

Article

# Mapping Forest Characteristics at Fine Resolution across Large Landscapes of the Southeastern United States Using NAIP Imagery and FIA Field Plot Data

John Hogland <sup>1,\*</sup>, Nathaniel Anderson <sup>1</sup>, Joseph St. Peter <sup>2</sup>, Jason Drake <sup>3</sup> and Paul Medley <sup>3</sup>

<sup>1</sup> Rocky Mountain Research Station, U.S. Forest Service, Missoula, MT 59801, USA; nathanielmanderson@fs.fed.us

<sup>2</sup> College of Forestry and Conservation, University of Montana, Missoula, MT 59801, USA; jstpeter@fs.fed.us

<sup>3</sup> U.S. Forest Service, Tallahassee, FL 32321, USA; jasondrake@fs.fed.us (J.D.); pbmedley@fs.fed.us (P.M.)

\* Correspondence: jshogland@fs.fed.us; Tel.: +1-406-329-2138

Received: 15 February 2018; Accepted: 29 March 2018; Published: 3 April 2018



**Abstract:** Accurate information is important for effective management of natural resources. In the field of forestry, field measurements of forest characteristics such as species composition, basal area, and stand density are used to inform and evaluate management activities. Quantifying these metrics accurately across large landscapes in a meaningful way is extremely important to facilitate informed decision-making. In this study, we present a remote sensing based methodology to estimate species composition, basal area and stand tree density for pine and hardwood tree species at the spatial resolution of a Forest Inventory Analysis (FIA) program plot (78 m by 70 m). Our methodology uses textural metrics derived at this spatial scale to relate plot summaries of forest characteristics to remotely sensed National Agricultural Imagery Program (NAIP) aerial imagery across broad extents. Our findings quantify strong relationships between NAIP imagery and FIA field data. On average, models of basal area and trees per acre accounted for 43% of the variation in the FIA data, while models identifying species composition had less than 15.2% error in predicted class probabilities. Moreover, these relationships can be used to spatially characterize the condition of forests at fine spatial resolutions across broad extents.

**Keywords:** NAIP; FIA; remote sensing; forest measurements

## 1. Introduction

Forest management is a complex, integrated process that combines multiple objectives to accomplish a predefined set of goals as they relate to forested lands [1]. Since the United States National Forest Management Act of 1976, the federal definition of forest management has expanded well beyond timber management to include economic and social goals as components of management choices, the consideration of broader multiple use management challenges, and the need to quantitatively justify forest management plans and decisions [1]. This expansion in scope fundamentally changed not only the values for which forests are managed, but also how managers justify forest management decisions, emphasizing the need for effective, information-driven natural resource planning for diverse values in broad spatial, ecological, social, and economic contexts.

For forests of varying ownership, complexity, size, and extent, forest plans guide management activities and steer silviculture to meet both private and public objectives and goals. Effective development and implementation of those plans require knowledge of the biotic and abiotic conditions of a forest and an understanding of how such factors interact and change within the context of the objectives and goals defined [2,3]. To gain an understanding of the existing structure and composition of forests, practitioners have been implementing well-established mensuration

techniques for over a century [4]. Generally, these techniques can be described as aggregating a sample of field plots for a given forest stand or geographic area to estimate means and variation in forest characteristics within the sampled area (i.e., direct estimation). In this context, direct estimation approaches use plot samples to estimate parameters for a predefined region (i.e., domain). In contrast, indirect estimation techniques use the relationship(s) between field plots and some other data, such as the spectral values in remotely sensed imagery, to predict and estimate values of variables and parameters using models (e.g., regression). While direct techniques are well described and widely applied, especially in forest inventory, they can be extremely expensive and problematic to implement at fine spatial resolutions across large areas, mostly because of the labor costs associated with field data collection and the need for repeated measures as the landscape changes over time. Due to the costs associated with quantifying basic forest characteristics such as species composition, tree diameter, basal area and density, as well as more complex features such as the spatial juxtapositions of forest and non-forest ecosystems, habitat connectivity, biodiversity, water quality, esthetics, and carbon flux, limited information is known at the spatial resolution desired to inform management decisions at the project scale [5]. Much of the existing information is focused on forestlands managed for commercial timber, which tends to be the subject of more intensive, more frequent inventory than lands managed primarily for other values.

As a solution to these challenges, remotely sensed data can be related to plot data and the associated models can be used to estimate landscape characteristics across large extents [6] (i.e., indirect estimation). For many years, remotely sensed data have been used to study the surface of the earth [7] and stratify the terrestrial environment [8,9]. The opportunities to integrate field data and remote sensing data in new ways to measure and monitor landscapes have expanded greatly with recent advances in technology, mathematics, statistics, machine learning, and computer science. Relationships between remotely sensed reflected portions of the electromagnetic spectrum and the earth's terrestrial surface have been documented and exploited to build a wide range of data products depicting characteristics such as topography [10], land use and cover [11], vegetative indices [12], vegetation communities [13,14], fire severity [15], land cover change [16], and temperature [17]. Additionally, remotely sensed data have been used to estimate common forest metrics at fine spatial scales across relatively small extents [18,19] or at medium spatial scales across broad extents [20]. However, few studies have used fine scale (<30 m) passive remotely sensed data to predict key forest metrics such as species composition, basal area and stand density across broad extents (>10,000 km<sup>2</sup>) and across multiple ownership boundaries. To achieve this goal, we developed and deployed a remote sensing based approach that relates summarized U.S. Forest Service Forest Inventory and Analysis (FIA) [8] field plot data to spectral and textural metrics derived from the United States Department of Agriculture (USDA) National Agriculture Imagery Program (NAIP) [21] imagery in four regions of the southeastern United States (US). The models and outputs derived from these relationships provide the spatially detailed information needed to locate and prioritize forest management and conservation activities across all ownerships within these regions.

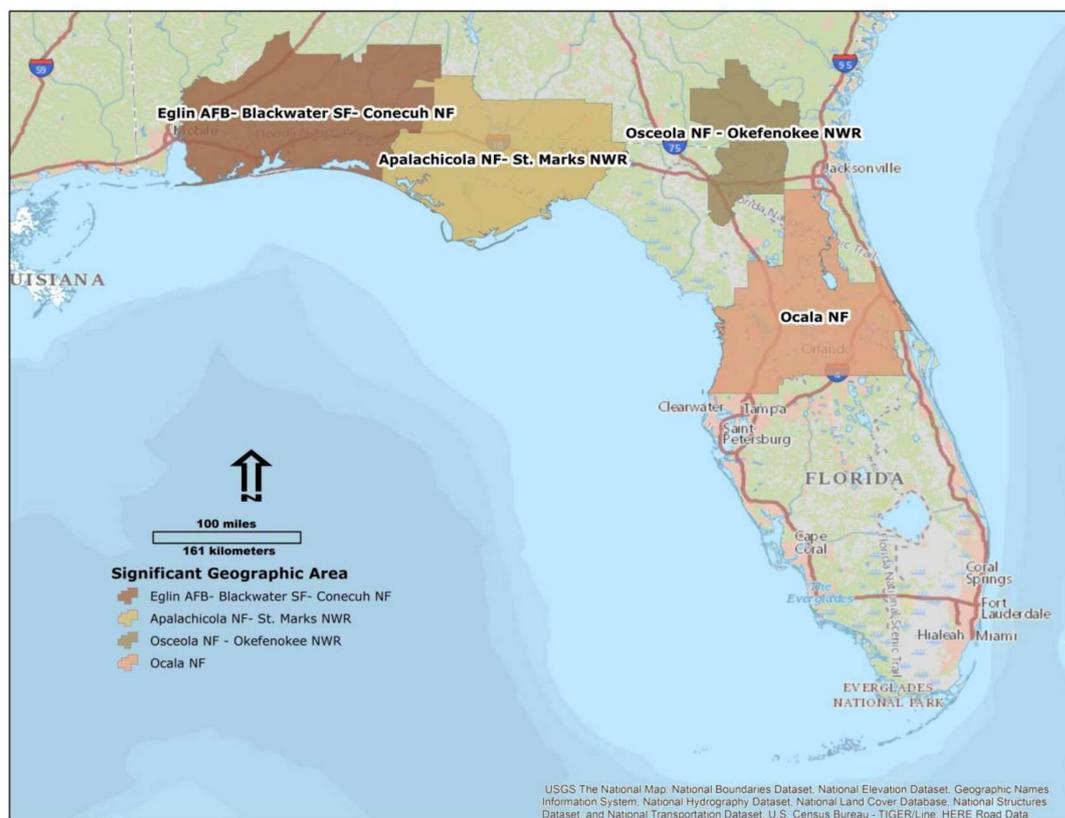
## 2. Materials and Methods

### 2.1. Overview

Across the southeastern US, multiple state, private, and federal organizations are actively promoting the restoration and conservation of longleaf pine (*Pinus palustris*) ecosystems [22]. Longleaf pine ecosystems, which were once the dominant forest type across the southeast US, are critically endangered and occupy less than five percent of their historic range [23]. These ecosystems, which support a wide diversity of plants and wildlife, are important to many environmental, social, and economic aspects of the region. Because of their importance, the restoration of longleaf pine has received significant attention. However, resources for restoration activities are limited, which forces agencies and organizations to evaluate management options and prioritize activities that have the

greatest potential for impact. To aid in this endeavor, spatially detailed information describing the condition of forest ecosystems, including longleaf pine forests, is needed across large portions of the southeast. Our remote sensing modeling approach provides this detail by relating summarized FIA field plot data to spectral and textural metrics of NAIP imagery to estimate and map forest characteristics at high resolution.

Our approach uses a suite of geospatial tools, spatially explicit probabilistic models, and associated raster surfaces to estimate and visualize several variables of interest at the spatial scale of an FIA field plot. These variables include a probabilistic classification of the dominant forest types, a measure of seedling dominance and dominance of longleaf pine, estimation of basal area in  $\text{ft}^2$  per acre (BAA), and stand density in trees per acre (TPA) for all tree species and for pine species explicitly. We focused our study on four specific Significant Geographic Areas (SGA) previously delineated by stakeholders in the Range-Wide Conservation Plan for Longleaf Pine [22], ranging in area from 10,113 to 27,172  $\text{km}^2$  each (Figure 1).



**Figure 1.** Extent of the study and corresponding significant geographic areas.

## 2.2. Data

The FIA program of the US Forest Service is a forest inventory program designed to measure, analyze, monitor, and report on the conditions of US forest and rangeland resources. To achieve its mission, a substantial component of the FIA program is dedicated to collecting forest information across the US on a systematic grid at a sampling intensity of approximately one field plot for every 2428 hectares [24]. For each field plot, FIA field technicians navigate to the fixed plot location, and then identify, measure, and record vegetative information within four subplots according to a standardized protocol [8]. Subplots have radii of 7.3 or 2 m depending on tree diameter and are located at the center of the plot and at 36.6 m from the center of the plot at  $0^\circ$ ,  $120^\circ$ , and  $240^\circ$  azimuth [8]. Once collected,

tree data are processed through a series of quality assurance and quality control routines and made publicly available through the FIA DataMart website [25].

FIA tree and plot data were downloaded as a Microsoft Access database from FIA's DataMart website, and confidential spatial locations were obtained directly from FIA following FIA's spatial data request protocol [26]. Tree species within FIA plots were grouped and summarized based on the species groupings shown in in Table 1, and based on tree diameter at breast height measured in inches (DBH; measured 4.5 ft (1.37 m) above the ground) for two size classes: between 1 in (2.54 cm) and 5 in (12.7 cm) DBH and greater than or equal to 5 in (12.7 cm) in DBH. Trees less than 1 in (2.54 cm) in DBH were simply tallied. The names of species are not included here for brevity, but the codes in Table 1 can be used to replicate our species groupings because they are standardized across all FIA data [8].

**Table 1.** Forest Inventory analysis (FIA) tree species groupings, description, and FIA species codes. Each code number corresponds to a tree species in the FIA protocol, which can be referenced online [8]. For this study, species were placed into three species groupings: Pine, Longleaf, and Hardwood.

Group Name	Description	FIA Species Codes
Pine	Mixed Pine	115, 128, 131, 132, 107, 110, 111
Longleaf	Longleaf Pine	121, <i>Pinus palustris</i>
Hardwood	Mixed Hardwood	461, 462, 491, 521, 531, 544, 555, 591, 611, 621, 652, 653, 682, 691, 692, 693, 694, 711, 721, 762, 858, 922, 931, 972, 975, 993, 999, 316, 373, 391, 311, 313, 345, 471, 500, 501, 520, 541, 545, 548, 551, 662, 681, 701, 722, 731, 744, 860, 912, 915, 920, 925, 951, 953, 970, 971, 973, 994, 998, 451, 552, 742, 766, 318, 356, 367, 421, 422, 651, 660, 760, 761, 581, 654, 502, 995, 68, 221, 222, 43, 67, 802, 812, 813, 820, 825, 827, 831, 837, 822, 824, 834, 828, 832, 819, 835, 838, 842, 840, 841, 402, 403, 404, 409, 401, 407, 410, 602

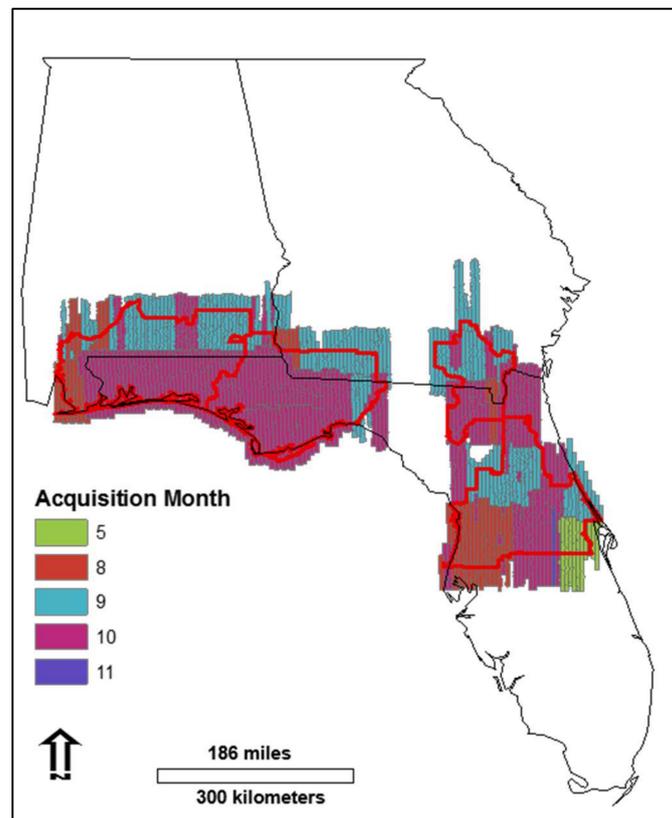
Summarized FIA plot values include BAA and TPA for pine and for all tree species combined. Seedlings and trees less than 5 in (12.7 cm) DBH were sampled with a smaller plot radius than trees greater than or equal to 5 in (12.7 cm) DBH, and expansion factors were applied to bring all plot measurements to appropriate density estimates [8]. Expansion factor values account for the difference in plot area measured on the ground and converted plot tallies to per acre estimates for trees in the different diameter classes.

Of the 3538 FIA plots downloaded and summarized, 1386 plots were removed based on a measurement date before 2010 or on a determination made after manual inspection of the plot location against a reference NAIP image. Manual inspections of FIA plots were performed by comparing summarized TPA estimates to the imagery. For locations where plot summaries clearly differed from what was visible in the imagery, such as a summary characterizing a dense forest falling on a non-forest area on the image, plots were removed from the dataset. The most common cause of incongruence between the imagery and the plot data was removal of the trees between the time the plot was measured and the time the image was acquired. This incongruence also occurred where the plot was measured at a location near, but not at the plot coordinates. However, plots were not removed because of registration errors. Even when plots were located at the exact plot coordinates, there are potentially fine spatial offsets between the imagery and the field plots due to registration errors in the imagery and the global positioning systems used to locate plot coordinates (around 6 m) and this is an added source of error in our models. After removal of old and incongruent plots, there were 458, 1451 and 243 FIA plots that we used to develop predictive models in Alabama, Florida, and Georgia, respectively.

NAIP images were downloaded from United States Geological Survey (USGS) file transfer protocol (FTP) site [27]. NAIP imagery was acquired in 2013 at a one-meter ground sample distance with a horizontal accuracy that matches within six meters of ground control points, with four band color

infrared (CIR) spectral resolution (red, green, blue and near infrared). Imagery was downloaded as digital ortho quarter quad tiles, each covering a  $3.75 \times 3.75$  min quarter quadrangle plus a 300 m buffer on all four sides. In total, 4986 NAIP tiles were downloaded and stored locally. Once downloaded, these images were assembled into three state-based Mosaic Raster datasets [28]. Though many datasets were evaluated as potential predictors of forest characteristics (e.g., soils, digital elevation models, ecoregions), few provided the detail and consistency needed to produce SGA-wide estimates of BAA and TPA. Of the datasets that did provide such detail and consistency, NAIP imagery was chosen as the primary predictive dataset because it is digital, quantifies reflected electromagnetic energy in the visible and CIR portions of the electromagnetic spectrum, has been preprocessed and color balanced [29] for each state, is readily available without charge, has a fine spatial resolution ( $1 \text{ m}^2$ ), is reacquired every two years, and has complete coverage within the continental US [21].

There are some drawbacks to using NAIP imagery across large extents that cross multiple states (as in this study). One of the drawbacks is the relatively small spatial footprint of NAIP imagery (approximately  $7 \text{ km} \times 8 \text{ km}$ ). Because of this small footprint, multiple flight paths must be flown to completely cover the area within an SGA. This results in spectral differences among images based on a camera's configuration and the time of image acquisition (Table 2, Figure 2). We did not try to correct for this variation but used NAIP imagery in its native format as provided [27].



**Figure 2.** Acquisition month for National Agricultural Imagery Program (NAIP) flight lines within the extent of the Significant Geographic Areas (SGAs).

However, flight line information is useful to examine potential impacts of this variation on outputs. To acquire flight line information, we built a set of routines that automatically downloaded vector data from map services and stored that information locally on a personal computer. In total, there are 1435 flight lines across Alabama, Georgia, and Florida with 391 flight lines overlapping the SGAs. NAIP imagery within SGAs was collected from May to November 2013 and varied within similar months by date and time of day of collection.

Given the fine grain size of NAIP imagery, acquisition, storage, and processing for the large extents of the SGAs presented logistical challenges. To address these challenges, we built an automated download tool that systematically selected and downloaded each NAIP image for a defined area of interest from USGS's FTP site [27]. A similar tool was used to automate downloads of FIA data, and both are included in the Rocky Mountain Research Station (RMRS) Raster Utility toolbar [30].

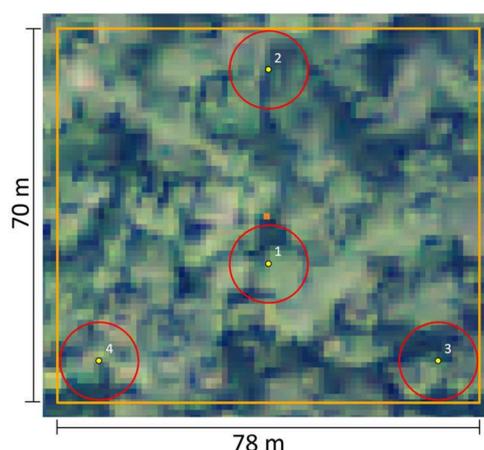
**Table 2.** Digital camera model, firmware, and range of acquisition months for National Agricultural Imagery Program (NAIP) images.

State	Model	Firmware	Months	Images
Alabama	Leica Geosystems	ADS40 SH51 F/W VER 3.23 ADS40 SH91 F/W VER 3.23	August–October	784
Georgia	Leica Geosystems	ADS40 SH51 F/W VER 3.23 ADS40 SH91 F/W VER 3.23	August–October	1831
Florida	Leica Geosystems	ADS40 SH51 F/W VER 3.23 ADS40 SH81 F/W VER 3.23 ADS40 SH91 F/W VER 3.23	May–November	2371

### 2.3. Data Analysis and Modeling

All analyses in this project were performed using Environmental Systems Research Institute's (ESRI's) geographic information system (GIS), R statistical software (version 3.4.3) [31], and the RMRS Raster Utility toolbar [30]. To facilitate the tabular, spatial, statistical, and GIS analyses performed over the course of the project, we developed a suite of spatial modeling tools that take advantage of Function Modeling (FM) [30,32] and parallel processing. These tools work within ESRI's ArcGIS and are available for free download [33]. If ESRI software is not available, GIS users may still find utility in the C# code for the RMRS Raster Utility, which is publicly available in a .NET subversion library [32].

Layers of predictor variables for models of forest characteristics were created by performing a series of textural transformations of NAIP imagery that quantified varying aspects of spectral values and image texture [34]. NAIP digital number (DN) values were transformed at the spatial extent of an FIA field plot (78 m by 70 m) using the "Focal Analysis" and "Create GLCM" tools within the RMRS Raster Utility toolbar. Each of these tools uses a moving window of specified size (78 by 70 cells in our study) to summarize cell values within the window and attribute that summarized value to the center cell of the window (Figure 3). By performing these moving window transformations on the NAIP imagery, every 1 m<sup>2</sup> cell represents the location of a potential FIA plot with summarized spectral values (texture) for the spatial extent of the FIA plot. Textural values located at each cell can be extracted based on the spatial coordinates of an FIA plot to train predictive models or used as predictor surfaces to estimate forest characteristics across the study area once a model has been trained. Textural values created for each of the four NAIP bands include mean, standard deviation, and horizontal gray level co-occurrence matrix (GLCM) contrast of DN values, and provide the base predictor variables used in subsequent models. To relate FIA field data to NAIP and texture derivatives, we used the "Sample Raster" tool within the RMRS Raster Utility toolbar.



**Figure 3.** Extent (orange box), central pixel (solid orange square), and spatial layout of an Forest Inventory Analysis (FIA) plot (red circles) with National Agricultural Imagery Program (NAIP) imagery in the background.

Models of forest characteristics consisted of a suite of different models for each state. Models that estimate class probabilities (Table 3) were developed using a softmax neural network and the “Softmax Nnet” tool [33]. Models to predict BAA and TPA were developed using random forest regressions and the “Random Forest” tool [33]. Softmax neural networks are a machine learning, artificial neural network classification technique that uses linear outputs, no hidden networks, and a final softmax normalization layer for classification. Softmax neural networks were chosen to classify different categorical aspects of the forest because of their high efficiency, flexibility, and the probabilistic interpretation of model predictions [35,36]. Random forest regression trees, which create multiple regression tree models using subsets of predictor variables and observations and then assemble those regression trees into one combined model, were chosen to predict continuous variables of forest characteristics due to their similar efficient and flexible nature [36,37].

**Table 3.** Logical rules used to categorize DomType, Regen, and Longleaf classes within Forest Inventory Analysis (FIA) plot summaries.

Model	Class	Code	Query
DomType	Hardwood	1	(Hardwood BAA > Pine BAA) and NOT (Total TPA < 8 and TPA less than 5 in DBH < 225)
	Pine	2	(Pine BAA ≥ Hardwood BAA) and not (Total TPA < 8 and TPA less than 5 in DBH < 225)
	NONFOR	3	Total TPA < 8 and TPA less than 5 in DBH < 225
Regen	Regen	1	(DomType < 3 and Total TPA < 20 and TPA less than 5 in DBH > 300)
LongDom	Longleaf	1	DomType = 2 AND (((Longleaf BAA)/Pine BAA) ≥ 0.5)

BAA = Basal area per acre; TPA = Trees per acre.

The quality of random forest models was evaluated using three metrics: the training root mean square error (Train RMSE), the out-of-bag root mean square error (OOB RMSE), and the standard deviation (SD) of the plot summaries. Train RMSE reports the average RMSE of the regression tree models when evaluated with the data used to calibrate the models. OOB RMSE reports the average RMSE of the regression tree models when evaluated against data held out of the calibration process.

The quality of softmax neural network models was evaluated using two metrics: average error and most likely class map accuracy. Average error in this case represents the mean error in estimated probability ( $\hat{\mu}_i$ ) of the class being predicted:

$$\frac{1}{n} \sum_{i=1}^n (1 - \hat{\mu}_i). \quad (1)$$

Most likely class map accuracy is calculated by labeling each observation within a sample using a most likely class rule across all response categories given  $\hat{\mu}_{ki}$ , summing up the number of times that rule correctly labels each observation, and then dividing by the total number of observations within a sample. To independently evaluate classification error, we developed and performed a bagging routine that randomly selected 80% of the data to train 50 separate models and then used the remaining 20% of the data to evaluate those models. Final models used to perform classifications were generated using all the data.

Each state has a set of models that relate remotely sensed texture derivatives of NAIP imagery to FIA plot data to predict three class probabilities: (1) the probability of a plot being composed of one of three dominant forest types based on BAA (abbreviated “DomType”); (2) the probability of a plot being composed primarily of regeneration based on TPA and BAA (abbreviated “Regen”); and (3) the probability of a plot being dominated primarily by longleaf pine basal area (abbreviated “LongDom”). Additional models predict the value of pine basal area (BAA, abbreviated “PineBAA”), the density of pine trees (TPA, abbreviated “PineTPA”), the basal area of all species (abbreviated “AllBAA”), and density of all species (abbreviated “AllTPA”).

Using these models, texture derivatives of NAIP imagery, and the “Build Raster From Model” tool, we built 21 raster surfaces (seven for each state) that estimate the probability of each 1 m<sup>2</sup> being a given class (DomType, Regen, and LongDom), along with estimates of BAA and TPA for pine specifically and for all tree species combined. Model predictor variables used for DomType and Regen consisted of all 12 NAIP texture derivatives (mean, standard deviation, and contrast for each of the four NAIP bands), while LongDom, PineBAA, PineTPA, AllBAA, and AllTPA consist of subsets of those variables and modeled outputs from DomType and Regen. Predictor variable selection for LongDom, PineBAA, PineTPA, AllBAA, and AllTPA was based on an iterative stepwise selection process using general additive modeling procedures developed within R statistical software (Appendix A). Predictor variables that reduced deviance and were statistically significant at  $\alpha = 0.1$  were used as inputs into random forest regressions and softmax neural network classifications. For random forest models, the number of classification trees allowed and the proportion of observations used to train the models remained constant at 100 trees and 66% of observations, while the number of variables randomly chosen varied depending on the total number of variables in the model, but ranged from 1 to 4.

#### 2.4. Processing

With the exception of predictor variable selection, which was done in R, all statistical modeling and data processing functions were performed using ESRI’s ArcGIS and the RMRS Raster Utility toolbar. To take advantage of multiple CPUs, newly coded tools were added to the RMRS Raster Utility that allowed for tiling of raster surfaces while saving small subsets of an otherwise large output in a parallel fashion. Once saved, tiled outputs were mosaicked into a “Mosaic Raster Dataset” and used as one continuous dataset. To minimize the storage footprint and streamline processing, we used FM [32] in the RMRS Raster Utility toolbar. The underlying concept behind FM is delayed reading that postpones processing data until absolutely necessary. Using this concept, multiple transformations to data occurred without writing intermediate outputs to disk, which saved both processing time and storage space. Outside of ArcMap, FM “batch” files were used to process data within the RMRS Raster Utility batch executable.

Outputs from the modeling process consisted of multiple batch files for each state that transform NAIP spectral and texture variables into DomType, Regen, LongDom, PineBAA, PineTPA, AllBAA, and AllTPA raster surfaces at the spatial resolution of 1 m<sup>2</sup>. To aid in handling and using these outputs, we saved modeled outputs into a series of tiled Erdas Imagine files (.img). Storage space for these datasets required multiple terabytes. Therefore, outputs were resampled to a spatial resolution of 10 m<sup>2</sup> and compressed using .TIF file format, resulting in much smaller and more manageable datasets ranging in size from 44.3 megabytes to 1.3 gigabytes. Resampling procedures were performed using RMRS Raster Utility's nearest neighbor "Resampling" tool.

### 3. Results

#### 3.1. Data Acquisition

The time required to download data from a website is a function of network speeds and server usage and can be calculated based on the amount of data requested and the speed of a connection. For this study, we used a connection speed of 6 Mbps and downloaded NAIP imagery for the extent of the SGAs and FIA state-wide tree lists. NAIP imagery was compressed in Jpeg 2000 format and amounted to 105 gigabytes of data in total. FIA tree lists were stored within an Access database and amounted to 3.8 gigabytes of data. In theory, this should require approximately 42 h to download the NAIP data and 1.5 h to download the FIA data. While it is fairly straightforward to manually download FIA data, finding, manually selecting, and downloading each NAIP tile within the extents of the SGAs could take up to one month. Our automated download tool took approximately 48 h to download all NAIP imagery for the four SGAs.

#### 3.2. Modeling

Texture variables for a moving window the size of an FIA plot were developed directly from NAIP imagery using the DN values for the four bands. Texture values were then used as predictors in our models. Models were built on a state-by-state basis creating a total of 21 unique models (seven for each state). Modeling error varied by state and response variable and is reported in Tables 4 and 5. In all cases, BAA and TPA model OOB RMSE error was substantially less than the standard deviation of the plots (Table 4). With regards to variation (RMSE<sup>2</sup> and standard deviation<sup>2</sup>), on average, our models accounted for 43% of the variation within the data. Similarly, Domtype, Regen, and LongDom classification models had substantially less deviation from actual probabilities than using the plot proportions alone, and, on average, had less than 15.2% error in predicted class probabilities. While modeled relationships were a significant improvement over classical spatial aggregation techniques that use FIA state level SGA summaries and their standard deviation estimates, there was more error in the relationships than we anticipated. Generally, relationships between pine BAA and pine TPA were stronger with less error than relationships between all tree BAA and all tree TPA (Table 4). Predictor variables used for BAA and TPA models varied from model to model, but generally included both texture and DomType values.

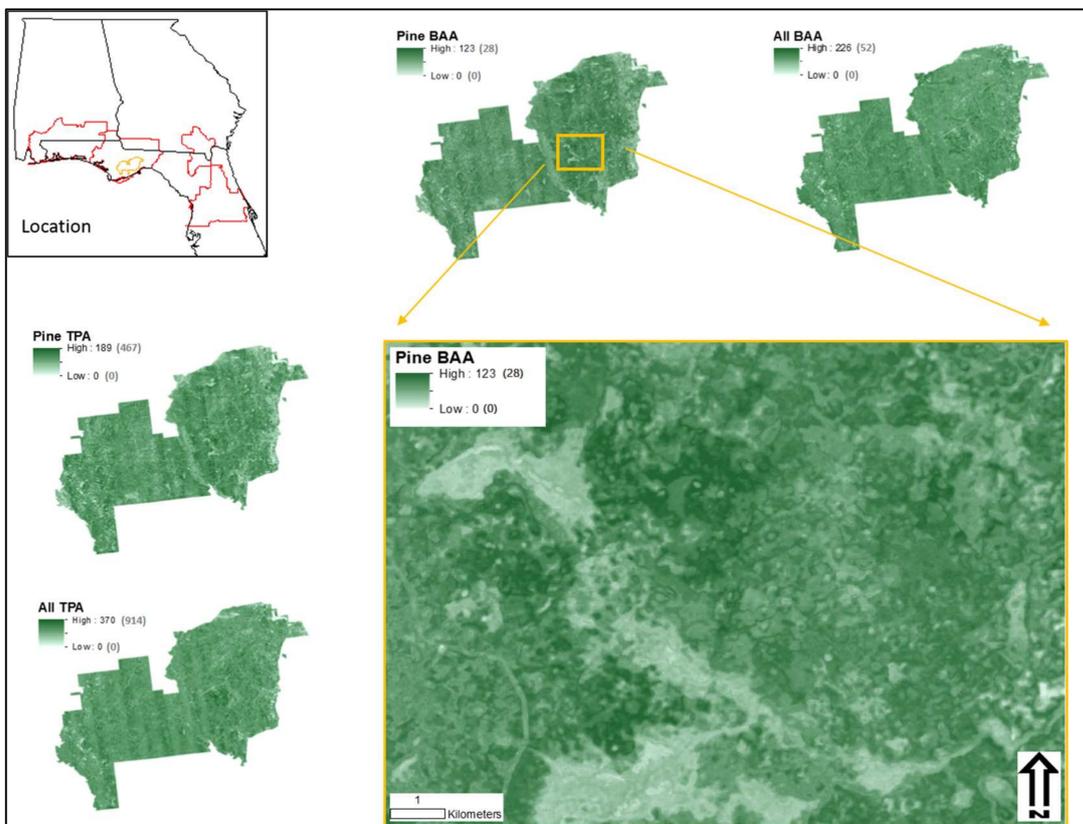
**Table 4.** Random Forest model predictor variables used, fit statistics for BAA and TPA, and standard deviation of FIA plots.

State	Model	Predictors	Train RMSE	OOB RMSE	Std. Dev.
Alabama	Pine BAA	Mean; NAIP1   Standard Deviation; NAIP4   GLCM; NAIP4   DomType; NonForest, Pine, Hardwood   Regen	9.66	28.03	35.14
	Pine TPA	Mean; NAIP1, NAIP3   Standard Deviation; NAIP4   GLCM; NAIP2, NAIP4   DomType; Pine   Regen	25.34	72.35	89.87
	All BAA	Mean; NAIP1, NAIP3   GLCM; NAIP2, NAIP3   DomType; NonForest, Pine	12.46	36.51	52.69
	All TPA	Mean; NAIP1, NAIP3   Standard Deviation; NAIP4   GLCM; NAIP2, NAIP4   DomType; Pine   Regen	25.21	71.79	99.05
Georgia	Pine BAA	Mean; NAIP1, NAIP4   Standard Deviation; NAIP4   GLCM; NAIP4   DomType; Hardwood   Regen	7.95	23.17	30.29
	Pine TPA	Standard Deviation; NAIP4   GLCM; NAIP4   DomType; Pine   Regen	18.86	54.8	67.23
	All BAA	Mean; NAIP2, NAIP3, NAIP4   GLCM; NAIP3   DomType; NonForested	13.62	40.2	52.86
	All TPA	Standard Deviation; NAIP4   GLCM; NAIP2, NAIP3   DomType; Hardwood   Regen	24.98	72.36	93.59
Florida	Pine BAA	Mean; NAIP1, NAIP2   Standard Deviation; NAIP1-NAIP4   GLCM Contrast; NAIP4   DomType; Pine, Hardwood	8.00	23.38	31.37
	Pine TPA	Mean; NAIP1, NAIP2   Standard Deviation; NAIP1-NAIP3   GLCM Contrast; NAIP1-NAIP4   DomType: Pine, Hardwood	20.53	59.55	82.27
	All BAA	Mean; NAIP1, NAIP4   Standard Deviation; NAIP2, NAIP3   GLCM Contrast; NAIP1, NAIP2, NAIP4   DomType: Hardwood   Regen	15.82	46.12	58.91
	All TPA	Mean; NAIP1, NAIP4   Standard Deviation; NAIP2-NAIP4   GLCM Contrast; NAIP1, NAIP2   DomType: Pine	27.17	77.87	103.60

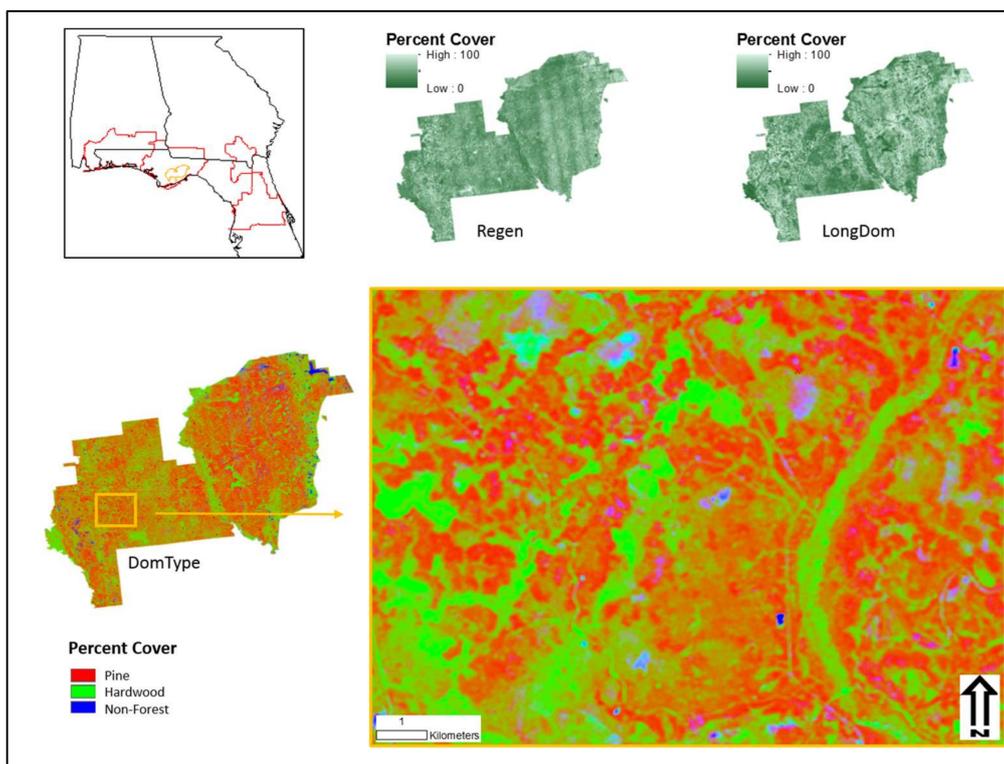
**Table 5.** Soft-max neural net predictor variables used and fit statistics for DomType and Regen models.

Model	State	Predictor variables	Average Error (Bagged)	Most Likely Class Map Accuracy (Bagged)
DomType	Alabama	All Mean   All Standard Deviation   All GLCM Horizontal Contrast	23% (35%)	76% (73%)
	Georgia	All Mean   All Standard Deviation   All GLCM Horizontal Contrast	20% (34%)	76% (70%)
	Florida	All Mean   All Standard Deviation   All GLCM Horizontal Contrast	23% (35%)	74% (73%)
Regen	Alabama	All Mean   All Standard Deviation   All GLCM Horizontal Contrast	10% (11%)	94% (94%)
	Georgia	All Mean   All Standard Deviation   All GLCM Horizontal Contrast	16% (17%)	89% (89%)
	Florida	All Mean   All Standard Deviation   All GLCM Horizontal Contrast	11% (11%)	94% (94%)
LongDom	Alabama	Mean; NAIP2, NAIP3   GLCM Contrast; NAIP2, NAIP3, NAIP4	13% (13%)	92% (92%)
	Georgia	Mean; NAIP2, NAIP3, NAIP4   Standard Deviation; NAIP3, NAIP4   GLCM Contrast; NAIP1, NAIP4   DomType; Hardwood, Pine   RegenProb	6% (7%)	96% (95%)
	Florida	Mean; NAIP1, NAIP3   Standard Deviation; NAIP4   GLCM Contrast; NAIP2   DomType; Pine   RegenProb	12% (12%)	92% (92%)

Similar to BAA and TPA models, Domtype, Regen, and LongDom classification models showed more error than expected given previous investigations [38]. However, like BAA and TPA models, these models significantly improved upon classical spatial aggregation approaches, which simply use the proportion of times a given category is encountered within the plot data and then extrapolates that proportion across the study area. Independent estimates of model error were almost identical to modeled error using all the data. This indicates that the decision to use all the data when calibrating the final model produces unbiased results and that the estimates of model error from those data were stable and not overly optimistic. Predictors in DomType and Regen classification models used all NAIP texture variables, while LongDom classification predictor variables varied by state (Table 5). Using these models, the appropriate NAIP and classification predictors, and the “Save parallel” tool, we built estimates of BAA and TPA (Figure 4) and DomType, Regen, and LongDom class probabilities for every 1 m<sup>2</sup> within the four SGAs (Figure 5). Total processing time and storage space for associated raster surfaces are summarize in Table 6.



**Figure 4.** Example of basal area per acre (BAA; (m<sup>2</sup> per ha)) and trees per acres (TPA (trees per ha)) surfaces generated for the Apalachicola National Forest, which is a subset of the Apalachicola National Forest-St. Marks National Wildlife Refuge SGA.



**Figure 5.** Example of DomType, Regen, and LongDom percent cover surfaces for an enlarged sub-section of the Apalachicola National Forest. DomType is displayed as three band color composite with red, green, and blue colors corresponding to pine, hardwood, and non-forest classes, respectively, while Regen and LongDom are displayed as a single band color stretch.

**Table 6.** Total processing time and size of outputs.

Model	Processing Time (h)	Storage Space (gigabytes)
ALL BAA	94	143
ALL TPA	94	163
PINE BAA	94	99
PINE TPA	94	145
DomType	104	250
Regen	104	38
LongDom	78	32

### 3.3. Aggregation and Resampling

To aid in using and distributing these outputs, we resampled our 21 raster surfaces to a spatial resolution of 10 m<sup>2</sup> [39]. The resampling routine used a nearest neighbor algorithm that takes the value from a finer resolution raster surface found closest to the center of a coarser resolution pixel, and populates that coarser resolution pixel with that value. The nominal spatial resolution of our raster surfaces relate to the spatial footprint of the FIA plot (78 m × 70 m) and resampling our outputs to a spatial scale of 10 m<sup>2</sup> can be interpreted as centering estimates of FIA plot tree summaries to every 10 m<sup>2</sup> across the four SGAs. Processing time associated with resampling was approximately 12 h. Total storage space associated with resampling was approximately 20 gigabytes.

## 4. Discussion

We were able to significantly improve estimates of mean forest characteristics for the four SGAs using NAIP imagery and FIA field plot data. By using our modeling approach, changes in

forest characteristics occurring between field plots were accurately depicted based on NAIP texture. Compared to spatially aggregating field plots (the classical method), this level of detail can uncover spatial patterns previously undetected and provide needed information related not only to the forested condition for the area as a whole, but also at the spatial scale of projects. In addition, we were able to demonstrate that this type of analysis can be done using readily available datasets and conventional personal computer hardware in a practical and relatively inexpensive manner. The models and associated raster outputs can be used in multiple ways to effectively plan forest management activities and longleaf pine restoration and conservation. To demonstrate how these raster datasets can be created and used, we have developed four tutorials covering topics ranging from model building to prioritization of silvicultural treatments [33]. Moreover, within these tutorials, we demonstrate how to use modeling error to accurately account for variation within the modeled outputs.

While our current models and outputs significantly reduce error and provide substantially more spatial detail when compared to classical aggregation techniques, as shown in Tables 4 and 5, there was more error in the models than we expected based on previous studies in the western US. Though the exact causes of the difference between using this approach in Colorado and Montana compared to the southeastern US have not been rigorously tested, we believe that the larger error in this region is primarily due to more heterogeneity across the landscape, more diverse species composition, seasonal variation, and bigger errors from the striping phenomenon in the NAIP imagery. Additionally, a small portion of model error for LongDom, BAA, and TPA models can also be attributed to using DomType and Regen as predictor surfaces. However, there are potentially many ways to further improve estimates and reduce the amount of error within the outputs for this region.

To better understand how we can accomplish those improvements, it helps to look at two primary sources of variability within the data: (1) differences in the acquisition time of NAIP imagery and (2) the FIA sampling protocol. First, within the NAIP imagery, the mosaicking process employed by NAIP appears to cosmetically bring images to a common radiometric scheme. On further inspection, the mosaicking process presents significant complications. Specifically, the NAIP mosaicking process and the variation in acquisition dates manifest in the models as increased error and in the outputs as striping (Figures 4 and 5). Moreover, the variation and striping within Florida, Alabama, and Georgia appear to be more dramatic than in other parts of the nation [32,38]. Using unaltered NAIP imagery, and applying a relative normalization routine such as “Aggregate No Change Regression” [40] across all image tiles would help solve this problem to some degree. However, it is difficult to acquire unaltered imagery under current NAIP data protocols, and this step would significantly increase the amount of processing time to generate outputs for the same spatial extent. Furthermore, given the time difference in image acquisition, which was as long as six months in this study (Figure 2), a relative normalization technique may not produce the desired outcome.

In an ideal situation, the imagery used in this type of analysis should be acquired in such a manner that differences in reflectance are minimized across acquisition flight lines. Using alternative fine resolution imagery, such as IKONOS or WorldView, which can be purchased, or using coarser resolution imagery that has broader coverage and finer temporal resolution such as Sentinel 2 and Landsat 8, has strong potential to improve results. While purchasing fine resolution imagery is appealing, it presents some logistical challenges and can be extremely expensive for large landscapes like the SGA extents, especially if they must be mapped repeatedly to keep up to date with forest conditions and land use change. A less expensive approach is to relate fine resolution outputs created from models calibrated without the NAIP striping impact (i.e., models developed using samples located between the stripes of NAIP imagery) to coarser free imagery such as Landsat 8.

The second source of variation relates to the spatial accuracy of FIA plot locations and the FIA field sampling protocol. As previously mentioned, the actual locations of the plots are kept confidential, but are made available for research and other applications through special agreements with FIA [24,26]. In this study, we used the actual coordinates of each FIA plot to relate NAIP texture variables to plot response variables, and used the resulting models to estimate and map those response variables across

the entire landscape. FIA plots are monumented, and revisited periodically, and coordinates of plot locations were developed based on a systematic grid, with monuments placed based on navigation grade GPS. This suggests that there can be locational variation in plot coordinates and monument locations. For the extent of an FIA plot, there can also be substantial variation in what is measured and therefore substantial variation in summarized values and estimates, especially for trees less than five inches in diameter. Each FIA plot covers approximately 1.3 acres (78 m by 70 m extent) and is composed of four, 1/24 acre subplots for trees greater than or equal to 5 in DBH and four 1/300 acre subplots for trees less than 5 in DBH [8]. Within an FIA plot, this means that approximately 12% of the extent of the plot is measured for trees greater than or equal to 5 in DBH and less than 1% of the extent is measured for trees less than 5 in DBH. To draw stronger relationships between plot estimates and NAIP texture derivatives, more of the plot extent should be sampled with an equal sampling area for all trees, regardless of size. The authors are currently testing such a sampling design in the Apalachicola SGA.

Given the error associated with our models and the striping phenomena associated with mosaicked NAIP imagery, care should be taken when using the outputs, and model error should be taken into consideration when making management decisions, especially at the stand level. For small areas (i.e., less than 100 acres, 40.5 hectares), summarized mean values can vary substantially. However, for larger areas (i.e., greater than 1000 acres, 405 hectares), mean estimates are relatively accurate and provide a better depiction of the true condition of the landscape than summarized plot data alone. This reinforces the idea that remote sensing models such as these are complimentary to field plot sampling, rather than a replacement for direct measurement. Even so, modeling error can be effectively incorporated into a GIS analysis to help understand and guide appropriate use of these data products, and we suggest that, when quantifying project-scale information, model error should be used to build lower and upper confidence limits for any estimates. Moreover, future research that attempts to relate field plot data with fine resolution imagery should consider both the striping of the mosaicked NAIP imagery and the field sampling design to better account for error in estimating forest characteristics.

## 5. Conclusions

Our models and outputs accurately link remotely sensed imagery and FIA plot data to describe forest characteristics in a spatial manner that is an improvement over simply aggregating plot information to large areas such as state, county, or SGA boundaries. The models and the outputs created in this study can be used not only to describe forests within specific areas, but also to provide information about the condition and spatial pattern of longleaf pine at much finer resolution over smaller spatial extents than the SGA, such as ownerships, forest stands and project areas. This makes them valuable for project-scale decision support. Moreover, these outputs can be combined, transformed, and summarized to address a wide range of issues related to forest management, including prioritizing treatments across large landscapes.

**Acknowledgments:** This research was funded by the National Fish and Wildlife Foundation through a grant to the Longleaf Alliance (Grant ID #50661) and by the U.S. Forest Service, State and Private Forestry and the Rocky Mountain Research Station. Additional support for the development of the RMRS Raster Utility was provided by the Agriculture and Food Research Initiative Competitive Grant 2013-68005-21298 from the USDA National Institute of Food and Agriculture through the Bioenergy Alliance Network of the Rockies (BANR).

**Author Contributions:** John Hogland proposed, built, and applied the tools developed within the study. He also designed and implemented parts of the case studies as described. He was the lead author of the manuscript. Nathaniel Anderson provided project oversight and contributed to writing portions of the manuscript. Joseph St. Peter provided technical assistance with developing the raster surfaces and writing portions of the manuscript. Jason Drake and Paul Medley evaluated outputs and contributed to writing portions of the manuscript.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Appendix A

The R scripting language source code used to perform general additive modeling (GAM) variable selection. This routine uses the MGCV, R library to iteratively perform GAM procedures and identify variables that significantly improve model fit. This routine was used in this study to identify predictive model explanatory variables.

```
#GAM variable selection routine
#indata = dataframe
#resp = text value identifying response variable
#pred = text vector of predictor variable names
#alpha = numeric values identifying significance level of variables keep in the model
#fam = optional parameter identifying the modeling family to use
#improveby = the amount a new variable must improve an existing model to consider using
  that variable
library(mgcv)
getSigFldNames <- function(indata, resp, pred, alpha, fam = gaussian(), improveby = 0)
{
  sigVar <- c()
  pdiv <- -0
  pr2 <- -0
  for(i in seq(length(pred)))
  {
    vars <- c(sigVar, pred[i])
    fm <- as.formula(paste(resp, "~ ", paste("s(", vars, ")", collapse = " + ")))
    md <- gam(fm, data = indata, family = fam)
    smry <- summary(md)
    div <- -smry$dev.expl
    if(div > (pdiv + improveby))
    {
      print(paste("Adding variable", pred[i], collapse = " "))
      pvalues <- -c(smry$s.pv)
      sigVar <- c()
      nonSigVar <- c()
      for (j in seq(length(vars)))
      {
        pv <- -pvalues[j]
        if(pv ≤ alpha)
        {
          sigVar <- c(sigVar, vars[j])
        }
        else
        {
          nonSigVar <- c(nonSigVar, vars[j])
        }
      }
      if(length(nonSigVar) > 0)
      {
        for (k in nonSigVar)
        {
          print(paste(cat("t"), "Rechecking non significant variables", k, collapse = " "))
        }
      }
    }
  }
}
```

```

vars2 <- c(sigVar, k)
cfm <- as.formula(paste(resp, "~ ", paste("s(", vars2, ")", collapse = " + ")))
nmd <- gam(cfm, data = indata, family = fam)
nsmry <- summary(nmd)
ndiv <- nsmry$dev.expl
pvalues <- c(nsmry$s.pv)
if(pvalues[length(vars2)] <= alpha)
{
#ndiv <- nsmry$dev.expl
#pdiv <- nsmry$dev.expl
sigVar <- c(sigVar, k)
print(paste(cat("t"), "adding", k, "back to the model", collapse = " "))
}
}
ndiv <- pdiv
if(length(sigVar) > 0)
{
cfm <- as.formula(paste(resp, "~ ", paste("s(", sigVar, ")", collapse = " + ")))
nmd <- gam(cfm, data = indata, family = fam)
nsmry <- summary(nmd)
ndiv <- nsmry$dev.expl
}
if(ndiv < (pdiv + improveby))
{
print(paste(cat("t"), "No improvement. Changing sig variables back to previous model"))
sigVar <- vars[1:length(vars)-1]
}
else
{
pdiv <- ndiv
}
}
else
{
pdiv <- ndiv
}
print(paste(cat("t"), "sig var for iter ", i, "(%Div = ", pdiv, "):",
paste(sigVar, collapse = " ")))
}
}
return(sigVar)
}

```

## References

1. Davis, L.; Johnson, N. *Forest Management*, 3rd ed.; McGraw Hill: New York, NY, USA, 1987; p. 790.
2. Smith, D.; Larson, B.; Kelty, M.; Ashton, M. *The Practice of Silviculture: Applied Forest Ecology*, 9th ed.; John Wiley & Sons: New York, NY, USA, 1997; p. 537.
3. Kimmins, J. *Forest Ecology: A Foundation for Sustainable Management*; Prentice Hall: Upper Saddle River, NJ, USA, 1997; p. 596.
4. Avery, T.; Burkhardt, H. *Forest Measurements*, 4th ed.; McGraw Hill: Boston, MA, USA, 1994; p. 408.

5. White, J.; Coops, N.; Wulder, M.; Varstaranta, M.; Hilker, T.; Tompaški, P. Remote Sensing Technologies for Enhancing forest Inventories: A Review. *Can. J. Remote Sens.* **2016**, *42*, 619–641. [[CrossRef](#)]
6. Rao, J.; Molina, I. *Small Area Estimation*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2015; p. 441.
7. Jensen, J. *Remote Sensing of the Environment: An Earth Resource Perspective*; Prentice Hall: Upper Saddle River, NJ, USA, 2000; p. 544.
8. U.S. Department of Agriculture Forest Service (USFS). *Forest Inventory and Analysis National Core Field Guide: Field Data Collection Procedures for Phase 2 Plots. Version 6.0. Vol. 1*; Internal Report, 2012; U.S. Department of Agriculture Forest: Washington, DC, USA, 2012. Available online: [http://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/2013/Core%20FIA%20P2%20field%20guide\\_6-0\\_6\\_27\\_2013.pdf](http://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/2013/Core%20FIA%20P2%20field%20guide_6-0_6_27_2013.pdf) (accessed on 5 June 2014).
9. Omernik, J.; Griffith, G. Ecoregions of the conterminous United States: Evolution of a hierarchical spatial framework. *Environ. Manag.* **2014**, *54*, 1249–1266. [[CrossRef](#)] [[PubMed](#)]
10. Gesch, D.; Oimoen, M.; Greenlee, S.; Nelson, C.; Steuck, M.; Tyler, D. The National Elevation Dataset. *Photogr. Eng. Remote Sens.* **2002**, *68*, 5–11.
11. Homer, C.; Dewitz, J.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N.; Wickham, J.; Megown, K. Completion of the 2011 National Land Cover Database for the conterminous United States—Representing a decade of land cover change information. *Photogr. Eng. Remote Sens.* **2015**, *81*, 345–354.
12. Gandhi, G.; Parthiban, S.; Thummalu, N.; Christy, A. Ndvi: Vegetation Change Detection Using Remote Sensing and Gis—A Case Study of Vellori District. *Procedia Comput. Sci.* **2015**, *57*, 1199–1210.
13. LANDFIRE. Existing Vegetation Type Layer, LANDFIRE 1.1.0, U.S. Department of the Interior, Geological Survey. 2008. Available online: <http://landfire.cr.usgs.gov/viewer/> (accessed on 28 October 2010).
14. Lowry, J.; Ramsey, R.; Boykin, K.; Bradford, D.; Comer, P.; Falzarano, S.; Kepner, W.; Kirby, J.; Langs, L.; Prior-Magee, J.; et al. *Southwest Regional Gap Analysis Project: Final Report on Land Cover Mapping Methods*; RS/GIS Laboratory, Utah State University: Logan, UT, USA, 2005.
15. Escuin, S.; Navarro, R.; Fernandez, P. Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. *Int. J. Remote Sens.* **2008**, *29*, 1053–1073. [[CrossRef](#)]
16. Davids, C.; Douglis, A. Unsupervised change detection of multitemporal Landsat imagery to identify changes in land cover following the Chernobyl accident. In *Proceedings of the Geoscience and Remote Sensing Symposium, Barcelona, Spain, 2008*; pp. 3486–3489.
17. Weng, Q.; Fu, P.; Gao, F. Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sens. Environ.* **2014**, *145*, 55–67. [[CrossRef](#)]
18. Popescu, S.; Wynne, R.; Scrivani, J. Fusion of Small-footprint Lidar and Multispectral Data to Estimate Plot-Level Volume and Biomass in Deciduous and Pine forests in Virginia, USA. *For. Sci.* **2004**, *50*, 551–565.
19. Naesset, E. Prediction forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **2002**, *80*, 88–99. [[CrossRef](#)]
20. Tomppo, E.; Olsson, H.; Stahl, G.; Nilsson, M.; Hagner, O.; Katila, M. Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sens. Environ.* **2008**, *112*, 1982–1999. [[CrossRef](#)]
21. National Agriculture Imagery Program (NAIP). National Agriculture Imagery Program (NAIP) Information Sheet. 2012. Available online: [http://www.fsa.usda.gov/Internet/FSA\\_File/naip\\_info\\_sheet\\_2013.pdf](http://www.fsa.usda.gov/Internet/FSA_File/naip_info_sheet_2013.pdf) (accessed on 14 May 2014).
22. America's Longleaf. Range-Wide Conservation Plan for Longleaf. Available online: [http://www.americaslongleaf.org/media/86/conservation\\_plan.pdf](http://www.americaslongleaf.org/media/86/conservation_plan.pdf) (accessed on 12 May 2015).
23. Noss, R.; LaRoe, E.; Scott, J. *Endangered Ecosystems of the United States: A Preliminary Assessment of Loss and Degradation*; Biological Report 28; National Biological Service: Washington, DC, USA, 1995.
24. O'Connell, B.; LaPoint, E.; Turner, J.; Riley, T.; Pugh, S.; Wilson, A.; Waddell, K.; Conkling, B. The Forest Inventory and Analysis Database: Database Description and User Guide for Phase 2 (Version 6.0.1). 2014. Available online: <http://www.fia.fs.fed.us/library/database-documentation/> (accessed on 26 May 2015).
25. Forest Inventory and Analysis Database (FIA Database). U.S. Department of Agriculture, Forest Service, Northern Research Station: St. Paul, MN, USA, 2016. Available online: <https://apps.fs.usda.gov/fia/datamart/datamart.html> (accessed on 28 March 2018).

26. Forest Inventory and Analysis Spatial Data Request (FIA SDR) Requesting Spatial Data, 2016. Available online: <http://www.fia.fs.fed.us/tools-data/spatial/requests/index.php> (accessed on 16 September 2016).
27. United States Geological Survey File Transfer Protocol [USGS FTP] Staged NAIP, 2016. Available online: <ftp://rockyftp.cr.usgs.gov/vdelivery/Datasets/Staged/NAIP/> (accessed on 16 September 2016).
28. Environmental Systems Research Institute (ESRI). What Is a Mosaic Raster Dataset? ArcPress. 2015. Available online: <http://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/what-is-a-mosaic-dataset.htm> (accessed on 26 May 2015).
29. Florida Geographic Data Library Documentation (FGDL). Florida (UTM 17) 1-m NAIP Digital Ortho Photo Image. 2015. Available online: [https://www.fla-etat.org/est/metadata/naip\\_utm17\\_2015.htm](https://www.fla-etat.org/est/metadata/naip_utm17_2015.htm) (accessed on 28 March 2018).
30. Hogland, J.; Anderson, N. Function modeling improves the efficiency of spatial modeling using big data from remote sensing. *Big Data Cogn. Comput.* **2017**, *1*, 1–14. [CrossRef]
31. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2014. Available online: <http://www.R-project.org/> (accessed on 12 January 2018).
32. Hogland, J.; Anderson, N. Improved analyses using function datasets and statistical modeling. In Proceedings of the 2014 ESRI Users Conference, San Diego, CA, USA, 14–18 July 2014; Environmental Systems Research Institute: Redlands, CA, USA, 2014. Available online: [https://www.fs.fed.us/rm/pubs\\_other/rmrs\\_2014\\_hogland\\_j001.pdf](https://www.fs.fed.us/rm/pubs_other/rmrs_2014_hogland_j001.pdf) (accessed on 28 March 2018).
33. RMRS. RMRS Raster Utility. 2014. Available online: <http://www.fs.fed.us/rm/raster-utility> (accessed on 3 June 2014).
34. Haralick, R.; Shanmugam, K.; Dinstein, I. Texture features for image classification. *IEEE Trans. Syst. Man. Cybern.* **1973**, *3*, 610–621. [CrossRef]
35. Bishop, C. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006; p. 738.
36. Sergey, B. ALGLIB. 2007. Available online: <http://www.alglib.net/> (accessed on 3 June 2014).
37. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
38. Hogland, J.; Anderson, N. Estimating FIA plot characteristics using NAIP imagery, function modeling, and the RMRS Raster Utility coding library. In *General Technical Report (GTR), Proceedings of the Pushing Boundaries: New Directions in Inventory Techniques and Applications: Forest Inventory and Analysis (FIA) Symposium, Portland, Oregon, 8–10 December 2015*; Stanton, S.M., Christensen, G.A., Eds.; PNW-GTR-931; U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station: Portland, OR, USA, 2015; pp. 340–344.
39. Hogland, J.; St. Peter, J.; Anderson, N. Raster Surfaces Created from the Longleaf Mapping Project. Fort Collins, 2017, CO: Forest Service Research Data Archive. Available online: <https://doi.org/10.2737/RDS-2017-0014/> (accessed on 2 April 2018).
40. Hogland, J. Creating Spatial Probability Distributions for Longleaf Pine Ecosystems across East Mississippi, Alabama, The Panhandle of Florida and West Georgia. Master’s Thesis, Auburn University, Auburn, AL, USA, 15 December 2005.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).