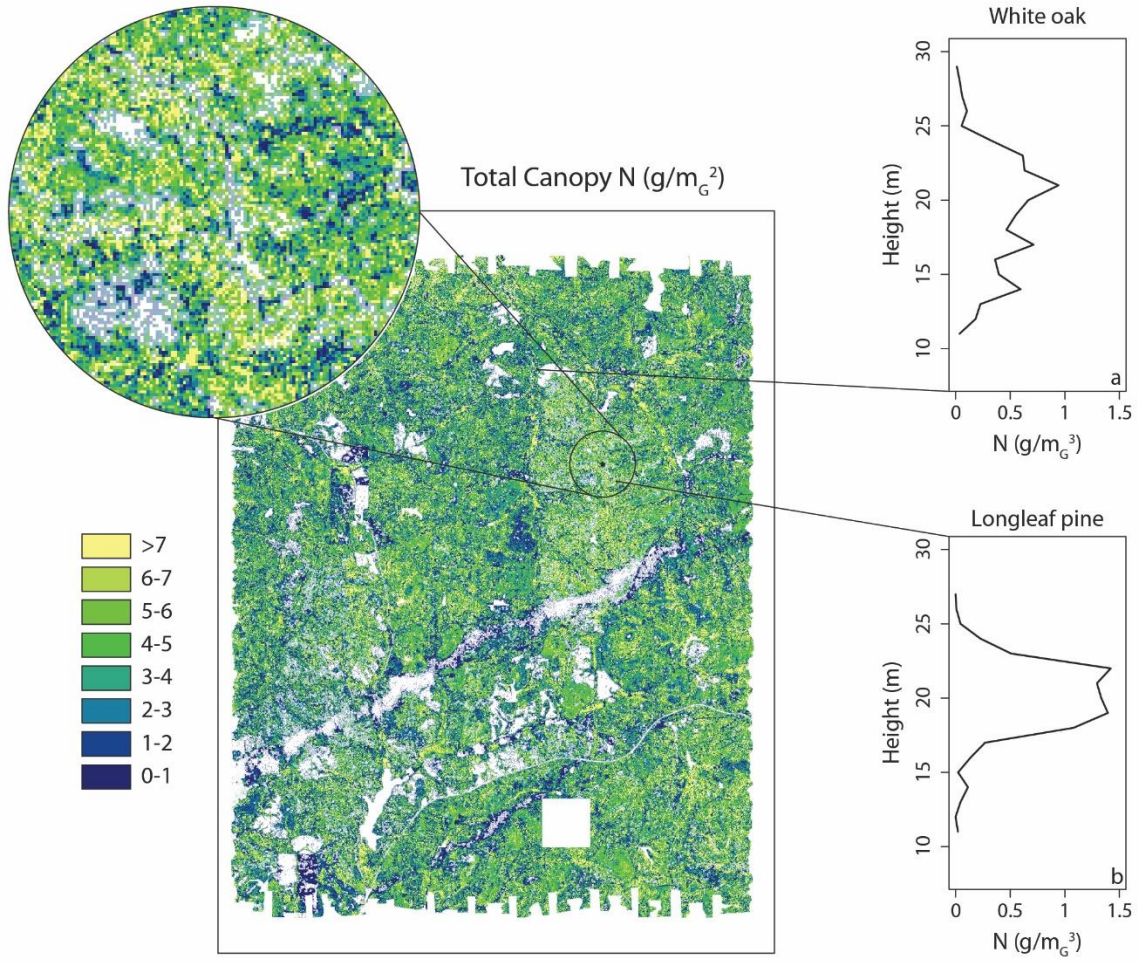

Total Canopy N (g/m$_G^2$)
>7
6-7
5-6
4-5
3-4
2-3
1-2
0-1
White oak
Height (m)
10
15
20
25
30
0
0.5
1
1.5
N (g/m$_G^3$)
a
Longleaf pine
Height (m)
10
15
20
25
30
0
0.5
1
1.5
N (g/m$_G^3$)
b

Fig. 4

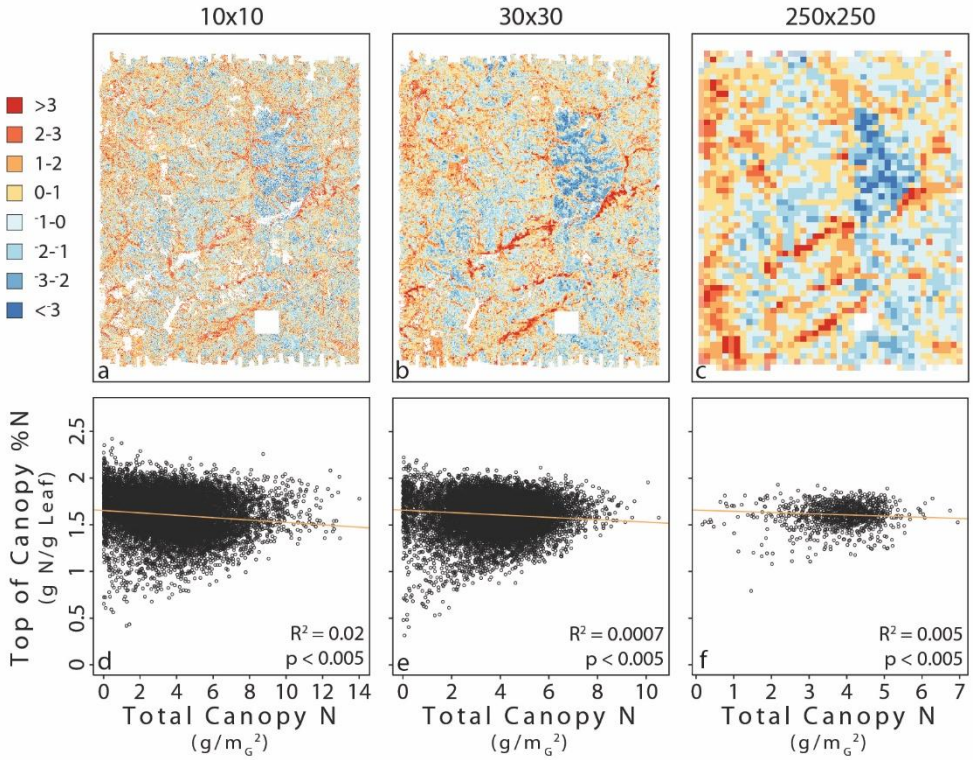

10x10
30x30
250x250
>3
2-3
1-2
0-1
-1-0
-2--1
-3--2
<-3
a
b
c
Top of Canopy %N
(g N/g Leaf)
0
0.5
1
1.5
2
2.5
R² = 0.02
p < 0.005
R² = 0.0007
p < 0.005
R² = 0.005
p < 0.005
d
e
f
0 2 4 6 8 10 12 14
0 2 4 6 8 10
0 1 2 3 4 5 6 7
Total Canopy N
(g/m_G²)
Total Canopy N
(g/m_G²)
Total Canopy N
(g/m_G²)

Fig. 5

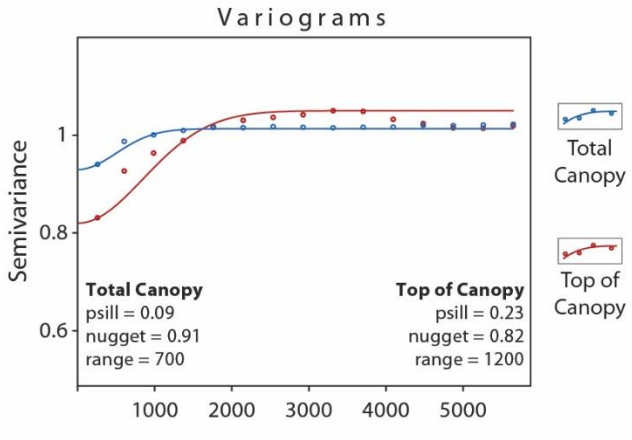

Variograms
Semivariance
1
0.8
0.6
1000
2000
3000
4000
5000
Total Canopy
psill = 0.09
nugget = 0.91
range = 700
Top of Canopy
psill = 0.23
nugget = 0.82
range = 1200
Total Canopy
Top of Canopy

Fig. 6

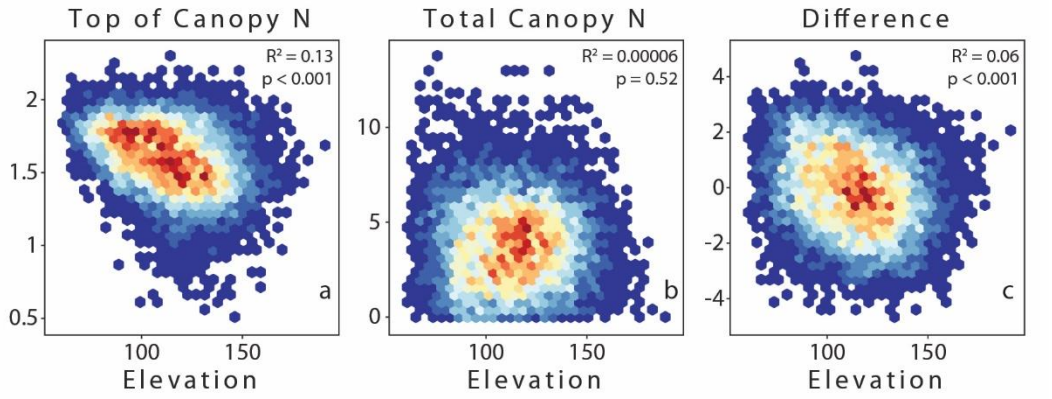

Top of Canopy N
R² = 0.13
p < 0.001
a
Total Canopy N
R² = 0.00006
p = 0.52
b
Difference
R² = 0.06
p < 0.001
c
Elevation
Elevation
Elevation

Fig. 7

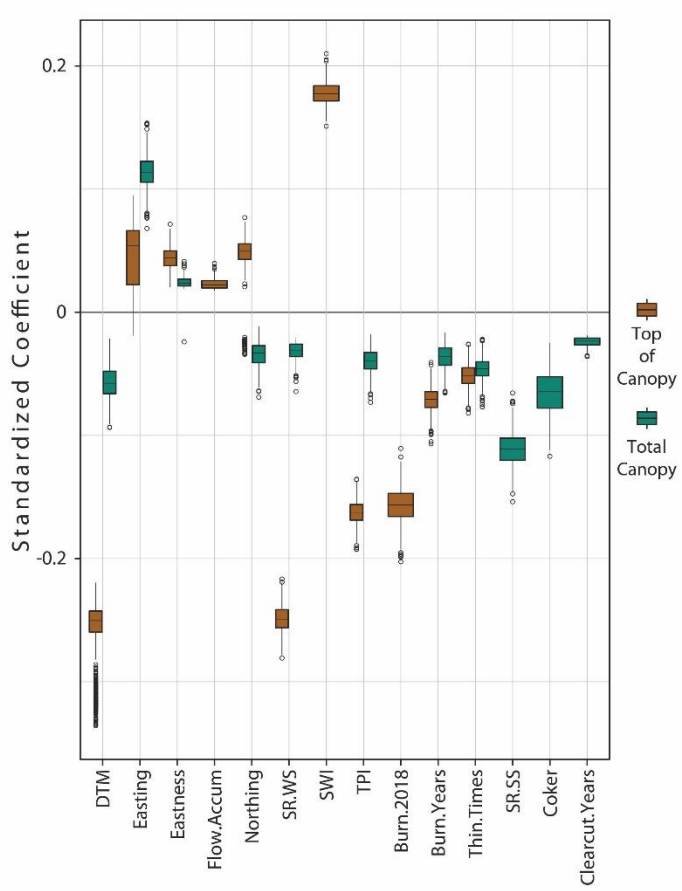

Standardized Coefficient
0.2
0
-0.2
DTM
Easting
Eastness
Flow.Accum
Northing
SR.WS
SWI
TPI
Burn.2018
Burn.Years
Thin.Times
SR.SS
Coker
Clearcut.Years
Top of Canopy
Total Canopy

Fig. 8

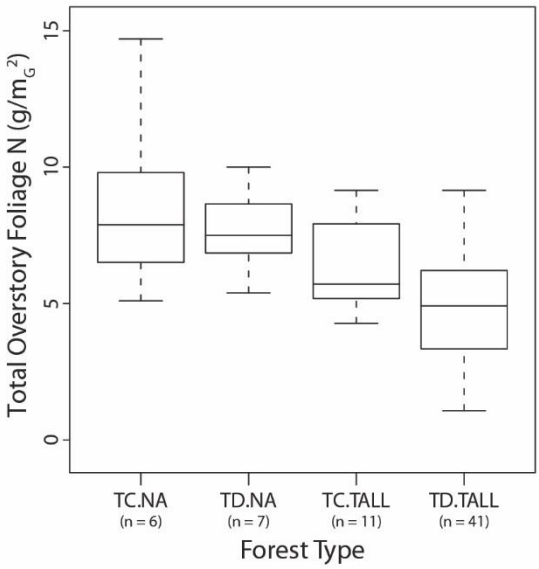

Total Overstory Foliage N (g/m$_G^2$)
0
5
10
15
TC.NA
(n = 6)
TD.NA
(n = 7)
TC.TALL
(n = 11)
TD.TALL
(n = 41)
Forest Type

Fig. 9

Appendix S1

Manuscript Title

Leaf traits and canopy structure together explain canopy functional diversity: an airborne remote sensing approach

Author names and affiliations

Aaron G. Kamoske[a] (corresponding author), Kyla M. Dahlin[a,b], Shawn P. Serbin[c], Scott C. Stark[d]

[a] *Michigan State University, Department of Geography, Environment, & Spatial Sciences*

673 Auditorium Rd. #116, East Lansing, MI 48824

[b] *Michigan State University, Program in Ecology, Evolutionary Biology, & Behavior*

103 Giltner Hall, 293 Farm Lane #103, East Lansing, MI 48824

[c] *Brookhaven National Laboratory, Environmental and Climate Sciences Department*

98 Rochester St, Upton, NY 11973

[d] *Michigan State University, Department of Forestry*

480 Wilson Rd #126., East Lansing, MI 48824

Journal Name

Ecological Applications

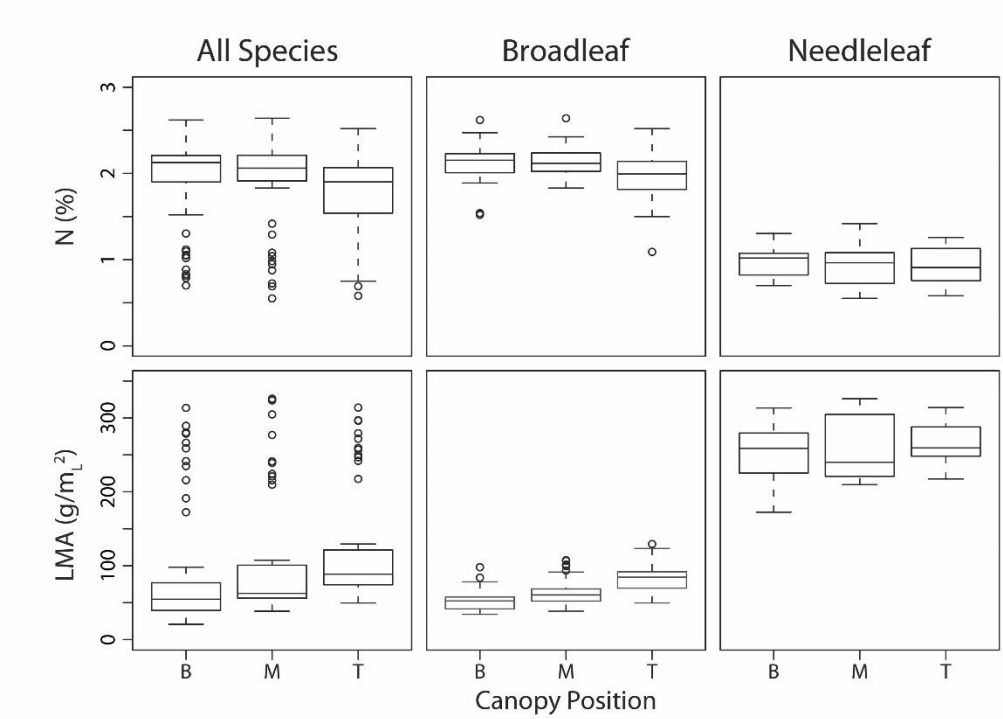


Fig. S1. Field data from TALL showing within canopy variation of LMA and %N. Canopy positions (Bottom, Middle, Top) were designated via visual assessment in the field.

	Bottom		Middle		Top		All	
	%N	*LMA*	*%N*	*LMA*	*%N*	*LMA*	*%N*	*LMA*
***Carya glabra* - 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
***Carya tomentosa* - 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
***Liquidambar styraciflua* - 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
***Liriodendron tulipifera* - 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
***Pinus palustris* - 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
***Pinus taeda* - 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
***Quercus alba* - 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
***Quercus falcata* - 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
***Quercus marilandica* - 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
***Quercus montana* - 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

Fig. S2. Mean and SD for field samples (%N and LMA), categorized by general position in the canopy which was determined by visual assessment.

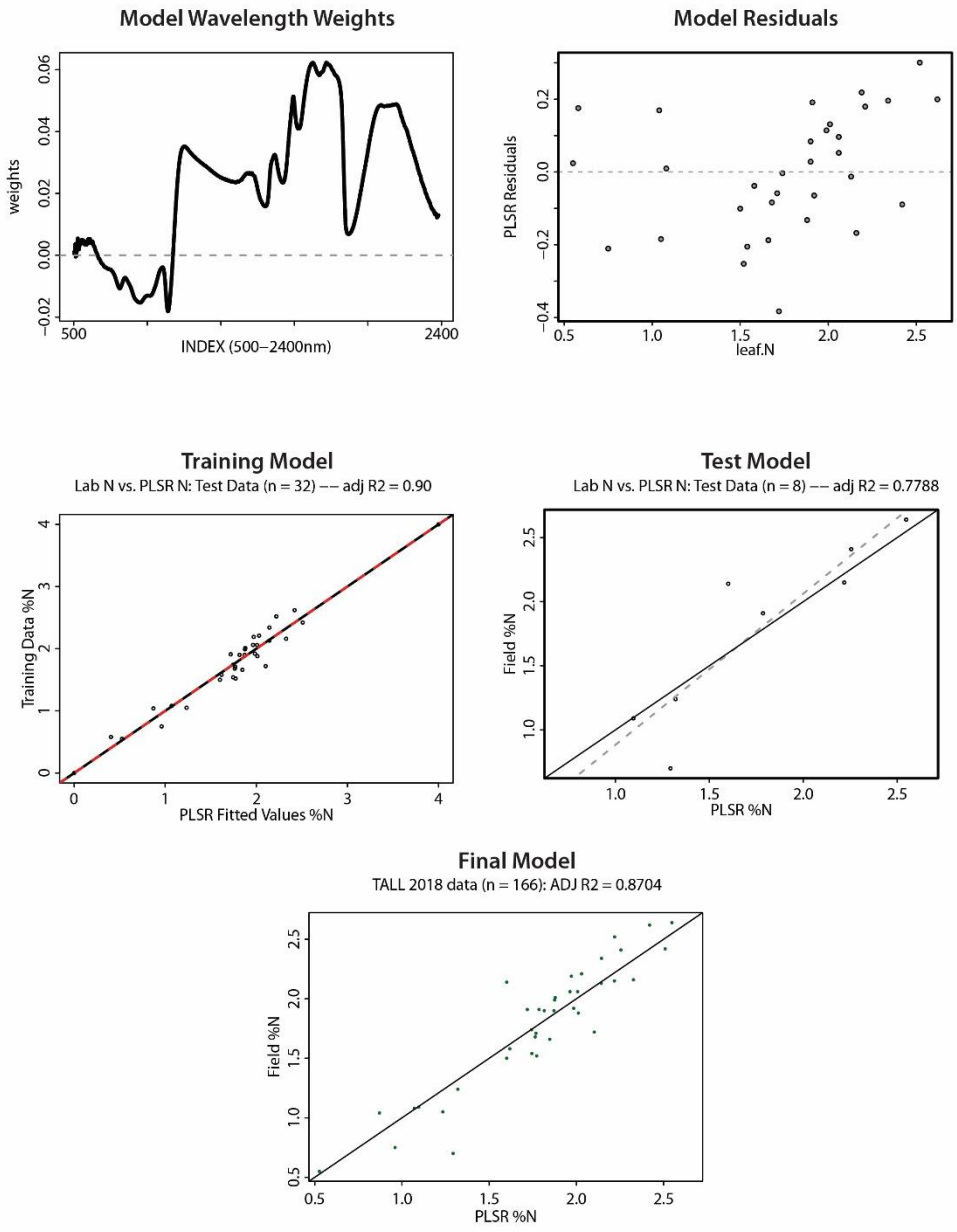


Fig. S3. PLSR output from laboratory %N estimation.

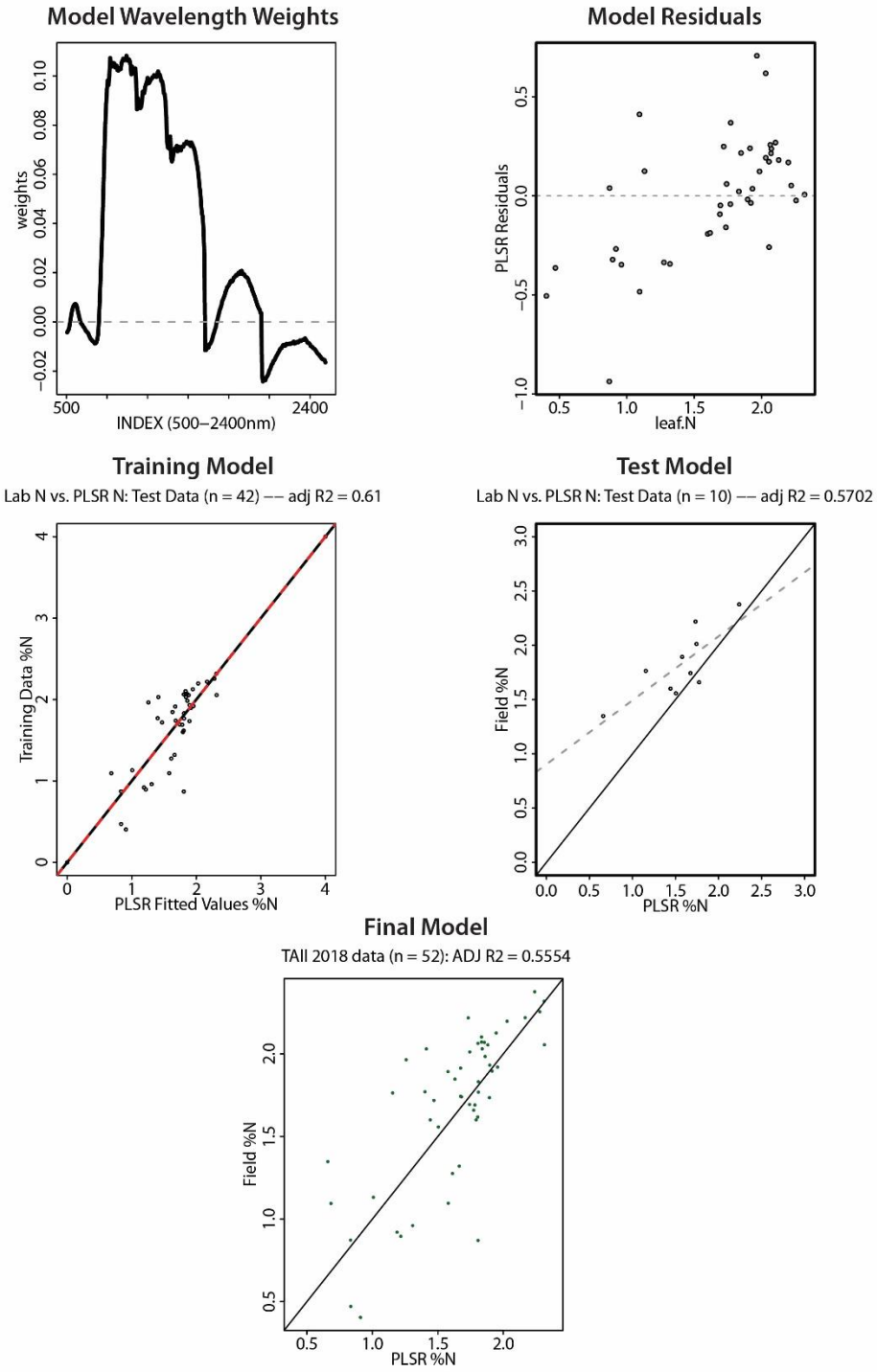


Fig. S4. PLSR output from HSI %N estimation.

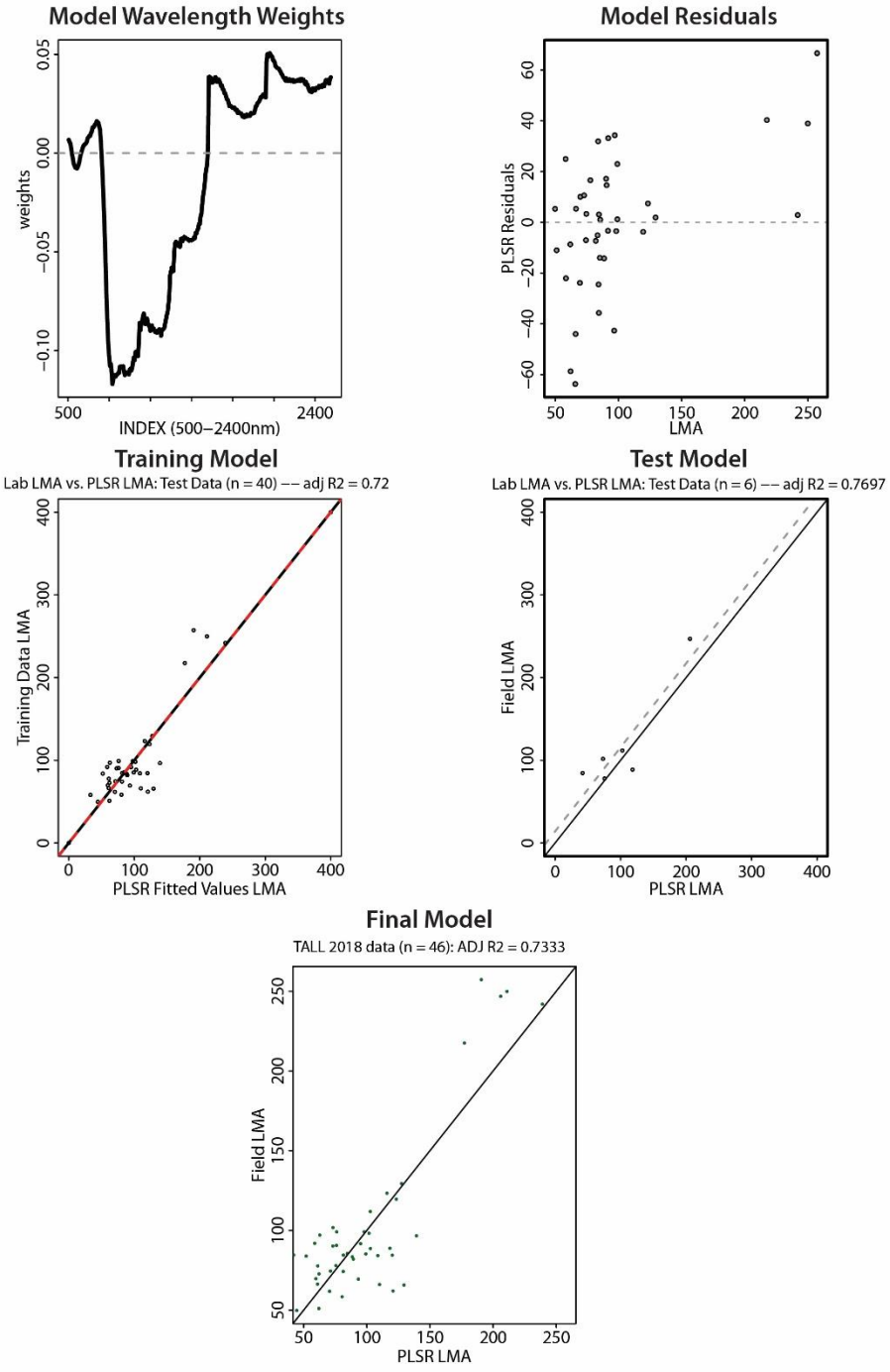


Fig. S5. PLSR output from HSI LMA estimation.

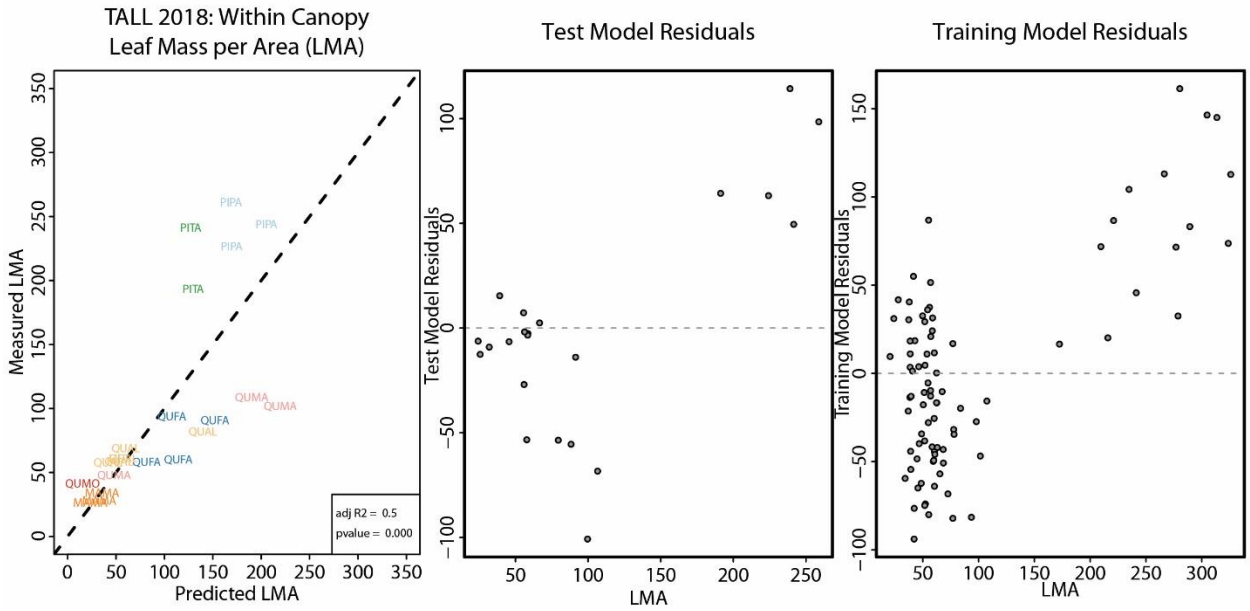
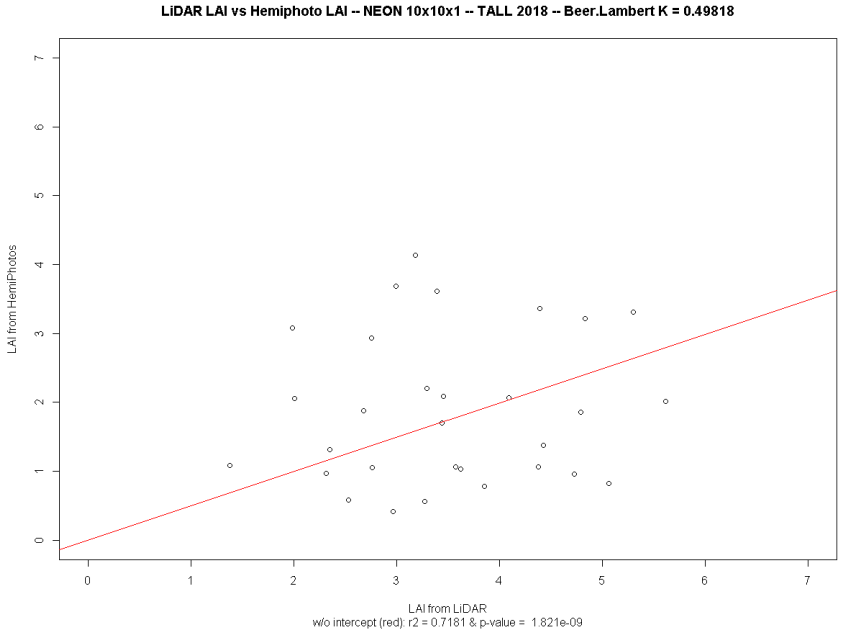


Fig. S6. Output from within-canopy trait prediction model: observed vs. predicted.

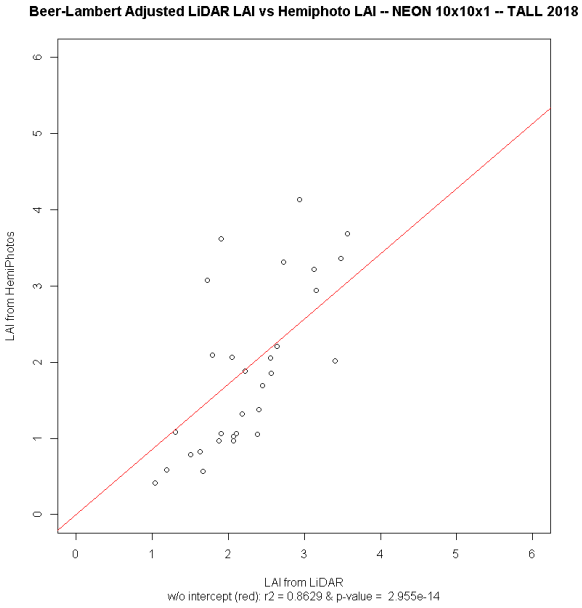
	Variable	Source
Topographic Variables	Digital Terrain Model	NEON AOP LiDAR; R programming language
	Eastness - Aspect	NEON AOP LiDAR; QGIS programming language
	Flow Accumulation	NEON AOP LiDAR; ArcGIS programming language
	Meters from Northern Collection Boundary	NEON AOP LiDAR; R programming language
	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 Substate	Horton 2017; https://doi.org/10.5066/F7WH2N65
Management Variables	Area burned in 2018 before NEON AOP flights	US Forest Service: https://data.fs.usda.gov/geodata/edw/datasets.php
	Times burned since 2007 (first year of data)	US Forest Service: https://data.fs.usda.gov/geodata/edw/datasets.php
	Times chemically treated since 2011 (first year of data)	US Forest Service: https://data.fs.usda.gov/geodata/edw/datasets.php
	Times clear cut since 1991 (first year of data)	US Forest Service: https://data.fs.usda.gov/geodata/edw/datasets.php
	Times thinned since 1993 (first year of data)	US Forest Service: https://data.fs.usda.gov/geodata/edw/datasets.php
	Years since last chemical treatment	US Forest Service: https://data.fs.usda.gov/geodata/edw/datasets.php
	Years since last clear cut	US Forest Service: https://data.fs.usda.gov/geodata/edw/datasets.php
	Years since last forest thinning treatment	US Forest Service: https://data.fs.usda.gov/geodata/edw/datasets.php
	Years since last prescribed burn	US Forest Service: https://data.fs.usda.gov/geodata/edw/datasets.php

Fig. S7. Names and references for abiotic and management variables.

81



82



83

84

85

86

Fig. S8. 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/kΔz 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.