THE EFFECT OF USING COMPLETE AND PARTIAL FORESTED FIA PLOT DATA ON BIOMASS AND FORESTED AREA CLASSIFICATIONS FROM MODIS SATELLITE DATA

Dumitru Salajanu, Forester
Dennis M. Jacobs, Research Forester
USDA Forest Service
Forest Inventory and Analysis
4700 Old Kingston Pike
Knoxville, TN 37919
dsalajanu@fs.fed.us
djacobs@fs.fed.us

ABSTRACT

Authors' objective was to determine at what level biomass and forest area obtained from partial and complete forested plot inventory data compares with forested area and biomass estimates from the national inventory data. A subset of 3819 inventory plots (100% forested, 100% non-forested, mixed-forest/non-forest) was used to classify the land cover and model the biomass in South Carolina. Forest inventory data have been further subdivided into three datasets containing 1) mixed plots at least 50% forest or a 50% non-forest, 2) mixed plots that are at least 75% forested or 75% non-forested, and 3) 100% forested/non-forested plots. Separately, each of these three datasets was used in a decision tree classification process applied to MODIS satellite (250-meter resolution) and ancillary data to classify the land cover and model the forest biomass. The satellite, ancillary, and plot data have been subdivided into three mapping zones (54, 58, and 59) for processing in See5 and Cubist software. Classification results for trials with 100% forested/non-forested and mixed (multi-condition) plots show that overall classification accuracy for the percent of pixels correctly classified (%PCC) increased from 75.4% to 79.2%. Comparison between classified forest area with mixed (75% forest, 75% non-forest) plots and the inventory forest area shows a 10% increase. The forest/non-forest single layer classification from each trial was used to mask out non-forested areas for the forest biomass classification. Accuracy of modeled forest biomass was compared with plot data estimates of forest biomass. Biomass obtained from Cubist models with 100% forested and mixed forest inventory plots when compared to the biomass from the published plot data estimates show a difference of less than 2%.

INTRODUCTION

Satellite remote sensing is an important tool for forest management and for surveying vast areas of forestland. Over the years, forest type classifications have been derived from an assortment of satellite data sensors with a variety of spatial and spectral resolutions. Results varied according to the classification algorithm, site location, number of classes used in each classification, etc.

Early classifications of forest and land cover/use from satellite data were produced based on spectral information in the image. Spectral and spatial resolutions were the primary elements that dictated classification accuracy and what could be achieved (Ma, 1985; Ma and Olson, 1989; Chavez et al., 1991; Salajanu, 1992; Lunetta et al., 1998).

Improvements in technology and classification algorithms allow ancillary data (slope, soil type, vegetation indices, merged information from sensors with different resolutions, etc.) to be resampled and incorporated into the original satellite data as new channels of raster data. Classification accuracy was improved when the original spectral channels were combined with ancillary data as additional channels in the classification process (Ricchetti, 2000; Chavez, 1986; Borry et al, 1990; Pellemans et al, 1993; Vogelmann et al 1998; Salajanu and Olson, 2001). In the last several years, classification and regression tree analysis have been implemented in several software programs and were used in many remote sensing applications (Huang and Jensen, 1997; Lawrence and Wright, 2001; Cooke and Jacobs, 2005).

The inventory design of the Forest Inventory and Analysis National Program of the United States Department of Agriculture Forest Service (FIA) requires annual measurements on a portion of all land in order to form rotating panels. The Southern States strive for a 20% sample each year as part of a 5-year cycle. Forest inventory plot data are used as ancillary information when classifying forest types and forest area, as well as modeling the forest biomass.

The main objective of this study is to determine whether there is significant improvement in the accuracy of forest/non-forest and forest biomass classifications from Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data by incorporating partially forested FIA plot data as opposed to using only complete forested plots (100% forested and 100% non-forested) from a complete 5-year cycle of FIA data.

STUDY AREA

The test site for this study covers the entire state of South Carolina. South Carolina was selected because it was the only state in the South with a complete 5-panel cycle of FIA field data available at commencement of this study. The test site consists of a large diversity of landforms (sandy beaches to coastal plains and hills to low mountains), soil types and land cover/use types. Hardwood forests are mixed throughout the State and are the dominant forest type followed by southern yellow pines (Smith et. al. 2004). The study area includes portions of United States Geological Survey (USGS) mapping zones 54, 57, 58, and 59 that fall within the state of South Carolina.

FOREST INVENTORY AND ANALYSIS PLOTS

The national inventory design of the Forest Inventory and Analysis program requires annual measurements on a proportion of all lands and 5-year reports. The field plot design consists of four subplots approximately 1/24 acre in size, and are used to collect data on trees with a diameter at breast height of 5 inches or greater (Figure 1). Each subplot contains a microplot of approximately 1/300 acre in size. Microplots are used to collect information data on seedlings and saplings.

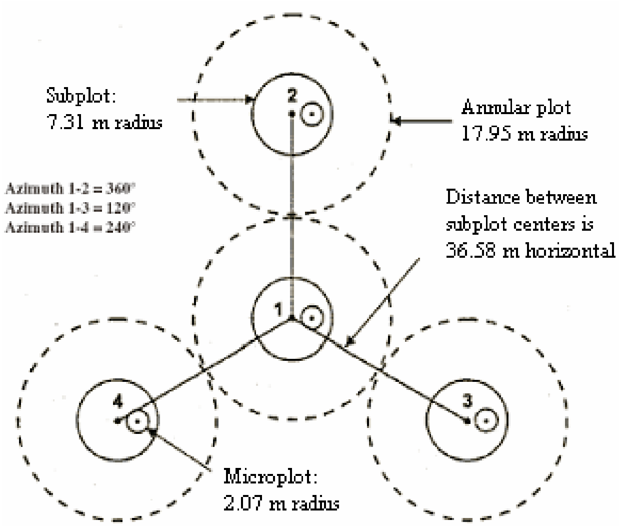


Figure 1. FIA field plot design.

An attempt is made on all forested plots to collect coordinates with a Global Positioning System (GPS) receiver at the center subplot. Some non-forest plots may have GPS coordinates, but most do not. Hence, not all of the plots used in the study have accurate GPS coordinates. For this study a subset of 3819 plots (2601 forested, 1204 non-forested, and 14 water) was filtered from a complete 5-panel dataset of South Carolina. This subset includes all of the 100% forested and 100% non-forested plots, and mixed plots containing both forest and non-forest conditions that are at least 50% forested and non-forested (multi-condition).

DATA BASE DESCRIPTION

The database consists of raster and vector data that fall within portions of the four USGS mapping zones indicated above. Since mapping zone 57 was such a small portion of the State it was merged with zone 54, and the database considered as three mapping zones. Each mapping zone contains 269 layers of data: a large number of ancillary and remote sensing layers re-sampled to a spatial resolution of 250 meters and projected to the Albers Equal Area projection. This set was used for modeling the forest/non-forest and forest biomass classifications. The data layers in Table 1 either existed as 250-meter resolution data or were re-sampled to 250-meters and projected to

Albers Equal Area projection by personnel at the USFS Remote Sensing Applications Center in Salt Lake City (RSAC). The database contains continuous and categorical variables.

Table 1. List of Layers Used to Map Forest, Non-forest and Forest Biomass.

Database Layers Description
MODIS 32 day composite imagery between 2001 and 2003
Conus MODIS32-2001097 - Bands 1 to 7
Conus MODIS32-2001193 - Bands 1 to 7
Conus MODIS32-2002129 - Bands 1 to 7
Conus MODIS32-2002225 - Bands 1 to 7
Conus MODIS32-2002257 - Bands 1 to 7
Conus MODIS32-2002321 - Bands 1 to 7
Conus MODIS32-2003161 - Bands 1 to 7
Conus Bailey's Ecoregions image layer
MODIS Vegetation Indicies Layers
Conus EVI- 2002097 image
Conus EVI- 2002225 image
Conus EVI- 2002321 image
Conus NDVI- 2002097 image
Conus NDVI- 2002225 image
Conus NDVI- 2002321 image
MODIS Vegetation Layer: MODIS –percent tree cover image
Reflectance layers from spring, summer and fall of 2002
Conus Reflectance – 2002097 – Bands 1 to 7
Conus Reflectance – 2002225 – Bands 1 to 7
Conus Reflectance – 2002321 – Bands 1 to 7
NLCD layers;
Conus NLCD – Percent conifer forest image
Conus NLCD – Percent deciduous forest image
Conus NLCD – Percent mixed forest image
Conus NLCD – Percent shrub land image
Conus NLCD – Percent woody wet land image
Terrain information; Conus dominant aspect, Conus mean elevation, stream density
Conus MODIS fire points from 2001 and 2002
Soil data layers; available water capacity, permeability, soil bulk density, soil ph, soil plasticity, soil porosity, rock volume and soil texture.
USGS mapping zone images
Precipitation – annual and for each month
Temperature layers – averages, minimum and maximum temperatures.

CUBIST AND SEE5

Cubist and See5 are regression tree software used to create decision tree classifications (forest/non-forest map) and models for modeling the forest biomass. See5 was used to classify/model categorical variables, forest, non-forest and water, while Cubist was used to model the biomass continuous variable. Two files are essential for running Cubist or See5, and several others are optional. The first essential file is the "names" file that lists the names and describes the classes and attributes/predictors (Table 2).

The first row/entry in the names file is the attribute (forest/non-forest, or biomass) that contains the target value to be classified/modeled based on values of the other predictors. Predictors contained in the names file are labeled as continuous or are discrete and defined by numeric values. The final entry in the names file specifies the list of predictors that are excluded from the classifier/model. The exclusion list can be empty. The second essential file is the data file that provides information on the training data used to construct the decision tree model. The entry for each case consists of one or more lines that give the values for all the predictors. A comma separates the values and the entry terminates with a period (http://www.rulequest.com). The test file is one of the optional files, and it is used to evaluate the performance of the classifier/model. There are several ways for assessing model predictive ability such as: collection of new data to check the model and its predictive ability; comparison of results with earlier empirical results; and use of a hold out sample when using a large data set. In this study FIA plot data have been split randomly 60% and 40% into data training and test files. Test file has the same structure as data file.

Table 2. Types of information contained in the name file

FNF FNF: 1, 2, 3, 4. awc-250m.img-band1: continuous. bdgrid-250m.img-band1: continuous. conus-dvi-2002225.img: continuous. conus-evi-2002097.img: continuous. conus-modis32-2001097-albers.img-band1: continuous. us_ppt01_jan.img: continuous. us_ppt02_feb.img: continuous. us_ppt03_mar.img: continuous. us_tavg301_albers.img: continuous. us_tavg302_albers.img: continuous. us_tavg303_albers.img: continuous. usgs_mapping_zones.img: 0, 54, 58, 59. ustmax01_albers.img: continuous. ustmax02_albers.img: continuous. : : attributes excluded: conus_modis32_2001097_albers.img_band5. conus_modis32_2001193_albers.img_band5.

FOREST NON-FOREST CLASSIFICATION

There are several types of algorithms and methods to classify satellite data, such as supervised and unsupervised classification, neural network, decision tree, etc. The decision tree algorithm in See5 was used for this study. A subset of plots with associated quality location coordinates was pulled from a complete 5-year cycle of FIA data. This dataset was partitioned into the following three subsets of plots for forest/non-forest labeling purposes: 1) plots that were either 100% forest or 100% non-forest, 2) plots that were at least 75% forest or at least 75% non-forest, and 3) plots that were at least 50% forest or at least 50% non-forest. Separately, each of these three datasets was used in an iterative decision tree classification process applied to MODIS satellite data (250 m resolution) and ancillary data to classify the forest, non-forest, and water. See5 cannot process geospatial data such as geographic information system data (GIS) or remote sensing layers in their inherent geospatial format. Prior to the data mining process, satellite, ancillary and plot data for each mapping zone were processed with tools developed at the Remote Sensing Applications Center (RSAC) in Salt Lake City for ERDAS Imagine to convert remote sensing and GIS layers to See5 and Cubist tabular data file formats.

The “Prepare FIA Data for Cubist/See5” tool extracts geospatial image information using FIA points. The program then creates three data files for See5 and three additional files for Cubist (data file, names file, and test file), and randomly selects a dataset to be set aside for accuracy assessment. Once the name, data, and test files have been produced, See5 program is used to create decision tree models. See5 offers several options (e.g., rulesets, boost) to build a decision tree model, and each option produces a different type of classifier/decision tree based on the way it is constructed.

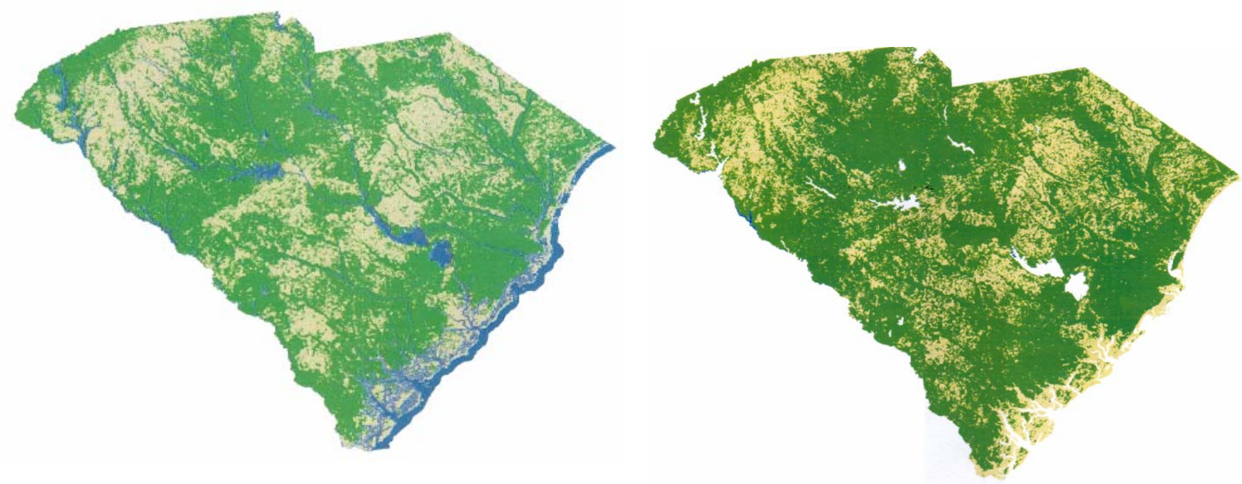
The boosting option set to ten trials was the only one used in this study to model the forest/non-forest categorical variable. The boosting option was selected because it creates several classifiers/decision trees. Each classifier/decision tree produced by the boosting option will be different from the previous. Each decision tree tries to correct the prediction error from the previous decision tree. This process continues for a pre-determined number of trials. The data file from each mapping zone was used in See5 to create forest/non-forest (fnf) decision tree models. Forty percent of the data was set-aside in each data set for accuracy assessment and 60% retained for processing. A sample of the output file from the See5 software program (Table 3) reports classification errors based on a confusion matrix produced for both training and test datasets.

Table 3. Sample of the See5 output showing the misclassifications

```
Options:
    10 boosting trials
Class specified by attribute `fnf'
Trial 9:  Decision tree:
SubTree [S1]
conus_modis32_2001097_albers.img_band3 <= 429: 2 (29)
conus_modis32_2001097_albers.img_band3 > 429: 1 (3.8/0.8)
SubTree [S2]
conus_reflectance_2002097.img_band4 > 2224: 2 (5.8)
conus_reflectance_2002097.img_band4 <= 2224:
:...conus_modis32_2002225_albers.img_band2 <= 2591: 2 (8.4/0.8)
    conus_modis32_2002225_albers.img_band2 > 2591:
    :...conus_modis32_2001097_albers.img_band6 > 2220: 2 (3)
        conus_modis32_2001097_albers.img_band6 <= 2220:
        :...ustmin10_albers.img <= 777: 2 (2.5)
            ustmin10_albers.img > 777:
            :...conus_ndvi_2002321.img <= 8052: 1 (42.2/3.2)
                conus_ndvi_2002321.img > 8052: 2 (2)
Evaluation on training data (695 cases):      Evaluation on test data (462 cases):
Trial     Decision Tree                        Trial     Decision Tree
          Size    Errors                                 Size   Errors
  0       70    21( 3.0%)                      0         70     126(27.3%)
  1       36    81(11.7%)                      1         36     105(22.7%)
  2       44    66( 9.5%)                      2         44     140(30.3%)
  3       51    61( 8.8%)                      3         51     152(32.9%)
  4       50    77(11.1%)                      4         50     137(29.7%)
  5       40    85(12.2%)                      5         40     142(30.7%)
  6       52    74(10.6%)                      6         52     130(28.1%)
  7       48    65( 9.4%)                      7         48     122(26.4%)
  8       48    78(11.2%)                      8         48     135(29.2%)
  9       52    78(11.2%)                      9         52     127(27.5%)
  boost          0( 0.0%)  <<                  boost             86(18.6%)
(a)  (b)   (c)   (d)   (e)  classified as         (a)   (b)   (c)   (d)   (e)  classified as
                            (a): class 0                                       (a): class 0
     477                    (b): class 1                301    30              (b): class 1
           215              (c): class 2                 53    73           1  (c): class 2
                            (d): class 3                                       (d): class 3
                        3   (e): class 4                  2                 2  (e): class 4
```

The “Apply See5 Results Spatially” tool developed for ERDAS Imagine by RSAC was used to create a spatial forest/non-forest data layer from the See5 decision tree models. The classification tree obtained from boosting was used in Apply See5 software module to model forest, non-forest and water classes as a function of the modeling dataset in each mapping zone. The final product is a single layer forest/non-forest image map (predicted output image) with values representing the variables (forest/non-forest) that were modeled (Figure 2) and a confidence image that shows spatial distribution of the correct and misclassified areas. Confidence values range from zero to one. A value of or near one indicates a more confident prediction for forest area, while values approaching zero show an increase in confidence prediction for non-forest area.

Pixels classified as forested have been converted to hectares and total forested area was compared to the total forestland area (U.S. Survey acres converted to hectares) reported in the Forest Resources of the US, 2002 report (Smith et. al. 2004).



a) b)

Figure 2. Forest non-forest MODIS classification modeled with FIA plot inventory data.
a) using 100% forest and 100% non-forest; b) using ≥50% forest and non-forest FIA plot data.

BIOMASS CLASSIFICATION

The procedure for preparing the data for Cubist and classifying/modeling the forest biomass is similar to the forest/non-forest classification procedures for See5. Cubist, similar to See5, can process a large volume of data but cannot read remote sensing and GIS data layers. The RSAC-developed tools within ERDAS Imagine were used to convert remote sensing image layers to Cubist data files. Before modeling forest biomass, a forest mask was produced for each mapping zone. Classified forest maps have been used to mask non-forested area (Figure 3). Plots with location coordinate falling outside the forest area mask were eliminated from the biomass modeling datasets.
Forest biomass estimates (total dry weight) from FIA plot data and hundreds of continuous predictor layers were used in Cubist to produce biomass predictor models. Cubist, like See5, offers several options (rules alone, let Cubist decide, etc.) to build decision tree models. A model consists of a collection of rules. Two of the available options in Cubist were used to produce decision tree biomass models – “rule alone” and “committee of 5 members.” Committee option, like boosting in See5, creates several rule-based models. Each member of the committee predicted a value for a class and the members’ predictions have been averaged into a final prediction. There were five committee members and each member of a committee model tries to correct the predictions of the previous member (www.rulequest.com). A biomass model was produced for each mapping zone using the complete subset of 100% forested/non-forested and one model for each subset of mixed (50% threshold, or 75% threshold) FIA plots. Again, for each data set, a random sample of 40% of the data was set aside for

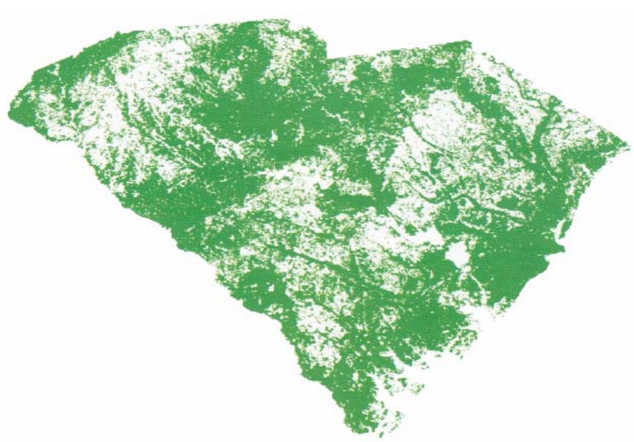


Figure 3. Forest mask for modeling forest biomass.

accuracy assessment and the remaining 60% of the data was used to build the model. For each mapping zone, several iterations of decision tree biomass models were performed and analyzed. With each step, predictor layers poorly correlated with the biomass estimates were excluded during the next iteration. The Cubist output file (decision tree model) reported the errors (average and relative error), and the correlation coefficient for both training and test data sets.

Forest biomass models from each mapping zone were used in the ERDAS Imagine tool “Apply Cubist Results Spatially” to create a spatial biomass image map (predicted image) with predicted values representing the biomass variable and an error image file showing the predicted misclassifications.

RESULTS AND DISCUSSIONS

Forest/non-forest land cover was classified for the entire state of South Carolina from multitemporal MODIS satellite data (acquired during the spring, summer, and fall of 2001, 2002, 2003), ancillary data layers, and FIA plot data. Single and multiple condition plots (greater than or equal to 50% forested/non-forested condition), and the See5 option of boosting with 10 trials were used to classify the land cover for South Carolina into forest, non-forest and water. Forest/non-forest classifications (from See5) showed an increase in overall classification accuracy (%PCC – percent of pixels correctly classified) within each mapping zone when the mixed condition plots were included rather than modeling with only the 100%-single-condition plots. Accuracy assessment is based on analysis of a contingency table produced by the See5 program for the 40-percent set-aside dataset. The overall classification accuracy (%PCC) varied in each mapping zone with the degree of land fragmentation and the types of FIA plot data (single condition - 100% forest/non-forest, multi-condition forest/nonforest) used in classification. Table 3 summarizes the producer, user, and overall accuracy (%PCC) for each mapping zone, and for the entire State obtained from classifications with each of the three sets of plots: 100% forested/non-forested single-condition plots, the single- and multi-condition plots that are at least 75% forested/non-forested, and the single- and multi-condition plots that are at least 50% forested/non-forested.

The lowest and highest accuracies occurred in the smallest zone (59), where the low accuracy of 63.2% was associated with plots having 50- to-100% forest or non-forest condition, and the high accuracy (84.5%) associated with the 100%-single-condition plots. At the state level, overall forest/non-forest accuracy (%PCC) increases from 75.4%, to 79.2%. Mapping zone 58 has lower accuracy values compared to mapping zones 54 and 59. This is explained by the fact that zone 58 contains the highest diversity of land cover in the entire state (sand beach, wetlands, small lakes and estuaries, forested wetlands, agriculture, etc.). Both producer and user accuracy show a much higher classification accuracy of forest class (in each mapping zone, and state level) than for non-forest and water classes, especially when multi-condition plots were used in the model.

Table 3. Classification accuracy of forest/non-forest by mapping zone

Mapping zone #	Percent of Forest/nonforest	Overall %PCC	Producer accuracy %			User accuracy %		
			Forest	Non-forest	Water	Forest	Non-forest	Water
Zone 54	≥50% forest/non-forest	81.38	90.94	59.35	50.00	84.55	70.87	66.67
	≥75% forest/non-forest	78.73	90.48	55.22	50.00	81.52	73.27	20.00
	100% forest/non-forest	80.70	88.07	70.00	33.33	82.45	77.08	100.00
Zone 58	≥50% forest/non-forest	73.43	86.27	47.92	100.00	76.57	63.89	100.00
	≥75% forest/non-forest	73.83	85.52	51.87	100.00	76.74	66.13	100.00
	100% forest/non-forest	78.22	82.69	71.68	66.67	81.13	73.80	66.67
Zone 59	≥50% forest/non-forest	63.16	83.33	29.63		66.67	50.00	
	≥75% forest/non-forest	78.95	87.50	66.67		80.77	78.26	
	100% forest/non-forest	84.48	91.43	72.73	100.00	84.21	84.21	100.00
State-South Carolina	≥50% forest/non-forest	76.88	88.00	52.87	66.67	80.40	66.84	50.00
	≥75% forest/non-forest	75.39	84.31	58.07	100.00	79.73	65.64	33.33
	100% forest/non-forest	79.25	84.69	71.25	57.14	81.69	75.13	80.00

Another way to assess classification accuracy of forest/non-forest from FIA data is to determine the total area in hectares of all the pixels classified as forest and compare that total with the total forested area (U.S. Survey acres converted to hectares) reported in the Forest Resources of the United States, 2002 report (Smith et. al., 2004). This publication is an updated report required by the Forest and Rangeland Renewable Resources Planning Act of 1974 (RPA). The results are summarized in Table 4.

Table 4. Comparison of forested area from MODIS classification with inventory forested area

Percent Forest	Classified as Forest (ha)	RPA Forest (ha)	Difference (+/-)	Percent (%)
≥50% forest/non-forest	5,627,527	5,024,380	+ 603,147	12.00
≥75% forest/non-forest	5,558,611	5,024,380	+ 534,231	10.63
100% forest/non-forest	5,015,420	5,024,380	- 8,960	0.18

At the State-level, map-based estimates of forest/non-forest area compare very well with RPA forested area when single-condition-only plots (100% forest, 100% non-forest) are used in the model. Forest/non-forest classifications obtained from multi-condition plots (≥75percentage or ≥50% forest/non-forest) show an increase in forest area of 10.6% and 12.0% respectively, when compared to RPA forested area. Analysis of producer and user classification accuracy by mapping zones and at the state level suggest that more pixels modeled by multi-condition plots (≥50% forest, ≥50% non-forest) that have plot center in the non-forest area are classified as forest then non-forest. It is important to note that 85% of the multi-condition FIA plots have their plot center (GPS coordinate) in the forested condition and only 15% have their plot center in the non-forested condition.

A forest/non-forest classification map within each mapping zone was used to mask out the non-forested area and retain only the forested area for further work. This forest area mask was then used as the area over which forest biomass was modeled with the same geospatial predictors. Models for each mapping zone have been applied to their corresponding zonal predictor dataset to produce forest biomass predictions on a pixel-by-pixel basis. A mosaic of the three mapping zones produced the biomass map shown in Figure 4. The decision tree models developed to model forest biomass varied in their ability to predict plot level biomass and the results show variation of the total biomass in each mapping zone as well as variations due to the status of FIA plots (single or multi-condition plots) used in the model. Biomass trend shows an increase of forest biomass and accompanies the increase in the number of FIA plots in the model, but it is not always the case since Cubist randomly sets aside a different set of plots for the test file in each trial.

An accuracy assessment of the biomass results obtained from Cubist models (single- and multi-condition FIA plots) and comparison with RPA biomass is shown in Table 5. Predicted biomass compared well with RPA.

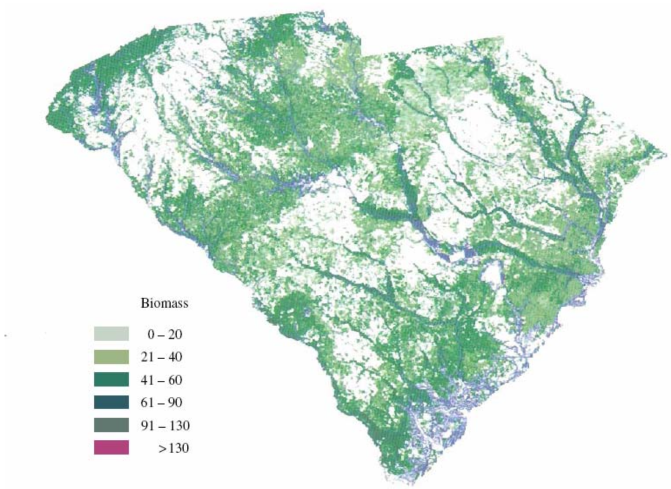


Figure 4. Predicted biomass for South Carolina (dry tons/ac).

Table 5. Comparison of modeled biomass with RPA biomass (total dry weight)

Percent of forest or non-forest in the plot	Classified/modeled biomass (tons)	Timberland biomass (tons)	Error %	Forest land biomass (tons)	Error %
≥50% Forest, ≥50% Non-forest	472,940,000	458,128,000	3.23	463,836,000	1.96
≥75% Forest, ≥75% Non-forest	472,719,000	458,128,000	3.18	463,836,000	1.91
100% Forest, 100% Non-forest	456,202,000	458,128,000	0.42	463,836,000	1.64

Results in table 5 show less than 2% difference between modeled biomass and RPA biomass even though the forest area increased by 10 to 12% when mixed (multi-condition) plots were included in the model. This is because the biomass value used to model the biomass for the same pixel that was classified as forest by a mixed condition plot (even though not the entire pixel was forested), represents only the biomass of the forested part of the same plot in the pixel.

Biomass classification accuracy can be evaluated by analysis of the average error, relative error, and the correlation coefficient produced by Cubist models for each mapping zone (Table 6). High values of relative errors correspond to areas with a high proportion of forest fragmentation that are hard to classify with MODIS (250 m pixel ground resolution).

Table 6. Biomass classification accuracies

Mapping Zone	Percent of forest or non-forest in the plot	Number of Test Plots	Average Error	Relative Error	Correlation Coefficient
State - South Carolina	50% forest, 50%non-forest	995	23.0078	0.91	0.37
	75% forest, 75%non-forest	868	22.8514	0.89	0.38

An analysis of the predicted biomass values show that Cubist models have overestimated the biomass for FIA plots containing low amounts of biomass, and underestimated biomass for FIA plots with high biomass values.

CONCLUSIONS

This dataset is a part of a product that was developed with the intention of using a full five-year cycle of FIA data. That luxury is currently not available in neighboring states, thus providing partial cycles of data across state lines within each mapping zone. Accuracy of the forest/non-forest map is a very important factor when modeling the correct area for forest biomass.

Based on the work described, it can be concluded that mixed (multi-condition) FIA plot information increased total area of forest/non-forest classifications. The study also suggests that forest biomass can be modeled with comparable results when mixed (multi-condition) FIA plots are used. Biomass at the state level provides information on how forest biomass is spatially distributed throughout the entire statewide landscape. The spatial pattern allows a small-area visual assessment of biomass distribution to help show areas of high to low forest biomass; finer detail than the usual county-level minimum-mapping unit associated with FIA tabular data.

FIA plot information ties See5 and Cubist models to actual FIA phase 2 plot measurements on the ground.

Even though the increases in accuracy were not too large, they are meaningful when used in different forest applications. Results suggest that FIA plot information can be used with good results in classifying large areas of land cover.

REFERENCES

Borry, F. C., B. P DeRoover, R. R. DeWulf, and R. E. Goossens (1990). Assessing the value of monotemporal SPOT-1 imagery for forestry applications under Flemish conditions. Photogrammetric Engineering and Remote Sensing, 56(8):114–1153.

Chavez, P .S., Jr. (1986). Digital merging of Landsat TM and digitized NHAP data for 1:24,000 scale image mapping. Photogrammetric Engineering and Remote Sensing, 52(10):1637-1646.

Chavez, P. S Jr., S. C. Sides, and J. A. Anderson (1991). Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic. Photogrammetric Engineering and Remote Sensing, 57(3):295–303.

Conner, R. C., T. Adams, B. J. Butler, W. A. Bechtold, T. G. Johnson, S. N. Oswalt, G. Smith, S. W. Wolf, and C.W. Woodall (2004). The state of South Carolina's forests, 2001, pp. 67.

Cooke, W. H., D. M. Jacobs (2005). Rapid classification of Landsat TM imagery for phase 1 stratification using the automated NDVI threshold supervised classification (ANTSC) methodology. In: Proceedings of the fourth annual forest inventory and analysis symposium, 2002 November 19–21 New Orleans, LA. Gen. Tech. Rep. NC-252. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station, pp. 81-86.

Guo, X., K. P. Price, and J. Stiles (2003). Grasslands discriminant analysis using Landsat TM single and multitemporal data. Photogrammetric Engineering and Remote Sensing, 69(11):1255–1262.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan (2004). Development of a 2001 national land-cover database for the United States. Photogrammetric Engineering and Remote Sensing, 70(7):829–840.

Lawrence, R. L., and A. Wright (2001). Rule-based classification system using classification and regression tree (CART) analysis. Photogrammetric Engineering and Remote Sensing, 67(10):1137–1142.

Lunetta, R.S., J. G. Lyon, B. Guindon, and C. D. Elvidge (1998). North American landscape characterization dataset development and data fusion issues. Photogrammetric Engineering and Remote Sensing, 64(8):821-829.

Ma, Z., and C. E. Olson Jr. (1989). A measurement of spectral overlap among cover types. Photogrammetric Engineering and Remote Sensing, 55(10):1441-1444.

Pellemans, A. H. J. M., R. W. L. Jordans, and R. Allewijn (1993). Merging multispectral and panchromatic SPOT images with respect to the radiometric properties of the sensor. Photogrammetric Engineering and Remote Sensing, 59(1):8 –87.

Ricchetti, E. (2000). Multispectral satellite image and ancillary data integration for geological classification. Photogrammetric Engineering and Remote Sensing, 66(4):429–435.

Salajanu, D. (1992). Utilization of SPOT panchromatic data for forest cover/land use classification. A Master's Thesis in School of Natural Resources, The University of Michigan, Ann Arbor, MI., pp. 88.

Salajanu, D., and C. E. Olson Jr. (2001). The significance of spatial resolution: Identifying forest cover from satellite data. Journal of Forestry, 99(6):32-38.

Skidmore, A. K., B. J. Turner, and E. Knowles (1997). Performance of a neural network: Mapping forests using GIS and remotely sensed data. Photogrammetric Engineering and Remote Sensing, 63(5):501-514.

Smith, W. B., P. D. Miles, J. S. Vissage, and S. A. Pugh (2004). Forest resources of the United States, 2002. Gen. Tech. Rep. NS-241. St. Paul MN: U.S. Department of Agriculture, Forest Service, North Central Research Station, pp. 137.

Volgemann, J. E., T. Sohl, and S. M. Howard (1998). Regional characterization of land cover using multiple sources of data. Photogrammetric Engineering and Remote Sensing, 64(1):45-57.

Zambon, M., R. Lawrence, A. Bunn, and S. Powell (2006). Effect of Alternative Splitting Rules on Image Processing Using Classification Tree Analysis. Photogrammetric Engineering and Remote Sensing, 72(1):25-30.

Wayman, J. P., R. H. Wynne, J. A. Scrivani, and G. A. Reams (2001). Landsat TM-based forest area estimation using iterative guided spectral class rejection. Photogrammetric Engineering and Remote Sensing, 67(10):1155–1166.