Canadian Journal of Remote Sensing, 42:411–427, 2016

ISSN: 0703-8992 print / 1712-7971 online
DOI: 10.1080/07038992.2016.1217482

Taylor & Francis
Taylor & Francis Group

Mapping Forest Structure and Composition from Low-Density LiDAR for Informed Forest, Fuel, and Fire Management at Eglin Air Force Base, Florida, USA

Andrew T. Hudak[1,*], Benjamin C. Bright[1], Scott M. Pokswinski[2], E. Louise Loudermilk[3], Joseph J. O'Brien[3], Benjamin S. Hornsby[3], Carine Klauberg[1], and Carlos A. Silva[4]

[1]*USDA Forest Service, Rocky Mountain Research Station, Moscow, ID 83843, USA*
[2]*University of Nevada, Biology Department, Reno, NV 89557, USA*
[3]*USDA Forest Service, Southern Research Station, Athens, GA 30602, USA*
[4]*University of Idaho, Department of Forest, Rangeland, and Fire Sciences, Moscow, ID 83843, USA*

Abstract. Eglin Air Force Base (AFB) in Florida, in the United States, conserves a large reservoir of native longleaf pine (*Pinus palustris* Mill.) stands that land managers maintain by using frequent fires. We predicted tree density, basal area, and dominant tree species from 195 forest inventory plots, low-density airborne LiDAR, and Landsat data available across the entirety of Eglin AFB. We used the Random Forests (RF) machine learning algorithm to predict the 3 overstory responses via univariate regression or classification, or multivariate *k*-NN imputation. Ten predictor variables explained ∼ 50% of variation and were used in all models. Model accuracy and precision statistics were similar among the various RF approaches, so we chose the imputation approach for its advantage of allowing prediction of the ancillary plot attributes of surface fuels and ground cover plant species richness. Maps of the 3 overstory response variables and ancillary attributes were imputed at 30-m resolution and then aggregated to the management block level, where they were significantly correlated with each other and with fire history variables summarized from independent data. We conclude that functional relationships among overstory structure, surface fuels, species richness, and fire history emerge and become more apparent at the block level where management decisions are made.

Résumé. Eglin Air Force Base (AFB) en Floride, aux États-Unis, possède un grand réservoir de peuplements de pins des marais (*Pinus palustris* Mill.) indigènes que les gestionnaires des terres entretiennent en utilisant des feux fréquents. Nous avons prédit la densité des arbres, la surface terrière et les espèces dominantes d'arbres à partir de 195 parcelles d'inventaire forestier, du LiDAR aéroporté à basse densité et de données Landsat disponibles dans l'ensemble d'Eglin AFB. Nous avons utilisé l'algorithme d'apprentissage automatique des forêts aléatoires « Random Forests » (RF) pour prédire les 3 réponses de l'étage supérieur par régression ou classification unidimensionnelles, ou par une imputation *k*-NN multidimensionnelle. Dix variables prédictives expliquent environ 50 % de la variabilité et ont été utilisées dans tous les modèles. La précision du modèle et les statistiques de précision ont été similaires entre les différentes approches FA, nous avons donc choisi l'approche d'imputation, car elle a l'avantage de permettre la prédiction des attributs auxiliaires des parcelles telles que les combustibles de surface et la richesse des espèces de plantes couvre-sol. Les cartes des 3 variables réponses de l'étage supérieur et les attributs auxiliaires ont été imputés à 30 m de résolution et ensuite agrégés au niveau du bloc de gestion, où ils étaient significativement corrélés les uns avec les autres et avec les variables de l'historique des feux synthétisées à partir de données indépendantes. Nous concluons que les relations fonctionnelles entre la structure de l'étage supérieur, les combustibles de surface, la richesse des espèces, et l'historique des feux émergent et deviennent plus apparentes au niveau du bloc où les décisions de gestion sont prises.

INTRODUCTION

Longleaf Pine Ecology

Native longleaf pine (*Pinus palustris* Mill.) forests of the southeastern United States are dependent on fire for healthy ecosystem structure and function (O'Brien et al. 2008). Frequent fire (fire return intervals are 1 year–10 years; Christensen 1981, 1988; Glitzenstein et al. 1995) arrests encroachment of understory shrubs to maintain a healthy overstory, prevents accumulation of surface fuels and duff to expose mineral soil and thus promote tree recruitment (Mitchell et al. 2006; Mitchell et al. 2009), and sustains an exceptionally diverse understory flora (Kirkman et al. 2001; Iacona et al. 2010). Although longleaf pine forest overstories are often monospecific, understory and

Received 12 October 2015. Accepted 4 July 2016.
*Corresponding author e-mail: ahudak@fs.fed.us.

ground cover plant species can number up to 50 per square meter, making longleaf pine forests a biodiversity hotspot (Mitchell et al. 2006; Palmquist et al. 2014).

A policy of fire exclusion provides a competitive advantage to fire-averse tree species that compete with and eventually replace longleaf pine in the absence of fire. This factor, coupled with extensive logging in the 19th century for dense, high-quality longleaf pine wood, have reduced formerly widespread longleaf pine forests to $< 5\%$ of their historic range (Frost 1993; Noss et al. 1995). Some of the largest remnant stands of longleaf pine can be found on Eglin Air Force Base (AFB) in the Florida Panhandle. Eglin AFB managers use fire as the principal tool to maintain longleaf pine stand structure and health, usually using prescribed fire, but also by incorporating wildfires into their operational fire program goal of burning 36,500 ha of longleaf pine forest per year (Williams 2012).

Longleaf pine grows slowly as it establishes during the grass seedling stage, and after the brief period when it bolts to escape the flame zone (Platt et al. 1988). Longleaf pine trees growing in the poor, sandy soils at Eglin AFB can obtain a maximum height of ~ 25 m in ~ 50 years but can continue slow lateral growth for hundreds of years. Trees might be randomly distributed (Platt et al. 1988; Gelfand et al. 2010) but can also be observed to clump and form gaps, where saw palmetto (*Serenoa repens*), turkey oak (*Quercus cerris* L.), and other oaks, broadleaf trees, and shrubs are more likely to establish. There is little understory growth in healthy, fire-maintained longleaf pine forests; however, forest undergrowth thickens after years of fire exclusion, resulting in succession to other forest types less valued by Eglin AFB managers.

Predictive Modeling

Airborne scanning light detection and ranging (LiDAR) is the preferred remote sensing technology for mapping forest canopy structure (Lefsky et al. 2001, 2002; Næsset 2002, 2004; Popescu et al. 2003; Dean et al. 2009; Lee et al. 2010), because 2-dimensional optical data tend to lose sensitivity to forest structure variation in dense forest conditions (Lefsky et al. 1999; Harding et al. 2001; Hudak et al. 2006). Because LiDAR data are 3-dimensional, forest canopy height measures can be derived directly from LiDAR point cloud data. Provided the LiDAR collection has been ground calibrated by the vendor, as is customary, canopy height measures derived from LiDAR after normalizing for topography do not require field height measures for calibration. Canopy cover and density are 2 other forest structure attributes that can be calculated from the LiDAR data without field calibration data. Canopy cover is calculated as the number of LiDAR first returns above a given height threshold, divided by the total number of first returns. Canopy density is calculated as the same ratio but based on all (not just first) returns.

LiDAR processing for forest inventory applications usually entails reducing the 3-dimensional distribution of canopy elements captured in the LiDAR point cloud data to statistical metrics of canopy height and density that can serve as more digestible inputs into predictive models. By calculating these metrics within defined areas, such as fixed-radius plots, the LiDAR metrics can serve as predictor variables to be empirically related to traditional forest structure attributes that serve as the response variables of interest. The area-based approach to forest inventory, using LiDAR, has been demonstrated in many studies (Lim et al. 2003; Hyyppä et al. 2008; Hollaus et al. 2009; Koch et al. 2009; Wulder et al. 2012; Hudak et al. 2008).

Many predictive modeling methods exist, including regression and *k* nearest neighbor (kNN) imputation (Eskelson et al. 2009). In this study, we used Random Forests (RF), a machine learning algorithm developed by Breiman (2001) and implemented by Liaw and Wiener (2002) and Crookston and Finley (2008) in R packages (R Core Team 2013). By bootstrapping through various predictor variable combinations while randomly withholding 1/3 of the data in an out-of-bag sample, the RF algorithm randomly generates a forest of classification and regression trees (CART) to predict a given response variable. Both continuous and categorical variables can be predicted using RF, which compares well with other predictive modeling methods (Hudak et al. 2008, 2009; Latifi and Koch 2012; Hayashi et al. 2015).

Objectives

Overstory structure will change slowly in response to multiple fires over many years, whereas surface fuel conditions are highly dynamic at fine temporal and spatial scales. Therefore, our conceptual approach was to map the relatively static overstory structure and composition that constrains comparatively dynamic surface fuel patterns and fire processes at finer scales.

Our goal was to apply airborne LiDAR and proven area-based modeling methods to map 3 overstory response variables of benefit to Eglin AFB managers. The first response, tree density measured in number of trees per hectare (TPH), determines competition among trees, constrains the understory environment, and is a principal determinant of habitat selection by wildlife such as Red-cockaded Woodpeckers (Smart et al. 2012). The second response, plot-level basal area (BA), is a standard forestry attribute, used for timber management and planning. The third response, a categorical variable indicating forest composition, was dominant species (DomSpp), defined as the tree species having the most basal area within the plot.

Our first objective was to apply the RF method to predict these 3 responses, using 3 independent, univariate models and a single, multivariate *k*-NN imputation model. Our second objective was to evaluate the alternative RF models and then apply a single modeling approach to map the 3 responses across Eglin AFB. Our third objective was to summarize the 3 mapped overstory responses and associated surface fuel and species richness measures at the management block (i.e., stand) level and to compare them to historical fire management records summarized at the same level, in order to reveal functional relationships. We hypothesized that the effects of management interventions as re-

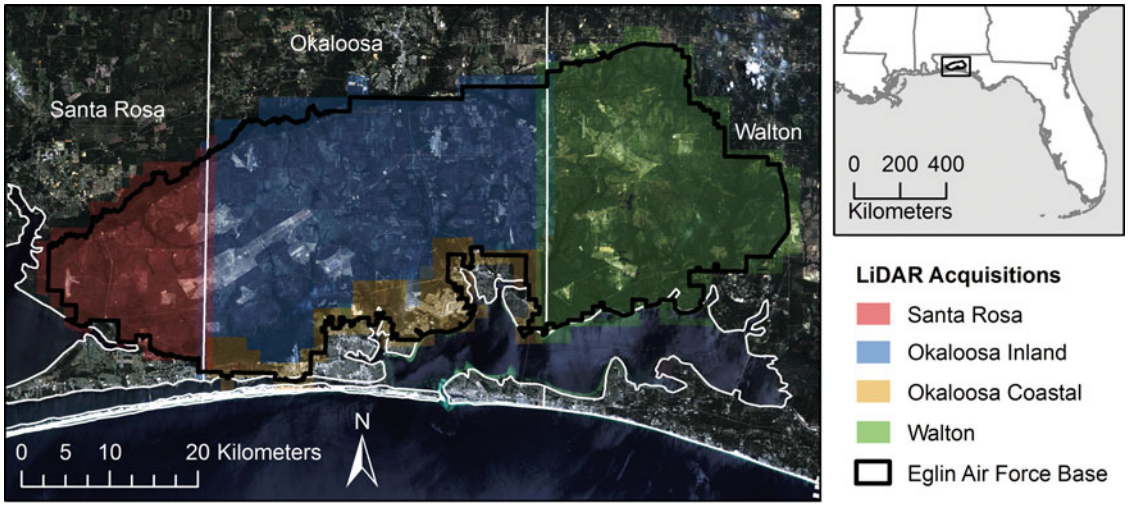

FIG. 1. Eglin Air Force Base study area in the Florida Panhandle, United States.

vealed through such functional relationships should emerge and be most apparent at the same scale at which the management actions are applied.

METHODS

Study Area

Eglin AFB covers 186,350 ha in the flat Gulf Coastal Plain of the Florida Panhandle, in the United States (Figure 1). Climate is subtropical with mean annual temperature of 19.8 °C and mean annual precipitation of 158 cm. Most rain falls from June through September, and the sandy soils are well-drained deposits of Quartzipsamments of the Lakeland series (Overing et al. 1995). The sandhills at Eglin AFB fall under the high pine characterization by Myers (1990), referring to the hilly undulating xeric terrain dominated by an open longleaf pine (*Pinus palustris*) canopy with a hardwood midstory made up of turkey oak (*Quercus laevis*), bluejack oak (*Q. incana*), and persimmon (*Diospyros virginiana*). Flatwoods at Eglin AFB are characterized by flat topography with mesic, poorly drained soils as described by Abrahamson and Hartnett (1990). Overstory is dominated by longleaf pine with a midstory and an understory dominated by saw palmetto (*Serenoa repens*), dwarf live oak (*Quercus minima*), and gallberry (*Ilex glabra*).

Eglin AFB managers consider longleaf pine "reference" stand conditions most desirable because of their characteristically open understories, maintained by frequent prescribed fires and conducive to military training exercises. Less desirable stand conditions are termed "restoration" stands and are in need of mechanical treatment, a management intervention more intensive and expensive than prescribed fire. Restoring longleaf pine stands to reference conditions is a long-term management goal and was a primary justification for establishing monitoring plots (Hiers et al. 2007).

Plot Data

Eglin AFB monitors vegetation status at 201 monitoring plots randomly located across the base as described in Hiers et al. (2007). Plots were established in 2001 to be monitored each summer following management treatments that included prescribed fire, herbicide, and harvesting. All overstory trees >10 cm diameter at breast height (DBH; 1.37 m) were stem mapped within a 61-m x 106-m (0.65 ha) rectangular plot. The 4 corners were geolocated to within 5 m without differential correction using Trimble Nomads connected to SXBlue receivers. Trees were tallied and stem mapped in all but 6 plots situated in plantations, so these 6 plots were dropped from consideration, leaving 195 for analysis. We assessed the spatial pattern and scale of tree stems mapped within plots in reference and restoration stands of predominantly longleaf pine using Ripley's L statistic (Clark and Evans 1954), using the "splancs" package (Bivand et al. 2014) in R (R Core Team 2013).

Surface fuels were sampled nondestructively along 2 parallel 50-m Brown's transects situated north and south of plot center as described in Hiers et al. (2007). Measures at each meter included: counts of 1-hr, 10-hr, 100-hr, and 1000-hr fuels (the time lag for 2/3 of the fuel to dry in response to atmospheric moisture, which increases with fuel size), where 1-hr and 10-hr fuels were subsampled along the north 50 m of the transect; duff, litter, and fuelbed depths averaged from 10 measures along the 100-m of both north and south transects; percent cover of the litter components: oak, conifer (long and short needle), grasses, forbs, shrubs, saw palmetto, and bare mineral soil. Ground cover plant species richness data were collected in 8 1-m^2 square subplots systematically located within a 20-m x 20-m area immediately west of plot center. Wherever repeat measures of trees or other attributes existed, only the most recent measures were used (Table 1), which varied by plot from 2003–2012.

TABLE 1
Plot-level overstory tree, surface fuel, and plant species richness summary statistics

Attributes	Minimum	Maximum	Mean	Standard Deviation
Tree density (trees ha^{-1})	0.00	433.50	120.87	80.66
Basal area (m^2 ha^{-1})	0.00	20.49	6.19	4.46
1-hr (counts)	0.00	3306.00	107.30	259.62
10-hr (counts)	0.00	466.00	27.48	43.80
100-hr (counts)	0.00	217.00	12.78	21.27
1000-hr (counts)	0.00	38.00	4.07	5.89
Litter depth (cm)	0.12	3.86	1.13	0.73
Duff depth (cm)	0.00	3.30	0.48	0.63
Fuelbed depth (cm)	2.58	54.85	20.58	10.55
Oak litter cover (%)	0.00	95.00	39.11	24.81
Long-needle conifer litter cover (%)	0.00	99.00	46.97	25.77
Short-needle conifer litter cover (%)	0.00	95.00	14.50	24.48
Grass litter cover (%)	0.00	47.00	9.92	11.82
Forb litter cover (%)	0.00	38.00	7.42	8.08
Shrub litter cover (%)	0.00	69.00	7.17	9.72
Saw palmetto litter cover (%)	0.00	21.00	1.46	3.14
Mineral soil cover (%)	0.00	71.00	22.62	18.21
Plant-species richness (species m^{-2})	0.13	13.50	6.52	2.32

LiDAR Data

LiDAR data collected circa 2006 and 2008 were downloaded from a public archive[1] with files organized at the county level; 620 tiles in .las format that ranged from 225 ha–625 ha in area were processed to cover the entirety of Eglin AFB, which spans portions of 3 counties (Fig. 1; Table 2).

Binary .las files were processed using LAStools (Isenburg 2015). The "lasground" utility was used to classify returns as ground or nonground, and normalize Z values of absolute elevation to Z values of height above ground for each return. The "blast2dem" tool was used to create a 2-m digital terrain model (DTM) across Eglin AFB. The proportion of classified ground returns in Santa Rosa County was 3 to 4 times greater than that in the 2 other counties, so the "lasthin" tool was used to reduce the ground return density in Santa Rosa County by a factor of 3 to 4 to match the proportion of ground returns in neighboring Okaloosa County (Figure 1), which, if not remedied, would lead to an artifact in the subsequent canopy density metrics along the Santa Rosa–Okaloosa County line. Canopy height and density metrics were calculated from all returns using the "lascanopy" utility (Table 3). Using the "lasclip" utility, canopy height and density metrics were calculated within the rectangular plot footprints for training and testing the RF models. Grids of the same metrics were created at a 30-m x 30-m binning resolution for mapping.

Rumple, a measure of canopy surface rugosity or roughness, is defined as the ratio of canopy surface area over the underlying ground area, and has been demonstrated to be a useful measure of forest canopy structure (Parker et al. 2004; Kane et al. 2010). Because the resolution of some of the LiDAR collections was low, we created a grid of rumple across Eglin AFB from a 5-m canopy height model (CHM), generated with the "lascanopy" tool of LAStools, and then used the "GridSurfaceStats" tool within the FUSION software package to calculate rumple at 30-m resolution (McGaughey 2015). The GridSurfaceStats tool creates a canopy surface by deriving a triangular irregular network (TIN) from the input CHM.

Ancillary Data

Landsat Thematic Mapper (TM) surface reflectance imagery of Eglin AFB (Path 20, Row 39) from January 21 and July 15, 2008, was downloaded.[2] Normalized Difference Vegetation Index (NDVI) and middle-infrared corrected NDVI (NDVIc; Nemani et al. 1993) images were created for each date. NDVIc is defined as

$$\frac{NIR - Red}{NIR + Red} \times \left[1 - \frac{(MIR - MIR_{min})}{(MIR_{max} + MIR_{min})}\right],$$

where Red is the red band (Landsat TM band 3), NIR is the near-infrared band (Landsat TM band 4), and MIR is the middle-infrared band (Landsat TM band 5). Incorporation of the

[1] http://www.nwfwmdlidar.com/

[2] http://earthexplorer.usgs.gov/

TABLE 2
LiDAR data acquisition specifications in the 3 counties spanned by Eglin AFB

Acquisition Area	Acquisition Time	Horizontal Accuracy	Vertical Accuracy	Sensor	Contractor	Mean Return Density (points m^{-2})
Santa Rosa	Jan, Feb. 2006	Not available	RMSE 18.5 cm	Leica ALS-50	Photo Science, Inc.	0.5
Okaloosa Inland	Feb. 2008	1 m RMSE or better	RMSE 15 cm	Leica ALS50-II MPiA	EarthData International, Inc.	1.3
Okaloosa Coastal	Not available	Not available	Not available	Not available	Not available	4.7
Walton	Jul., Aug. 2006	1 m	RMSE 13 cm	Not available	Sanborn	2.9
Eglin, inside Walton	Summer 2008	1 m	RMSE 15 cm	Optech 3100	Sanborn	2.9

middle-infrared band has been found to improve the relationship between NDVI and forest structure (Nemani et al. 1993; Pocewicz et al. 2004).

Eglin AFB managers provided geographic information system (GIS) layers pertinent to our objectives, including past prescribed fire and wildfire boundaries, date of burn, and fuel type. For burn histories starting in the 1990s, burn-area boundaries were based on the land management blocks that were prescribe burned, whereas the boundaries of any prescribed fire escapes, lightning ignitions, or other wildfires were delineated with a global positioning system (GPS). For fire records prior to the 1990s, a combination of burn documentation, fire maps, and Landsat image data (MSS and TM) was used (Laine 2015). These fire history records (beginning February 1, 1972) were summarized into 3 GIS layers: number of fire occurrences, number of years since last burn, and fuel type; in areas where no burns were recorded, a value of 50 was assigned as the number of years since last burn. A polygon layer was also provided to enable comparisons, at the management block level, among these fire/fuel variables, the mapped overstory responses, and the surface fuel variables associated with the overstory variables at the plot level. The size of the management blocks (n = 425) averaged 439 ha and ranged from 0.48 ha to 3,872 ha.

Analysis

Overview

Plot-level LiDAR metrics (predictor variables) were associated with the plot-level TPH, BA, and DomSpp attributes (response variables) at all 195 inventory plots. These data were randomly divided into training and testing datasets comprising 2/3 and 1/3 of the data, respectively. The testing dataset was reserved to evaluate all predictive RF models built from the training dataset. Based on this evaluation, the best RF model was used with all available plot data to generate predictive maps.

Predictive Modeling

The RF algorithm implemented in the RandomForest package of R (R Core Team 2013) can operate in either regression or classification mode to predict either continuous or categorical variables, respectively. For each of 3 response variables, a preliminary set of predictor variables was selected based on the Model Improvement Ratio (MIR), which is a scaled measure of the percent increase in mean square error used by the RF algorithm to assign importance values to predictor variables. The MIR variable selection tool (Evans 2015) iterates through randomly selected predictor variable combinations to identify the best suite of predictors for the given response (Evans and Cushman 2009; Evans et al. 2011; Murphy et al. 2010). The maximum Pearson correlation allowed between any 2 selected LiDAR metrics was 0.85. In cases when 2 candidate metrics exceeded $r = 0.85$, the metric having the lesser importance value was excluded from consideration. The 3 preliminary sets of selected variables were then concatenated into a single list of candidate predictors to consider for simultaneous imputation of all 3 responses, using the RF nearest neighbor selection method implemented in the *k*-NN imputation package, "yaImpute" (Crookston and Finley 2008) available in R (R Core Team 2013). By concatenating the random forests generated by the RF algorithm ($n = 500$ by default) for each response and then tallying nodes per variable across all random forests, a distribution of variable importance measures was derived. A final set of predictors with consistently higher importance was then selected for use in all models.

Model Evaluation

Univariate RF regression models predicting TPH or BA were evaluated for precision with root mean squared error (RMSE) and for accuracy with mean bias error (MBE). The RF classification model predicting DomSpp was evaluated with a con-

TABLE 3
Description of LiDAR metrics and Landsat indices considered as predictor variables; selected variables indicated in boldface

Metric	Description
min	Minimum canopy height
max	**Maximum canopy height**
avg	Mean canopy height
std	**Standard deviation of canopy heights**
ske	Skewness of canopy height
kur	Kurtosis of canopy height
p10	**10th percentile of canopy height**
p25	25th percentile of canopy height
p50	50th percentile of canopy height
p75	75th percentile of canopy height
p90	**90th percentile of canopy height**
rum	**Rumple (surface roughness or rugosity)**
cov	**Percentage of first returns > 1.37 m in height**
dns	Percentage of all returns > 1.37 m in height
d01	Percentage of returns > 1.37 m and < 5 m in height
d02	**Percentage of returns > 5 m and < 10 m in height**
d03	**Percentage of returns > 10 m and < 20 m in height**
d04	**Percentage of returns > 20 m and < 30 m in height**
NDVI.Jan	Normalized Difference Vegetation Index, 21 Jan. 2008
NDVI.Jul	Normalized Difference Vegetation Index, 15 Jul. 2008
NDVIc.Jan	**Middle-infrared corrected NDVI, 21 Jan. 2008**
NDVIc.Jul	Middle-infrared corrected NDVI, 15 Jul. 2008

fusion matrix and user's, producer's, and overall classification accuracies. The same 3 responses predicted with the multivariate RF *k*-NN imputation model were all evaluated in the same manner, with the exception that root mean squared difference (RMSD) was used instead of RMSE to assess the precision of the continuous TPH and BA responses. This RMSD statistic is more appropriate for evaluating imputations when the number of nearest neighbors (k) is limited to one (as in this study, and the default in yaImpute), because the imputed prediction is itself an observation rather than a unique value, as is the case with regression predictions (Crookston and Finley 2008; Stage and Crookston 2007). This causes regression predictions to be shifted toward the mean and have less variance compared to imputations, which preserve the variance in the observations when $k = 1$ (McRoberts et al. 2002). As a result, an imputation model RMSD will be larger than a regression model RMSE, based on the same input data, when $k = 1$.

Mapping

The RF modeling approach (i.e., univariate regression/classification versus imputation) producing the best evaluation statistics was chosen to map all 3 responses. All model evaluation results reported in this article are based on the independent testing dataset. However, the full dataset based on all available information was used for mapping responses, in order to provide more accurate maps for Eglin AFB managers. Nonforest areas within Eglin AFB were masked in the map; the nonforest mask was defined as having < 1% canopy cover greater than breast height (1.37 m), as calculated from the LiDAR data.

Map Evaluation

Map cells were summarized at the land management block (i.e., stand) level by mean or modal value, depending on whether they were continuous or categorical variables, respectively. These summary statistics were then compared to 2 surrogates for fire frequency—number of fire occurrences and years since most recent fire—and between fuel types. Because many of the data distributions were not normal but skewed, Spearman rank correlations (ρ) between fire management variables and the mapped responses, along with surface fuel and species richness attributes associated with the mapped overstory responses at the inventory plots, were calculated and evaluated for significance.

RESULTS

Ripley's L statistic was calculated from plot stem map data to characterize the spatial pattern of trees in longleaf pine stands classified as having either reference conditions maintained by frequent prescribed fires and considered more desirable, or restoration conditions in need of more intensive management intervention. Ripley's L trends comparing reference and restoration conditions showed a similar clustered pattern at distances of 20 m–50 m, with clustering often being a significant departure from randomness (Figure 2).

There was much overlap between variables selected to independently predict TPH, BA, and DomSpp based on the MIR, and the predictors selected for multivariate *k*-NN imputation (Table 3). Therefore, 10 predictors (Table 3) were selected for use in all RF models in order to standardize the data inputs and simplify comparison of RF model results. Three metrics—avg, dns, and p75—had high importance values based on the MIR but were dropped from consideration because they correlated highly ($r > 0.85$) with selected metrics.

The relative RMSD measures of precision for imputing TPH and BA (49% and 41%, respectively) were slightly larger than the corresponding RMSE measures from the RF models

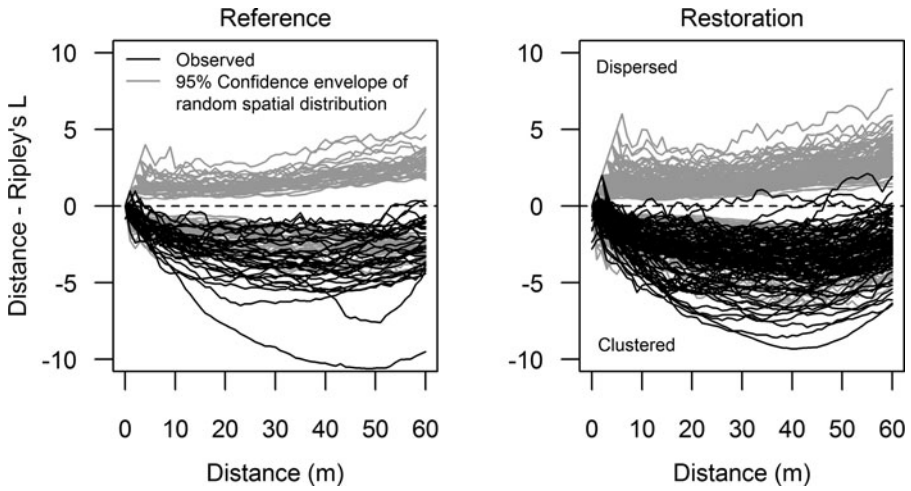

FIG. 2. Ripley's L summarizing the spatial pattern of trees from 61-m x 106-m (0.65 ha) plots randomly placed within longleaf pine stands classified by Eglin AFB managers as representing reference (left, $n = 35$) or restoration (right, $n = 93$) conditions. Only plots with a minimum of 20 trees are included.

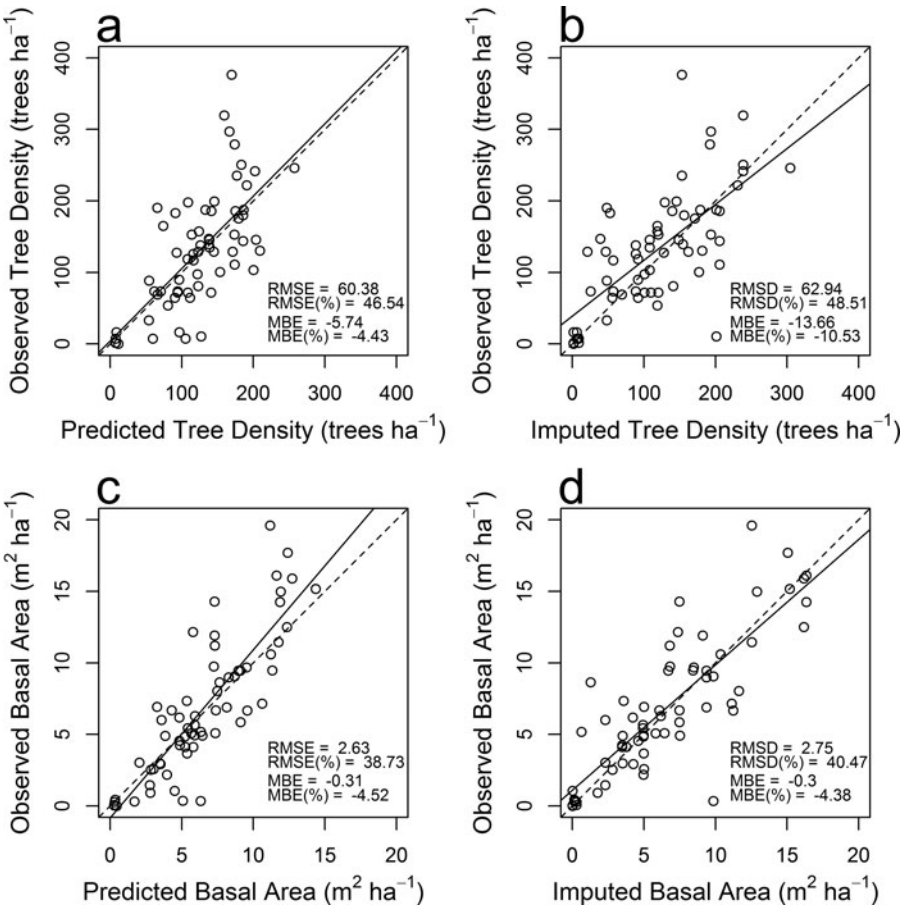

FIG. 3. Predicted versus observed results using RF models to predict TPH (a, b) or BA (c, d) as either a univariate response in regression mode (a, c) or via multivariate *k*-NN imputation (b, d). Models were trained with a random selection of 2/3 of the plot data ($n = 130$) and tested with the remaining 1/3 of the plot data ($n = 65$); these graphs illustrate the testing results. The same 10 predictor variables (Table 3) were used in all models. Solid lines in each graph indicate best linear fit; dashed lines indicate 1:1.

TABLE 4
Predicted versus observed results using Random Forests to predict DomSpp as (A) a univariate response in RF classification mode or as (B) one of 3 responses predicted using RF imputation

DomSpp	Black Tupelo	Slash Pine	Longleaf Pine	Sand Live Oak	Swamp Laurel Oak	Turkey Oak	Total Predicted	User's Accuracy (%)	Commission Error (%)
(A) Univariate response in RF classification mode									
Black Tupelo	0	0	0	0	0	0	0	100.0	0.0
Slash Pine	0	0	0	0	0	0	0	100.0	0.0
Longleaf Pine	1	2	46	3	0	0	52	**88.5**	**11.5**
Sand Live Oak	0	0	4	0	2	0	6	0.0	100.0
Swamp Laurel Oak	0	0	0	0	0	0	0	100.0	0.0
Turkey Oak	0	1	3	2	0	1	7	14.3	85.7
Total Observed	1	3	53	5	2	1	65		
Producer's Accuracy (%)	0.0	0.0	**86.8**	0.0	0.0	100.0		Overall Accuracy = 72.3%	
Omission Error (%)	100.0	100.0	**13.2**	100.0	100.0	0.0			
(B) One of 3 multivariate responses using RF k-NN imputation ($k = 1$)									
Black Tupelo	0	0	0	0	1	0	1	0.0	100.0
Slash Pine	0	0	2	0	0	0	2	0.0	100.0
Longleaf Pine	1	3	42	3	0	0	49	**85.7**	**14.3**
Sand Live Oak	0	0	6	0	1	0	7	0.0	100.0
Swamp Laurel Oak	0	0	2	0	0	1	3	0.0	100.0
Turkey Oak	0	0	1	2	0	0	3	0.0	100.0
Total Observed	1	3	53	5	2	1	65		
Producer's Accuracy (%)	0.0	0.0	**79.2**	0.0	0.0	0.0		Overall Accuracy = 64.6%	
Omission Error (%)	100.0	100.0	**20.8**	100.0	100.0	100.0			

Models were trained with a random selection of 2/3 of the plot data ($n = 130$) and tested with the remaining 1/3 of the plot data ($n = 65$); these tables illustrate the testing results. The same 10 predictor variables (Table 3) were used in all models. Longleaf pine classification accuracy and errors, of primary interest in this study, are indicated in boldface.

predicting TPH and BA (47% and 39%, respectively) independently in regression mode (Figure 3). Based on the relative MBE statistic, imputed TPH (−10.5%) was less accurate than TPH predicted via RF regression (−4.4%), while imputed BA was only slightly more accurate (relative MBE: −4.4%) than BA predicted via RF regression (relative MBE: −4.5%; Figure 3). With regard to the DomSpp categorical response, overall classification accuracy was somewhat higher if predicted by RF as a univariate response in classification mode (72%) than by imputation (65%) (Table 4). Similarly, classification of longleaf pine, the tree species of interest, was slightly more accurate by univariate classification (87%–88%) than by imputation (79%–86%).

Based on the similarity in these RF model results and for the sake of utility, we chose the imputation model in order to preserve the covariance relationships among the response variables and other plot attributes of interest to managers. The imputed maps illustrated in Figure 4 have RMSD and MBE statistics of 57.7 and −8.8 for imputing TPH and 2.45 and −0.26 for imputing BA. The overall accuracy of the imputed DomSpp map is 75%, with user's accuracy of 82.7% and producer's accuracy of 89.3% for classifying longleaf pine. Because these statistics are based on all plot data (i.e., training and testing datasets combined), they show an expected (but not dramatic) improvement over the more conservative evaluation of the models trained from 2/3 of the plot data and tested against the other 1/3 (Figure 3, Table 4).

Plot ID was imputed as an ancillary variable (Figure 4d). Since k was set equal to 1 in this analysis, plot ID as a categorical variable could be mapped by virtue of its association with the nearest neighbor plot having the most similar multivariate association of TPH, BA, and DomSpp responses and their relationship to the combination of 10 LiDAR predictors collectively weighting the model. Although plot ID had no weight in the model, stand-level patterns are still evident in Figure 4d, as in Figures 4a–c, because of the greater similarity in overstory canopy structure within stands than between stands, making imputation of the same nearest neighbor plots more likely locally, even though the model makes no accounting for spatial dependency between adjacent map cells.

FIG. 4. Imputations of (a) TPH, (b) BA, and (c) DomSpp from airborne LiDAR across Eglin AFB, and (d) Plot ID imputed as an ancillary variable (i.e., having no weight in the model). This model used for mapping was based on all plots ($n = 195$) and the same 10 predictor variables (Table 3) for all RF models.

Plot-level surface fuel and species richness measures were mapped by association with the imputed Plot ID map (Figure 4d). All mapped continuous variables were averaged by land management block ($n = 425$) and compared to explore functional relationships. Maps of plant-species richness and duff depth are illustrated as examples (Figure 5a–b) and were, themselves, highly correlated (Spearman $\rho = -0.81$). The TPH and BA overstory responses were compared to 2 indicators of fire frequency: number of fires and years since last fire (Figure 5c–d), which were, themselves, highly correlated (Spearman $\rho = -0.86$). The relationships shown in Figure 6 have highly significant correlations but do not reveal any underlying mechanisms.

Besides duff depth (Figure 5b), many other surface fuel attributes were significantly correlated with TPH and BA, as well as the fire frequency indicators (Table 5). Although the significant correlations indicate potentially meaningful relationships, the functional form of these relationships is revealed only on viewing scatterplots. For example, the variables underlying 4 of the stronger correlations in Table 5 are plotted in Figure 7.

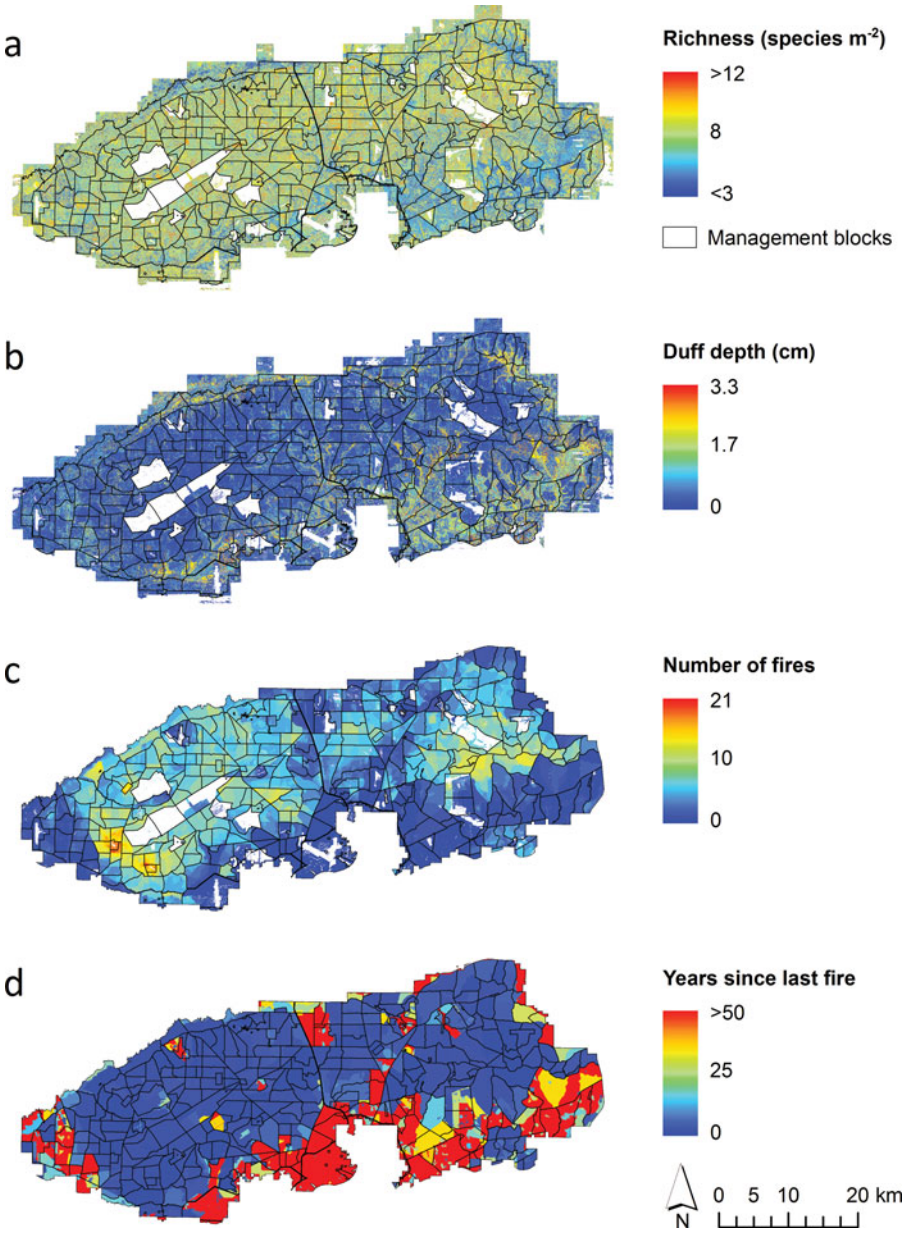

FIG. 5. (a) Plant-species richness and (b) duff depth related to the imputed overstory responses via plot ID (Figure 4d). Also shown are the fire history variables (c) number of fires and (d) years since last fire extracted from Eglin AFB fire management records.

Note that although longleaf pine litter cover is linearly related to TPH (Figure 7b), total fuelbed depth is clearly nonlinearly related (Figure 7c), at least in part due to the shrub component (Figure 7a). Plant-species richness is inversely but strongly correlated to TPH (Figure 7d).

Accompanying the fire history variables in the fire management data, yet unmentioned in our results up to this point, was the categorical variable of fuel type. We found that TPH and BA were approximately 7 and10 times lower, respectively, in the longleaf pine sandhill fuel type managed with a $<$ 3-yr fire return interval, compared to the other major fuel types. This result agrees with ground observations of more open forest structure in the relatively xeric and unproductive sandhills that are frequently burned. Surface fuel components in the other fuel types also differed noticeably from this fuel type, which was characterized by a higher fuelbed depth with high shrub and grass litter cover, but with lower oak leaf and pine needle cover, fewer 1-hr, 10-hr, 100-hr, and 1000-hr fuel counts, and little duff. These findings are also consistent with high-frequency fire effects on surface fuels. Plant-species richness was also markedly higher on this frequently burned fuel type compared to other fuel types.

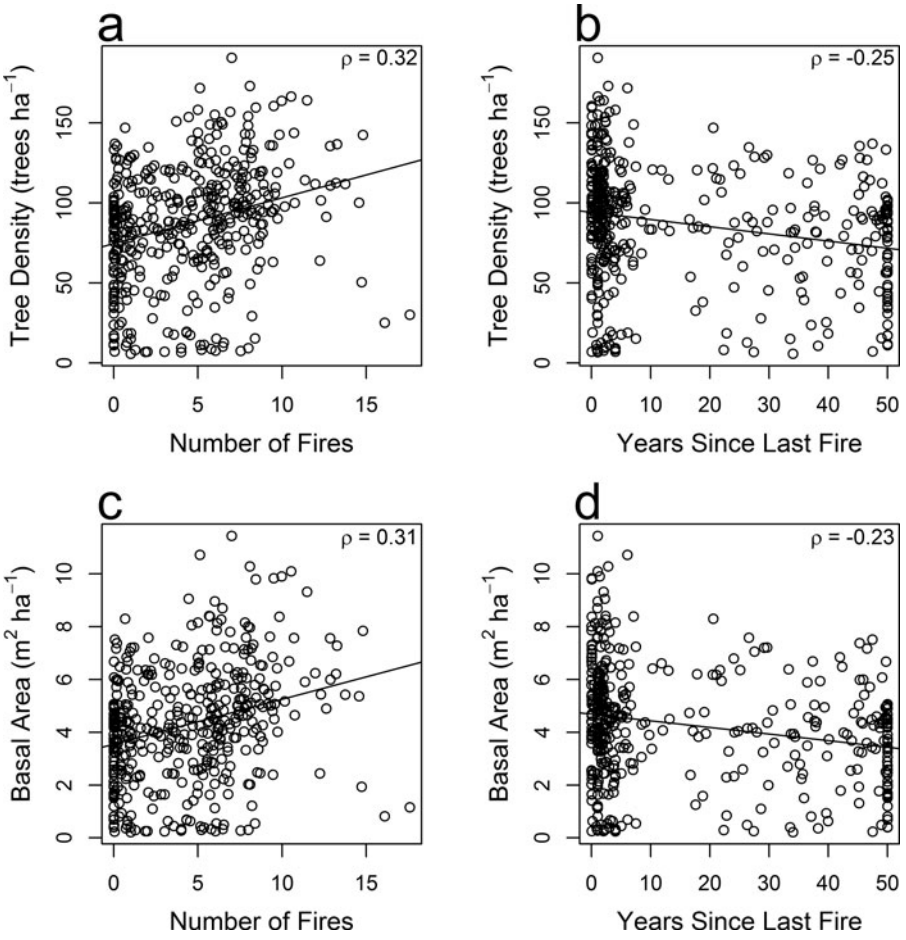

FIG. 6. Relationships between TPH and (a) number of fires and (b) years since last fire, and between BA and (c) number of fires and (d) years since last fire. Each plot symbol represents the mean of the 30-m map cells within an Eglin AFB management block ($n = 425$). Spearman correlations (ρ) are all highly significant ($p < 0.0001$). Solid lines in each graph indicate best linear fit.

DISCUSSION

The prevailing trend in Ripley's L shows tree clustering is most pronounced at a distance of 30 m–40 m (Figure 2), which supports our strategy to bin the LiDAR returns and map the responses at a 30-m × 30-m (0.09 ha) resolution, which is also convenient given the 30-m resolution of Landsat-derived spectral vegetation indices that supplemented the LiDAR metrics as predictor variables. The ∼ 30 m scale at which overstory structure predominantly varies does constrain surface fuels (Table 5, Figure 7), which vary at submeter scales in longleaf pine forests, as shown by Loudermilk et al. (2009).

The LiDAR data used in this study were not all collected at the same time, nor at the same return density, but this does not preclude their utility for broad-scale, observational studies such as ours. The 2006–2008 range in LiDAR data collection dates, situated midway within the 2003–2012 range of the most recent field plot data collections, would affect the predicted TPH and BA responses little in slow-growing longleaf pine. Scatterplots of plot-level TPH imputations against either point density or scan angle were random, with no discernible trends to suggest that either variable could have biased the models. Furthermore, we saw no evidence that differences between the LiDAR surveys (Table 2) had any subsequent effect on the gridded LiDAR metrics, other than the higher proportion of classified ground returns in Santa Rosa County affecting the density metrics (which was remedied by applying lasthin to just the ground returns, as was already described). The DTM is a much more sensitive indicator of a height bias than the CHM, neither of which showed any artifacts on visual inspection.

Binning the metrics at 30-m resolution ensured that there were at least 450 returns/cell, even at the lowest return density of 0.5 returns/m2 in Santa Rosa County, in order to maintain a stable height distribution for generating metrics. Hudak et al. (2012) showed that height and density metrics such as those used in this study could be compared between 2 LiDAR collections collected 6 years apart, to quantify biomass change due to forest

TABLE 5
Spearman rank correlations between mapped overstory responses or fire history variables and surface fuel or plant species richness measures associated with the Plot ID map (Figure 4d) and aggregated within Eglin AFB land management blocks ($n = 425$); significant ($p < 0.05$) correlations indicated in boldface

Attributes	Tree Density (trees ha^{-1})	Basal Area (m^2 ha^{-1})	Number of Fires	Years Since Last Fire
1-hr (counts)	**0.27**	**0.28**	**−0.50**	**0.43**
10-hr (counts)	**0.37**	**0.37**	**−0.28**	**0.25**
100-hr (counts)	**0.46**	**0.49**	**−0.09**	**0.14**
1000-hr (counts)	**0.41**	**0.44**	**−0.36**	**0.35**
Litter depth (cm)	**0.58**	**0.64**	**−0.15**	**0.18**
Duff depth (cm)	**0.49**	**0.52**	**−0.42**	**0.38**
Fuelbed depth (cm)	**−0.45**	**−0.42**	**0.08**	**−0.09**
Oak litter cover (%)	**0.14**	**0.12**	**−0.55**	**0.46**
Long-needle conifer litter cover (%)	**0.62**	**0.61**	**0.69**	**−0.60**
Short-needle conifer litter cover (%)	**−0.11**	**−0.11**	**−0.65**	**0.56**
Grass litter cover (%)	**−0.14**	**−0.13**	**0.61**	**−0.52**
Forb litter cover (%)	0.02	−0.01	**0.64**	**−0.57**
Shrub litter cover (%)	**−0.57**	**−0.54**	−0.03	0.03
Saw palmetto litter cover (%)	**0.35**	**0.35**	**−0.41**	**0.31**
Mineral soil cover (%)	**−0.47**	**−0.48**	**0.44**	**−0.39**
Plant-species richness (species m^{-2})	**−0.55**	**−0.59**	**0.31**	**−0.32**

growth, despite a 12-fold difference in point density. The difference between point densities among these LiDAR collections (Table 2) is no greater.

An unusual aspect of this study was the large size (0.65 ha) of the plots. Larger plots are more immune to poor geolocation accuracy and suffer less from edge effects, given their lower perimeter/area ratio, because tree crowns spanning the plot edge add noise to the relationship with the LiDAR metrics (Frazer et al. 2011). Frazer et al. (2011) used synthetic LiDAR with a sparse return density of 0.47 points m^{-2} and simulated Douglas-fir forest canopies to evaluate the influence of plot size on total aboveground biomass predictions. They found that model precision and accuracy increased as plot size increased from 0.0314 ha to 0.1964 ha, tending toward an asymptote at 0.25 ha. Gobakken and Næsset (2009) reported similar but less pronounced improvement for a smaller, 0.02 ha to 0.04 ha range of plot sizes in Norwegian forest composed of Norway spruce and Scots pine, using a LiDAR return density of 0.9 points m^{-2} more generally in line with the range of return densities used in our study (Table 2). The stand-level estimates of Gobakken and Næsset (2009) were derived from sample plots within stands, whereas our plots included a census of all trees within the unusually large area of 0.65 ha. These 2 characteristics of our plots should mitigate bias caused by their poor geolocation accuracy, and Gobakken and Næsset (2009) also state that larger plot sizes can compensate for sample plot position errors.

The 30-m × 30-m (.09 ha) grid cell size we used for mapping is larger than most inventory plots, yet much smaller than the 0.65 ha plots used for modeling. Plots of this size are more typical of tropical rainforests, where large plots are necessary to overcome edge effects from huge trees (Mauya et al. 2015). Such trees do not exist at Eglin AFB. Thus, it is apparent, based on the 30-m scale of longleaf pine clustering revealed by the Ripley's L analysis (Figure 2), that monitoring could be more efficiently accomplished on smaller plots. The 0.09-ha map cells capture this scale of clustering. It is advisable for the grid cell size of the maps to match the size of the sample plots (Magnussen and Boudewyn 1998, Næsset and Bjerknes 2001, Hudak et al. 2012). We recommend standard, 16-m fixed-radius inventory plots for monitoring, which are only slightly smaller (0.08 ha) but also are more efficient practically because they are round instead of square (Bormann 1953). Previous LiDAR-based forest inventory studies in coniferous forests (e.g., Hudak et al. 2006, 2014) found it more difficult to achieve accurate and precise estimates of TPH compared to BA, which is more correlated with LiDAR-derived canopy height and density metrics than TPH. Although tree height–diameter relationships tend to be stronger with coniferous than deciduous trees, the height–diameter allometry for longleaf pine breaks down after reaching a diameter of $\sim$ 25 cm, when height growth asymptotes at $\sim$ 25 m (Gonzalez-Benecke et al. 2014). This growth habit coupled with the open canopy structure that is typical of managed longleaf pine forests are probably why TPH and BA in this study were themselves highly correlated (Pearson $r = 0.89$) and predicted with comparable relative precision and accuracy (Figure 3).

Individual trees are more obviously apparent in higher density LiDAR point clouds; thus, more accurate tree counts can

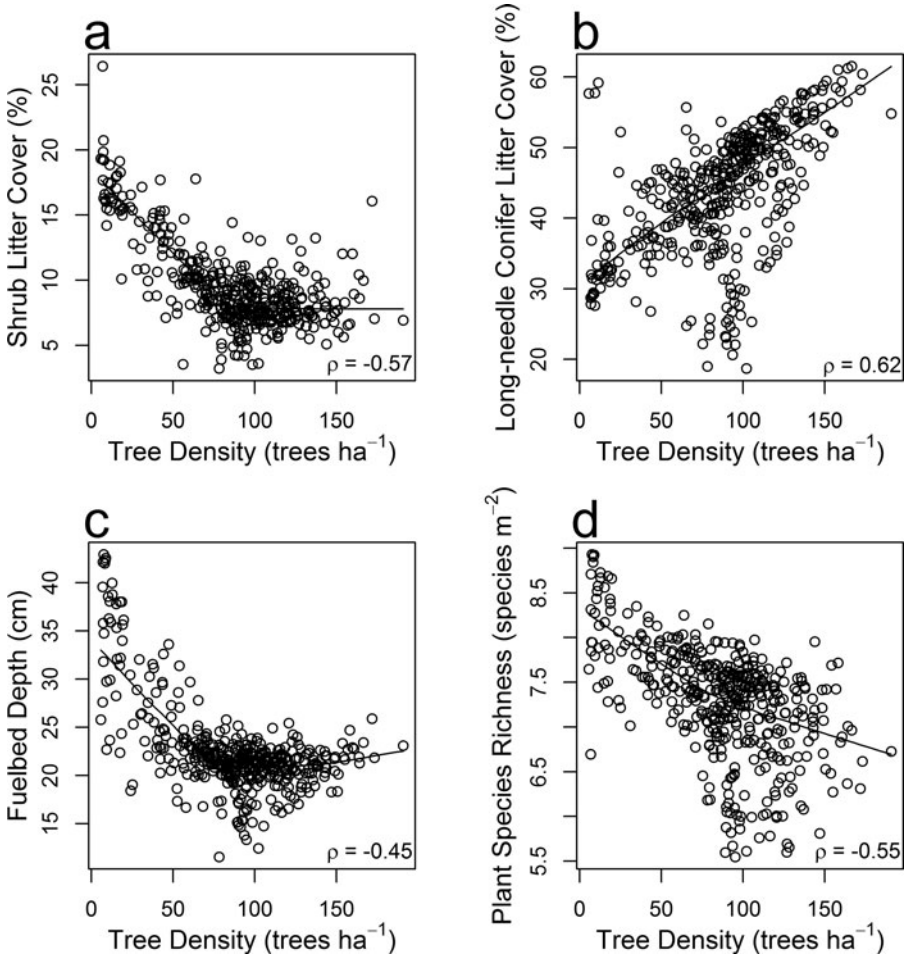

FIG. 7. Functional relationships between TPH and (a) shrub litter cover, (b) long-needle conifer litter cover, (c) fuelbed depth, and (d) plant-species richness. Each plot symbol represents the mean of the 30-m map cells within an Eglin AFB management block ($n = 425$). Spearman correlations (ρ) are all highly significant ($p < 0.0001$). Solid lines in each graph indicate loess smooth fit.

generally be obtained from higher-density LiDAR (e.g., Lee et al. 2010) than the low-density LiDAR used in this study. Silva et al. (2016) developed an automated individual tree detection and delineation approach that is sensitive to crown area coverage. Given the difficulty in implementing an automated tree identification algorithm across such a large area as Eglin AFB, it was deemed much more practical in this study to apply an area-based modeling approach to predicting TPH, as has been successfully demonstrated in previous studies (e.g., Hudak et al. 2006; Yu et al. 2010).

Basal area estimates are even more difficult to estimate at the individual tree level because of the additional need for stem diameter measures. LiDAR data are much more sensitive to tree height and crown diameter than to stem diameter, given that few LiDAR returns actually reflect off of the vertically oriented tree stems. Furthermore, the $\sim$ 25-m cap on longleaf pine height growth, regardless of tree age, makes for a noisy height–diameter relationship for longleaf pine at Eglin AFB. The addition of crown dimension attributes to a biometric model can help but would require accurate individual tree crown delineation from the LiDAR returns. Although there are automated tools available to do this at the plot (Silva, Crookston, Hudak, et al. 2015) and stand levels (Silva, Hudak, Crookston et al. 2015), methods are still lacking for automating such tools at the landscape level, particularly as large and structurally diverse a landscape as Eglin AFB. Thus, operational utility also argued for an area-based modeling approach for predicting basal area.

Our area-based approach to predicting DomSpp is easiest to defend, because to define DomSpp at the individual tree level is not even possible conceptually. In this study, DomSpp was defined in each plot as the tree species having the most BA and is akin to forest type. Knowing which tree species comprises most of the basal area in a given management block increases the practical value of the TPH and BA maps to Eglin AFB managers, who are most concerned with maintaining healthy stands of longleaf pine, and for which we obtained higher Dom-

Spp classification accuracies (79%–86%) than overall (65%; Table 4). DomSpp varies more in the spectral and textural domains, increasing the utility of NDVIc and rumple as predictors. The NDVIc image from January was a better predictor than the NDVIc image from July (Table 3), possibly because ground vegetation would have been predominantly senesced in January to provide starker contrast to the predominantly green, coniferous, overstory canopy reflectance.

The management blocks (i.e., stands) provide an intermediate scale of analysis for bridging the divide between the Eglin AFB landscape and finer-scale ecological processes. With the possible exception of hurricanes, fire is the principal disturbance driving the ecosystem, and the primary tool used to manage longleaf pine ecosystems. Prescribed fires are applied at the management block level, and most wildfires are of similar or only slightly larger size. These disturbances drive ecological dynamics at the stand level; it is therefore an appropriate scale to analyze functional relationships between overstory structure and surface fuels (Figure 7), as influenced by prescribed fire, harvesting, or other management tools. We found forest and fire management strongly impacts plant-species diversity as well (Figure 7d), but further research is needed to understand the ecological mechanisms interacting to drive this functional relationship.

Multivariate imputation modeling is advantageous for preserving the covariance structure between forest attributes (Moeur and Stage 1995; Tuominen et al. 2003). Furthermore, imputation provides managers with a practical means, grounded in actual measurements, to associate surface fuel and species diversity characteristics to overstory structure attribures, even if knowledgeable explanations for such associations are incomplete (Fehrmann et al. 2008; McRoberts 2008). Regression models, because they are limited to univariate responses, break these associations (Tomppo et al. 2008). Although such associations might be weak, as manifested in maps that can appear "noisy," they become more apparent on aggregating so many mapped predictions within management units. Moreover, aggregation overcomes the uncertainty associated with map noise because managers make decisions at the scale of management polygons, not pixels. It is by aggregating so many mapped pixel-level predictions that significant and useful relationships might emerge at the stand level that are neither obvious to a careful observer in a 0.08 ha field plot nor from a 100,000+ ha vegetation map. By aggregating the data to the same scale that prescribed fire and other management interventions are applied, the stand scale might be most appropriate for revealing functional constraints imposed by the overstory canopy (Table 5, Figure 7), tied in longleaf pine forests to a strongly management-driven disturbance regime.

CONCLUSION

Overstory height and density metrics derived from airborne LiDAR, despite having lower point density in this analysis than is typically used for forestry applications, explained $\sim$ 50% of the variation in overstory structure and composition across Eglin AFB. The RF method produced comparable predictions of TPH, BA, and DomSpp by both the univariate regression/classification and multivariate k-NN imputation approaches, with only slightly less precise and accurate validation statistics from the imputation model, as expected by theory when $k = 1$. This slight disadvantage of imputation was outweighed by the advantage of imputation for preserving the covariance relationships between these responses as well as with ancillary plot measurements. The accuracy in predicting such ancillary measurements is contingent on the strength of their association with the response variables actually weighting the model.

This study demonstrates a method by which remote sensing data (LiDAR, in this case), seemingly unrelated plot data (surface fuel and plant-species richness data, in this case), and management data (fire management records, in this case) can be used in conjunction to reveal functional relationships between ecosystem structure and disturbance history. Stand-level forest, fuel, and fire management strategies and decisions influence not only the tree overstory but also ground cover plant composition and structure, and presumably the midstory and understory canopy layers in between. Eglin AFB managers are diligent at maintaining accurate, consistent, and comprehensive fire and monitoring records. We encourage other land management organizations to do the same, to complement and add value to investments in both field and remote sensing data collections.

ACKNOWLEDGMENTS

We thank Stephen Laine for Eglin AFB land management GIS data, Brett Williams and Nathan Price for descriptions of management techniques and goals. We also thank 3 anonymous reviewers for their constructive feedback.

FUNDING

This research was funded by the Department of Defense Strategic Environmental Research and Development Program (#RC-2243).

REFERENCES

Abrahamson, W.G., and Hartnett, D.C. 1990. "Flatwoods and dry prairies." In *Ecosystems of Florida*, edited by R.L. Myers and J.J. Ewel. Orlando, FL: University of Central Florida Press.

Bivand, R., Rowlingson, B., Diggle, P., Petris, G., and Eglen, S. 2014. *Splancs: Spatial and Space-Time Point Pattern Analysis*. R package version 2.01-36, accessed January 2016, http://cran.r-project.org/web/packages/splancs/index.html.

Bormann, F.H. 1953. "The statistical efficiency of sample plot size and shape in forest ecology." *Ecology*, Vol. 34: pp. 474–487.

Breiman, L. 2001. "Random forests." *Machine Learning*, Vol. 45: pp. 5–32.

Christensen, N.L. 1981. "Fire regimes in southeastern ecosystems." In *Fire regimes and ecosystem properties*, edited by H.A. Mooney, T.M. Bonnickson, N.L. Christensen, J.E. Lotan, and W.A. Reiners, pp. 112–130. General Technical Report WO-26. Washington, DC: US Department of Agriculture, Forest Service.

Christensen, N.L. 1988. "Vegetation of the southeastern coastal plain." In *North American Terrestrial VegetationI*, edited by M.G. Barbour and W.D. Billings, pp. 317–363. New York: Cambridge University Press.

Clark, P.J., and Evans, F.C. 1954. "Distance to nearest neighbor as a measure of spatial relationships in populations." *Ecology*, Vol. 35: pp. 445–453.

Crookston, N.L., and Finley, A.O. 2008. "yaImpute: an R package for kNN imputation." *Journal of Statistical Software*, Vol. 23(No. 10): pp. 1–16.

Dean, T.J., Cao, Q.V., Roberts, S.D., and Evans, D.L. 2009. "Measuring heights to crown base and crown median with LiDAR in a mature, even-aged loblolly pine stand." *Forest Ecology and Management*, Vol. 257: pp. 126–133.

Eskelson, B.N.I., Temesgen, H., Lemay V., Barrett, T.M., Crookston, N.L., and Hudak, A.T. 2009. "The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases." *Scandinavian Journal of Forest Research*, Vol. 24 (No. 3): pp. 235–246. DOI: 10.1080/02827580902870490

Evans, J.S. 2015. "Random forests model selection (downloadable tool)." In *Quantitative Methods in Spatial Ecology*, accessed Sept. 2015, http://evansmurphy.wix.com/evansspatial#!random-forests-model-select/cksm.

Evans, J.S., and Cushman, S.A. 2009. "Gradient modeling of conifer species using random forest." *Landscape Ecology*, Vol. 5: pp. 673–683.

Evans, J.S., Murphy, M.A., Holden, Z.A., and Cushman, S.A. 2011. "Modeling species distribution and change using random forests." In *Predictive Modeling in Landscape Ecology* edited by C.A. Drew, Y.F. Wiersma, and F. Huettmann, pp. 139–159. New York: Springer.

Fehrmann, L., Lehtonen, A., Kleinn, C., and Tomppo, E. 2008. "Comparison of linear and mixed-effect regression models and a *k*-nearest neighbour approach for estimation of single-tree biomass." *Canadian Journal Forest Research*, Vol. 38: pp. 1–9.

Frazer, G.W., Magnussen, S., Wulder, M.A., and Niemann, K.O. 2011. "Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass." *Remote Sensing of Environment*, Vol. 115: pp. 636–649.

Frost, C.C. 1993. "Four centuries of changing landscape patterns in the longleaf pine ecosystem." In *Tall Timbers Fire Ecology Conference 18*, edited by S.M. Hermann, pp. 17–44. Tallahassee, Florida: Tall Timbers Research Station.

Gelfand, A.E., Diggle, P.J., Fuentes, M., and Guttorp, P. 2010. *Handbook of Spatial Statistics*. Boca Raton, FL, USA: Chapman & Hall/CRC.

Gobakken, T., and Næsset, E. 2009. "Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data." *Canadian Journal of Forest Research*, Vol. 39: pp. 1036–1052.

Gonzalez-Benecke, C.A., Gezan, S.A., Samuelson, L.J., Cropper Jr., W.P., Leduc, D.J., and Martin, T.A. 2014. "Estimating *Pinus palustris* tree diameter and stem volume from tree height, crown area and stand-level parameters." *Journal of Forestry Research*, Vol. 25(No. 1): pp. 43–52.

Glitzenstein, J.S., Platt, W.J., and Streng, D.R. 1995. "Effects of fire regime and habitat on tree dynamics in north Florida longleaf pine savannas." *Ecological Monographs*, Vol. 65: pp. 441–476.

Harding, D.J., Lefsky, M.A., Parker, G.G., and Blair, J.B. 2001. "Laser altimeter canopy height profiles: methods and validation for closed-canopy, broadleaf forests." *Remote Sensing of Environment*, Vol. 76: pp. 283–297.

Hayashi, R., Kershaw Jr., J.A., and Weiskittel, A. 2015. "Evaluation of alternative methods for using LiDAR to predict aboveground biomass in mixed species and structurally complex forests in northeastern North America." *Mathematical and Computational Forestry & Natural Resource Sciences*, Vol. 7(No. 0) pp. 1–17. In press.

Hiers, J. K., O'Brien, J.J., Will, R.E., and Mitchell, R.J. 2007. "Forest floor depth mediates understory vigor in xeric *Pinus paulustris* ecosystems." *Ecological Applications*, Vol. 17: pp. 806–814.

Hollaus, M., Dorigo, W., Wagner, W., Schadauer, K., Höfle, B., and Maier, B. 2009. "Operational wide-area stem volume estimation based on airborne laser scanning and national forest inventory data." *International Journal of Remote Sensing*, Vol. 30: pp. 5159–5175.

Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M.S., Gessler, P., and Morgan, P 2006. "Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return LiDAR and multispectral satellite data." *Canadian Journal of Remote Sensing*, Vol. 32: pp. 126–138. doi:10.5589/M06-007.

Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., and Falkowski, M.J. 2008. "Nearest neighbor imputation of species-level, plotscale forest structure attributes from LiDAR data." *Remote Sensing of Environment*, Vol. 112(No. 5): pp. 2232–2245.

Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., and Falkowski, M.J. 2009. "Corrigendum to nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data" *Remote Sensing of Environment*, Vol. 113(No. 1): pp. 289–290.

Hudak, A.T., Strand, E.K., Vierling, L.A., Byrne, J.C., Eitel, J.U.H., Martinuzzi, S., and Falkowski, M.J. 2012. "Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys." *Remote Sensing of Environment*, Vol. 123: pp. 25–40.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougon, F., Yu, X., and Maltamo, M. 2008. "Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests." *International Journal of Remote Sensing*, Vol. 29: pp. 1339–1336.

Iacona, G.D., Kirkman, L.K., and Bruna, E.M. 2010. "Effects of resource availability on seedling recruitment in a fire-maintained savanna." *Oecologia*, Vol. 163(No. 1): pp. 171–180.

Isenburg, M. 2015. *LAStools – efficient tools for LiDAR processing, version 150304*, accessed September 2015, http://lastools.org.

Kane, V.R., McGaughey, R., Bakker, J.D., Gersonde, R.F., Lutz, J.A., and Franklin, J.F. 2010. "Comparisons between field- and LiDAR-based measures of stand structure complexity." *Canadian Journal of Forest Research*, Vol. 40(No. 4): pp. 761–773.

Kirkman, L.K., Mitchell, R.J., Helton, R.C., and Drew, M.B. 2001. "Productivity and species richness across an environmental gradient in a fire-dependent ecosystem." *American Journal of Botany*, Vol. 88(No. 11): pp. 2119–2128.

Koch, B., Straub, C., Dees, M., Wang, Y., and Weinacker, H. 2009. "Airborne laser data for stand delineation and information extraction." *International Journal of Remote Sensing*, Vol. 30: pp. 935–963.

Laine, Stephen. 2015. Personal communication, September 2015.

Latifi, H., and Koch, B. 2012. "Evaluation of most similar neighbor and random forest methods for imputing forest inventory variables using data from target and auxiliary stands." *International Journal of Remote Sensing*, Vol. 33(No. 21): pp. 6668–6694.

Lee, H., Slatton, K.C., Roth, B.E., and Cropper Jr., W.P. 2010. "Adaptive clustering of airborne LiDAR data to segment individual tree crowns in managed pine forests." *International Journal of Remote Sensing*, Vol. 31: pp. 117–139.

Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., and Harding, D. 1999. "LiDAR remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests." *Remote Sensing of Environment*, Vol. 70: pp. 339–361.

Lefsky, M.A., Cohen, W.B., Parker, G.G., and Harding, D.J. 2002. "LiDAR remote sensing for ecosystem studies." *Bioscience*, Vol. 52: pp. 19–30.

Lefsky, M.A., Cohen, W.B., and Spies, T.A. 2001. "An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon." *Canadian Journal of Forest Research*, Vol. 31: pp. 78–87.

Liaw, A., and Wiener, M. 2002. "Classification and regression by random-forest." *R News*, Vol. 2: pp. 18–22.

Lim, K., Treitz, P., Wulder, M., St-Onge, B., and Flood, M. 2003. "LiDAR remote sensing of forest structure." *Progress in Physical Geography*, Vol. 27: pp. 88–106.

Loudermilk, E.L., Hiers, J.K., O'Brien, J.J., Mitchell, R.J., Singhania, A., Fernandez, J.C., Cropper, W.P., Slatton, K.C. 2009. "Ground-based LIDAR: a novel approach to quantify fine-scale fuelbed characteristics." *International Journal of Wildland Fire*, Vol. 18: pp. 676–685.

Magnussen, S., and Boudewyn, P. 1998. "Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators." *Canadian Journal of Forest Research*, Vol. 28: pp. 1016–1031.

Mauya, E.W., Hansen, E.H., Gobakken, T., Bollandsås, O.M., Malimbwi, R.E., and Næsset, E. 2015. "Effects of field plot size on prediction accuracy of aboveground biomass in airborne laser scanning-assisted inventories in tropical rain forests of Tanzania." *Carbon Balance and Management*, Vol. 10: p. 10.

McGaughey, R. 2015. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization, version 3.50*, accessed December 2016, http://forsys.cfr.washington.edu/fusion/fusionlatest.html.

McRoberts, R.E. 2008. "Using satellite imagery and the *k*-nearest neighbors technique as a bridge between strategic and management forest inventories." *Remote Sensing of Environment*, Vol. 112: pp. 2212–2221.

McRoberts, R.E., Nelson, M.D., and Wendt, D.G. 2002. "Stratified estimation of forest area using satellite imagery, inventory data, and the k-nearest neighbors technique." *Remote Sensing of Environment*, Vol. 82: pp. 457–468.

Mitchell, R.J., Hiers, J.K., O'Brien, J.J., Jack, S.B., and Engstrom, R.T. 2006. "Silviculture that sustains: the nexus between silviculture, frequent prescribed fire, and conservation of biodiversity in longleaf pine forests of the southeastern United States." *Canadian Journal of Forest Research*, Vol. 36: pp. 2724–2736.

Mitchell, R.J., Hiers, J.K., O'Brien, J.J., and Starr, G. 2009. "Ecological forestry of the southeast: understanding the ecology of fuels." *Journal of Forestry*, Vol. 107: pp. 391–397.

Moeur, M., and Stage, A.R. 1995. "Most similar neighbor: an improved sampling inference procedure for natural resource planning." *Forest Science*, Vol. 41: pp. 337–359.

Murphy, M.A., Evans, J.S., and Storfer, A.S. 2010. "Quantify *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics." *Ecology*, Vol. 91: pp. 252–261.

Myers, R.L. 1990. "Scrub and high pine." In *Ecosystems of Florida*, edited by R.L. Myers, and J.J. Ewel, pp. 150–193.Orlando, FL: University of Central Florida Press.

Næsset, E. 2002. "Predicting forest stand characteristics with airborne scanning laser using a practical 2-stage procedure and field data." *Remote Sensing of Environment*, Vol. 80: pp. 88–99.

Næsset, E., and Bjerknes, K.O. 2001. "Estimating tree heights and number of stems in young forest stands using airborne laser scanner data." *Remote Sensing of Environment*, Vol. 78: pp. 328–340.

Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, Å., and Söderman, U. 2004. "Laser scanning of forest resources: the Nordic experience." *Scandinavian Journal of Forest Research*, Vol. 19: pp. 482–499.

Nemani, R., Pierce, L., Running, S., and Band, L. 1993. "Forest ecosystem processes at the watershed scale: sensitivity to remotely-sensed leaf area index estimates." *International Journal of Remote Sensing*, Vol. 14(No. 13): pp. 2519–2534.

Noss, R.F., LaRoe I, E.T., and Scott, J.M. 1995. *Endangered Ecosystems of the United States: A Preliminary Assessment of Loss and Degradation*. National Biological Service Biological Report 28. Washington, DC: U.S. Department of the Interior.

O'Brien, J.J., Hiers, J.K., Callaham Jr., M.A., Mitchell, R.J., and Jack, S.B. 2008. "Interactions among overstory structure, seedling life-history traits, and fire in frequently burned neotropical pine forests." *Ambio*, Vol. 37(No. 7–8): pp. 542–547.

Overing, J.D., Weeks, H.H., Wilson, J.P., Sullivan, J., and Ford, R.D. 1995. *Soil Survey of Okaloosa County, Florida*. Washington, DC: USDA Natural Resource Conservation Service.

Palmquist, K.A., Peet, R.K., and Weakley, A.S. 2014. "Changes in plant species richness following reduced fire frequency and drought in one of the most species-rich savannas in North America." *Journal of Vegetation Science*, Vol. 25(No. 6): pp. 1426–1437.

Parker, G.G., Harmon, M.E., Lefsky, M.A., Chen, J., Van Pelt, R., Weiss, S.B., Thomas, S.C., et al. 2004. "Three-dimensional structure of an old-growth *Pseudotsuga-tsuga* canopy and its implications for radiation balance, microclimate, and gas exchange." *Ecosystems*, Vol. 7(No. 5): pp. 440–453.

Pocewicz, A.L., Gessler, P., and Robinson, A.P. 2004. "The relationship between effective plant area index and Landsat spectral response across elevation, solar insolation, and spatial scales in a northern Idaho forest." *Canadian Journal of Forest Research*, Vol. 34: pp. 465–480.

Platt, W.J., Evans, G.W., and Rathbun, S.L. 1988. "The population dynamics of a long-lived conifer (*Pinus palustris*)." *The American Naturalist*, Vol. 131(No.4): pp. 491–525.

Popescu, S.C., Wynne, R.H., and Nelson, R.F. 2003. "Measuring individual tree crown diameter with LiDAR and assessing its influence on estimating forest volume and biomass." *Canadian Journal of Remote Sensing*, Vol. 29: pp. 564–577.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, accessed September 2015, http://www.R-project.org.

Silva, C.A., Crookston, N.L., Hudak, A.T., and Vierling, L.A. 2015. *Web-LiDAR Forest Inventory Applications*, http://forest.moscowfsl.wsu.edu:3838/csilva/Web-LiDAR/.

Silva, C.A., Hudak, A.T., Crookston, N.L., and Vierling, L.A. 2015. *rLiDAR: An R Package for Reading, Processing and Visualiz-*

ing LiDAR (Light Detection and Ranging) data, version 0.1, accessed October 2015, http://cran.r-project.org/web/packages/rLiDAR/index.html.

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.A., Lee, H., Falkowski, M.J., and Khosravipour, A. 2016. "Imputation of individual longleaf pine forest attributes from field and LiDAR data." *Canadian Journal of Remote Sensing*, Vol. 42(No. 5): pp. 554–573. doi: 10.1080/07038992.2016.1196582

Smart, L.S., Swenson, J.J., Christensen, N.L., and Sexton, J.O. 2012. "Three-dimensional characterization of pine forest type and Red-cockaded Woodpecker habitat by small-footprint, discrete-return LiDAR." *Forest Ecology and Management*, Vol. 281: pp. 100–110.

Stage, A.R., and Crookston, N.L. 2007. "Partitioning error components for accuracy assessment of near-neighbor methods of imputation." *Forest Science*, Vol. 53(No. 1): pp. 62–72.

Tomppo, E., Olsson, H., Ståhl, G., Nilsson, M., Hagner, O., and Katila, M. 2008. "Combining national forest inventory field plots and remote sensing data for forest databases." *Remote Sensing of Environment*, Vol. 112: pp. 1982–1999.

Tuominen, S., Fish, S., and Poso, S. 2003. "Combining remote sensing, data from earlier inventories, and geostatistical interpolation in multisource forest inventory." *Canadian Journal of Forest Research*, Vol. 33: pp. 624–634.

Williams, Brett. 2012. Personal communication, October 2012.

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., and Gobakken, T. 2012. "LiDAR sampling for large-area forest characterization: a review." *Remote Sensing of Environment*, Vol. 121: pp. 196–209.

Yu, X., Hyyppä, J., Holopainen, M., and Vastaranta, M. 2010. "Comparison of area-based and individual tree-based methods for predicting plot-level forest attributes." *Remote Sensing*, Vol. 2: pp. 1481–1495.