Rethinking the fundamental unit of ecological remote sensing: Estimating individual level plant traits at scale

Sergio Marconi[1]*, Sarah J. Graves[2], Ben. G. Weinstein[3], Stephanie Bohlman[4], Ethan P. White[3]

[1]School of Natural Resources and Environment, University of Florida, Gainesville, FL, 32611 USA

[2]University of Wisconsin-Madison, Nelson Institute for Environmental Studies, Madison, WI, United States, 53706

[3]Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32603, USA

[4]School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32603, USA

Corresponding author: Sergio Marconi, sergio.marconi@weecology.org

Abstract

Functional ecology has increasingly focused on describing ecological communities based on their traits (measurable features affecting individuals' fitness and performance). Analyzing trait distributions within and among forests could significantly improve understanding of community composition and ecosystem function. Historically, data on trait distributions in ecosystems are generated by (1) collecting few leaves from a small number of trees, which suffers from limited sampling but produces information at the fundamental ecological unit (the individual); or (2) using remote sensing images to infer traits, producing information continuously across large regions, but as plots (containing multiple trees of different species) or pixels, not individuals.

Remote sensing methods that identify individual trees and estimate their traits would provide the benefits of both approaches, producing continuous large-scale data linked to biological individuals. We used data from the National Ecological Observatory Network (NEON) to develop a method to scale up functional traits measured on 160 trees to the millions of trees within the spatial extent of two NEON sites. The pipeline consists of three stages: 1) image segmentation, to identify individual trees and estimate their structural traits; 2) ensemble of models to infer leaf mass area (LMA), nitrogen, carbon, and phosphorus content using hyperspectral signatures; and 3) making predictions for segmented crowns for the full remote sensing footprint at the NEON sites.

Nitrogen, LMA and phosphorus models R^2 values ranged from 0.50 to 0.75 on held out test data. The ensemble approach performed better than single partial least squares generalized linear regressions models. Models for carbon performed poorly compared to other traits (R^2 of 0.22). Crown segmentation produced the highest uncertainty in the pipeline, generally due to over-segmentation. Despite this, crown traits predictions performed significantly better than comparable $1m^2$ pixel predictions, resulting in improvement of R^2 on test data of between 0.07 to 0.20 points.

We used the pipeline to produce individual level trait data for ~5 million individual crowns, covering a total extent of ~360 km^2. This large data set allows testing ecological questions on landscapes scales, for example revealing leaf chemical traits were correlated with physical traits and environmental conditions including height and elevation.

Keywords:

1. Introduction

Functional traits are biochemical, physiological and structural characters that influence organism performance or fitness (Nock et al., 2016). They are central to how organisms perform under different environmental conditions, interact with other species, and influence the ecosystems to which they belong (McGill 2006, Dwyer et al., 2017, Collalti et al., 2019). For individual organisms, traits influence core demographic parameters including survival and reproduction. At the species level, traits can influence species distributions and how species respond to changes in land use and climate). At the ecosystem level, organismal traits influence biogeochemical cycles and habitat availability for other species (e.g., Fisichelli et al., 2015). Given their central importance across multiple levels of organization, understanding how traits vary within and among species, across environmental gradients, and through time is essential to understanding many areas of ecology and predicting how ecological systems will change in the future (McGill 2006, Lawler et al. 2010, Valladares et al., 2014).

In plants, two primary categories of plant traits are leaf (e.g. nitrogen content, leaf mass per area, carotenoid, Fe, Zn) and tree structural traits (e.g. height, dbh, canopy size). Historically, structural traits are collected for thousands of trees per targeted area, while studies of leaf chemical traits have relied on collecting samples of a few leaves from a small number of trees to estimate the average trait values for each species. These average values are then used to explore how ecosystem level leaf traits vary biogeographically or through time by assuming that all individuals of a species in a region share the same trait value (Swenson et al., 2010, Clark et al., 2016). This approach is necessary because it is expensive and time consuming to collect individual level leaf trait data, but it fails to describe trait variation within species driven by evolution and plastic responses to the conditions an individual or population experiences

(Messier et al., 2017, Niinemets et al, 2017, Muller et al. 2010, Nicotra et al. 2010, Albert et al. 2010, Callaway et al. 2003). Moreover, a large discrepancy between leaf and tree structural trait sample size may affect the generality of relationships observed between them. This limitation is magnified when studying changing environments (across space or time) because of bias in where the data for each species is collected. Data are typically collected in small subsets of the full range of conditions that species experience and are often selected in a biased manner that fits the purpose of the original studies (e.g., selecting individuals of a particular health status, size or species). As a result, it has been suggested that trait data contains less information on the relationship between plant communities and their environment than species data (Clark, 2016). Measuring traits systematically across geographic gradients would address this limitation, but this is challenging with traditional methods (Anderson-Teixeira 2015).

An alternative approach that allows continuous estimation of traits across the landscape is to use remote sensing data (Homolova et al. 2013, Houborg et al., 2015, Kerr & Ostrovsky, 2003). Hyperspectral remote sensing imagery can be used to estimate the chemical composition of sunlit leaves, and laser imaging, detection, and ranging (LiDAR) can infer vertical and horizontal vegetation structure. Together, these two data sources can be used to estimate many of the standard leaf and structural tree traits for trees (Serbin et al., 2014, Singh et al., 2015, Asner et al., 2017, Barbosa et al., 2017). However, remote sensing applications generally use either the pixel (the smallest resolution component of the image) or the plot (a region of space typically containing multiple individuals) as the fundamental unit. This is a natural result of the structure of the remote sensing data combined with the difficulty of linking individual crowns in remote sensing images to field data. However, it results in a disconnect between the remote sensing analysis and the fundamental biological unit: the individual (Liu et al., 2016, De Angelis, 2018,

Marconi et al. 2019). Individuals reproduce, interact with their neighbors, and exhibit plastic responses to environmental conditions. Populations of individuals evolve in response to natural selection. As a result, our understanding of many biological processes is grounded in the individual and many field-based survey methods focus on collecting data with individual trees as the primary unit of measurement. While the plot is a scale widely used in forestry (Newnham et al, 2015), it loses important information about the relative distribution of individuals and their traits within the plot and hence much of the embedded information about community structure. To fully understand how traits vary across space, change through time, and are determined by biological processes, it is important to develop approaches linking these fundamental characteristics to individual trees that can be applied at scales of hundreds of km^2.

Despite its importance for biological research, few studies have attempted to model leaf traits using remote sensing at the crown level (but see Chadwick & Asner 2016, Martin et al., 2018). Even when crown level models are developed, the resulting leaf trait predictions are made for pixels, not crowns, because of the challenges in crown segmentation, hyperspectral data resolution, lack of co-registered or potential misalignment between LiDAR and hyperspectral data (Blaschke, 2010). Similarly, there are few studies estimating structural traits (like crown height and area) at crown level, with traditional methods predicting tree height and cover at plot level. Perhaps as a result of these differences, structural and chemical leaf traits are not currently predicted together at large scales. Consequently, ecology lacks the large scale individual level trait estimates that are necessary to fully understand tradeoffs between between leaf and tree structural traits, as like as their relationship with species and the environment. Data from NEON Airborne Observatory Platform NEON (AOP) provide co-registered high resolution LiDAR and

hyperspectral data, whose integration represents a great opportunity to circumvent these challenges.

To address this gap, we developed a pipeline for making crown level trait predictions at scales of 500 km^2 with associated uncertainties on both crown segmentation and trait estimation. Building on Chadwick & Asner (2016) and Martin et al. (2018), we: 1) identify individual crowns in remote sensing imagery that can be associated with field-based trait measurements; 2) build models relating the remote sensing measurements to field based traits; and 3) apply those models to estimate trait values and examine patterns of tree structural and chemical traits from individual to landscape scales. We advance the state of the art (Chadwick & Asner 2016, Martin et al. 2018) in all three steps in this pipeline. For identifying individual tree crowns, we adopted a digital field-mapping approach to accurately locate individual crowns in remote sensing images of closed canopy forests (Graves et al. 2018). The field-mapped individual tree crowns make it possible to calibrate, validate, and compare different crown segmentation algorithms, and identify uncertainty in the first step of the pipeline. When modeling traits based on remote sensing data, we use new methods that allow uncertainty to be accurately estimated by producing likelihoods and assess whether crown-level modeling produces more accurate predictions than pixel level modeling. Most importantly we estimate traits at the crown level instead of at the pixel level. This combination of approaches allows us to produce and distribute derived data products on the location, size, shape, and leaf traits of millions of individual trees distributed over tens of thousands of hectares. Providing data as a downloadable data product will facilitate individual level research at scales over 100 times greater than current field-based datasets allow.

2. Methods

In our pipeline, predicting crown level leaf and structural traits from remote sensing we used: 1) field measurements of functional traits for building and evaluating models; 2) data on the shape and location of individual tree crowns (ITCs), for building accurate models and assessing uncertainty in crown segmentation algorithms; and 3) high resolution remote sensing LiDAR (for crown segmentation and estimation of structural traits) and hyperspectral data (for estimation of leaf chemical traits). To obtain these components, we built on the National Ecological Observatory Network's (NEON) infrastructure by combining NEON's LiDAR and hyperspectral data with field data we collected at NEON sites on leaf traits as well the location and shape of individual tree crowns.

2.1. Site descriptions

The study was conducted at two core terrestrial NEON sites; Ordway Swisher Biological Station in Florida (OSBS, NEON Domain 03) and the Oakmulgee Management District of Talladega National Forest in Alabama (TALL, NEON Doman 08). The two sites (Figure S1) have a mix of deciduous, evergreen, and mixed forest types (Homer et al., 2012). Upland areas at both sites are dominated by fire-tolerant oaks and pine species, primarily longleaf pine (*Pinus palustris*). The longleaf pine at OSBS forms open stands whereas the longleaf pine canopy is closed at TALL. Lowland areas near lakes or wetlands (OSBS), and riparian areas (TALL) are dominated by closed canopy hardwood forests (Beckett and Golden 1982, Cox and Hart 2015).

2.2. Remote sensing data

All aerial remote sensing data products were provided by NEON Airborne Observation Platform (NEON-AOP, Table 1). We used data from the May 2014 flight for OSBS, and the June 2015 for TALL. We used the raw L1 data products: (1) “classified LiDAR point cloud”, and (2) “hyperspectral surface reflectance” data, orthorectified and atmospherically corrected. The LiDAR data consist of 3D spatial point coordinates (4-6 points/m^2) which provides high resolution information about crown shape and height. These data are released in 1 km x 1 km tiles. Hyperspectral reflectance data consist of 1m^2 spatial resolution images with 426 channels (or bands), each one collecting the magnitude of reflectance in 5 nm wide interval of wavelengths, ranging from visible to near infrared light (from 350 to 2500 nm). These images were provided as multiple, ~15 km x 0.8 km, flight lines with a total area of ~215 km^2 in OSBS, and ~145 km^2 in TALL. The hyperspectral images were provided as “prototype” data, pre-processed differently than post 2017 data, and delivered on hard drives (Table 1).

2.3. Field Data

During this project leaf trait data collected by the NEON Terrestrial Observation System (TOS) were not available. For this reason, we collected leaf samples for 226 trees near NEON inventory plots, which are randomly located across the study site and stratified by land cover type. The sampled trees were located away from major roads (since roads can affect the spectral reflectance of the remote sensing data), had crowns that were visible from above and could be identified in the image, and had sunlit branches that were accessible for leaf collection

(https://www.usa-china-biodiversity.org). Trees were not sampled within NEON plots to avoid disturbing NEON's long-term monitoring efforts.

Field data were collected in the early part of the peak growing season in 2015 and 2016. In particular, 87 individual trees (for 16 species) were sampled from OSBS in May-June 2015, and 137 individuals (29 species) in July 2015 from TALL. Sunlit leaves were sampled with a shotgun. Immediately after collection, the leaves were placed in a labeled plastic bag and stored in a cooler. Leaf samples were further processed within 2-4 hours of collection. Samples were analyzed for nitrogen (%N), carbon (%C), phosphorus (%P) by weight, and leaf mass per area (LMA) according to standard protocols (Murphy and Riley 1962; Cornelissen et al. 2003). For LMA, a random sample of leaves was weighed and scanned on a flatbed scanner to determine leaf area. The leaves were then dried at 65 C for at least 72 hours and reweighed. A separate sample of leaves was dried and used for measurements of %N, %C, and %P.

The year following leaf and airborne data collection, individual trees were mapped in the remote sensing images using a field tablet and GIS software. This process involved mapping individual tree crowns on the hyperspectral image in the field to ensure the sampled trees matched directly with image pixels (Graves et al., 2018). This individual tree crown (ITC) data is necessary for accurately linking field measurements with pixels from remote sensing spectral data and for quantifying uncertainty in crown the segmentation algorithms.

2.4. The algorithm pipeline

We developed a modular pipeline based on three steps: (1) build and evaluate crown segmentations from LiDAR data; (2) develop an ensemble of statistical models to infer tree structural traits diameter at breast height, DBH (cm), crown area, CA (m^2) and stem height H (m)

from LiDAR data, as like as leaf traits mass per area (LMA, g/m^2), nitrogen (%N), phosphorus (%P), and carbon (%C) for individual trees from hyperspectral data; (3) make predictions for each individual tree crown in each NEON site. We predicted individual tree height and crown size (structural traits) by calculating the polygon area and the maximum height within polygon respectively. For each crown we also extracted values of elevation, slope, aspect and the LAI determining Index (Delialieux 2008), provided as NEON AOP data products (Table 2), aiming to build a comprehensive dataset including topographic, leaf chemical and tree structural traits for any tree detected within the AOP footprint. The structure of the pipeline presented in this paper is summarized in Figure 1.

2.5. Leaf chemistry model

We extracted the hyperspectral pixels from within each crown for sampled trees, whose crown boundaries were delineated in the field (ITC). Data were normalized and preprocessed to reduce noisy, highly mixed, or unvegetated pixels (Figure S.2). We filtered out shadowed or low vegetation pixels using thresholds for both the near infrared band (reflectance in 860 nm < 0.3) and the Normalized Difference Vegetation Index (NDVI < 0.7) (Colgan et al., 2012, Graves et al., 2016). We did not use a height threshold (Asner et al., 2017), since height was indirectly used to select pixels belonging to each crown in the field mapping process. We normalized the spectral values for each pixel by dividing each spectral vector by its root sum of squares. We used this method to reduce the effect of peripheral light and shadows within each crown (Singh et al., 2015).

Field data were split at the tree level, and stratified by species, into training (n = 168), validation (n = 42), and test sets (n = 15). Since the two sites have similar species composition, we

aggregated the two datasets to build a joint model. As is common for trait studies, our field data on foliar traits had a single value for each individual tree. However, individual crowns contain multiple pixels, and crowns vary in the number and quality of these pixels. In each crown, some pixels are better than others for linking traits to hyperspectral signatures because some register light reflected primarily by leaves, others from branches, understory, or ground. To capture the range of possible models from different pixels in each crown, we randomly sampled one pixel from each training crown 1000 times, and used the resulting 1000 vectors of pixels (one pixel for each crown in the dataset) to build 1000 independent partial least squares generalized linear regression (PLS-GLR) models (Bastien et al., 2005, Bertrand et al, 2014). Hereafter we we refer to these models as "Single Pixel Models" (SPM). We used PLS-GLR (instead of regular PLS regressions) because it calculates component vectors using maximum likelihood estimation, which allows the calculation of AIC values to compare model goodness of fit and provides a measure of uncertainty for each model prediction. We treated site (OSBS versus TALL) as a fixed effect to account for site specific ancillary conditions. The number of components included in each model were determined using 10-fold cross-validation (CV) using the PRESS statistic on the training set. Following Feilhauer et al. (2015), we used the log transformation of the first derivative of the spectra for %P and %C.

To compare pixel and crown based approaches to modeling leaf traits, we compared three modeling strategies: 1) A traditional pixel based approach with no information on tree crown identity using the 100 best SPMs; 2) A pixel based approach that included information on crown identity using an ensemble of the 100 best SPM models with individual models weighted based on their performance on validation data using delta AIC based multi-model averaging (Burnham & Anderson 2002) that we refer to as the "Ensemble Pixel Based Model" (EPBM); 3) A crown

based ensemble model that focuses on the crown as the fundamental unit of remote sensing by calculating an average trait value across all pixels in a crown using the EPBM calculated for each pixel that we refer to as the "Crown Ensemble Aggregation Model" (CEAM). We built these models from two different pixels datasets: (1) pixels belonging to ground delineated crowns (Graves et al., 2018), or (2) pixels belonging to algorithmically delineated crowns (Silva et al., 2016). We compared PBMs, EPBMs, and CEAMs for the two datasets to quantify how errors in estimating polygons produced by the crown segmentation would affect our ability to predict leaf traits.

All models were tested on each pixel of the 15 crowns withheld in the test dataset. The test data were not used at any phase of the fitting or the ensemble process. Accuracy was evaluated using the predictive coefficient of determination (R^2) and the root-mean-square error (RMSE). The coefficient of determination produces values between 1 and negative infinity, where negative values indicate that the model predictability is lower than the sample average. As such, negative R^2 values indicate that the statistical model did not learn any meaningful information from the data. A value of 1 indicates that predicted values perfectly match observations. We evaluated the uncertainty of predictions for each model using the coverage of the 95% prediction interval (95PI). The prediction interval is the range values that is expected to contain 95% of the observed data points, and therefore a model with good estimates of uncertainty should have approximately 95% of the test data falling within this range. Since the CEAM was generated by the ensemble of the 100 best SPMs, we estimated the 95PI for CEAM predictions by averaging the error functions for the best 100 SPMs. We used the same data split, data transformation and PLS-GLR parameterization for all models. For pixel-based estimations, we compared ground measures of LMA, N, C, and P with predictions from each pixel in the test dataset. For crown-

based estimations, we averaged pixel-based predictions belonging to all of the pixels in the crown. The same rationale was used for comparing pixel and crown-based uncertainty.

2.6. Tree structural traits and Crown Segmentation

We used pre-classified point cloud LiDAR data to create a 0.5 m^2 resolution canopy height model (CHM). Building a higher resolution CHM than the one provided by NEON data products allowed us to better capture the canopy structure and produce better predictions of individual tree crowns. We used the CHM to determine the number of trees in the scene using a method based on local maxima filtering (Popesco et al., 2004), then we applied three unsupervised crown segmentation algorithms (Dalponte & Coomes 2016, Silva et al., 2016, and a watershed algorithm as in Barnes et al., 2014)**.** These algorithms were chosen because they were fast enough to scale up to hundreds of km^2, and they performed well when applied on the same data in a similar context (Marconi et al, 2019). We used a single parametrization (Table S1) for each method at both sites to facilitate cross-model and cross-sites comparison. Data were processed using the lidR R package (Roussel & Auty, 2017). We used the mean pairwise Jaccard index coefficient (Real & Vargas, 1996) to choose the most suitable algorithm, and quantify uncertainty in individual tree crowns (ITCs). The Jaccard index was calculated by comparing ITCs collected in the field with the single most overlapping predicted crown. Field delineated crowns not overlapping with any crown segmented by the algorithm were labelled as undetected. The Jaccard index was chosen to evaluate the segmentation accuracy because it is fully automatable, has good continuity, and can be used when only a subset of ITCs in the image have been measured on the ground (Zhen et al., 2016, Marconi et al. 2019). The Silva et al. (2016) segmentation algorithm performed the best on our data (Table S1, see results section), consistent

with results from a multi-group data science competition (Marconi et al. 2019). Therefore, this algorithm was used to extract ITCs from the CHM data for the full remote sensing footprint for both NEON sites.

We estimated tree structural traits from the derived polygons and the CHM. Crown area (CA) was calculated from the polygon geometry using the geoPandas python package (https://geopandas.readthedocs.io/). Tree height (H) was extracted from the CHM as the maximum height within each ITC. Diameter at breast height (DBH) was calculated using an allometric regression model relating the log-transformed DBH to log-transformed height and canopy area. The parameters of the allometric relationships were determined by fitting the relationships to the NEON terrestrial vegetation structure data (Table 1).

2.7. Building individual-level derived data for full flight paths

Each remote sensing image was split into 1 km^2 tiles. We pre-processed each tile using the same filters as used for developing models. To make predictions we ensembled only the best 20 models (not the 100 used in algorithm development) to reduce the computational processing for each tile to <48 core hours (~ 76,800 hours across all tiles). We predicted trait average values with associated uncertainty for each suitable pixel in the tile independently. We also calculated the standard deviation of each PLS-GLR prediction as a measure of uncertainty. Crown-based predictions were achieved by averaging the values of all suitable pixels lying within the corresponding predicted ITC boundaries. For those areas where the ITC overlapped with more than one flight path (flight paths overlap by ~30%), we averaged the crown-based predictions.

The resulting dataset was then compiled as a comma delimited file containing all the geometry information to rebuild polygons shape and location. The data is distributed as both raw data files

in a Zenodo archive (http://doi.org/10.5281/zenodo.3232978).The resulting files can be imported into GIS systems and spatial analysis packages in languages like R and Python for analysis. To demonstrate the potential uses of this data product we analyzed the correlation between terrain, leaf, and structural traits to summarize the general patterns observed at the scale of an entire NEON remote sensing footprint.

3. Results

The crown segmentation algorithm described in Silva et al. (2016) yielded the highest overlap with the field-delineated ITC crowns (Table S.1). For individual trees, the Jaccard Index ranged between 0 (undetected trees) and 0.81, with an average of 0.35 and median of 0.34. The algorithmic approach produced individual tree crowns overlapping with 88% of the ground delineated crowns. Crowns identified by the algorithm were generally larger than those delineated in the field, which resulted in overestimated crown areas and weak correlations between field data and algorithmic crown areas (Figure 3; R^2 = 0.17 for OSBS and 0.07 for TALL). Estimates of other structural traits from the algorithmic tree crowns varied were better than crown area. Height showed the highest correspondence between field and remotely sensed measures (R^2 = 0.44 for OSBS and 0.63 for TALL) in part because LIDAR directly measures height. Estimates for DBH were intermediate (R^2 of 0.17 for OSBS and 0.43 for TALL) reflecting the fact that the allometric models depend on both height and crown area as predictors. Visual assessment of paired field and algorithmically delineated crowns shifts by a couple of meters relative that are likely a result of imperfect alignments between LiDAR and Hyperspectral data (Figure S.3; Marconi et al. 2019).

Of the three methods - the single pixel methods (SPM), the ensemble pixel methods (EPBM) and the crown based methods (CEAM) - CEAM performed the best, explaining 72% of the variance in LMA, 67% of the variance in %N, 44% of the variance in %P, and 23% of the variance in %C (Figure 2, Table S.2). Except for %C, these results were comparable to those obtained by Martin et al. (2018) for trees in Borneo (71%, 46%, 44%, 0.48% respectively for LMA, %N, %P, %C). Predictions of %P showed lower accuracy than %N and LMA despite being highly correlated to observed values (Pearson correlation = 0.80), due to a strong relationship that deviated from the 1:1 line. This suggests that there is potential to improve predictions for %P accounting of the presence of a non-linear relationship between the leaf traits and the hyperspectral information. Except for %C, RMSE was always between 10 and 15% of the range of observed values, meeting the quality threshold recommended by Singh et al. (2015).

These crown level results (CEAM) performed better than methods using pixels as the fundamental unit (SPM and EPBM). Within the pixel-based methods the ensembles that considered crown identity (EPBM) performed better than the simple pixel-based approach (SPM) (Table S.2; Figure 2). Few individual PBMs performed better than the EPBM ensembles (Figure S.4) and when individual pixel models did outperform the ensemble, they usually were not among the best PBM models (i.e., the models with the lowest delta AIC in the validation). As such, these pixel models could not have been selected and used to provide better fits - they just happened to provide better fits on the test data. This suggests that EPBM provide the best method for making out-of-sample predictions at the pixel level. CEAM models outperformed EPBM with higher predicted R^2 values for all traits except %P, where we found no significant difference between crown and pixel-based ensemble. The main improvements in predictability from the crown-based approach were for LMA (from 0.60 to 0.73) and %N (0.60 to 0.67) (Table

S.2; Figure 2). While predictability of %C also increased (from 0.13 to 0.23), we considered the model's performance too low to provide useful individual-level information, and we discourage the use of %C products to address meaningful ecological questions. Crown based prediction also produced the best estimates of uncertainty. CEAM 95PI showed an average coverage of 83% of held out observations, EPSM 81%, and SBM 74%, with the ideal value being 95% (Figure 2B).

Extracting pixels algorithmically, using the automatically delineated crowns from LIDAR rather than ground delineated crowns, did not negatively affect the CEAM models (Figure 2C). R^2 from held out predictions using CEAMs improved with automatically delineated crowns by 0.05 for both LMA and %N and the RMSE was reduced by 3.37 g m^{-2} and 0.1 %N respectively. While SPMs and EPBMs performed significantly worse for %P and %C (0.44 to -0.21 and -0.17 to -0.51 respectively), predictions for the CEAM were comparable to those built from ground delineated crowns (0.44 to 0.37 for %P and 0.23 to 0.22 for %C). This suggests that scaling remote sensing predictions to crowns can reduce the effect of uncertain crown segmentation on the final models' performance (Table S.2).

Scaling algorithmic crown segmentation and trait estimation to the full extent of the NEON remote sensing data yielded trait predictions for approximately 5 million ITCs for the two sites combined (Figure 4). Despite some striping patterns associated with the NEON flight lines, landscape patterns in traits are evident, including east-west gradients in LMA, %N, and %P at OSBS. At TALL, lower LMA and higher %N and %P are found in a dendritic pattern associated with the stream network (Figure S.7). On average, OSBS showed lower LMA and higher %N than TALL (Figure 5). Assessing correlations between estimated structural traits, leaf traits, and abiotic environmental conditions showed strong correlations between LMA, %N, and %P, consistent with the leaf economic spectrum (Wright et al., 2004). Results showed correlations of

leaf traits with elevation and LAI (Figures S.9, S.15). Leaf traits and tree structure were also correlated (e.g. LMA with tree H, figure 5B), but correlations were generally stronger for OSBS where the forest stature is about 10 m lower (Figures S.11, S.9, S.11), the canopy is more open, and the water table more superficial (Bonoli et al., 1993, Motz et al., 1992).

4. Discussion

The individual organism is one of the fundamental units of biology. As a result, studying the distribution of individuals and their traits across space and through time is central to many aspects of ecology. However, collecting individual level data at the large scales required for many ecological questions is challenging, and larger scale studies based on remote sensing typically focus on either plots or pixels, not individual biological organisms. To address this limitation we develop a fully automated modular pipeline to link remote sensing products from the National Ecological Observatory Network (NEON) to data collected in the field, convert the remote sensing data into estimates of the leaf and structural traits for all canopy trees at landscape scales, and distribute this derived data on millions of individual trees in an open and accessible format (Cassey et al., 2006, Hampton et al., 2016) for use by the broader scientific community.

In addition to the value of focusing on individuals for the application of remote sensing based data products, we found that modeling and predicting leaf traits at the individual crown level resulted in improved accuracy and reduced uncertainty in the predictions compared to pixel based approaches (Figure 2). Linking pixels to crowns allows the ensembling of models built from the different pixels making up the crowns. Different pixels contain different combinations of leaves, branches, and ground, which affects the underlying chemometric relationship between

foliar traits and their spectral signature. Weighted ensembling allows the models to identify the best combinations of pixels for relating traits and hyperspectral signatures, which produced more generalizable and accurate relationships. Aggregating pixel predictions from the EPBMs to the crown level may also reduce the relative weight of outlier pixels. This appears to have been particularly important when using algorithmically-delineated crowns. The crown-based method (CEAM) showed that using algorithmic crowns that included errors in crown segmentation generally did not negatively affect model predictability (Figure 2C, Table S.2). In contrast, single pixel methods sometimes result in major decreases in accuracy when using algorithmically-delineated crowns compared to field delineated crowns, with the accuracy of %P predictions dropping from explaining 44% of the variation to being worse than using the average value of %P at the site (R^2 = -0.21). The robustness of the crown-based method may be related to the primary form of error in the algorithmic crown-delineations, which tended to overestimate the size of crowns. CEAM may be robust to this overestimation because the weighted ensembles can ignore bad pixels outside of the true crown. In cases where these extra pixels do include components of the true crown they may improve estimates by allowing more shaded or mixed pixels to be discarded from the crown interior. In contrast, for pixel-based methods, erroneously including non-crown pixels will reduce the average quality of fit, especially if those pixels are major outliers (e.g, leaves of another species). This suggests that crown-based approaches to trait estimation will be particularly valuable when linking field data to remote sensing without field delineated crowns, which is the most common case when developing remote sensing based trait models. Our method produced models with predictive power comparable to two other crown-based estimation methods (Barbosa and Asner, 2017; Martin et al., 2018), suggesting that the performance of these approaches may generalize beyond the current study. However, neither of

these studies compared their approaches to pixel-based methods, so additional research will be necessary to assess the generality of improved predictive power at the crown level.

Crown delineation is central to estimating structural traits and the algorithmic delineations only overlapped by 34% on average with field delineated crowns. While this level of overlap suggests that improved methods for crown delineation are needed, it is also the result of limitations in the resolution and alignment of the remote sensing products. The resolution of the remote sensing data (1m^2 pixels) relative to the scale of most tree crowns (average tree crown area in this study = 35 m^2) means that even single pixel errors can have a large influence. For example, if the extent of a 4x4 m crown (area = 16 m^2) is overestimated by a single pixel in each direction (6x6m = 36 m^2), the Jaccard index is only 0.44 and CA more than doubled. In addition, remote sensing data from multiple sensors often have imperfect alignments between products. The prototype NEON AOP data used in this study may have misalignments between LiDAR and Hyperspectral products on the scale of 1-2 meters (Marconi et al. 2019). Misalignment of this extent will primarily affect pixels at crown borders. When estimating plant traits, the key question is how the errors in delineation driven by models, resolution, and alignment influence the accuracy of structural and leaf trait models. While crown-based leaf trait models are robust to these errors (see above) the structural trait estimates are more directly related to crown delineation and are therefore more likely to be influenced by errors in that delineation. Fortunately estimates of DBH and height were well correlated with field values. This is likely because height is directly measured LIDAR and height was the most important factor in the allometric models used to predict DBH. As long as the top of the tree falls within the algorithmically delineated crown this should produce a good measurement of tree height and a reasonable estimate of DBH. This suggests that data derived from our pipeline could be used to

assess estimates of individual tree biomass by applying allometric functions linking biomass to DBH and H (Graves et al. 2018, Jucker et al. 2016).

While our crown-based methods were effective for estimating a number of leaf and structural traits, there is substantial uncertainty even for the best performing traits. Quantifying this uncertainty provides information on the range of likely trait values for each individual and allows this uncertainty to be propagated when using these derived data values to test scientific hypotheses (Miller-Gulland and Shea, 2017). We used methods for leaf trait estimation that allowed us to estimate uncertainty (pls-GLR) and the crown-based method (CEAM) provided the best uncertainty estimates (Figure 2). Current methods for delineating crowns do not include explicit measures of uncertainty (Dalponte & Coomes 2016, Silva et al., 2016). It is important for future methods to address this limitation because comparisons to field data suggested high uncertainty in the delineations.

Generating derived individual level data on leaf and structural traits at landscape scales allows trait patterns at these scales to be effectively assessed. This is valuable for many ecological questions (even non-spatial questions) because of the limited sample sizes, limited geographic extents, and potential for bias (e.g., selecting easily accessible trees with certain attributes, e.g., healthy individuals) in field collected data (Yang et al., 2018). A complete analysis of the spatial and species patterns is beyond the scope of this methods paper, but we provide some examples of analysis that can be done here. Distributions of leaf trait estimates for the two NEON sites in this study are generally multimodal consistent with mixed forests present at these sites that contain both coniferous and deciduous trees. Coniferous species (all *Pinus* at these two sites) have higher LMA and lower %N than the deciduous species, although this pattern does not hold for %P. Comparing the distributions between these two sites, TALL shows lower %N and higher LMA

than OSBS (Figure 4). LMA, %N and %P at both sites showed similar patterns to the Leaf Economic Spectrum (LES; Wright et al. 2004), despite the large difference in the spatial extent from which these datasets were derived (Figure S.12). Assessing the spatial variation in trait values shows that higher %N and %P is generally associated to trees closer to lakes and rivers (where deciduous species predominate), while LMA increases with distance to water (Figure S.7). This difference could be attributable to the edaphic conditions of soil that affect both species distribution and leaf traits (Walter & Gerlach, 2013). For example, OSBS sandhills are characterized by thin and well drained sandy soil, with potentially limited nutrient availability that favor *Pinus* and thick leaves with low %N. However, the water table is superficial at OSBS (Varner et al., 2009). At lower elevations soil is marshy and richer in nutrients that could favor the establishment of N and P rich species (Bodker et al., 2015). Similar patterns can be found in TALL, where higher deposition of nutrients can be expected in watershed valleys (Bannister et al., 2015).

Producing derived data at the individual level also facilitates landscape scale assessments of relationships between leaf and structural traits that could not be compared at the pixel level, since structural traits (e.g., crown area, stem height) are properties of individual trees. For example, our results suggest that the relationship between leaf and structural traits for individual trees differs across sites, even forests with relatively similar species and trait composition such as the two studied here. At OSBS there were significant correlations between crown area, height, and the predicted chemical traits (LMA, %N, %P), while no strong correlation to chemical traits was found in TALL (Figure S.15). This difference could be attributable to differences in edaphic conditions (Walter & Gerlach, 2013). Upland sites at OSBS have lower water and nutrient holding capacities that limit height growth of the Pinus and Quercus species that grow in this

habitat and tend to have higher LMA and lower %N. At TALL, the soils of upland pine dominated sites are potentially less restrictive of height growth and nutrient content of trees and leaves. OSBS has on average 10 m lower than TALL (Figure S.11). Furthermore, the different management histories of TALL and OSBS may impact relationships among traits. TALL is actively managed for timber production, OSBS is managed for conservation and restoration. Ontogeny strongly affects allometric strategies and correlation between structural and leaf traits (Prado-Junior et al., 2016). Observed patterns may be affected by different developmental stages and forest structure in different patches of the landscape, which suggests that estimation of crown age may be a fundamental feature for further analyses.

Our crown-based approach to modeling and predicting tree traits produces individual data similar to that collected in the field. This approach has a number of benefits. First, it will make data integration with field-based forest and trait surveys easier because both derived remote sensing data and field surveys will be composed of the same fundamental unit. Second, crown-based approaches are likely better for aligning trait data across years. The same pixel in two consecutive years could vary significantly in a trait because of small errors in spatial alignment of pixels through time, whereas large crown-level regions will be more robust to small errors and can be aligned to address larger errors as long as the same crown is identified in all images. Finally, this approach allows a more compact representation of derived trait data as spatial polygons instead of rasters. While this will not be the best representation for all analyses, for individual level analyses it results in vastly reduced storage computational requirements compared to raster data. The data produced by our individual level pipeline could be extended by including predictions for other leaf and structural traits, environmental variables, management, or disturbance. Moreover, our pipeline can potentially be used to extract ecological information for

every tree at all NEON AOP sites for the full life of the observatory. This will produce a publicly available, spatially explicit database of detailed ecological information for hundreds of millions of trees across the US that could be fused with other continental data (e.g. Forest Inventory and Analysis) to address cross scale functional ecological (http://doi.org/10.5281/zenodo.3232978). Such data could be used to further understand the biology behind trait tradeoffs, and investigate cross scale ecological processes and patterns from individual to landscape to continental scale.

5. *Description of author's responsibilities*

Sergio Marconi and Sarah Graves designed the experiment, Sarah Graves carried out field work, Sergio Marconi Sarah Graves and Ethan White developed the methods, Sergio Marconi performed the analysis, Ethan White and Stephanie Bohlman supervised the work as lab leaders, helped with experimental design and advices on data analysis. Ben Weinstein helped with technical aspects and manuscript editing. All authors contributed to editing the manuscript. We acknowledge Dr. Adytia Singh for his suggestions in the initial phase of the project.

Funding: This work was supported by the Gordon and Betty Moore Foundation's Data-Driven Discovery Initiative through grant GBMF4563 to Ethan P. White; by the NSF Dimension of Biodiversity program grant (DEB-1442280) and USDA/NIFA McIntire-Stennis program (FLA-FOR-005470) to Stephanie Ann Bohlman; by the University of Florida Biodiversity Institute Graduate Fellowship to Sergio Marconi. There was no additional external funding received for this study.

References

1. Albert, C.H., Thuiller, W., Yoccoz, N.G., Soudant, A., Boucher, F., Saccone, P. and Lavorel, S., 2010. Intraspecific functional variability: extent, structure and sources of variation. Journal of Ecology, 98(3), pp.604-613.

2. Anderson□Teixeira, K.J., Davies, S.J., Bennett, A.C., Gonzalez□Akre, E.B., Muller□Landau, H.C., Joseph Wright, S., Abu Salim, K., Almeyda Zambrano, A.M., Alonso, A., Baltzer, J.L. and Basset, Y., 2015. CTFS□ForestGEO: a worldwide network monitoring forests in an era of global change. Global change biology, 21(2), pp.528-549.

3. Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C.B., Sinca, F., Vaughn, N.R. and Llactayo, W., 2017. Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. Science, 355(6323), pp.385-389.

4. Bannister, J.M., Herbert, E.R. and Craft, C.B., 2015. Spatial variability in sedimentation, carbon sequestration, and nutrient accumulation in an alluvial floodplain Forest. In The Role of Natural and Constructed Wetlands in Nutrient Cycling and Retention on the Landscape (pp. 41-55). Springer, Cham.

5. Barbosa, J.M. and Asner, G.P., 2017. Prioritizing landscapes for restoration based on spatial patterns of ecosystem controls and plant–plant interactions. Journal of Applied Ecology, 54(5), pp.1459-1468.

6. Barnes, R., Lehman, C., Mulla, D., 2014. Priority-flood: An optimal depression-filling and watershed-labeling algorithm for digital elevation models. Computers & Geosciences 62, 117–127. doi:10.1016/j.cageo.2013.04.024

7. Bastien, P., Vinzi, V.E. and Tenenhaus, M., 2005. PLS generalised linear regression. Computational Statistics & data analysis, 48(1), pp.17-46.

8. Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS journal of photogrammetry and remote sensing, 65(1), pp.2-16.

9. Bodker, J.E., Turner, R.E., Tweel, A., Schulz, C. and Swarzenski, C., 2015. Nutrient-enhanced decomposition of plant biomass in a freshwater wetland. Aquatic Botany, 127, pp.44-52.

10. Bonoil, D., M. Williams, and D. Munch. 1993. Mapping recharge to the Floridan aquifer unsing a geographic information system. St. Johns River Water Management District, Technical Publication SJ93-5, Palatka, Florida. pp 41.

11. Callaway, R.M., Pennings, S.C. and Richards, C.L., 2003. Phenotypic plasticity and interactions among plants. Ecology, 84(5), pp.1115-1128.

12. Cassey, P. and Blackburn, T.M., 2006. Reproducibility and repeatability in ecology. BioScience, 56(12), pp.958-959.

13. Chadwick, K.D. and Asner, G.P., 2016. Organismic-scale remote sensing of canopy foliar traits in lowland tropical forests. Remote Sensing, 8(2), p.87.

14. Clark, J.S., Gelfand, A.E., Woodall, C.W. and Zhu, K., 2014. More than the sum of the parts: forest climate response from joint species distribution models. Ecological Applications, 24(5), pp.990-999.

15. Clark, J.S., 2016. Why species tell more about traits than traits about species: predictive analysis. Ecology, 97(8), pp.1979-1993.

16. Colgan, M.S., Baldeck, C.A., Féret, J.B. and Asner, G.P., 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(11), pp.3462-3480.

17. Collalti, A., P. E. Thornton, A. Cescatti, A. Rita, M. Borghetti, A. Nole, C. Trotta, P. Ciais, and G. Matteucci., 2019. The sensitivity of the forest carbon budget shifts across processes along with stand development and climate change. Ecological Applications 00(0): e01837. 10.1002/eap.1837

18. Cornelissen, J. H. C., S. Lavorel, E. Garnier, S. Diaz, N. Buchmann, D. E. Gurvich, P. B. Reich, H. ter Steege, H. D. Morgan, M. G. A. van der Heijden, J. G. Pausas, and H. Poorter. 2003. A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany 51:335 – 380.

19. Dalponte, M. and Coomes, D.A., 2016. Tree□centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. Methods in ecology and evolution, 7(10), pp.1236-1245.

20. De Angelis, D.L., 2018. Individual-based models and approaches in ecology: populations, communities and ecosystems. CRC Press.

21. Dwyer, J.M. and Laughlin, D.C., 2017. Constraints on trait combinations explain climatic drivers of biodiversity: The importance of trait covariance in community assembly. Ecology letters, 20(7), pp.872-882.

22. Feilhauer, H., Asner, G.P. and Martin, R.E., 2015. Multi-method ensemble selection of spectral bands related to leaf biochemistry. Remote Sensing of Environment, 164, pp.57-65.

23. Frederic Bertrand, Nicolas Meyer and Myriam Maumy-Bertrand (2014). Partial Least Squares Regression for Generalized Linear Models, R package version 1.1.1.

24. Graves, S.J., Caughlin, T.T., Asner, G.P. and Bohlman, S.A., 2018. A tree-based approach to biomass estimation from remote sensing data in a tropical agricultural landscape. Remote sensing of environment, 218, pp.32-43.

25. Graves, S.J., Asner, G.P., Martin, R.E., Anderson, C.B., Colgan, M.S., Kalantari, L. and Bohlman, S.A., 2016. Tree species abundance predictions in a tropical agricultural landscape with a supervised classification model and imbalanced data. Remote Sensing, 8(2), p.161.

26. Graves S, Gearhart J, Caughlin TT, Bohlman S. 2018. A digital mapping method for linking high-resolution remote sensing images to individual tree crowns. PeerJ Preprints6:e27182v1 https://doi.org/10.7287/peerj.preprints.27182v1

27. Hampton, S.E., Anderson, S.S., Bagby, S.C., Gries, C., Han, X., Hart, E.M., Jones, M.B., Lenhardt, W.C., MacDonald, A., Michener, W.K. and Mudge, J., 2015. The Tao of open science for ecology. Ecosphere, 6(7), pp.1-13.

28. Houborg, R., Fisher, J.B. and Skidmore, A.K., 2015. Advances in remote sensing of vegetation function and traits.

29. Jucker, T., Avăcări□ei, D., Bărnoaiea, I., Duduman, G., Bouriaud, O. and Coomes, D.A., 2016. Climate modulates the effects of tree diversity on forest productivity. Journal of Ecology, 104(2), pp.388-398.

30. Lawler, J.J., Tear, T.H., Pyke, C., Shaw, M.R., Gonzalez, P., Kareiva, P., Hansen, L., Hannah, L., Klausmeyer, K., Aldous, A. and Bienz, C., 2010. Resource management in a changing and uncertain climate. Frontiers in Ecology and the Environment, 8(1), pp.35-43.

31. Li, X., Messina, J.P., Moore, N.J., Fan, P. and Shortridge, A.M., 2017. MODIS land cover uncertainty in regional climate simulations. Climate Dynamics, 49(11-12), pp.4047-4059.

32. Liu, X., Swenson, N.G., Lin, D., Mi, X., Umaña, M.N., Schmid, B. and Ma, K., 2016. Linking individual□level functional traits to tree growth in a subtropical forest. Ecology, 97(9), pp.2396-2405.

33. Marconi, S., Graves, S.J., Gong, D., Nia, M.S., Le Bras, M., Dorr, B.J., Fontana, P., Gearhart, J., Greenberg, C., Harris, D.J. and Kumar, S.A., 2018. A data science challenge for converting airborne remote sensing data into ecological information (No. e26966v1). PeerJ Preprints.

34. Martin, M. E., L. C. Plourde, S. V. Ollinger, M. L. Smith, and B. E. McNeil. 2008. A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. Remote Sensing of Environment 112:3511–3519.

35. Martin, R.E., Chadwick, K.D., Brodrick, P.G., Carranza-Jimenez, L., Vaughn, N.R. and Asner, G.P., 2018. An approach for foliar trait retrieval from airborne imaging spectroscopy of tropical Forests. Remote Sensing, 10(2), p.199.

36. McGill, B.J., Enquist, B.J., Weiher, E. and Westoby, M., 2006. Rebuilding community ecology from functional traits. Trends in ecology & evolution, 21(4), pp.178-185.

37. Messier, J., McGill, B.J., Enquist, B.J. and Lechowicz, M.J., 2017. Trait variation and integration across scales: is the leaf economic spectrum present at local scales?. Ecography, 40(6), pp.685-697.

38. Milner□Gulland, E.J. and Shea, K., 2017. Embracing uncertainty in applied ecology. Journal of applied ecology, 54(6), pp.2063-2068.

39. Motz, L.H., J.P. Heaney, W.K. Denton, M.S. Fowler, and G. Leiter. 1992. Upper Etoniah Creek Hydrologic Study, Phase II Final Report. St. Johns River Water Management District, Technical Publication SJ92-SP18, Palatka, Florida. pp 414.

40. Müller, I., Schmid, B. and Weiner, J., 2000. The effect of nutrient availability on biomass allocation patterns in 27 species of herbaceous plants. Perspectives in plant ecology, evolution and systematics, 3(2), pp.115-127.

41. Murphy, J., and J. P. Riley. 1962. A modified single solution method for the determination of phosphate in natural waters. Analytica Chimica Acta 27:31 – 36.

42. National Ecological Observatory Network. 2014. Data Products NEON.DP1.10098, NEON.DP1.30010, NEON.DP3.30015, NEON.DP1.30008. Provisional data

downloaded from http://data.neonscience.org on 26 Jan 2016. Battelle, Boulder, CO, USA

43. Newnham, G.J., Armston, J.D., Calders, K., Disney, M.I., Lovell, J.L., Schaaf, C.B., Strahler, A.H. and Danson, F.M., 2015. Terrestrial laser scanning for plot-scale forest measurement. Current Forestry Reports, 1(4), pp.239-251.

44. Nicotra, A.B., Atkin, O.K., Bonser, S.P., Davidson, A.M., Finnegan, E.J., Mathesius, U., Poot, P., Purugganan, M.D., Richards, C.L., Valladares, F. and van Kleunen, M., 2010. Plant phenotypic plasticity in a changing climate. Trends in plant science, 15(12), pp.684-692.

45. Niinemets, Ü., Keenan, T.F. and Hallik, L., 2015. A worldwide analysis of within□canopy variations in leaf structural, chemical and physiological traits across plant functional types. New Phytologist, 205(3), pp.973-993.

46. Nock, C. A., Vogt, R. J. and Beisner, B. E. (2016). Functional Traits. In eLS, John Wiley & Sons, Ltd (Ed.). doi:10.1002/9780470015902.a0026282

47. Popescu, Sorin & Wynne, Randolph. (2004). Seeing the Trees in the Forest: Using Lidar and Multispectral Data Fusion with Local Filtering and Variable Window Size for Estimating Tree Height. Photogrammetric Engineering and Remote Sensing. 70. 589-604. 10.14358/PERS.70.5.589.

48. Prado-Junior, J.A., Schiavini, I., Vale, V.S., Raymundo, D., Lopes, S.F. and Poorter, L., 2016. Functional traits shape size-dependent growth and mortality rates of dry forest tree species. Journal of Plant Ecology, 10(6), pp.895-906.

49. Real, R. and Vargas, J.M., 1996. The probabilistic basis of Jaccard's index of similarity. Systematic biology, 45(3), pp.380-385.

50. Roussel, J.R. and Auty, D., 2017. lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications.

51. Sarstedt, M., Hair, J.F., Ringle, C.M., Thiele, K.O. and Gudergan, S.P., 2016. Estimation issues with PLS and CBSEM: Where the bias lies!. Journal of Business Research, 69(10), pp.3998-4010.

52. Sexton, J.O., Noojipady, P., Anand, A., Song, X.P., McMahon, S., Huang, C., Feng, M., Channan, S. and Townshend, J.R., 2015. A model for the propagation of uncertainty from continuous estimates of tree cover to categorical forest cover and change. Remote Sensing of Environment, 156, pp.418-425.

53. Silva, C.A., Hudak, A.T., Vierling, L.A., Loudermilk, E.L., O’Brien, J.J., Hiers, J.K., Jack, S.B., Gonzalez-Benecke, C., Lee, H., Falkowski, M.J. and Khosravipour, A., 2016. Imputation of individual Longleaf Pine (Pinus palustris Mill.) Tree attributes from field and LiDAR data. Canadian journal of remote sensing, 42(5), pp.554-573.

54. Singh, A., Serbin, S.P., McNeil, B.E., Kingdon, C.C. and Townsend, P.A., 2015. Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. Ecological Applications, 25(8), pp.2180-2197.

55. Swenson, N.G. and Weiser, M.D., 2010. Plant geography upon the basis of functional traits: an example from eastern North American trees. Ecology, 91(8), pp.2234-2241.

56. Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M.B., Balaguer, L., Benito□Garzón, M., Cornwell, W., Gianoli, E., Kleunen, M., Naya, D.E. and Nicotra, A.B., 2014. The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. Ecology letters, 17(11), pp.1351-1364.

57. Varner, J. Morgan, et al. "Post-fire tree stress and growth following smoldering duff fires." Forest Ecology and Management 258.11 (2009): 2467-2474.

58. Walters, Michael B., and John P. Gerlach. "Intraspecific growth and functional leaf trait responses to natural soil resource gradients for conifer species with contrasting leaf habit." Tree physiology 33.3 (2013): 297-310.

59. Weinstein, B., Marconi, S., Bohlman, S., Zare, A. and White, E., 2019. Individual tree-crown detection in RGB imagery using self-supervised deep learning neural networks. bioRxiv, p.532952.

60. Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H., Diemer, M. and Flexas, J., 2004. The worldwide leaf economics spectrum. Nature, 428(6985), p.821.

61. Yang, J., Cao, M. and Swenson, N.G., 2018. Why functional traits do not predict tree demographic rates. Trends in ecology & evolution, 33(5), pp.326-336.

62. Zhen, Z., Quackenbush, L.J. and Zhang, L., 2016. Trends in automatic individual tree crown detection and delineation—evolution of lidar data. Remote Sensing, 8(4), p.333.

Table 1. Data products and sources (National Ecological Observatory Network, 2016). Information about data products can be found on the NEON data products catalogue (http://data.neonscience.org/data-product-catalog).

Name	NEON data product ID	Data date	How it was used
Spectrometer orthorectified at-sensor radiance	NEON.DP1.30008	2014 and 2015	Hyperspectral images used to model foliar chemical and physical properties
Discrete return LiDAR point cloud	NEON.DP1.30003.001	2014 and 2015	Crown segmentation and calculation of tree height
AOP L2 and L3 data products (LAI, Albedo, Elevation, Slope, Aspect)	NEON.DP2.30012.001, DP3.30024.001, DP3.30025.001	2014 and 2015	Link modeled tree crown properties with other NEON-AOP data products
Woody plant vegetation structure	DP1.10098.001	2014-2018	Build allometric relationship to infer DBH from Crown Area and Tree Height
Field ITC	Graves et al. 2018	2017	Validate crown segmentation, define tree objects.

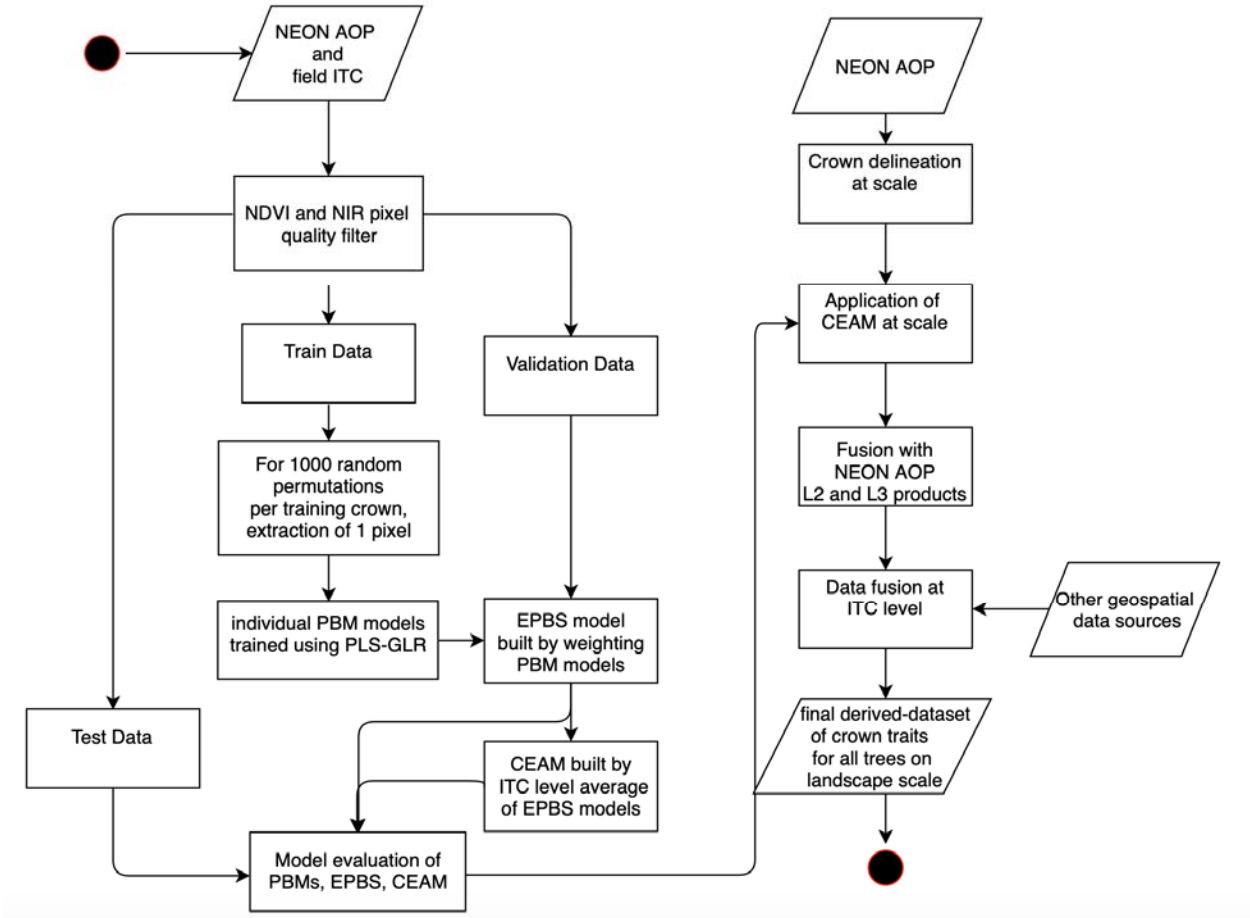

Figure 1. Workflow of the pipeline following Unified Modeling Language (UML). Left side shows the method to build and test the three modeling approaches: PBM (pixel-based model), EPBS (ensemble of pixel-based models), CEAM (crown ensemble averaged model). Right side represents the part of the pipeline dealing with scaling, fusion, and public distribution of derived data. NEON L2 and L3 products fused to the derived traits data are: Aspect (DP3.30025.001), Albedo (DP2.30011.001), Elevation (DP3.30024.001), Slope (DP3.30025.001) and LAI (DP3.30012.001).

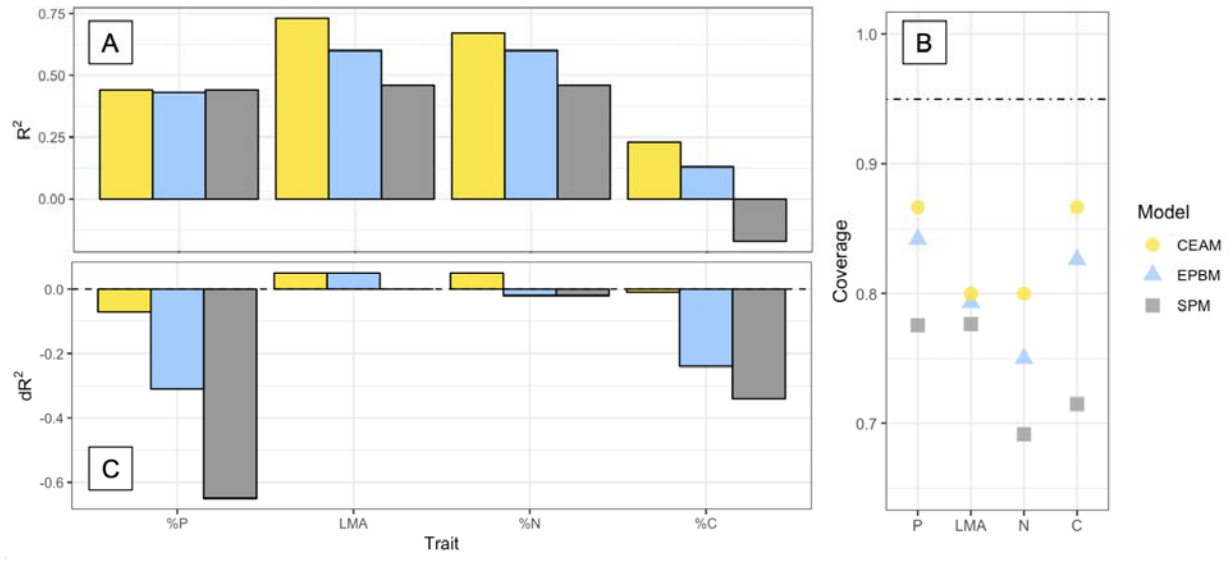

Figure 2. Model evaluation and comparison of pixel based (SPM, gray), ensemble pixels (EPBM, azure), and crown ensemble average (CEAM, yellow) predictions on an independent test set of 15 observed crowns for %P, LMA, %N and %C. (A) Evaluation and comparison of models built on pixels extracted from ground delineated crowns; (B) Coverage of the 95 predictions intervals on held out data for the four models (LMA, N, P and C). Dash-dotted line represent the ideal coverage value of 95%; (C) difference in performance (predicted R^2) between models built on pixels extracted from automatically delineated crowns and models built on ground delineated crowns. Positive values mean that for that comparison model built on automatically delineated crowns performed better.

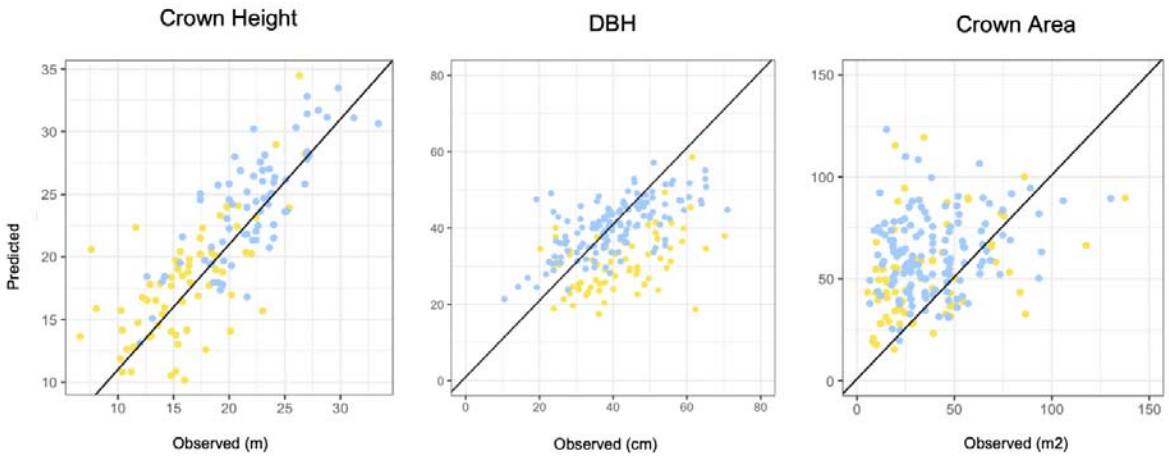

Figure 3. Comparison between observed and predicted tree structural traits for algorithmically delineated crowns corresponding to ground delineated ITCs: (A) tree crown height (m), (B) tree diameter at breast height (cm), and (C) tree crown area (m^2). Yellow and blue points represent OSBS and TALL site data points respectively. Black diagonal is the 1:1 line.

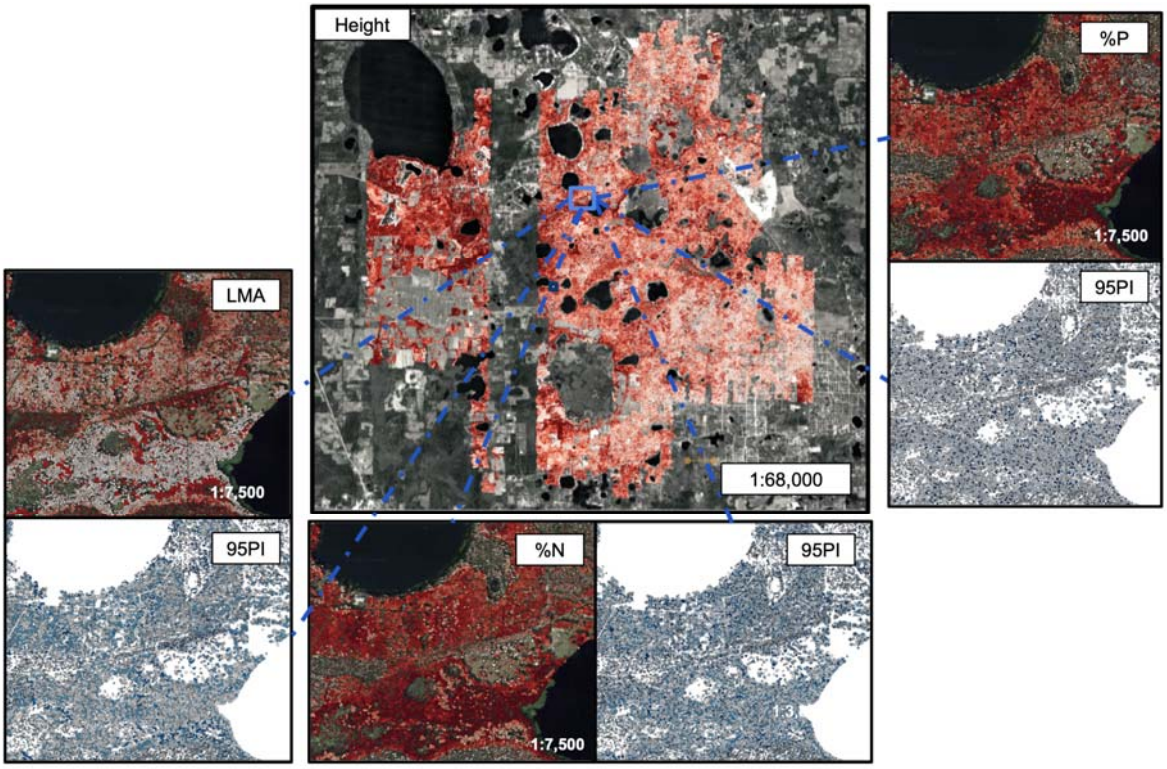

Figure 4. Example of predictions at the landscape scale for individual tree crowns (ITC) leaf traits at Ordway Swisher Biological Station (OSBS).(A) Predictions of tree height for ~2.5 million trees at the scale of NEON AOP site's footprint (215 km^2). Blue square represent a 2.2 km^2 spatial subset for estimations of LMA, %N, and %P. Expected values values are presented on intensity scale of reds (with lower values in white, and higher values in deep red). 95% probability intervals for ITC predictions of leaf chemical traits presented for B, C, and D: values are presented on intensity scale of blue (with lower values in white, and higher values in deep blue).

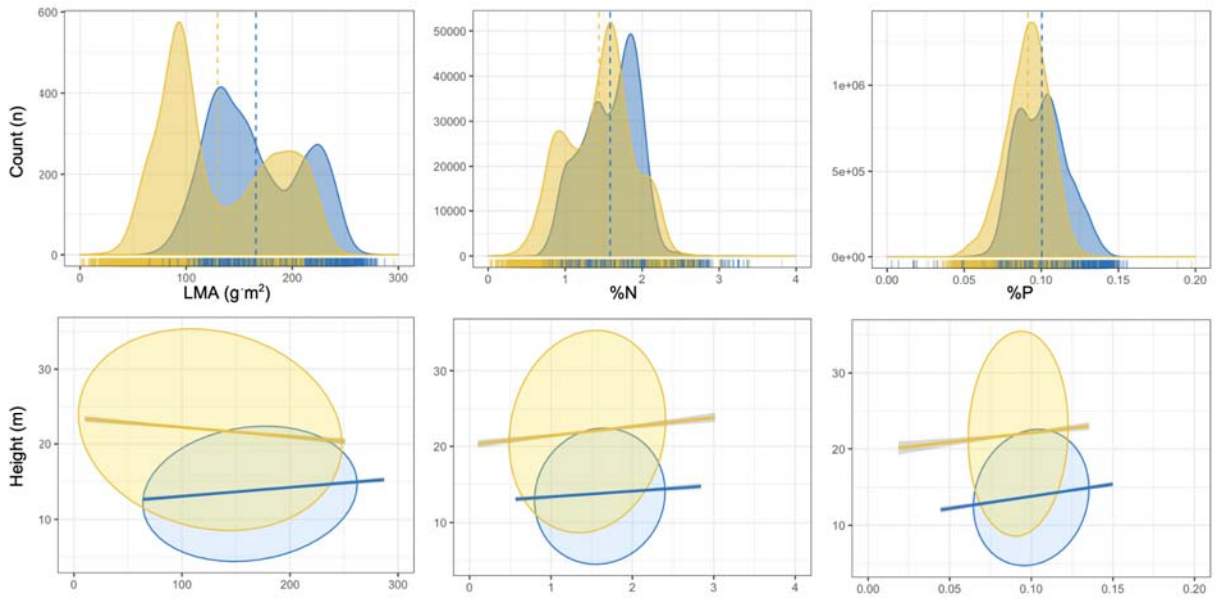

Figure 5. Leaf chemical distributions and relationship with tree height on a sample of 100,000 derived individual tree crowns (ITCs). On the top row, comparison between distributions of LMA, %N, and %P for the two sites, OSBS (blue), and TALL (yellow). Vertical dotted lines represent the average for the site. Rug plots on the x-axis represent the marginal distribution of OSBS and TALL data between the minimum and maximum range of derived observations. On the lower row, example of relationship between tree height and the same three leaf chemical traits: from left to right, LMA, %N, %P. Linear trends and 95CI ellipses are represented for each relationship and site, following the same color scheme as above.