

United States
Department of
Agriculture



Forest Service
Southern Research Station

e-General Technical
Report 157

May 2012

Monitoring Across Borders: 2010 Joint Meeting of the Forest Inventory and Analysis (FIA) Symposium and the Southern Mensurationists



Editors:

Will McWilliams and Francis A. Roesch

Disclaimer:

The use of trade or firm names in this publication is for reader information and does not imply endorsement by the U.S. Department of Agriculture of any product or service.

May 2012

Southern Research Station
200 W.T. Weaver Blvd.
Asheville, NC 28804

These proceedings represent the range of topics covered during the 2010 Joint Meeting of the Forest Inventory and Analysis (FIA) Symposium and the Southern Mensurationists, October 5-7, 2010 in Knoxville, TN. The meeting was a gathering of forest scientists with a quantitative leaning and, as such, the papers discuss the aspects of the observation, estimation, modeling and monitoring of forest resources that are of contemporary interest. Papers included in this publication have been sorted into a number of general topic areas. Those areas include International Forest Monitoring, Biometrics, Forest Ecosystems, Forest Health, Data Integrity, Cover Estimation, and Carbon and Biomass.

Keywords: statistics, estimation, sampling, modeling, remote sensing, forest health, data integrity, environmental monitoring, cover estimation, international forest monitoring.

Papers published in these proceedings were submitted by authors in electronic media. Some editing was done to ensure a consistent format. Authors are responsible for content and accuracy of their individual papers and the quality of illustrative materials and equations.

TABLE OF CONTENTS

International Forest Monitoring	1
HOW IS FIA HELPING OTHER COUNTRIES MONITOR THEIR FORESTS? Charles T. Scott	3
CURRENT FOREST AND WOODLAND CARBON STORAGE AND FLUX IN CALIFORNIA: AN ESTIMATE FOR THE 2010 STATEWIDE ASSESSMENT Timothy A. Robards	7
EVALUATING THE COMPATIBILITY OF AMERICAN AND MEXICAN NATIONAL FOREST INVENTORY DATA Todd A. Schroeder, Sean P. Healey, and Gretchen G. Moisen	15
Biometrics	23
MODELING ALASKA BOREAL FORESTS WITH A CONTROLLED TREND SURFACE APPROACH Mo Zhou and Jingjing Liang	25
CURIOUS OR SPURIOUS CORRELATIONS WITHIN A NATIONAL-SCALE FOREST INVENTORY? Christopher W. Woodall and James A. Westfall	39
ESTIMATING TREE CROWN WIDTHS FOR THE PRIMARY ACADIAN SPECIES IN MAINE Matthew B. Russell and Aaron R. Weiskittel	45
AN EVALUATION OF THE PROPERTIES OF THE VARIANCE ESTIMATOR USED BY FIA John P. Brown and James A. Westfall	53
A MULTIVARIATE MIXED MODEL SYSTEM FOR WOOD SPECIFIC GRAVITY AND MOISTURE CONTENT OF PLANTED LOBLOLLY PINE STANDS IN THE SOUTHERN UNITED STATES Finto Antony, Laurence R. Schimleck, Alex Clark III and Richard F. Daniels	59
Forest Ecosystems	65
MODELING FOREST ECOSYSTEM CHANGES RESULTING FROM SURFACE COAL MINING IN WEST VIRGINIA John Brown, Andrew J. Lister, Mary Ann Fajvan, Bonnie Ruefenacht, and Christine Mazzarella.....	67
A PRELIMINARY TEST OF AN ECOLOGICAL CLASSIFICATION SYSTEM FOR THE OCONEE NATIONAL FOREST USING FOREST INVENTORY AND ANALYSIS DATA W. Henry McNab, Ronald B. Stephens, Richard D. Rightmyer, Erika M. Mavity, and Samuel G. Lambert.....	77
CHANGES IN EARLY-SUCCESSIONAL HARDWOOD FOREST AREA IN FOUR BIRD CONSERVATION REGIONS ACROSS FOUR DECADES Sonja N. Oswald, Kathleen E. Franzreb, and David A. Buehler.....	87
RELATIONSHIPS BETWEEN HARVEST OF AMERICAN GINSENG AND HARDWOOD TIMBER PRODUCTION Stephen P. Prisley, James Chamberlain, and Michael McGuffin.....	95

Forest Health **103**

QUANTIFYING CHANGE IN RIPARIAN ASH FORESTS FOLLOWING THE INTRODUCTION OF EAB IN MICHIGAN AND INDIANA
Susan J. Crocker and Dacia M. Meneguzzo 105

USING INVENTORY DATA TO DETERMINE THE IMPACT OF DROUGHT ON TREE MORTALITY
Greg C. Liknes, Christopher W. Woodall, and Charles. H. Perry 109

POTENTIAL IMPACTS OF YEAR-ROUND SAMPLING ON MONITORING PRESENCE-ABSENCE OF INVASIVE FLORA IN THE SOUTHERN UNITED STATES
Christopher M. Oswald, Sonja N. Oswald, W. Keith Moser 113

WHITE ASH (*Fraxinus americana*) HEALTH IN THE ALLEGHENY PLATEAU REGION, PENNSYLVANIA: EVALUATING THE RELATIONSHIP BETWEEN FIA PHASE 3 CROWN VARIABLES AND A CATEGORICAL RATING SYSTEM
Alejandro A. Royo, Kathleen S. Knight, Jamie M. Himes, and Ashley N. Will 123

ESTIMATION OF INVASIVE PROBABILITY OF MULTIFLORA ROSA IN THE UPPER MIDWEST
Weiming Yu, Zhaofei Fan, W. K. Moser, M. H. Hansen and M. D. Nelson 131

Data Integrity **137**

A NATIONAL ANALYTICAL QUALITY ASSURANCE PROGRAM: DEVELOPING GUIDELINES AND ANALYTICAL TOOLS FOR THE FOREST INVENTORY AND ANALYSIS PROGRAM
Phyllis C. Adams and Glenn A. Christensen 139

RESOLVING THE PULPWOOD CANVASS WITH INVENTORY HARVEST INFORMATION
Joseph M. McCollum and Tony G. Johnson 147

ALGORITHMIC DECISION RULES FOR ESTIMATING GROWTH, REMOVALS, AND MORTALITY WITHIN A NATIONAL-SCALE FOREST INVENTORY (USA)
William H. McWilliams, Carol L. Alerich, William A. Bechtold, Mark Hansen, Christopher M. Oswald, Mike Thompson, and Jeff Turner 155

APPLICATION OF AN ASSESSMENT PROTOCOL TO EXTENSIVE SPECIES AND TOTAL BASAL AREA PER ACRE DATASETS FOR THE EASTERN COTERMINOUS UNITED STATES
Rachel Riemann, Ty Wilson, and Andrew Lister 161

A DATABASE STRATEGY FOR NEW VARIABLES
B. Tyler Wilson, Ali Conner, Glenn Christensen, John Shaw, Jason Meade, and Larry Royer 175

Cover Estimation **179**

POTENTIAL APPLICATIONS OF PREFIELD LAND USE AND CANOPY COVER DATA: EXAMPLES FROM NONFOREST AND NONSAMPLED FOREST INVENTORY PLOTS
Sara A. Goeking 181

REPEATABILITY IN PHOTO-INTERPRETATION OF TREE CANOPY COVER AND ITS EFFECT ON PREDICTIVE MAPPING
Thomas A. Jackson, Gretchen G. Moisen, Paul L. Patterson, and John Tipton 189

A FRAMEWORK FOR REPORTING TREE COVER ATTRIBUTES IN AGRICULTURAL LANDSCAPES Dacia M. Meneguzzo and Greg C. Liknes	193
CHOOSING APPROPRIATE SUBPOPULATIONS FOR MODELING TREE CANOPY COVER NATIONWIDE Gretchen G. Moisen, John W. Coulston, Barry T. Wilson, Warren B. Cohen, and Mark V. Finco	195
SAMPLING INTENSITY AND NORMALIZATIONS: EXPLORING COST-DRIVING FACTORS IN NATIONWIDE MAPPING OF TREE CANOPY COVER John Tipton, Gretchen Moisen, Paul Patterson, Thomas A. Jackson, and John Coulston	201
ASSESSING ALTERNATIVE MEASURES OF TREE CANOPY COVER: PHOTO-INTERPRETED NAIP AND GROUND-BASED ESTIMATES Chris Toney, Greg Liknes, Andy Lister, and Dacia Meneguzzo	209
A TOOL TO DETERMINE CROWN AND PLOT CANOPY TRANSPARENCY FOR FOREST INVENTORY AND ANALYSIS PHASE 3 PLOTS USING DIGITAL PHOTOGRAPHS Matthew F. Winn and Philip A. Araman	217
AN ALTERNATIVE METHOD FOR ESTIMATING CROWN CHARACTERISTICS OF URBAN TREES USING DIGITAL PHOTOGRAPHS Matthew F. Winn and Philip A. Araman	223
COMPARISON OF LIDAR- AND PHOTOINTERPRETATION-BASED ESTIMATES OF CANOPY COVER Demetrios Gatziolis	231
COMPARING ALTERNATIVE TREE CANOPY COVER ESTIMATES DERIVED FROM DIGITAL AERIAL PHOTOGRAPHY AND FIELD-BASED ASSESSMENTS Tracey S. Frescino and Gretchen G. Moisen	237
Carbon and Biomass	245
ASSESSING THE ACCURACY OF CROWN BIOMASS EQUATIONS FOR THE MAJOR COMMERCIAL SPECIES OF THE INTERIOR NORTHWEST: STUDY PLAN AND PRELIMINARY RESULTS David L.R. Affleck and Brian R. Turnquist	247
FOREST BIOMASS SUPPLY FOR BIOENERGY IN THE SOUTHEAST: EVALUATING ASSESSMENT SCALE Christopher S. Galik and Robert C. Abt	255
LONG-TERM SIMULATIONS OF FOREST MANAGEMENT IMPACTS ON CARBON STORAGE FROM LOBLOLLY PINE PLANTATIONS IN THE SOUTHERN U.S. Huei-Jin Wang, Philip J. Radtke, and Stephen P. Prisley	265
STAND DENSITY INDEX AS A TOOL TO ASSESS THE MAXIMIZATION OF FOREST CARBON AND BIOMASS Christopher W. Woodall, Anthony W. D'Amato, John B. Bradford, and Andrew O. Finley	293
THE ZERO INFLATION OF STANDING DEAD TREE CARBON STOCKS Christopher W. Woodall and David W. MacFarlane	297

International Forest Monitoring

HOW IS FIA HELPING OTHER COUNTRIES MONITOR THEIR FORESTS? Charles T. Scott	3
CURRENT FOREST AND WOODLAND CARBON STORAGE AND FLUX IN CALIFORNIA: AN ESTIMATE FOR THE 2010 STATEWIDE ASSESSMENT Timothy A. Robards	7
EVALUATING THE COMPATIBILITY OF AMERICAN AND MEXICAN NATIONAL FOREST INVENTORY DATA Todd A. Schroeder, Sean P. Healey, and Gretchen G. Moisen	15

HOW IS FIA HELPING OTHER COUNTRIES MONITOR THEIR FORESTS?

Charles T. Scott

ABSTRACT

The demand for forest monitoring is growing rapidly with emphasis on carbon dynamics, due in part by incentives being negotiated under the United Nation's Reducing Emissions from Deforestation and Forest Degradation (REDD+) process. While much of the temperate and boreal forest in developed countries is being monitored as part of national forest inventories, tropical forests are the least monitored and most at risk of deforestation. The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service is working through its National Inventory and Monitoring Applications Center (NIMAC) to coordinate technical assistance in forest monitoring for other countries. NIMAC has developed a 15-step approach and an inventory planning and design tool. Examples from Honduras, Peru, Guyana, and the Democratic Republic of the Congo are given. Challenges remain in the areas of efficient plot configurations and sampling designs for remote areas with high biodiversity and in maximizing the use of remote sensing to enhance ground-based estimation.

INTRODUCTION

The demand for forest monitoring is growing rapidly around the globe with an emphasis on carbon dynamics. This is driven in large part by incentives being negotiated under the United Nation's Reducing Emissions from Deforestation and Forest Degradation (REDD+) process. Others reasons include sustainable forest management for timber products, biodiversity conservation, wildlife habitat, and other ecosystem services. The United States government is interested because REDD+ is a cost-effective opportunity for reducing carbon emissions while providing for sustainable use of forests to help keep them as forests. While much of the temperate and boreal forests in developed countries are being monitored as part of national forest inventories (NFI), tropical forests are the least well monitored and most at risk of deforestation.

WHAT IS FIA DOING ABOUT IT?

The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service is working through its National Inventory and Monitoring Applications Center (NIMAC) to coordinate technical assistance in forest monitoring to other countries. Requests for assistance come through the U.S. Department of State's Agency for International Development and International Programs of the U.S. Forest

Service. NIMAC identifies individuals from within and outside NIMAC with the skills needed for each project, thus improving efficiency and providing more consistent assistance. NIMAC staff has also developed documentation and software tools to enhance the quality of the advice and delivery of technical assistance in a collaborative manner.

WHAT IS NIMAC DOING?

NIMAC takes a 15-step approach (Scott 2009) that starts by determining each country's objectives and monitoring needs. This focuses the effort on the goals and ensures that each step supports achievement of the goals. To facilitate the first seven steps, NIMAC developed an inventory planning and design tool called the Design Tool for Inventory and Monitoring (DTIM). NIMAC also offers advice on the use of other FIA tools, such as data recorder software, databases, and data analysis. While some in-country assistance is critical, FIA has adopted a capacity-building approach by providing training that can be offered in the United States, the host country, as multi-country workshops, as webinars, or as online training tools.

The 15 monitoring steps are grouped into four categories:

- A. Design phase
- B. Data collection
- C. Database and data processing
- D. Analysis, reporting, and decisionmaking

DESIGN PHASE

The first seven steps focus on planning and designing the inventory. DTIM was developed to lead users through these steps. DTIM is continually being updated; a Spanish language version is now available.

1. *Identify stakeholders and their objectives.* Typically there are national and international stakeholders. Objectives often include measuring (or assessing) carbon sequestration, sustainable timber production, and biodiversity

conservation. DTIM provides a list of common objectives from which to choose.

2. Identify key monitoring questions for each objective.

Develop questions that will determine whether the chosen objectives are being met. This is often challenging, so DTIM provides a list of standard questions that are linked to the list of common objectives from which to choose.

3. Identify attributes needed to answer each question.

This is often best done by identifying the table of estimates needed to answer the question, such as estimating area by forest type and stage of development class when the objective is to restore forest composition and structure. DTIM presents a list of FIA attributes (or metrics) that could be used to answer each of the standard questions.

4. Evaluate existing data to determine if they are sufficient for answering the questions.

Identify available data sources for the area of interest. Determine whether data cover the entire area of interest, include all the required attributes, and have sufficient precision to answer the questions. If so, skip to the analysis phase. If not, then determine what is needed to fill the data gap. Typically, there are only data for parts of tropical countries and are rarely from remeasured plots.

5. Determine the precision required and funds available for the inventory.

For each of the key variables (e.g., forest area, area change, totals for each of the forest carbon pools, and total commercial volume), determine the confidence level (e.g., 90 percent) and confidence intervals (e.g., ± 15 percent). The decisionmaker must assess the risk of making an incorrect decision based on imprecise estimates. REDD+ payments are to be linked to the reliability of the estimates. Precision requirements are difficult to establish; cost limits are usually easier to determine, so the inventory planner can focus on optimizing precision for inventorying a suite of variables for a predetermined funding level. By specifying both in advance, the final requirements can be established iteratively. In many cases, external matching funds are being provided to begin the monitoring project.

6. Determine the optimal plot design based on the requirements and information needs, then determine the sampling design, including sample sizes.

For long-term monitoring (remeasurements), we recommend using proportional allocation to strata. Since strata boundaries tend to change over time, the selection probabilities can become very complicated to determine. The subject of plot design is ripe for further research in the tropics, but solutions usually involve large plots (e.g., 0.5 ha) due to high biodiversity and remoteness, and plots are rectangular (e.g., 20 x 250 m) due to difficulty of determining border trees.

7. Identify the plot locations. For new inventories, we recommend a spatially balanced design such as FIA's hexagonal grid (Reams 2009) or NIMAC's space-filling curve approach (Lister and Scott 2009). This method was utilized by NIMAC in the broadleaf inventory of the Rio Plátano region in Honduras.

DATA COLLECTION

8. Develop a field guide. We encourage the use of existing guides (or some of the attributes they contain) from within the country, nearby countries, or developed for international use, such as by FAO (United Nations Food and Agriculture Organization). Collect the targeted attributes in as consistent a way as possible for upward reporting purposes. An important part of this process is developing definitions for each variable, such as land use and land cover classes, and forest land. FIA's field guides are available at: <http://www.fia.fs.fed.us/library/field-guides-methods-proc/>

9. Plan the field logistics and provide training. Field logistics is an issue that is usually best left to local expertise and to international organizations, such as FAO's National Forest Monitoring and Assessment group, who routinely do this kind of work. The tropics present many logistical challenges so the solutions are different. For example, FIA uses one to two person field crews. But crews in the tropics typically range from 5 to 14 people, often including one or two trained foresters, someone skilled at identifying tree species, one or two crew members who wield machetes to clear the way, and sometimes a cook and someone to care for the crew's health. FIA has cooperated in conducting field training, including field safety.

10. Conduct data collection. With planning, preparation, and training completed, the field work begins. We encourage the use of portable data recorders (PDRs) to collect the data in the field. Since FIA's MIDAS (Mobile Integrated Data Acquisition System) program is very specific to FIA, NIMAC is working with Honduras on the development of a general data entry program called SIBP2. We also recommend performing Quality Control and Quality Assurance checks. While revisiting plots to check on data quality is expensive, not checking can be even more expensive, because forest management decisions could be based on inaccurate data.

DATABASE AND DATA PROCESSING

11. Store data in a database. Upload the data from the PDRs or enter the data from tally sheets, and store in a relational database. Run edit checks on the data – data validation and

crosschecks. If using a PDR, the edits should have been done during data entry in the field where the corrections are most easily and accurately made.

12. *Use models and formulas to estimate computed fields.* Once the data are clean, then computed variables, such as basal area, volume, biomass and carbon, are calculated and added to the database along with the field data. For countries in the REDD readiness phase, this can be a challenge since the carbon pool models often do not exist and no credit will be given when models from outside the region are used. Any stratification used (either prior to plot selection or afterwards) must be assigned to individual plots and the stratum weights stored in the database. These stratifications are typically obtained from classification of satellite imagery.

ANALYSIS, REPORTING AND DECISIONMAKING

13. *Perform the analysis.* Produce one-, two- or three-way tables to answer the questions asked in the second step. FIA has developed an MS Access program called EVALIDator. The program is stored with the data and can be downloaded at <http://www.fia.fs.fed.us/tools-data/default.asp> then click on FIA DataMart. It produces tables with associated sampling errors. Produce report(s) based on the tables. This step is often never reached, leaving the data greatly underutilized. Easy-to-use analytical tools help to overcome this problem. The program has proved flexible enough for recent use in a project in Honduras.

14. *Check that monitoring system met the need.* Did the data collected meet the information needs? Were the questions answered? Were the precision and cost limits met? Adjust the monitoring system as needed for application elsewhere and/or for remeasurement.

15. *Make management decisions.* For REDD+ countries, the results will be used to make decisions on payments and as feedback on the efficacy of policies, regulations, and programs.

EXAMPLES

FIA has provided assistance in the Democratic Republic of Congo, Peru, Guyana, and several other countries. The recent focus has been on REDD+, national forest inventories, inventories for sustainable forest management planning, and on estimating populations of endangered species, such as mahogany.

As the largest of the 10 countries in the Congo River basin, the Democratic Republic of Congo (DRC) has received significant attention to improve a forest monitoring in preparation for an eventual operational REDD+ mechanism. The country is seventh in forest area (321 million ha) and has the second largest tropical forest. Much like Alaska, it is remote with few roads. Inventories of some of the regions have been completed in the accessible areas, but a national forest inventory has never been completed. Aerial photography is outdated. Securing quality satellite imagery for the DRC is difficult due to persistent cloud coverage and other challenges. Historically, the deforestation rates are low. In addition to REDD+, DRC is interested in a full NFI and in intensifying the sample further to inform management of production forests (forest management concessions) and otherwise better inform land-use decisions. The expectation is that additional well managed forests would provide sorely needed economic development opportunities from national to the local levels, resulting in many benefits including providing a viable alternative to slash-and-burn agriculture. Peru has initiated three large inventory projects. In the first project, FAO and Peru are co-funding the first complete national forest inventory. NIMAC is providing some technical assistance with the planning and design. In the second project, Peru is preparing for a second round of offering forest concessions for roughly a third of the country's forest area (the remaining two-thirds is under various forms of protection). NIMAC is providing assistance on how to do a low-cost pre-concession inventory that provides improved spatial resolution for the bidders of individual concessions. Third, NIMAC is exploring options in Peru for enhancing the estimates of the mahogany and Spanish cedar populations, such as by increasing the NFI plot size for these species and/or intensifying the sample in areas with high probability of occurrence.

In Guyana, FIA has provided technical assistance that led to two requests for proposals, one for estimating carbon for REDD+ and the other on remote sensing to estimate land use change. NIMAC will continue to provide technical expertise on a national forest inventory in addition to Measuring, Reporting and Verification (MRV) for REDD+, but will focus on providing various kinds of training.

U.S. Forest Service International Programs and FIA plan to expand the training to provide more consistent and effective technical assistance to the many countries seeking such aid. By providing regional training sessions to multiple countries, web-based training, and hosting training sessions in the United States, the demand on FIA staff time should be reduced.

FUTURE DIRECTION

Although FIA has long provided technical assistance to other countries, the current demand is unprecedented. To meet this demand, FIA is collaborating with FAO, other international aid agencies, universities, and other partners on several topics. Research is needed on efficient plot configurations and sampling designs for remote areas with high biodiversity. The current methods are efficient for one-time inventories, but may not be cost-effective for long-term monitoring. Research is needed on maximizing the use of remote sensing to enhance ground-based estimation. Fortunately, considerable funding and efforts are being expended to address this important issue. Other topics include how to quantify forest degradation and how to determine the baseline. To assist with planning, design, electronic data collection, data management, analysis and reporting, NIMAC, FAO and others are developing software tools that are flexible for meeting individual countries needs. Finally, training materials and modules are needed in various forms, such as in-country, regional and U.S.-based training sessions, webinars, web-based tutorials, and documentation. Together, these efforts will help address the growing need for forest carbon monitoring.

LITERATURE CITED

- Lister, A.J.;** Scott, C.T. 2009. Use of space-filling curves to select sample locations in natural resource monitoring studies. *Environmental Monitoring and Assessment*. 149:71-80.
- Reams, G.A.;** Smith, W.D.; Hansen, M.H.; Bechtold, W.A.; Roesch, F.A.; Moisen, G.G. 2005. The forest inventory and analysis sampling frame. In: Bechtold, W.A.; Patterson, P.L., eds. *The enhanced Forest Inventory and Analysis program - national sampling design and estimation procedures*. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station: 11-26.
- Scott, C.T.** 2009. Overview of the National Inventory and Monitoring Applications Center (NIMAC). In: McWilliams, W.; Moisen, G.; Czaplowski, R., eds. *Forest Inventory and Analysis (FIA) Symposium 2008*. RMRS-P-56CD. Ft. Collins, CO: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station: 223-229.

CURRENT FOREST AND WOODLAND CARBON STORAGE AND FLUX IN CALIFORNIA: AN ESTIMATE FOR THE 2010 STATEWIDE ASSESSMENT

Timothy A. Robards

ABSTRACT

This study used USDA Forest Service Forest Inventory and Analysis (FIA) plot data, forest growth models, wildland fire emission estimates and timber harvest data to estimate the live tree carbon storage and flux of California's forests and woodlands. Approximately 30 Tg CO₂e per year was estimated as the annual flux for all California forests. The forest inventory components not analyzed here may reduce this to about 28 Tg CO₂e per year. Over 80 percent of the annual net sequestration was estimated to come from public forestlands; however the private lands forest growth was likely underestimated given the growth models that were used. Suggestions for continued improvements in forest carbon inventory estimates include more accurate projections, biomass function improvements, continued FIA data collection, and spatial data analysis of change from natural and anthropogenic disturbance.

INTRODUCTION

The forestry sector, in the global context of the forest industry and the forests themselves, was estimated by the IPCC to produce about 17 percent of global greenhouse gas (GHG) emissions (IPCC 2007). The majority of these emissions were from tropical deforestation. Temperate and boreal forests, while generally not under the socio-economic development pressures of some tropical forests, can also impact GHG accounting at the state and national levels. The EPA estimates that U.S. forests sequester approximately 600 megatonnes (Tg) of CO₂e per year (EPA 2004). Conversely, the recent mountain pine beetle (*Dendroctonus ponderosae*) outbreak in British Columbia was estimated to cause 990 Tg of CO₂e emission from 2000 to 2020, taking the forest from a sink to a large net carbon emitter (Kurz et al. 2008).

The EPA forest carbon estimates included live trees, understory vegetation, forest floor, down dead wood, soils, wood products in use, and landfilled wood products (EPA 2004). The California Energy Commission (CEC) commissioned a study of forest carbon in California that estimated 7.5 Tg of CO₂e per year were sequestered (Brown et al. 2004). The carbon pools included in that study were the on-site pools, excluding wood products. The

California Air Resources Board (CARB), in developing the Scoping Plan (CARB 2008) for implementation of The Global Warming Solutions Act of 2006 (AB 32), used a conservative target of annual forest sequestration that was derived from the CEC report. This sequestration estimate was 5.0 Tg of CO₂e per year.

CARB is required to periodically report on GHG emissions in California (CARB 2009). CARB uses an atmospheric flow approach to estimate net flux between pools. Refinements of forest carbon cycling will assist in ensuring that AB 32 targets are met. This study, which is summarized in the California Forest and Range Assessment (FRAP 2010), provides estimates of some elements of an inventory with a focus on areas that were most likely to be substantially different from existing estimates. This includes live tree and wood products pools with mortality losses from competition, pests and fire.

METHODS

A ten-year period was used to characterize sequestration in tree growth; emissions from tree mortality caused by fire, harvest and other agents; and storage in in-use and landfill wood product pools. The most recent 10-year period was used for each component to most accurately estimate current fluxes. The current economic recession was generally not included in these estimates, which likely overestimate 2009-2010 harvest levels and associated emissions and storage.

The USDA Forest Service's Forest Inventory and Analysis (FIA) data was relied on for estimates of current storage (FIA 2008). Stock change estimates were derived by applying forest growth simulations. The FIA data is generally measured on 10-year cycles in California although shorter cycles exist on some National Forests (FIA 2009a). Modeling simulations were necessary because the FIA plots were essentially relocated, with minor overlap of a subplot,

in 2001 so that insufficient re-measurements exist for reliable stock change estimates.

Each FIA plot cluster was grown using one of four variants of the USDA Forest Service's Forest Vegetation Simulator (FVS). A computer application called the California Forest and Range Analysis System (CFRAS) was developed by the author in Microsoft Visual Basic to serve as a menu-driven user interface to read and process FIA data, call FVS simulators, and process FVS output (Robards 2010). The FVS variants and the geographic areas they cover are listed in Table 1. The number of plots were evenly distributed in each year from 2001 to 2007 so that the 10-year projections of growth was averaged over a seven year period.

The CFRAS application processed the tree lists at time zero and ten years to calculate the above and below-ground live tree carbon. Above-ground biomass (bole, bark and crown limbs) used the USDA Forest Service FIA regional volume and biomass functions (FIA 2009b; FIA 2009c). The below-ground biomass was estimated using the following model from Cairns (1997).

$$BGB = e^{-0.7747+0.8836 \times \log(AGB)} \quad [1]$$

where,

AGB = above-ground biomass,

BGB = below-ground biomass.

Carbon was estimated by multiplying biomass by 0.5. Carbon dioxide was estimated by multiplying carbon by 3.67.

Simulations were made for four land bases in California:

- all forestland,
- public forestland only,
- private forestland only, and
- private timberland only.

Timberland is a subset of forestland and is defined as lands capable of producing in excess of 20 cubic feet/acre/year at its maximum production.

TREE GROWTH

The difference in tree size over the ten-year projection period was the tree growth, which was calculated in terms of carbon tonnes by plot. No harvesting or mortality was assumed (i.e. all trees survived). This was termed simply "growth."

NON-FIRE EMISSIONS FROM MORTALITY

Two projections of growth were made using the FIA data and FVS models; the first with no mortality simulated

(see Tree Growth above) and the second with background and density-related mortality enabled. The difference in carbon estimates was the amount of carbon associated with mortality, which was assumed to be an immediate emission. Since trees decay over several years, sometimes many decades, this is a conservative assumption.

The background mortality was simulated by default; by using the MORTMULT keyword (Van Dyck 2007) with a zero parameter the background mortality was turned off. The density-related mortality, which uses the stand density index (SDI) concept (Reineke 1933), is also simulated by default. The SDIMAX keyword was used to switch off density related mortality by setting the maximum SDI parameter to 9999 and the percentage of maximum density where mortality was invoked set to 95 percent. This essentially required a SDI value of 9,499 for mortality to be invoked, which is an order of magnitude above observed SDI's.

WILDFIRE RELATED EMISSIONS

Wildfire emissions were estimated from official state estimates of emissions associated with wildfires. The FIA data was not appropriate for this estimate because of the lack of a re-measurement and because the sparse cluster design will not be accurate for change detection without auxiliary data. Wildfire carbon monoxide emissions were retrieved for each county from the CARB online database of annual estimated average emissions (CARB 2010). Queries were made for each county for wildfire emissions of carbon monoxide (CO). A CO₂/CO ratio of 13 was used (Klaus Scott, ARB, personal communication) to estimate carbon dioxide (CO₂) from CO.

The acres of forested public and private lands in each county were estimated using FRAP vegetation data (2006). The proportion of public and private forestland was estimated by dividing by the number of total acres for a county. These proportions were then multiplied by the CO₂ emissions estimate for each county. Totalling the county estimates resulted in an estimate of the average statewide annual CO₂ emissions associated with wildfire.

WOOD PRODUCTS POOLS

Wood products pools, like the wildfire emissions, were estimated from a source independent of the FIA data. The lack of re-measurement data and therefore harvest estimates made the use of a separate data source necessary.

Harvest emissions from bole wood were estimated from 10-year average Board of Equalization data and DOE 1605(b) conversion factors. The average annual board foot production was 1.713 billion board feet. The conversion from board feet to metric tons of carbon was assumed to be 0.427 (DOE 2007, table 1.7). CO₂ was estimated from C

by multiplying by 3.67. Harvest amounts were prorated to private and public lands based on BOE averages and were 92.8 percent and 7.2 percent respectively.

Non-merchantable emissions were estimated using harvest efficiency along with top, stump and root relationships to the bole (Cairns et al. 1997; Christensen et al. 2008). The following proportions of tree biomass were assumed.

- Roots are 20.63 percent of live tree based on belowground to aboveground ratio of .26 (Cairns et al. 1997).
- Non-bole aboveground biomass is 28.54 percent based on ratio of tops, limbs, and stumps to merchantable bole (Christensen et al. 2008) equal to 0.562.
- Bole biomass is 50.82 percent, which is the remainder of the total live tree biomass.
- Total live tree biomass excluded foliage.

Storage due to wood products in-use and landfill were calculated based on the 10-year average storage from the DOE 1605(b) emission inventory technical guidelines for voluntary reporting of GHGs (DOE 2007, Part I). Softwood mill efficiency was estimated to be 0.675. The loss due to defect was estimated to be 6.15 percent (Morgan and Spoelma 2008). The average storage of wood products in in-use for the first 10 years was estimated to be 5.32 percent. The landfill storage estimate for the first decade was 6.7 percent.

Portions of harvests were of live trees and others were salvaged from dead or dying trees. The California Board of Equalization data distinguishes between green and dead wood. Dead wood was estimated to be 22.8 percent on average over ten years. This amount of harvest was removed from the emission portion, not storage, to avoid double counting with the wildfire and mortality emissions.

INVENTORY COMPONENTS NOT ANALYZED

Brown et al. (2004) identified eight components related to carbon flux in the baseline analysis for forest and range carbon. They were:

- fire (emission),
- harvest (emission),
- development (emission),
- unverified increases in stocks (sequestration),
- other increases in stocks (sequestration),
- pest-related (emission),
- seasonal, and
- regrowth (sequestration).

The CARB inventory analysis (CARB 2009, Table 2) used nine categories in the forestry sector accounting, which

followed the 2006 IPCC guidelines (IPCC 2006). They were:

- forest biomass growth,
- fire,
- other disturbances (such as insect pest damage),
- development,
- timber harvest slash,
- fuel wood,
- wood waste dumps,
- discarded wood and paper in landfills, and
- composting of wood waste materials.

Considering the factors from the two sources above, the following inventory elements were not analyzed in this paper.

- Development,
- fuel wood,
- wood waste dumps, and
- composting of wood waste materials.

No benefits from urban forests were estimated including sequestration or energy conservation benefits. No other biogenic emissions such as GHGs from urban trees or emissions from non-wildfires were estimated. Wood stored in landfills prior to the current analysis, and associated emissions from landfills, was not analyzed. Imports and exports of wood products and logs were not included in this paper, including leakage effects from California's high wood products demand and policy-constrained supply.

RESULTS

The results of the carbon stocks and sequestration analysis are presented by land base type in tables 2 through 5. The estimated annual sequestration rate for all California forestlands was about 30 Tg of CO₂e (Table 2). A third of the approximately 60 Tg of CO₂e per year that could be sequestered was lost to non-wildfire related mortality. Ten percent was estimated to be lost to wildfire-related mortality. About eight percent was lost to harvest-related emissions while less than three percent was estimated to be in wood product pools. This left about one half of the potentially sequestered live tree carbon after estimated emissions deductions. These percentages varied slightly for private and public landowner classes due to most harvesting being associated with private lands.

The estimate for private forestlands was about 5 Tg of CO₂e per year (Table 3). Public forestlands were estimated to sequester about 25 Tg of CO₂e per year (Table 4). Considering only private timberlands, rather than

forestlands, yielded an estimate of about half a Tg more per year of CO₂e (Table 5).

A summary of the total CO₂e tonnes by land class, along with other measures of forest stocking and change, is shown in Table 6. The annual change estimate does not include wildfire or harvest related emissions, only model mortality. Table 7 is expressed on a per acre basis and also includes SDI density. Estimates of per acre live tree carbon stocks were highest on private timberlands. Private forestlands were lowest, which is reasonable since this will include significant acreages of non-commercial hardwood and other forest lands. The SDI values for landowner classes were in the same ranking as carbon. On average across landowner types, there was about 160 tonnes per acre of CO₂e. This compares with about 3.5 thousand cubic feet (MCF) per acre and 14 thousand board feet (MBF) per acre.

The annual per acre stock change, net of modeled mortality, was estimated to be about 1¼ tonnes of CO₂e per year for all ownerships. Public forestland was estimated to be sequestering twice the amount of carbon as private forestland. When considering only private timberlands, however, the difference narrows to 20 percent. Interestingly, the annual per acre board foot production on private timberlands is 40 percent higher than public forestlands. For all ownership types, the projected number of trees per acre decreased while stand densities increased. Some of this increase in density will be countered by harvesting and wildfire emissions.

DISCUSSION

This analysis is an inventory compilation and modeling exercise with unknown error. The general realism of these estimates may be considered by comparing the estimates to the results from other studies. The per acre carbon stocks for all forestlands in California was estimated by Christensen et al., (2008) as 33.7 tons (30.6 tonnes) C per acre above-ground live tree carbon. The estimate of aboveground live tree carbon from this analysis was 31.1 tonnes C per acre, which compares favorably as a check on the analysis. The Christensen study was based on 2001-2005 FIA data, while this study included two additional years of FIA data. Hudiburg et al. (2009) estimated average stocks of 6.5 to 19 kg/m² across Northern California and Oregon, which equates to 96.5 to 282.2 tonnes CO₂e per acre. That estimate brackets the values in this report.

The FVS growth models used in this analysis were developed primarily from data on national forests and are used for long-term planning on national forests. Intensively managed forests, as found on many private timberlands,

will likely have growth underestimated and mortality overestimated. Coast redwood, which is primarily privately owned, is missing from FVS; the other softwoods category was used as a surrogate in this study. Therefore, the private lands estimates should be considered a lower range of possible results, particularly for the coast redwood region and for plantations.

The CARB (2009) forest inventory estimate contains components that were not included in this paper. Additional emissions of 0.021 Tg CO₂e per year from development, 1.514 Tg CO₂e per year from fuel wood use, and 0.808 Tg CO₂e per year of wood waste composting sums to 2.3 Tg CO₂e per year. Combining these additional sources of emission would reduce the statewide forest carbon flux from 30.4 Tg CO₂e per year to 28.1 Tg CO₂e per year.

The differences in the public and private lands may be a function of stand age as well as productivity. Hudiburg et al., (2009, figure 6) showed that there are marked differences in stand age distributions, with private lands having substantially younger stands. A USDA Forest Service analysis (Goines and Nechodom 2009) showed that while national forests are currently sequestering substantial amounts of carbon, there are long-term risks associated with storage given disturbance and management assumptions. Consideration should be given to both the amounts of carbon sequestered and the probability of long-term storage. Potential long-term sustainable carbon storage on private lands needs further analysis. Hudiburg et al. (2009) estimate that total landscape stocks in Oregon and Northern California could theoretically be increased 46 percent. The relative amount of current stocks in relation to long-term sustainable stocks is of considerable policy interest and needs further study.

This paper should be considered an interim step in moving towards a more accurate and consistent estimate of forest carbon flux in California. Effects from development and other disturbance will require monitoring in a spatial context that plot inventories alone cannot provide. Wood products decay rates will likely continue to rely on estimates from the national inventory, which is informed by USDA Forest Service research. This study focused on the live tree components of forests. Refined models of other forest plant species and the incorporation of dead wood decay and soil carbon models will provide a more complete forest carbon inventory. As additional FIA data is collected and re-measurements begin, then stock change measurements may begin to calibrate and supplant model predictions of current forest carbon flux. Finally, the biomass functions used have been observed to have anomalies in bark biomass for some species. Given the importance of biomass functions in carbon estimation, the evaluation and improvement of biomass functions should be a priority.

ACKNOWLEDGMENTS

This work was performed while the author was employed as a forest biometrician for the California Department of Forestry and Fire Protection, Fire and Resource Assessment Program.

LITERATURE CITED

- Brown, S.;** Pearson, T.; Dushku, A. [and others]. 2004. Baseline greenhouse gas emissions and removals for forest, range, and agricultural lands in California. Winrock International, for the California Energy Commission, PIER Energy-Related Environmental Research. 500-04069F, p. 80.
- Cairns, M.A.;** Brown, S.; Helmer, E.H.; Baumgardner, G.A. 1997. Root biomass allocation in the world's upland forests. *Oecologia*. 111:1-11.
- CARB.** 2008. Climate change scoping plan: A framework for change. California Air Resources Board, Sacramento, p. 133.
- CARB.** 2009. California's 1990-2004 greenhouse gas emissions inventory and 1990 emissions level: forest and rangelands methods excerpted from the full Technical Support Document. California Air Resources Board, Sacramento, p. 18.
- CARB.** 2010. 2009 almanac emission projection data. California Air Resources Board, Sacramento, p. <http://www.arb.ca.gov/app/emsinv/emssumcat.php>.
- Christensen, G.A.;** Campbell, S.J.; Fried, J.S. 2008. California's forest resources, 2001-2005: five-year Forest Inventory and Analysis report. Gen. Tech. Rep. PNWGTR-763. U.S. Department of Agriculture Forest Service, PNW Research Station, Portland, OR, p. 183.
- Dixon, G.E.** 2009a. Inland California and Southern Cascades (CA) variant overview, Forest Vegetation Simulator. U.S. Department of Agriculture Forest Service, Forest Management Service Center, Fort Collins, CO. p. 62.
- Dixon, G.E.** 2009b. South Central Oregon and Northeastern California (SO) Variant Overview, Forest Vegetation Simulator. USDA Forest Service, Forest Management Service Center, Fort Collins, p. 78.
- Dixon, G.E.** 2009c. Western Sierra Nevada (WS) Variant Overview, Forest Vegetation Simulator. U.S. Department of Agriculture Forest Service, Forest Management Service Center, Fort Collins, CO, p. 47.
- Dixon, G.E.;** Johnson, R. 2009. Klamath mountains (NC) Variant Overview, Forest Vegetation Simulator. U.S. Department of Agriculture Forest Service, Forest Management Service Center, Fort Collins, CO. p. 48.
- DOE.** 2007. Technical guidelines, voluntary reporting of greenhouse gases (1605(b)) program. Office of Policy and International Affairs U.S. Department of Energy, Washington, D.C. p. 280.
- EPA.** 2004. Inventory of U.S. greenhouse gas emissions and sinks: 1990-2002. U.S. Environmental Protection Agency, Washington, D.C. p. 304.
- FIA.** 2008. FIA Database Description and Users Guide, version 3.0, revision 1. U.S. Department of Agriculture Forest Service, p. 243.
- FIA.** 2009a. Forest Inventory and Analysis National Program, Data and Tools. U.S. Department of Agriculture Forest Service.
- FIA.** 2009b. Regional biomass equations used by FIA to estimate bole, bark and branches (dated 13 May 2009). U.S. Department of Agriculture Forest Service, Pacific Northwest Research Station.
- FIA.** 2009c. Volume estimation for the PNW-FIA Integrated Database (dated 13 May 2009). U.S. Department of Agriculture Forest Service, Pacific Northwest Research Station.
- FRAP.** 2010. California Forest and Range Assessment. California Department of Forestry and Fire Protection, Sacramento.
- Goines, B.;** Nechodom, M. 2009. National Forest Carbon Inventory Scenarios for the Pacific Southwest Region (California). U.S. Department of Agriculture Forest Service, p. 81.
- Hudiburg, T.;** Law, B.; Turner, D.P. [and others]. 2009. Carbon dynamics or Oregon and Northern California forests and potential land-based carbon storage. *Ecological Applications*. 19:163-180.
- IPCC.** 2006. 2006 IPCC guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme, Eggleston HAS. Volume 4: Agriculture, forestry and other land uses Eds.
- Biennia, L.;** Miwa, K.; Negara, T.; Tanabe, K. Kanagawa, Japan. IPCC 2007. Climate Change 2007: Synthesis Report. Intergovernmental Panel on Climate Change, Fourth Assessment Report, Valencia, Spain, p. 73.
- Kurz, W.A.;** Dymond, C.; Stinson, G [and others]. 2008. Mountain pine beetle and forest carbon feedback to climate change. *Nature*. 452: 987-990.
- Morgan, T.A.** Spoelma 2008. California logging utilization. *Western Journal of Applied Forestry*. 23:12-18.
- Reineke, L.H.** 1933. Perfecting a stand-density index for even-aged forests. *Journal of Agricultural Research*. 46:627-638.
- Robards, T.A.** 2010. California Forest and Range Analysis System (CFRAS). California Department of Forestry and Fire Protection, Sacramento, CA.
- Van Dyck, M.G.** 2007. Keyword reference guide for the forest vegetation simulator. U.S. Department of Agriculture Forest Service. Forest Management Service Center: 130.

Table 1 – Forest Vegetation Simulator variant information and geographic area where applied

FVS Variant Name	Reference	Latitude (Degrees)		Longitude (Degrees)	
		East	West	South	North
South Central Oregon and Northeast California (SO)	(Dixon 2009b)	-120.0	-122.5	41.2	42.0
		-120.0	-121.3	40.4	41.2
Klamath Mountains (NC)	(Dixon and Johnson 2009)	-123.3	-124.5	40.3	42.0
		-123.0	-124.5	39.4	40.3
		-121.4	-124.0	37.2	39.4
		-121.4	-122.5	35.0	37.2
Westside Sierra Nevada (WS)	(Dixon 2009c)	-114.0	-121.4	32.5	42.0
Inland California and Southern Cascades (CA)	(Dixon 2009a)	-122.5	-123.3	41.2	42.0
		-121.3	-123.3	40.4	41.2
		-121.3	-123.3	39.4	40.4

Table 2 – Results for all California forestlands (32,114,317 acres). Harvest emissions were reduced by 22.8% for to avoid double-counting with mortality and fire emissions

Source	Type	C (tonnes)	CO2e (tonnes)
Growth	Storage	-16,367,285	-60,067,936
Model Mortality	Emission	5,455,351	20,021,137
Wildfire	Emission	1,719,915	6,312,087
Harvest (merch)	Emission	565,315	2,074,706
Harvest (non-merch)	Emission	791,776	2,905,819
WP (in-use)	Pool	-389,436	-1,429,231
WP (landfill)	Pool	-48,796	-179,081
Net		-8,273,161	-30,362,499

Table 3 – Results for California private forestlands (12,646,761 acres). Harvest emissions were reduced by 22.8% for to avoid double-counting with mortality and fire emissions

Source	Type	C (tonnes)	CO2e (tonnes)
Growth	Storage	-3,708,104	-13,608,743
Model Mortality	Emission	1,136,233	4,169,977
Wildfire	Emission	304,478	1,117,436
Harvest (merch)	Emission	524,612	1,925,327
Harvest (non-merch)	Emission	734,768	2,696,600
WP (in-use)	Pool	-361,397	-1,326,326
WP (landfill)	Pool	-45,283	-166,188
Net		-1,414,691	-5,191,917

Table 4—Results for California public forestlands (19,467,566 acres). Harvest emissions were reduced by 22.8% for to avoid double-counting with mortality and fire emissions

Source	Type	C (tonnes)	CO2e (tonnes)
Growth	Storage	-12,660,007	-46,462,226
Model Mortality	Emission	4,319,121	15,851,175
Wildfire	Emission	1,415,436	5,194,651
Harvest (merch)	Emission	40,703	149,379
Harvest (non-merch)	Emission	57,008	209,219
WP (in-use)	Pool	-28,039	-102,905
WP (landfill)	Pool	-3,513	-12,894
Net		-6,859,292	-25,173,600

Table 5—Results for California private timberlands (7,647,009 acres). Harvest emissions were reduced by 22.8% for to avoid double-counting with mortality and fire emissions.

Source	Type	C (tonnes)	CO2e (tonnes)
Growth	Storage	-3,603,556	-13,225,049
Model Mortality	Emission	1,010,508	3,708,564
Wildfire	Emission	184,106	675,670
Harvest (merch)	Emission	524,612	1,925,327
Harvest (non-merch)	Emission	734,768	2,696,600
WP (in-use)	Pool	-361,397	-1,326,326
WP (landfill)	Pool	-45,283	-166,188
Net		-1,556,240	-5,711,402

Table 6—Summary table of total estimated carbon, volume and tree density stocking and annual change (net of mortality only) by landowner class

Landbase	Acres	Stocks				Change, Net of Mortality			
		CO2e (tonnes)	Cubic Vol. (MCF)	Bd. Ft. Vol (MBF)	No. Trees	CO2e (tonnes)	Total MCF	Merch MBF	No. Trees
All Forestlands	32,114,317	5,099,162,048	113,695,755	447,709,621	10,058,521,955	40,046,799	1,419,806	5,764,470	-58,328,612
Public Forestland	19,467,566	3,343,515,541	76,368,749	340,794,682	5,685,834,310	30,611,051	751,107	3,438,690	-38,089,971
Private Forestland	12,646,761	1,755,647,124	37,327,502	106,914,068	4,372,687,646	9,438,766	668,726	2,325,853	-20,237,568
Private Timberland	7,647,009	1,418,463,058	31,054,447	103,118,272	4,364,675,374	9,516,486	591,411	2,242,743	-17,094,787

Table 7—Summary table of per acre estimated carbon, volume, and density stocking and annual change (net of mortality only) by landowner class

Landbase	Stocks					Change, Net of Mortality				
	CO2e (tonnes)	Cubic Vol. (MCF)	Bd. Ft. Vol (MBF)	No. Trees	SDI	CO2e (tonnes)	Cubic Vol. (MCF)	Bd. Ft. Vol (MBF)	No. Trees	SDI
All Forestlands	158.8	3.5	13.9	313.2	214.1	1.247	0.044	0.179	-1.816	2.422
Public Forestland	171.7	3.9	17.5	292.1	225.1	1.572	0.039	0.177	-1.957	2.015
Private Forestland	138.8	3.0	8.5	345.8	197.1	0.746	0.053	0.184	-1.600	3.050
Private Timberland	185.5	4.1	13.5	570.8	258.0	1.244	0.077	0.293	-2.235	4.189

EVALUATING THE COMPATIBILITY OF AMERICAN AND MEXICAN NATIONAL FOREST INVENTORY DATA

Todd A. Schroeder, Sean P. Healey, and Gretchen G. Moisen

ABSTRACT

The international border region between the United States and Mexico represents a point of discontinuity in forest policy, land use management and resource utilization practices. These differences along with physical barriers which separate the two countries can interact to alter the structure and functioning of forest vegetation. One valuable source of information for analyzing potential effects of management on forest attributes is National Forest Inventory (NFI) data. Both Mexico and the United States have systematically designed NFI programs, the U.S. Forest Service Forest Inventory and Analysis (FIA) program and the Comisión Nacional Forestal (CONAFOR) Inventario Nacional Forestal y de Suelos (INFyS). However, data from NFIs are seldom harmonized with respect to reporting units, field procedures and estimation methods. Here we evaluate two important aspects of NFI data compatibility using seamless geospatial data. First, to gauge plot measurement and location accuracy we compared the elevations recorded in each countries NFI database with those taken from an independently derived digital elevation model (DEM). Second, basal area compatibility was determined by means of analysis of covariance (ANCOVA) using a seasonal time series of normalized difference vegetation index (NDVI) data from Landsat. The results showed that both countries have good location and measurement accuracy in relation to DEM elevations and in the majority of cases, statistically similar estimates of basal area per unit of NDVI. Despite finding a high level of plot data compatibility, our study uncovered key differences in inventory stratification between the two countries which prevented further statistical comparison of oak woodland stand densities. Suggestions for improving local and regional scale analysis compatibility of American and Mexican NFI data are provided.

INTRODUCTION

In response to interest concerning the effects of global climate change there is a growing need for information on the health, status, and biodiversity of the world's forest resources. In many countries, the current condition of forests is often estimated with data collected by national forest inventory (NFI) programs. NFIs typically collect detailed tree and stand measurements across a statistically designed, systematic layout of field plots. Although timber assessment has traditionally been a focus of many NFIs (Scott and Grove, 2001), measurement of forest attributes relating to ecosystem functioning and health is increasing. NFI data is frequently called upon to generate continental- and global-scale information on biological diversity, ecosystem

health and forest carbon pools. However, data collected by independent NFIs is seldom harmonized (i.e., in agreement) with respect to reporting units, field procedures, and estimation methods (Winter et al., 2008). Resulting discrepancies can produce large uncertainties when multiple NFIs are used to estimate attributes such as forest area and biomass change (Schoene, 2002; Cienciala et al., 2008).

As multinational NFI data represents a critical source of global information on greenhouse gases (e.g., United Nations Framework Convention on Climate Change (UNFCCC) and its Kyoto Protocol (1997)) and sustainability (e.g., Food and Agriculture Organization (FAO), Global Forest Resources Assessment (FRA, 2006)), promoting the harmonization of NFI definitions and measurement protocols will help reduce uncertainty, and facilitate the comparison of estimates across international boundaries.

Until recently, reporting efforts in North America have been hampered by the lack of systematically collected field data over much of the continent. Although the United States has been conducting a statistically based NFI since the late-1920's (Shaw, 2006), neither Canada nor Mexico had, until recently adopted systematically implemented national programs (Canada see Gillis, 2001 and Gillis et al., 2005; Mexico see Sandoval et al., 2008). Given different histories of the three countries, efforts to harmonize terminology and field measurement protocols are only beginning to take shape. Nonetheless, as Mexican and Canadian NFI data begin to come on-line, new methods will be needed to determine the extent to which plot data from the three North American NFIs are compatible for continental scale reporting. Here our objective is to evaluate the inter-compatibility of plot data collected in borderland oak woodland forests by the U.S. Department of Agriculture's Forest Inventory and Analysis (FIA) program and in Mexico by Comisión Nacional Forestal (CONAFOR).

Focusing on oak woodland on both sides of the Arizona/Sonora border, we evaluate two important aspects of data

compatibility through the analysis of seamless elevation and spectral geospatial data sets. First, to gauge a sense of plot measurement and location accuracy we compare the elevations recorded in each country's NFI database with those taken from an independently derived digital elevation model (DEM). Second, we use a seasonal time series of normalized difference vegetation index (NDVI) images derived from Landsat satellite data to assess the consistency of the relationship between plot and satellite forest measurements across the border.

METHODS

STUDY AREA

Lying equidistant between the United States and Mexico, the study area is the 74,655 km² Madrean Archipelago ecoregion (Omernik level III, CEC 1997; Fig. 1). The forests in this border region of southeastern Arizona (United States) and northeastern Sonora (Mexico) contain some of the most diverse temperate forest ecosystems in the world. The mountains here straddle two major faunal realms (Neotropic/Holarctic) and two climatic zones (Subtropical/Temperate). The confluence of these zones interacts with complex mountain topography to support high levels of endemic biodiversity (Coblentz and Riitters, 2004). The forests, which primarily reside on a series of disconnected mountain ranges, are surrounded by vast "seas" of desert vegetation. These valley seas inhibit new species colonization which serves to isolate the higher elevation "island" biotic communities (Warshall, 1994). The forest composition displays an altitudinal gradient; open oak woodlands are found at lower elevations, which cede to closed canopy pine and fir dominated forests as elevation increases.

Oak woodland forests were selected for this analysis because they represent the largest area of forested land within the ecoregion. Focusing on this large forested area ensured that a sufficient number of NFI plots from each country were available for analysis and, minimized the effect of different inventory stratification procedures used by the two countries. The oak woodland forest areas were defined by a geographic information system (GIS) coverage of biotic communities assembled by the U.S. Forest Service at a scale of 1:1,000,000 (Brown and Lowe, 1982). In the United States, the oak woodland forest type covers 16 percent of the landscape and captures 36 percent of the FIA inventory plots collected in the ecoregion (Arizona plots only, New Mexico plots are not included in this analysis). In Mexico, the oak woodland forest type covers 32 percent of the landscape and captures 75 percent of the collected CONAFOR inventory plots in the ecoregion (Sonora plots only, Chihuahua plots are not included in this analysis).

DATA

FIA

FIA data are collected on a nationally consistent hexagonal sampling frame where at least one plot is randomly selected within each 6,000 acre hexagon (Bechtold and Scott, 2005). Each plot consists of four fixed-radius circular subplots, which taken together represent an area approximately 1 acre in size. Data collected on each FIA plot includes land use, tree measurements (e.g., species, height, and diameter) as well as other tree and site related forest attributes.

For this study, we queried the FIA database to obtain the annual inventory data collected in Arizona between 2001 and 2007. Using the geographic coordinate locations of the plots, a GIS overlay operation was used to identify the Arizona plots falling within the oak woodland boundary. The measured live tree data from these plots ($n = 117$) was then used to calculate basal area. Most of the trees sampled in this region are defined by FIA as woodland species which are measured for diameter at the root collar (DRC) near ground line. Thus, to calculate basal area we first converted DRC to diameter at breast height (DBH) using Eq. 1 (Chojnacky and Rogers, 1999),

$$DBH = \beta_0 + \beta_1 DRC + \beta_2 stm + \beta_3 Pied + \beta_4 DRC_p + \beta_5 Quga + \beta_6 DRC_q \quad [1]$$

where DBH is diameter at 1.3m above groundline, DRC is diameter at root collar, stm is 1 for trees with 1 stem at DRC or 0 otherwise, Pied is 1 for pinyon pine species and 0 otherwise, Quga is 1 for oak species and 0 otherwise, DRC_p is DRC for pinyon pine species, and DRC_q is DRC for oak species. Constants for the β terms (in inches) are $\beta_0 = -2.6843$, $\beta_1 = 1.0222$, $\beta_2 = 0.7433$, $\beta_3 = 0.7469$, $\beta_4 = -0.0399$, $\beta_5 = 1.2244$, and $\beta_6 = -0.0689$. Equation 1 was formulated using 224 trees sampled in western Colorado for Pinyon pine (*Pinus edulis*), Utah juniper (*Junipers osteosperma*) and Gambel oak (*Quercus gambelii*). Here we applied the equations at the genus level (e.g., all oak species were converted to DBH using Quga in Eq. 1), which accounted for nearly 85 percent of the measured trees in the study area. The remaining trees were converted to DBH using the closest available matching equation (e.g., deciduous species were converted using the Quga equation, conifer species using the Pied equation). Although this extrapolation involves applying the equation outside of the range and species in which it was initially developed, it currently represents the best available option for converting FIA data from DRC-to-DBH.

After converting from DRC to DBH, basal area per tree was calculated for each measured live tree ≥ 5 inches using Eq. 2,

$$BA \text{ (ft}^2\text{)} = 0.005454 \times \text{DBH}^2 \quad [2]$$

where BA is basal area in ft^2 and DBH is in inches. Basal area per tree was multiplied by trees per acre (TPA) and condition proportion (COND_PROP) variables in the FIA database, and then summed across each plot to yield per plot estimates of basal area in ft^2/ac . The ft^2/ac estimates were then multiplied by 0.2296 to get basal area in m^2/ha .

CONAFOR

CONAFOR data are also collected on a nationally consistent sampling grid which consists of more than 24,000 plots covering all vegetation types. The grid spacing of plots depends on vegetation type (e.g., 5x5 km grid for temperate and high tropical forests, 10x10 km for shrub lands and low tropical forests, and 20x20 km for arid regions) which is taken from a national land use and vegetation cover map derived from Landsat data. Similar to FIA, CONAFOR data are collected on four circular subplots which cover an area approximately 1 acre in size. Data collected include topography, land use and disturbance as well as tree species and diameter measurements among others. For more information on the enhanced Mexican national forest inventory program see Sandoval et al. (2008).

Plot data for the Mexican state of Sonora were spatially queried in a GIS system to select the plots contained within the geographic extent of the oak woodland boundary. The measured live trees ≥ 12.7 cm DBH (or 5 inches, same minimum used for FIA) from the selected plots ($n = 142$) were used to calculate basal area per tree using Eq. 3,

$$BA \text{ (m}^2\text{)} = 0.00007854 \times \text{DBH}^2 \quad [3]$$

where BA is basal area in m^2 and DBH is in cm. The CONAFOR tree data is collected at DBH approximately 1.3 m above ground line, therefore no DRC conversion was necessary. The Mexican inventory data does not contain expansion factors. In order to obtain basal area on a per hectare basis, we used only the plots which contained 4 measured subplots. Because the plots have a fixed radius, this allowed the use of a constant 6.25 area expansion factor. Basal area per tree was multiplied by this constant expansion factor, then summed across each plot to yield per plot estimates of basal area in m^2/ha .

SRTM DEM

To help evaluate the location and measurement accuracy of the NFI plots (described below) we obtained digital elevation data from the Consultative Group for International Agriculture Research – Consortium for Spatial Information

(CGIAR-CSI; <http://srtm.csi.cgiar.org/>). Based on the unfinished 3 arc second data originally released by the National Aeronautics and Space Administration (NASA), the CGIAR-CSI version-4 data used here have been hydrologically corrected with a gap-filling algorithm to produce a smooth continuous raster surface at 90 m spatial resolution. The data were downloaded in separate $1^\circ \times 1^\circ$ degree grid tiles, which were mosaiced together in ArcInfo Grid to produce seamless coverage of the study area. Once mosaiced, the study area elevation grid was reprojected from geographic coordinates to UTM projection with WGS 84 datum.

SATELLITE IMAGERY

To evaluate the consistency of basal area measurements among the two countries, we compared plot measurements using NDVI data from Landsat (described in more detail below). NDVI is a satellite measure of green leaf area; therefore it can vary seasonally with changes in precipitation and background reflectance. To account for this we developed a series of images which covers nearly the full extent of the dry season, which ranges from mid-April to mid-July. To achieve complete seasonal coverage we acquired cloud-free Landsat TM data (LT1 processing) for path 35, rows 38 and 39 for six dates (4/24/2004, 5/8/2003, 5/13/2005, 6/11/2004, 6/25/2003, and 7/16/2005). Each date of path/row images (see Fig.1 for coverage) were mosaiced and then converted to surface reflectance using the COST model (Chavez, 1996). NDVI was calculated as the ratio of (Band 4 – Band 3) / (Band 4 + 3). The final set of processed NDVI images had 30 m spatial resolution, UTM projection and WGS 84 datum.

DATA COMPATIBILITY TESTS

PLOT LOCATION

One important indicator of data compatibility is that NFI plots are located where they are supposed to be in geographic space and that they accurately reflect the topography of the landscape. In general, if plots are properly located and measured, then we should be able to use each plot's geographic coordinates to derive independent estimates of topographic variables (e.g., elevation from a DEM) which closely match the records found in each country's NFI database. To test this idea we compare independent estimates of elevation extracted from an SRTM DEM with those found in each country's NFI database (FIA $n = 117$, CONAFOR $n = 142$). SRTM data was extracted for each NFI plot using the mean of a 3x3 window placed over plot center (for both FIA and CONAFOR we used actual plot coordinates, not publically available). It is possible that the level of agreement (based on R^2) of the two countries will differ because plot elevations in the CONAFOR data are taken from field measurements, whereas in FIA they are either taken from

field measurements, DEM or topographic map. Minor agreement differences aside, if the plots are reasonably located in geographic space, and in the case of CONAFOR are accurately measured, the plots should fall on or close to the 1:1 line when viewed in a two dimensional scatter plot. This test is intended only as a check for errors which might bring into question the general reliability of the location and measurement of the NFI plots, and is not intended to be a precise quantitative assessment of elevation accuracy.

BASAL AREA ESTIMATION

Barring differences in precipitation and back-ground effects, it is to be expected that NDVI (a satellite based measure of green leaf area) will increase as basal area increases. If the basal area estimates derived for each country are compatible, we should find no statistical difference between their fitted relationships with NDVI. To test this hypothesis we conducted an analysis of covariance (ANCOVA). The analysis was restricted to the range of basal area measured by both countries (i.e., 16.77 m²/ha). In addition to capping the range of basal area, the Landsat images do not cover the full extent of the study area, thus the number of plots available for the ANCOVA analysis (FIA n = 74, CONAFOR n = 121) is less than was used for the plot location and measurement test described above. For the plots qualifying for the analysis, NDVI was extracted from each of the six seasonal images using the mean of a 3x3 window placed over plot center (for FIA and CONAFOR we used actual plot coordinates, not publically available). We then tested the null hypothesis that the slopes of each countries fitted lines were equal using a standard F test. If the slopes are found equal, then each countries fitted mean basal area is “adjusted” according to the overall mean of NDVI. The null hypothesis of equal adjusted means is then tested with a second F test. If we do not reject the null hypothesis of equal adjusted means (i.e., p-value ≥ 0.05) then we can conclude that per unit NDVI, the sample plots collected on both sides of the border have statistically similar estimates of basal area.

RESULTS

PLOT LOCATION TEST

Scatter plots comparing the SRTM elevations and elevations from the NFI data revealed good agreement for FIA as indicated by all of the plots falling along the 1:1 line (Fig. 2). Although the majority of CONAFOR plots also fell on or near the 1:1 line, we did find seven plots (indicated by dashed oval and arrow in Fig. 2) which were not; all but one of these plots had recorded elevations almost exactly 1,000 m above the SRTM measurements. Given the small percentage of plots affected (4.9 percent) and the systematic nature of these deviations, it is likely these errors were the result of data entry mistakes rather than plot location or measurement inaccuracies. Removing the seven outliers

from the CONAFOR data we found that elevations from both NFI data sets were in similar agreement with the independent SRTM elevations (FIA $R^2 = 0.99$, CONAFOR $R^2 = 0.97$). Although the CONAFOR data displays higher residual variance (Figure 2), the R^2 results verified that both FIA and CONAFOR plots were placed on the landscape with sufficient accuracy that the topographic descriptors published in each database could be accurately reproduced using independent data.

BASAL AREA COMPATIBILITY TEST

The ANCOVA results revealed that the fitted lines for both countries were statistically similar for all six NDVI image dates (Table 1). While the fitted lines were not necessarily parallel (Figure 2), they were similar enough that the null hypothesis of equal slopes could not be rejected. The test of equal adjusted means revealed that for four of the six image dates the null hypothesis could not be rejected (Table 1). This indicates that once canopy conditions represented by NDVI were accounted for, the adjusted mean basal areas of the two countries were, in the majority of cases, not significantly different. Although two of the image dates (4/24/2004, 6/25/2003) produced results which were close to rejecting the null hypothesis, the small average difference in adjusted mean basal area (0.2 m²/ha) across the six image dates supports the conclusion that the basal area estimates from the two countries are similar enough to be deemed compatible.

DISCUSSION

In this study we evaluated the compatibility of American (FIA) and Mexican (CONAFOR) NFI data using sample plots collected across an area of borderland oak woodland forest. Given the discontinuous nature of the NFI data, the evaluation of compatibility focused on the analysis of geospatial data sets which seamlessly and consistently spanned the area of data collection. To gauge location and measurement compatibility we compared the similarity of each countries plot responses to topographic (i.e., elevation) and spectral based NDVI data.

As verification that the NFI plots were located on the landscape with sufficient spatial accuracy for joint analysis, we compared independently derived SRTM DEM data with elevations recorded in each countries NFI database. The test identified seven CONAFOR plots which had considerable deviation in recorded elevations. As these errors were systematic in nature they were most likely the result of data entry mistakes. Comparison with freely available SRTM data may in the future be an efficient quality control measure for NFI elevation data. Plots identified as erroneous can be re-inspected to verify coordinate, elevation measurement and data entry integrity. Aside from the identified outliers, we found all of the FIA and CONAFOR NFI elevations to be in good agreement with the SRTM data. This provided

evidence that the FIA and CONAFOR plots were reasonably located on the landscape and that elevation was accurately measured or estimated by each inventory program. In general, this test provided an indirect verification of plot location compatibility, as well as an effective means for identifying plots which might have potential misalignment or measurement errors.

Although the inventory programs have similar plot layout designs, differences in data collection protocols exist which must be accounted for before undertaking a more thorough assessment of data measurement compatibility. Here, efforts were taken to harmonize the calculation of basal area in order to assess the compatibility of stand density estimates derived from the two NFI data sets. Harmonization efforts included applying published equations to convert FIA DRC to DBH, using the same minimum DBH cutoff and converting basal area estimates to like units (m^2/ha). Although the DRC-to-DBH equations (Chojnacky and Rogers, 1999) used here were extrapolated well beyond their geographic and ecological boundaries, the adjustment was a critical step in harmonizing the NFI data.

In this study FIA basal area was reduced by roughly 32 percent after the DRC-to-DBH conversion was applied. This is similar in magnitude to the 10 to 25 percent reduction in basal area reported by Chojnacky and Rogers (1999) for ponderosa pine forests in the Gila National Forest, New Mexico. It should be noted that basal area reported in the FIA database (e.g., variable BALIVE) is calculated for woodland species without converting DRC-to-DBH. In addition, reported diameters for woodland species in the FIA database (e.g., variable DIA) are actually DRC measurements, thus when calculating basal area, volume or biomass with equations that require DBH as input (e.g., Jenkins et al., 2003) DRC-to-DBH conversion is required. Given the considerable difference between DRC-corrected and uncorrected basal area, future work should focus on improving the necessary equations required to make this critical adjustment. These improvements would also stand to benefit future studies which use FIA data to estimate carbon and biomass for woodland species. As CONAFOR measurements are taken at DBH, the conversion of FIA DRC-to-DBH was an important step in developing harmonized estimates of basal area for the two countries.

To test the compatibility of the basal area estimates we used NDVI data from Landsat. The reasoning behind this test comes from the fact that canopy conditions of forests in this region are open and highly variable. For example, a stand with the same unit basal area could have a relatively open canopy structure consisting of a few large but scattered trees or a closed canopy structure consisting of several small but densely clustered trees. Although in this situation basal area is the same, the different canopy conditions result in

very different measures of NDVI. Thus, if the relationship between basal area and NDVI were drastically different for each country, this might suggest that inventory stratification or systematic data collection differences might be affecting the compatibility of the basal area estimates. The ANCOVA analysis showed that the fitted relationships between basal area and NDVI were similar for each country, thus offering evidence that per unit NDVI the basal area estimates were in the majority of cases, statistically compatible.

Overall, both independent tests based on the analysis of seamless geospatial data indicated a high degree of plot level compatibility between American and Mexican NFI data. Given the high level of plot data compatibility we hoped to proceed with a joint analysis of the two NFI data sets with the purpose of investigating the ecological impacts of divergent management and land use practices on stand density in borderland oak woodland forests. Jointly analyzing the NFI data in this context could take two different approaches. One approach might be to use the plot data from both countries to analyze basal area distributions using tests of central tendency (e.g., looking for statistical differences in population means or medians). A second approach might collectively use the NFI plots from each country to derive statistical estimates of basal area for a particular region of interest (e.g., the Madrean archipelago ecoregion or the oak woodland forest type).

To perform these types of joint analyses requires accounting for differences in inventory stratification which exist between the two countries. FIA's sample grid extends with equal intensity to all lands, but only forested plots (as identified through photos or in the field) are surveyed in detail. In contrast, CONAFOR pre-stratifies their sample grid according to a land use map produced by the Mexican Agency INEGI (Instituto Nacional de Estadística, Geografía e Informática). Detailed plot measurements are taken on all lands with forests more heavily sampled than other wooded lands and deserts. While the Mexican plot data contain stratum identifiers, thus allowing calculation of average conditions by stratum, the stratification map itself is not publically available. Without this map, it is impossible to determine weights for a complex analysis unit such as the oak woodland forest type within the Madrean archipelago ecoregion. Publication of the Mexican land cover map, or at least development of factors communicating the area represented by each plot, would greatly increase the inventory's usefulness in local and cross-border analyses.

CONCLUSION

We determined that plot data from the two inventories are compatible: plots from both countries were accurately geolocated, and the relationship between measured basal area

and satellite imagery was consistent across the border. The following recommendations may be identified following our work.

ACKNOWLEDGEMENTS

The authors wish to thank Rigoberto Palafox Rivas, Carmen L.M. Tovar and Vanessa S. Mascorro of CONAFOR for their important contributions to this study.

CONAFOR	<ol style="list-style-type: none"> 1. Systematically check archived plot elevations against freely available global terrain data 2. Attach stratum-adjusted area equivalents (expansion factors) to plot data
FIA	<ol style="list-style-type: none"> 1. Develop systematic conversion from root collar diameter to diameter at breast height to improve regional basal area variables

LITERATURE CITED

- Bechtold, W.A.;** Scott, C.T. 2005. The forest inventory and analysis plot design. In: Patterson, P.L., ed. The enhanced forest inventory and analysis program – national sampling design and estimation procedures. Gen. Tech. Rep. SRS 80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. pp. 27 – 42.
- Brown, D.E.;** Lowe, C.H. 1982. Biotic communities of the Southwest [map]. Gen. Tech. Rep. RM 78. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. Map scale 1:1,000,000.
- Chavez, P.S., Jr.** 1996. Image-based atmospheric corrections – Revisited and improved. Photogrammetric Engineering and Remote Sensing. 62: 1025-1036.
- Chojnacky, D.C.;** Rogers, P. 1999. Converting tree diameter measured at root collar to diameter at breast height. Western Journal of Applied Forestry. 14: 14-16.
- Cienciala, E.;** Tomppo, E.; Snorrason, A. [and others]. 2008. Preparing emission reporting from forests: Use of national forest inventories in European countries. Silva Fennica. 41, 73 – 88.
- Coblentz, D.D.;** Riitters, R.H. 2005. A quantitative topographic analysis of the Sky Islands: a closer examination of the topography-biodiversity relationship in the Madrean Archipelago. In: DeBano, F.L.; Ffolliot, P.F.; Ortega-Rubio, A. [and others]. eds. Biodiversity and management of the Madrean Archipelago: The Sky Islands of Southwestern United States and Northwestern Mexico. Gen. Tech. Rep. RM 264. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. pp. 69-74.
- Commission for Environmental Cooperation (CEC).** 1997. Ecological regions of North America: toward a common perspective. Montreal, Quebec, Canada. 71pp. Map scale 1:12,500,000.
- Food and Agriculture Organization of the United Nations [FAO].** 2006. Global Forest Resources Assessment 2005 main report. FAO Forestry paper 147, Rome. (<http://www.fao.org/forestry/site/24691/en/>).
- Gillis, M.D.** 2001. Canada’s national forest inventory (responding to current information needs). Environmental Monitoring and Assessment. 67: 121-129.
- Gillis, M.D.;** Omule, A.Y.; Brierley, T. 2005. Monitoring Canada’s forests: The national forest inventory. The Forestry Chronicle. 81: 214-221.
- Jenkins, J.C.;** Chojnacky, D.C.; Heath, L.S. [and others]. 2003. National-scale biomass estimators for United States tree species. Forest Science. 49: 12-35.
- Kyoto Protocol** 1997. The Kyoto Protocol to the framework convention on climate change (<http://unfccc.int/resource/docs/convkp/kpeng.pdf>).
- Sandoval Uribe, A.;** Healey, S.P.; Moisen, G.G. [and others]. 2008. Mexican forest inventory expands continental carbon monitoring. EOS. 89: 485.
- Schoene, D.** 2002. Assessing and reporting forest carbon stock changes: a concerted effort? Unasylva. 210: 76-81.
- Scott, C.T.;** Gove, J.H. 2001. Forest Inventory. In: El-Shaarawi, A.H., and Piegorsch, W.W., eds. Encyclopedia of Environmetrics. 2: 814-820.
- Shaw, J.D.** 2006. Benefits of strategic national forest inventory to science and society: the USDA Forest Service Forest Inventory and Analysis program. Forest. 3: 46-53.
- Warshall, P.** 1994. The Madrean Sky Island archipelago: a planetary overview. In: DeBano, F.L.; Ffolliot, P.F.; Ortega-Rubio, A. [and others] eds. Biodiversity and management of the Madrean Archipelago: The Sky Islands of Southwestern United States and Northwestern Mexico. Gen. Tech. Rep. RM 264. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. pp. 6-18.
- Winter, S.;** Chirici, G.; McRoberts, R.E. [and others]. 2008. Possibilities for harmonizing national forest inventory data for use in forest biodiversity assessments. Forestry. 81: 33-44.

Table 1—ANCOVA results for basal area compatibility test

Image Date	Equal Slope		Adj. Mean BA (m ² /ha)			Equal Adj. Mean	
	F	p-value	FIA	CONAFOR	Diff	F	p-value
4/24/2004	1.06	0.306	6.3	5.4	0.9	3.81	0.052
5/8/2003	1.74	0.189	5.5	5.9	-0.5	0.84	0.360
5/13/2005	0.03	0.856	6.3	5.4	1.0	4.51	0.035
6/11/2004	1.68	0.197	5.5	5.9	-0.3	0.47	0.493
6/25/2003	0.47	0.492	5.1	6.2	-1.1	3.56	0.061
7/16/2005	0.04	0.845	6.5	5.2	1.3	7.55	0.007

*significant tests in bold

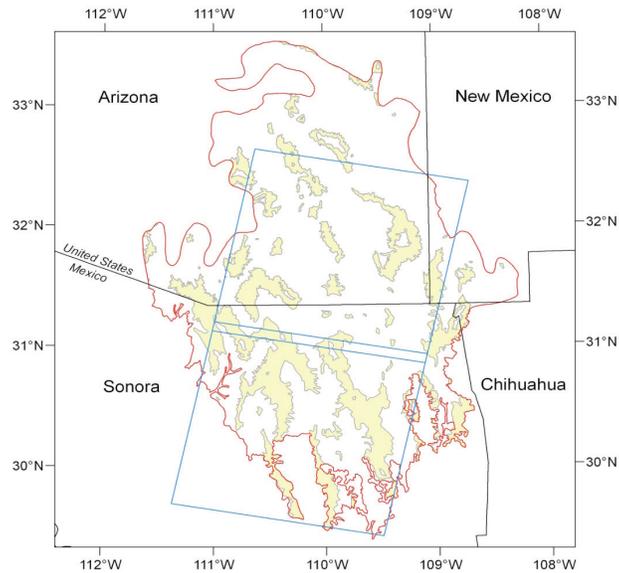


Figure 1—The Madrean Archipelago study area (red outline) showing location of oak woodland forest (yellow) and Landsat path/rows (blue outline).



Figure 2—Scatter plots of SRTM DEM elevation versus NFI database elevation for FIA (+) and CONAFOR (o).

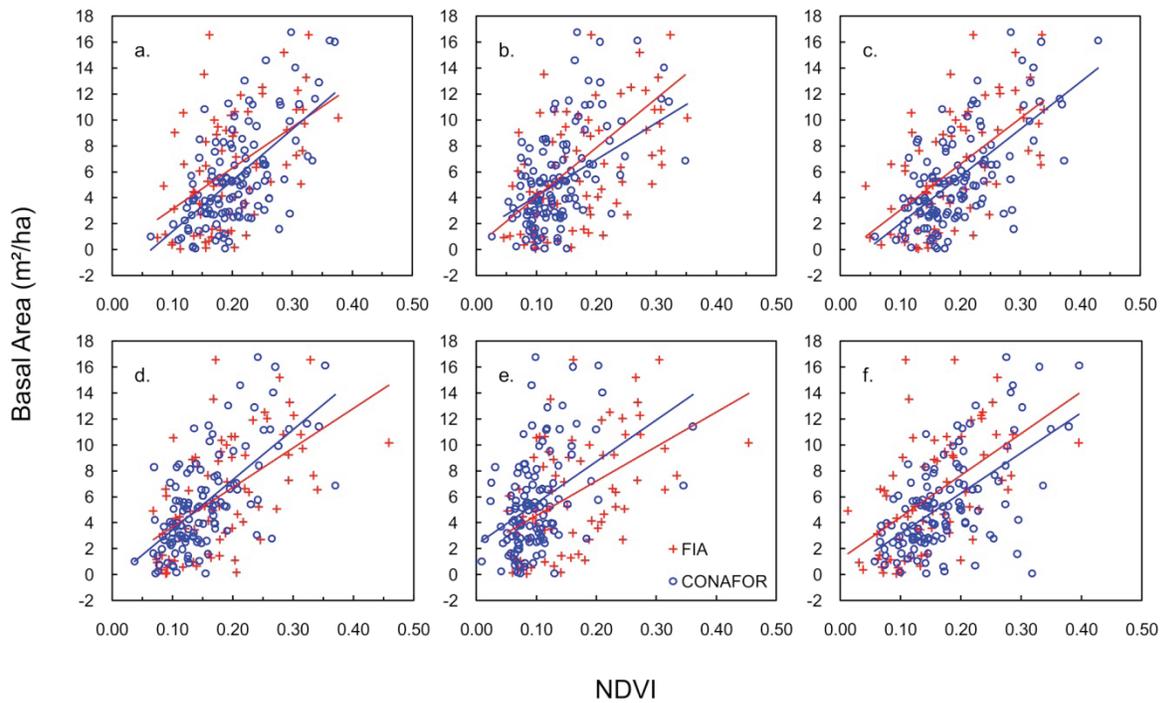


Figure 3—Fitted relationships between NDVI and basal area for FIA (+) and CONAFOR (o) for a.) 4/24/2004, b.) 5/8/2003, c.) 5/13/2005, d.) 6/11/2004, e.) 6/25/2003, and f.) 7/16/2005 image dates.

Biometrics

MODELING ALASKA BOREAL FORESTS WITH A CONTROLLED TREND SURFACE APPROACH Mo Zhou and Jingjing Liang	25
CURIOUS OR SPURIOUS CORRELATIONS WITHIN A NATIONAL-SCALE FOREST INVENTORY? Christopher W. Woodall and James A. Westfall	39
ESTIMATING TREE CROWN WIDTHS FOR THE PRIMARY ACADIAN SPECIES IN MAINE Matthew B. Russell and Aaron R. Weiskittel	45
AN EVALUATION OF THE PROPERTIES OF THE VARIANCE ESTIMATOR USED BY FIA John P. Brown and James A. Westfall	53
A MULTIVARIATE MIXED MODEL SYSTEM FOR WOOD SPECIFIC GRAVITY AND MOISTURE CONTENT OF PLANTED LOBLOLLY PINE STANDS IN THE SOUTHERN UNITED STATES Finto Antony, Laurence R. Schimleck, Alex Clark III and Richard F. Daniels	59

MODELING ALASKA BOREAL FORESTS WITH A CONTROLLED TREND SURFACE APPROACH^a

Mo Zhou and Jingjing Liang*

ABSTRACT

An approach of Controlled Trend Surface was proposed to simultaneously take into consideration large-scale spatial trends and nonspatial effects. A geospatial model of the Alaska boreal forest was developed from 446 permanent sample plots, which addressed large-scale spatial trends in recruitment, diameter growth, and mortality. The model was tested on two sets of validation plots and the results suggest that the controlled trend surface model was generally more accurate than both nonspatial and conventional trend surface models. With this model, we mapped the forest dynamics of the entire Alaska boreal region by aggregating predicted stand states across the region.

INTRODUCTION

Geospatial effects at large scales have been reported in many biological and ecological studies. The conventional trend surface analysis (e.g. Kuuluvainen and Sprugel 1996; Thomson 1986) has been developed to capture such trends in various disciplines (Gittins 1968) and there exist numerous studies attempting to explain these effects (e.g. Kuuluvainen and Sprugel 1996; Wilmsking and Juday 2005).

Existing spatial studies of forest dynamics have been mainly focusing on small-scale spatial effects, such as interactions of neighboring trees or stands (e.g. Franklin and others 1985; Larson and others 2006; Liu and Ashton 1998; Pacala and others 1996). Little has been done to identify large-scale spatial factors of forest dynamics and separate them from small-scale variations attributable to local effects (Schenk 1996), such as site and stand basal area (e.g. Bonan and Shugart 1989; Liang and others 2005).

The purpose of this paper was to propose an innovative method, controlled trend surface (CTS), to account for both large-scale spatial effects and well-recognized nonspatial factors in modeling. With this proposed method, a geospatial dynamics model of the Alaska boreal forest was developed based on the same data that were used to calibrate the

nonspatial model of Liang (2010). With remote sensing data and the Geographic Information System (GIS), stand-level predictions were aggregated to tentatively map forest dynamics of the entire region.

The Alaska boreal forest is generally defined as a biome characterized by coniferous forests. In this study, it represented a vast area composed of the following ecoregions: Interior Alaska-Yukon lowland Taiga, Cook Inlet Taiga, and Copper Plateau Taiga. Forestry is very important for the state of Alaska (AlaskaDNR 2006; Wurtz and others 2006), and is an indispensable component of rural economies (AlaskaDNR 2006). Liang (2010) develops the first Matrix Model for all major Alaska boreal tree species which is tested to be much more accurate than the two growth and yield tables. However, due to a lack of control for large-scale spatial patterns which “may cause substantial errors between actual and predicted stand states” (See Liang 2010, P.10), caution is advised when applying the Matrix Model on stands out of the sample area or on areas of considerable sizes.

METHODS

CONTROLLED TREND SURFACE (CTS)

The conventional trend surface analysis studies the spatial trend of given observations $Z(\mathbf{s})$ at location \mathbf{s} within the region D (Grant 1957; Ripley 1981; Watson 1971):

$$Z(\mathbf{s}) = \mu(\mathbf{s}) + \delta, \quad \mathbf{s} = (\mathbf{x}, \mathbf{y})' \in D \subset \mathbb{R}^2 \quad (1)$$

where

$$\mu(\mathbf{s}) = \sum_{i=1}^k f_{i-1}(\mathbf{s}) \rho_{i-1}$$

represents an unknown linear combination of known functions $f_i(\mathbf{s})$ of spatial coordinates $\mathbf{x} = (x_1, \dots, x_n)'$ and $\mathbf{y} = (y_1, \dots, y_n)'$ with unknown but fixed parameters ρ_i , $i=1, 2, \dots, k-1$. δ is a zero-mean, stationary error term with known covariogram (see Berke 1999, p.219).

Mo Zhou, School of Management, University of Alaska Fairbanks, Fairbanks, AK 99775-6080

Jingjing Liang, Department of Forest Sciences, University of Alaska Fairbanks, P.O.Box 757200, Fairbanks, AK 99775-7200, USA

Tel: 907-474-1831; Fax: 907-474-6184; Email: jliang@alaska.edu

^a This paper is condensed and modified from Liang and Zhou (2010).

*Corresponding author.

Now assume that δ was controlled by non-spatial factors \mathbf{n} , viz. the factors with distributions independent from location $\mathbf{s}=(\mathbf{x},\mathbf{y})'$, the model of controlled trend surface (CTS) was obtained as follows:

$$\mathbf{Z}(\mathbf{s})=\boldsymbol{\mu}(\mathbf{s})+\boldsymbol{\zeta}(\mathbf{n})+\boldsymbol{\zeta}, \quad \mathbf{s}=(\mathbf{x},\mathbf{y})' \in \mathbf{D} \subset \mathbb{R}^2 \quad (2)$$

with $\boldsymbol{\zeta}(\mathbf{n})$ being the nonspatial component— an unknown combination of functions of non-spatial factors and $\boldsymbol{\zeta}$ representing a zero-mean, stationary error term with known covariogram independent from both spatial and non-spatial factors. Apparently, under-parameterized conventional trend surface estimates were biased when nonspatial effects were present. In this case, CTS model (Eq. 2) was appropriate and provided unbiased estimates.

MODEL DESCRIPTION

A conventional Matrix Model (e.g. Buongiorno and others 1995; Liang and others 2005) predicts the forest stand state in Year $t+1$ based on the stand state in Year t :

$$\mathbf{y}_{t+1}=\mathbf{G}\mathbf{y}_t+\mathbf{R}+\boldsymbol{\varepsilon} \quad (3)$$

where $\mathbf{y}_t=[y_{ijt}]$ was a column vector representing the number of live trees per unit of land area of species i and diameter class j at time t . $\boldsymbol{\varepsilon}$ was a random error. \mathbf{G} and \mathbf{R} represented a spatial-independent growth matrix and recruitment vector, respectively.

The CTS Matrix Model extended Eq. 3 to control for the large-scale spatial trend by recognizing geographic location and terrain characteristics of the stand:

$$\mathbf{V}_{t+1}(\mathbf{s})=\mathbf{G}(\mathbf{s})\mathbf{V}_t(\mathbf{s})+\mathbf{R}(\mathbf{s})+\boldsymbol{\varepsilon} \quad (4)$$

$$\mathbf{s}=(\mathbf{x},\mathbf{y})' \in \mathbf{D} \subset \mathbb{R}^2$$

where $\mathbf{V}_t(\mathbf{s})=[v_{ijt}(\mathbf{s})]$ was a space-dependent column vector representing the number of live trees per unit of land area of species i ($i=1,\dots,4$) and diameter class j ($j=1,\dots,19$) at location \mathbf{s} and at time t . $\boldsymbol{\varepsilon}$ was a zero-mean, stationary process with known covariogram. \mathbf{x} and \mathbf{y} represented the plot coordinates within the Alaska boreal forest region \mathbf{D} in the plane (\mathbb{R}^2).

$\mathbf{G}(\mathbf{s})$ was a state- and space-dependent matrix that described how trees grew or died between t and $t+1$ at location \mathbf{s} . $\mathbf{R}(\mathbf{s})$ was a state- and space-dependent vector representing the recruitment of each species between t and $t+1$ at location \mathbf{s} .

The $\mathbf{G}(\mathbf{s})$ and $\mathbf{R}(\mathbf{s})$ matrices were defined as:

$$\mathbf{G}(\mathbf{s}) = \begin{bmatrix} \mathbf{G}_1(\mathbf{s}) & & & & \\ & \mathbf{G}_2(\mathbf{s}) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \mathbf{G}_m(\mathbf{s}) \end{bmatrix}, \mathbf{G}_i(\mathbf{s}) = \begin{bmatrix} a_{i1}(\mathbf{s}) & & & & \\ b_{i1}(\mathbf{s}) & a_{i2}(\mathbf{s}) & & & \\ & & \ddots & & \\ & & & b_{i,m-2}(\mathbf{s}) & a_{i,m-1}(\mathbf{s}) \\ & & & & b_{i,m-1}(\mathbf{s}) & a_{im}(\mathbf{s}) \end{bmatrix} \quad (5)$$

$$\mathbf{R}(\mathbf{s}) = \begin{bmatrix} \mathbf{R}_1(\mathbf{s}) \\ \mathbf{R}_2(\mathbf{s}) \\ \vdots \\ \mathbf{R}_m(\mathbf{s}) \end{bmatrix}, \mathbf{R}_i(\mathbf{s}) = \begin{bmatrix} R_i(\mathbf{s}) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $R_i(\mathbf{s})$ was the number of trees of species i recruited in the smallest diameter class (3.8cm) each year at location \mathbf{s} . Recruitment was zero in all the higher diameter classes. The probabilities of a tree of species i and diameter class j stayed alive in the same diameter class $a_{ij}(\mathbf{s})$, and stayed alive and move up a diameter class $b_{ij}(\mathbf{s})$ between t and $t+1$ at location \mathbf{s} were related by:

$$a_{ij}(\mathbf{s})=1-b_{ij}(\mathbf{s})-m_{ij}(\mathbf{s}) \quad (6)$$

where $m_{ij}(\mathbf{s})$ was the probability that a tree of species i and diameter class j died between t and $t+1$ at location \mathbf{s} . $b_{ij}(\mathbf{s})$ was calculated as the annual tree diameter growth $g_{ij}(\mathbf{s})$ divided by the width of the diameter class (2 cm except for the first diameter class of 1.2 cm width), assuming that trees were evenly distributed in a diameter class.

It was assumed that the large-scale spatial trend $\boldsymbol{\mu}(\mathbf{s})$ was represented by a second-order polynomial function of northing (y) and easting (x) coordinates:

$$\boldsymbol{\mu}(\mathbf{s}) = d_0 + d_1x + d_2y + d_3x^2 + d_4y^2 + d_5xy \quad (7)$$

where d 's were coefficients to be estimated in each equation. Northing and easting coordinates of the Universal Transverse Mercator system (UTM, see Snyder 1987) were used here to approximate the Cartesian system in which the distance between permanent sample plots could be easily calculated (Ripley 1981). The easting values were then set as the absolute distance from the center of that UTM zone to mitigate edge effects near borders.

The non-spatial component of the recruitment, $R_i(\mathbf{s})$, diameter growth $g_{ij}(\mathbf{s})$, and mortality $m_{ij}(\mathbf{s})$ was composed of a terrain function and stand basal area (B), permafrost (P), and the number of tree species present in the plot (H), as B and P have been employed as key predictors in many existing forest dynamics models (e.g. Boltz and Carter 2006;

Bonan and Shugart 1989; Liang and others 2005; Namaalwa and others 2005), and H represented marked differences in species life histories with effects of complementarity and niche facilitation that may change forest dynamics (Liang and others 2007). In addition, D_j , the midpoint of the DBH class j , was used in both diameter growth and mortality equations, as tree size is an important factor of diameter growth and mortality (Buongiorno and Michie 1980; Buongiorno and others 1995; Liang and others 2005). Stem density, the number of trees per hectare of the species of interest (N), was used in recruitment equation to represent the abundance of seeds and seedlings (Liang and others 2005; Liang and others 2007). Although the presence of permafrost (P) was significantly correlated with the northing ($\rho=0.18$, p -value=0.00), since the correlation coefficient was small and the effect of permafrost on forest growth is local (Chapin and others 2006), permafrost (P) was considered as a non-spatial variable. None of the other non-spatial variables was spatially correlated.

The terrain component (Eq. 8) represented the interacting effects of the slope (l), aspect (a), and elevation (z) on site productivity (see Stage and Salas (2007), p.487). The function has been tested a better proxy of site productivity than other existing terrain functions, and is considered as an inseparable entity, in which all the terms are conjoint and should be used together or not at all (Stage and Salas 2007).

$$\tau(l, a, z) = c_0 + l[c_1 + c_2 \cos(a) + c_3 \sin(a)] + \ln(z+1) \cdot [c_4 + c_5 \cos(a) + c_6 \sin(a)] + z^2 \cdot [c_7 + c_8 \cos(a) + c_9 \sin(a)] + c_{10}z + c_{11}z^2 \quad (8)$$

where c 's were parameters to be estimated in each equation.

The annual diameter growth $g_{ij}(\mathbf{s})$ was estimated by the following model:

$$g_{ij}(\mathbf{s}) = \gamma_0 + \gamma_1 D + \gamma_2 D^2 + \gamma_3 D^3 + \gamma_4 B + \gamma_5 P + \gamma_6 H + \tau(l, a, z) + \mu(\mathbf{s}) + \varepsilon_1 \quad (9)$$

where γ 's were parameters to be estimated, and ε_1 was a random error independent of spatial patterns.

The probability of annual mortality rate, $m_{ij}(\mathbf{s})$ was calculated by dividing $M_{ij}(\mathbf{s})$ by the elapsed time of T years between the two inventories. $M_{ij}(\mathbf{s})=1$ if a tree died between the two inventories, and $M_{ij}(\mathbf{s})=0$ otherwise. $M_{ij}(\mathbf{s})$ was estimated with a species- and size-dependent Probit function (Bliss 1935):

$$M_{ij}(\mathbf{s}) = \Phi(\delta_0 + \delta_1 D + \delta_2 D^2 + \delta_3 D^3 + \delta_4 B + \delta_5 P + \delta_6 H + \tau(l, a, z) + \mu(\mathbf{s})) + \varepsilon_2 \quad (10)$$

where Φ was the standard normal cumulative function, and δ 's were parameters and ε_2 was a random error independent of spatial patterns.

The expected recruitment of species i was estimated with the following model:

$$R_i(\mathbf{s}) = \beta_0 + \beta_1 N_i + \beta_2 B + \beta_3 P + \beta_4 H + \tau(l, a, z) + \mu(\mathbf{s}) + \varepsilon_3 \quad (11)$$

where β 's were the parameters, and ε_3 was a random error independent of spatial patterns.

DATA

The CTS Matrix Model presented here was calibrated with data from 446 remeasured permanent sample plots of the Cooperative Alaska Forest Inventory (CAFI) (Malone and others 2009). The sample area stretches over 500km from the Kenai Peninsula in the south to the Fairbanks area in the north, and represents a wide range of stand conditions and species composition (Fig. 1). The same data, except for geographic coordinates, have been used to calibrate the nonspatial Matrix Model of Liang (2010) (Table 1).

The species studied here were *Betula neoalaskana* Sarg. (birch), *Populus tremuloides* Michx. (aspen), *Picea glauca* (Moench) Voss (white spruce), and *Picea mariana* (Mill.) B.S.P. (black spruce). White spruce had the highest basal area of all the species (37 percent), followed by birch (28 percent), aspen (20 percent), and black spruce (5 percent). The other species, *Populus trichocarpa* Torr. & Gray, *P. balsamifera* L. *Larix laricina* (DuRoi) K.Koch, and *Betula kenaica* W.H. Evans, accounted for less than 10 percent of the total basal area (Table 2). Trees were grouped into 19 diameter classes by species, from 3.8 to 5.0 cm up to 39.0 cm and above. Tables 3 and 4 display the summary statistics of plot level and individual tree variables.

PARAMETER ESTIMATION AND MODEL VALIDATION

The recruitment $R_i(\mathbf{s})$ and diameter growth $g_{ij}(\mathbf{s})$ equations were estimated by the generalized least squares (GLS) method (Rao 1973), and a generalized coefficient of determination (Nagelkerke 1991) was calculated for each equation as a proxy for the common coefficient of determination. Mortality $m_{ij}(\mathbf{s})$ was a Probit function (Bliss 1935) estimated with maximum likelihood.

To avoid compromised type-I error rates and severe artifacts commonly associated with model selection procedures (Mac Nally 2000), predictive variables were selected with three criteria: the expected biological responses, the statistical significance, and the contribution to the model goodness-of-fit. In this study, we used the hierarchical partitioning or HP (Chevan and Sutherland 1991) to decompose the model goodness-of-fit represented by likelihood through incremental partitioning, and determined the average independent contribution of each variable to the overall goodness-of-fit. The HP analysis was conducted with the

hier.part package of the R program (Mac Nally and Walsh 2004).

The accuracy of this model was determined by the prediction errors, the differences between the observed stand states of the third inventory and the predicted ones, on two phases of validation plots. Phase I plots were 175 CAFI sample plots on which a third inventory has been conducted 10 years after the first inventory, solely for the validation purpose. Phase II was consisted of 40 Forest Inventory and Analysis (FIA) plots located on the boreal transitional zone on the Kenai Peninsula outside the current sample area, and no data from these plots were used to calibrate the CTS model. Phase I and II plots represented a temporal and spatial extension of the current sample coverage, respectively (Fig. 1). For each Phase I plot, the expected number of trees at the third inventory was predicted by setting the stand state at the first inventory as the initial state, and applying Eq. 4 iteratively over 10 years. For each Phase II plot, the expected number of trees at the second inventory was predicted by applying Eq. 4 iteratively over the specific interval of that plot, averaging 4.78 years across all the plots.

For comparison, we also predicted the stand states of both Phase I and II plots with the nonspatial Matrix Model (Liang 2010) and a conventional trend surface Matrix Model in which recruitment, diameter growth, and mortality were equations of second-order trend surfaces only (Eq. 7). Both models were calibrated with data from the same 446 sample plots. For each model, root mean squared errors (RMSE, see Wooldridge 2000, P.600) were calculated based on the difference between the predicted and observed basal area by diameter class and species as a measure of accuracy of that model over the validation plots.

RESULTS

MODEL PARAMETERS

For recruitment $R(\mathbf{s})$, the total number of trees (N) and stand basal area (B) were the most significant control variables, and their effects on recruitment were consistent over all the species (Table 5). When regarded as an entity, the spatial component was significant for all the species in recruitment, and so was the terrain component. Generally, N contributed most to the goodness-of-fit of recruitment (67~82 percent), followed by P and H (2~14 percent). The spatial component contributed 3~9 percent, and the terrain variables 8~16 percent. B contributed little to the goodness-of-fit (2~6 percent), albeit its high level of significance (Table 6).

In the diameter growth model, all the control variables were significant, except basal area (B) for aspen and black spruce, permafrost (P) for birch, and species diversity (H) for birch and white spruce (Table 7). The spatial and

terrain components were both highly significant (Table 6). Generally, the diameter (D) contributed the most to the overall goodness-of-fit of diameter growth (4~71 percent), followed by the terrain (12~33 percent) and spatial component (4~25 percent, Table 6).

In the mortality model, all the control variables were significant, except basal area (B) for aspen and black spruce, permafrost (P) for birch, and species diversity (H) for birch and white spruce (Table 8). Both the spatial and terrain components were highly significant (Table 6). Generally, the terrain component contributed the most to the overall goodness-of-fit of mortality (22~55 percent), followed by the diameter (3~51 percent) and spatial component (12~33 percent, Table 6).

VALIDATION AND RESIDUALS

Over the 175 Phase I validation plots, the stand basal area predicted by the CTS model was generally accurate over all species and size, as they all fell within 95 percent confidence interval of the observed ones, except for the smallest black spruce (Fig. 2). The nonspatial model was quite close to the CTS model in terms of predictions over the Phase I plots, and the conventional trend surface model underestimated aspen and overestimated black spruce in general. Compared to the nonspatial model, the CTS model was 7.88, 20.73, 22.28, and 11.00 percent more accurate in terms of RMSE for birch, aspen, white spruce, and black spruce, respectively. The CTS model was also 2.98, 18.41, 22.32, and 16.88 percent more accurate than the conventional trend surface model for the four species in terms of RMSE (Fig. 2).

The accuracy of the CTS model was more prevalent for deciduous species over the 40 Phase II validation plots. The CTS model was 21.41, 64.10, 7.24, and 3.70 percent more accurate in terms of RMSE than the nonspatial model for birch, aspen, white spruce, and black spruce, respectively. When compared with the conventional trend surface model, the CTS model was more than 60 percent more accurate for deciduous species, and 13.62 percent more accurate for white spruce. The CTS model was 7.98 percent less accurate than the conventional trend surface model for black spruce, but the difference was negligible especially for forest management purposes as most errors of the CTS model occur in the smallest diameter class (Fig. 3).

SPATIAL INFERENCE

Using the method in Liang and Zhou (2010), we created maps of the predicted future Alaska boreal forest. The map of the predicted stand basal area change in the Year 2011, 2051, and 2101 shows that without major disturbances and substantial changes of climate conditions, the total stand basal area would keep increasing over time for most of the

region (Fig. 4). The Yukon River Basin and Copper River Valley were predicted to have the best basal area growth. The Matanuska-Susitna Valley, Kuskokwim River Basin, and some sporadic areas, such as Nenana and Healy, on the contrary, would see a decline in the basal area. Overall, the stand basal area may increase in the central and eastern region, while some negative basal area change may occur in the southern and western region. The magnitude of changes over the entire region slightly increased over time. Between the year 2001, 2011, 2051, and 2101, the average annual basal area change was 0.20, 0.27 and 0.33 m²/ha/y, respectively. The prediction implies that under current conditions, the total basal area of the Alaska boreal forest may become higher at an increasing rate for the Twenty-First Century.

Current distribution of dominant species throughout the region was predicted to remain the same until the Year 2051, and a large portion of the deciduous forests may switch to coniferous forests thereafter (Fig. 5). In the Year 2101, without major disturbances and catastrophes, more than 90 percent of the forest located between 62°N and 66°N was predicted to be coniferous, while at present, most of the coniferous forests are clustered in the Copper River Valley and the area to the southeast of Fairbanks. The Porcupine River Valley and eastern Mat-Su Valley, however, may continue to be covered by deciduous forests in a century, according to the model (Fig. 5).

CONCLUSION

This paper proposes a method of Controlled Trend Surface to simultaneously account for large-scale spatial trends and nonspatial local effects. By incorporating well-recognized nonspatial factors, CTS would be particularly useful for studying biological and ecological processes, such as forest growth and fish habitat alteration, where spatial patterns and effects of local variables were both important, and predictions were needed over areas of considerable sizes. With this method, a geospatial model of forest dynamics was developed for the Alaska boreal forest, based on a large and representative dataset which covers a wide range of forests, from lowland monospecific coniferous stands to upland uneven-aged hardwood stands. The CTS model was in general more accurate for all the species than the nonspatial model (Liang 2010) and the conventional trend surface model, both of which were calibrated with the same data, over the 175 Phase I and 40 Phase II validation plots.

The CTS model was beyond traditional stand growth models because its geospatial component represented trends of forest dynamics on a large spatial scale, likely caused by the spatial variation of temperature and precipitation and other unknown factors. Therefore, this model would be more

useful than traditional stand growth models to predict forest dynamics over the entire Alaska boreal region. Although it was a bold extrapolation, of which the accuracy remained to be assessed for most locations outside the sample area, the predictions were generally consistent with previous knowledge and offered a striking illustration of the potential power of including spatial and topographic information in forest dynamics models.

LITERATURE CITED

- AlaskaDNR.** 2006. Alaska State Forester's Office Annual Report. Anchorage, AK: Alaska Department of Natural Resources, Division of Forestry. 70 p.
- Berke O.** 1999. Estimation and prediction in the spatial linear model. *Water, Air, and Soil Pollution* 110:215-237.
- Bliss CI.** 1935. The calculation of the dosage-mortality curve. *Annals of Applied Biology* 22:134-167.
- Boltz F,** and Carter DR. 2006. Multinomial logit estimation of a matrix growth model for tropical dry forests of eastern Bolivia. *Canadian Journal of Forest Research* 36(10):2623-2632.
- Bonan GB,** and Shugart HH. 1989. Environmental factors and ecological processes in boreal forests. *Annual Review of Ecology and Systematics* 20:1-28.
- Buongiorno J,** and Michie BR. 1980. A matrix model of uneven-aged forest management. *Forest Science* 26(4):609-625.
- Buongiorno J,** Peyron JL, Houllier F, and Bruciamacchie M. 1995. Growth and management of mixed-species, uneven-aged forests in the French Jura: Implications for economic returns and tree diversity. *Forest Science* 41(3):397-429.
- Chapin FSI,** Oswood MW, Cleve KV, Viereck LA, and Verbyla DL. 2006. Alaska's changing boreal forest. Chapin FSI, Oswood MW, Cleve KV, Viereck LA, and Verbyla DL, editors. Oxford: Oxford University Press.
- Chevan A,** and Sutherland M. 1991. Hierarchical partitioning. *American Statistician* 45:90-96.
- Franklin J,** Michaelsen J, and Strahler AH. 1985. Spatial analysis of density dependent pattern in coniferous forest stands. *Plant Ecology* 64(1):29-36.
- Gittins R.** 1968. Trend-surface analysis of ecological data. *Journal of Ecology* 56(3):845-869.
- Grant F.** 1957. A problem in the analysis of geophysical data. *Geophysics* 22:309-344.
- Kuuluvainen T,** and Sprugel DG. 1996. Examining age- and altitude-related variation in tree architecture and needle efficiency in Norway spruce using trend surface analysis. *Forest Ecology and Management* 88:237-247.
- Larson JA,** Franklin FJ, and Franklin J. 2006. Structural segregation and scales of spatial dependency in *Abies amabilis* forests. *Journal of Vegetation Science* 17(4):489-498.

- Liang J.** 2010. Dynamics and management of Alaska boreal forest: an all-aged multi-species Matrix growth model. *Forest Ecology and Management* 260(4):491-501.
- Liang J, Buongiorno J, and Monserud RA.** 2005. Growth and yield of all-aged douglas-fir/western hemlock stands: a Matrix model with stand diversity effects. *Canadian Journal of Forest Research* 35:2369-2382.
- Liang J, Buongiorno J, Monserud RA, Kruger EL, and Zhou M.** 2007. Effects of diversity of tree species and size on forest basal area growth, recruitment, and mortality. *Forest Ecology and Management* 243:116-127.
- Liu J, and Ashton PS.** 1998. FORMOSAIC: an individual-based spatially explicit model for simulating forest dynamics in landscape mosaics. *Ecological Modelling* 106:177-200.
- Mac Nally R.** 2000. Regression and model-building in conservation biology, biogeography and ecology: The distinction between – and reconciliation of – ‘predictive’ and ‘explanatory’ models. *Biodiversity and Conservation* 9 (5):655-671.
- Mac Nally R, and Walsh C.** 2004. Hierarchical partitioning public-domain software. *Biodiversity Conservation* 13:659-660.
- Malone T, Liang J, and Packee. EC.** 2009. Cooperative Alaska Forest Inventory. Portland, OR: Gen. Tech. Rep. PNW-GTR-785, USDA Forest Service, Pacific Northwest Research Station. 42 p.
- Nagelkerke N.** 1991. A note on a general definition of the coefficient of determination. *Biometrika* 78(3):691-692.
- Namaalwa J, Eid T, and Sankhayan P.** 2005. A multi-species density-dependent matrix growth model for the dry woodlands of Uganda. *Forest Ecology and Management* 213:312-327.
- Pacala SW, Canham CD, Saponara J, Silander JA, Kobe RK, and Ribbens E.** 1996. Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecological Monographs* 66:1-43.
- Rao CR.** 1973. *Linear statistical inference and its applications.* New York: John Wiley. 626 p.
- Ripley BD.** 1981. *Spatial statistics.* Hoboken, NJ. : John Wiley & Sons, Inc. . 252 p.
- Schenk HJ.** 1996. Modeling the effects of temperature on growth and persistence of tree species: A critical review of tree population models. *Ecological Modelling* 92:1-32.
- Snyder JP.** 1987. *Map projections - a working manual.* U.S. Geological Survey Professional Paper ed. Washington, D.C. : United States Government Printing Office. p 385.
- Stage AR, and Salas C.** 2007. Interaction of elevation, aspect, and slope in models of forest species composition and productivity. *Forest Science* 53(4):486-492.
- Thomson AJ.** 1986. Trend surface analysis of spatial patterns of tree size, microsite effects, and competitive stress. *Canadian Journal of Forest Research* 16:279-282.
- Watson GS.** 1971. Trend surface analysis. *Mathematical Geology* 3:215-226.
- Wilmking M, and Juday GP.** 2005. Longitudinal variation of radial growth at Alaska’s northern treeline-recent changes and possible scenarios for the 21st century. *Global and Planetary Change* 47:282-300.
- Wooldridge JM.** 2000. *Introductory econometrics: a modern approach.* Cincinnati, OH: South-Western College Publishing.
- Wurtz TL, Ott RA, and Maisch JC.** 2006. Timber harvest in interior Alaska. In: Chapin FS, Oswood MW, Van Cleve K, Viereck LA, and Verbyla DL, editors. *Alaska’s changing boreal forest.* New York, NY: Oxford University Press. p 302-308.

Table 1—Definition of variables

Variable	Definition
Tree-level variables	
D	Diameter at breast height (cm) of a live tree
g	Annual diameter increment (cm) of a live tree
Plot-level variables	
R_i	Annual recruitment, the number of trees grew into the smallest diameter class (3.8 to 5.0 cm) of species i in a year
N_i	Total number of trees per hectare of species i
B	Stand basal area (m^2/ha)
P	Permafrost. A coded variable representing the likelihood of permafrost on site, where one stands for 90% likely, two 60% likely, three 30% likely, and four most unlikely (0%)
H	Number of tree species present on a plot
z	Plot elevation (km)
l	Plot slope (%)
α	Plot aspect showing the direction to which the plot slope faces ($^{\circ}$). 0 means no slope, 180 and 360 represented south- and north-facing slopes, respectively.
x	Easting of UTM coordinates (10^6m)
y	Northing of UTM coordinates (10^6m)

Table 2—Distribution of total basal area by species in the sample plots

Common Name	Shortened Name	Scientific Name	Percentage
white spruce	white spruce	<i>Picea glauca (Moench) Voss</i>	37.40
Alaska birch	birch	<i>Betula neoalaskana Sarg.</i>	27.83
quaking aspen	aspen	<i>Populus tremuloides Michx.</i>	20.27
black spruce	black spruce	<i>Picea mariana (Mill.) B.S.P.</i>	4.99
Other species			9.51
Total			100.00

Table 3—Summary statistics of plot-level variables, based on 446 sample plots

	N (trees*ha ⁻¹)				B (m ² ha ⁻¹)	P	H
	Birch	Aspen	White spruce	Black spruce			
Mean	336.35	286.82	651.20	281.56	22.91	3.33	2.32
S.D.	31.73	32.13	44.36	51.69	0.49	0.04	0.04
Max	5955.03	4867.80	8771.93	12700.77	63.43	4.00	5.00
Min	0.00	0.00	0.00	0.00	0.00	1.00	0.00
	Recruitment (trees*ha ⁻¹ *y ⁻¹)				z (km)	s (%)	α (°)
	Birch	Aspen	White spruce	Black spruce			
Mean	5.60	2.49	27.92	32.30	0.36	10.17	146.41
S.D.	0.97	0.69	2.63	6.48	0.01	0.60	5.09
Max	197.68	222.39	444.77	1161.35	0.96	77.00	360.00
Min	0.00	0.00	0.00	0.00	0.02	0.00	0.00

Table 4—Summary statistics for individual tree data

	Birch	Aspen	White spruce	Black spruce
Diameter (cm)				
Mean	13.13	12.30	10.52	6.12
S.D.	7.57	6.03	7.23	3.90
Max	59.49	53.29	85.39	30.71
Min	3.80	3.80	3.80	3.80
n	6080	5206	11677	4862
Diameter growth (cm*y ⁻¹)				
Mean	0.10	0.08	0.11	0.09
S.D.	0.12	0.08	0.12	0.11
Max	1.55	0.81	2.50	1.82
Min	-3.99	-0.62	-2.27	-2.20
n	6080	5206	11677	4862
Mortality Rate (y ⁻¹)				
Mean	0.02	0.03	0.01	0.01
S.D.	0.06	0.07	0.04	0.03
Max	0.20	0.20	0.20	0.20
Min	0.00	0.00	0.00	0.00
n	6885	6011	12161	5014

Table 5—Parameters of the recruitment equation

Explanatory Variables	Species			
	Birch	Aspen	White spruce	Black spruce
<i>Constant</i>	-10278.00 **	3096.00	-5290.00	-511.00
β_1	0.01 ***	0.01 ***	0.03 ***	0.11 ***
β_2	-0.36 ***	-0.12 *	-0.68 ***	-1.05 ***
β_3	0.96	-900.30	1659.00	140.00
β_4	3.09 **	1384.00	-7453.00	1030.00
<i>Spatial component</i>				
d_1	2922.00 **	65.38	-129.00	-9.30
d_2	3555.00	-324.70	49.00	-2703.00
d_3	-207.66 **	-188.50	1036.20	-98.00
d_4	-793.20	-0.91	6.66 **	-0.25
d_5	-495.50	1.57 *	2.96	0.30
<i>Terrain component</i>				
c_1	-0.17	-0.07	0.75	1.38
c_2	-0.51	-0.02	0.14	0.15
c_3	-0.11	-0.23	-0.75	-1.75
c_4	1.04	-0.10	-4.78	-6.48
c_5	4.10 **	0.50	-1.26	0.03
c_6	0.40	0.71	4.26	8.02
c_7	-0.91	0.08	3.15	2.52
c_8	-3.62 **	-0.64	0.62	-1.66
c_9	-0.13	-0.34	-3.65	-5.29
c_{10}	3.28	-5.29	96.85 *	87.99
c_{11}	-9.50	6.49	-127.59 **	-95.73
R^2	0.19	0.19	0.35	0.79
n	446	446	446	446

Note:

-Dependent variable =stand recruitment in trees·ha⁻¹·y⁻¹.

- R^2 = generalized coefficient of determination.

- n = degrees of freedom.

-Level of significance: *: $P < 0.10$; **: $P < 0.05$; ***: $P < 0.01$.

-The complete model is:

$$\begin{aligned}
 R_i(\mathbf{s}) = & \beta_0 + \beta_1 N(\mathbf{s}) + \beta_2 B(\mathbf{s}) + \beta_3 P(\mathbf{s}) + \beta_4 H(\mathbf{s}) + (d_1 x + d_2 y + d_3 x^2 + d_4 y^2 + d_5 x \cdot y) \\
 & + l(\mathbf{s})[c_1 + c_2 \cos(\alpha(\mathbf{s})) + c_3 \sin(\alpha(\mathbf{s}))] + \ln(z(\mathbf{s}) + 1) \cdot l(\mathbf{s})[c_4 + c_5 \cos(\alpha(\mathbf{s})) + c_6 \sin(\alpha(\mathbf{s}))] \\
 & + z(\mathbf{s})^2 \cdot l(\mathbf{s})[c_7 + c_8 \cos(\alpha(\mathbf{s})) + c_9 \sin(\alpha(\mathbf{s}))] + c_{10} z(\mathbf{s}) + c_{21} z(\mathbf{s})^2
 \end{aligned}$$

Table 6—Percentage contribution (%) to the overall goodness-of-fit and the level of significance of variables and components

	Species							
	Birch		Aspen		White spruce		Black spruce	
	Recruitment (trees ha ⁻¹ *y ⁻¹)							
Spatial component	3.56	***	6.23	**	8.77	***	2.53	*
Stem density	68.44	***	68.95	***	67.46	***	82.00	***
Stand basal area	6.32	***	1.69	*	2.73	***	4.88	***
Terrain component	8.02	***	8.77	*	16.21	***	8.41	*
Others (<i>P, H</i>)	13.67	**	14.35	*	4.83	*	2.20	
All	100.00	***	100.00	***	100.00	***	100.00	***
	Diameter Growth (cm*y ⁻¹)							
Spatial component	10.18	***	4.30	***	12.28	***	24.89	***
Diameter	58.25	***	70.96	***	34.19	***	4.20	***
Stand basal area	13.97	***	5.38	***	10.72	***	4.60	***
Terrain component	12.45	***	13.80	***	17.85	***	32.59	***
Others (<i>P, H</i>)	5.14	***	5.55	***	24.96	***	33.71	***
All	100.00	***	100.00	***	100.00	***	100.00	***
	Mortality (y ⁻¹)							
Spatial component	25.32	*	13.48	***	32.62	***	12.26	***
Diameter	43.10	***	50.60	***	7.09	***	3.13	***
Stand basal area	1.70	***	8.27		7.30	***	2.22	
Terrain component	22.61	***	22.31	***	44.94	***	55.22	***
Others (<i>P, H</i>)	7.27		5.33	***	8.05	***	27.17	***
All	100.00	***	100.00	***	100.00	***	100.00	***

Note:

-Level of significance: *: $P < 0.10$; **: $P < 0.05$; ***: $P < 0.01$.

-Due to the limit of computing capacity, percentage contribution (%) to the overall goodness-of-fit is approximated with the following terms by the hierarchical partitioning method:

$D, B, P, H, x, y, l, l \cdot \cos(\alpha), l \cdot \sin(\alpha), z, z^2$.

Table 7—Parameters of the diameter growth equation

	Species			
	Birch	Aspen	White spruce	Black spruce
<i>Constant</i>	33.347 ***	-40.126 ***	-3.835	81.300 ***
γ_1	0.022 ***	0.017 ***	0.014 ***	0.014 ***
γ_2	-0.754 ***	-0.412 ***	-0.381 ***	-1.271 ***
γ_3	8.457 ***	4.236 ***	3.242 ***	28.367 ***
γ_4	-0.003 ***	-0.002	-0.003 ***	0.000
γ_5	0.006	0.004 **	0.032 ***	0.027 ***
γ_6	-0.003	-0.001 ***	0.001	0.014 ***
<i>Spatial component</i>				
d_1	-9.150 ***	11.184 ***	0.965	-22.802 ***
d_2	-19.931 **	35.540 ***	7.109 **	-38.110 ***
d_3	0.628 ***	-0.779 ***	-0.060	1.598 ***
d_4	3.992	-9.245 ***	-5.985 ***	9.880 ***
d_5	2.704 **	-4.906 ***	-0.888 **	5.167 **
<i>Terrain component</i>				
c_1	0.001 *	-0.001	-0.002 ***	-0.002
c_2	-0.001 ***	0.002 ***	-0.002 ***	-0.003 **
c_3	0.002 **	0.000	0.001 **	-0.005 ***
c_4	-0.014 ***	0.012 ***	0.007 ***	-0.012
c_5	0.003	-0.002	0.004 *	0.006
c_6	-0.010 **	-0.009 **	-0.005 *	0.023 **
c_7	0.009	-0.014 ***	-0.003	0.016 *
c_8	-0.002	-0.003	-0.004 *	0.000
c_9	0.012 ***	0.011 ***	0.003	-0.016 *
c_{10}	0.180 ***	-0.217 ***	0.024	0.190 **
c_{11}	-0.241 ***	0.275 ***	-0.147 ***	-0.255 **
R^2	0.16	0.30	0.19	0.09
n	6079	5205	11676	4861

Note:

-Dependent variable =diameter increment in $\text{cm}\cdot\text{y}^{-1}$.

- R^2 = generalized coefficient of determination.

- n = degrees of freedom.

-Level of significance: *: $P<0.10$; **: $P<0.05$; ***: $P<0.01$.

-The complete model is:

$$g_{ij}(\mathbf{s}) = \gamma_0 + \gamma_1 D(\mathbf{s}) + \gamma_2 D^2(\mathbf{s}) + \gamma_3 D^3(\mathbf{s}) + \gamma_4 B(\mathbf{s}) + \gamma_5 P(\mathbf{s}) + \gamma_6 H(\mathbf{s}) + (d_1 x + d_2 y + d_3 x^2 + d_4 y^2 + d_5 x \cdot y) + l(\mathbf{s})[c_1 + c_2 \cos(\alpha(\mathbf{s})) + c_3 \sin(\alpha(\mathbf{s}))] + \ln(z(\mathbf{s}) + 1) \cdot l(\mathbf{s})[c_4 + c_5 \cos(\alpha(\mathbf{s})) + c_6 \sin(\alpha(\mathbf{s}))] + z(\mathbf{s})^2 \cdot l(\mathbf{s})[c_7 + c_8 \cos(\alpha(\mathbf{s})) + c_9 \sin(\alpha(\mathbf{s}))] + c_{10} z(\mathbf{s}) + c_{21} z(\mathbf{s})^2$$

Table 8—Parameters of the mortality equation

	Species			
	Birch	Aspen	White spruce	Black spruce
<i>Constant</i>	-75.3843	-699.1010 ***	-327.0340 **	-212.7930
δ_1	-0.3089 ***	-0.3942 ***	-0.3292 ***	0.1429 *
δ_2	0.0118 ***	0.0142 ***	0.0195 ***	-0.0168 **
δ_3	-0.0001 ***	-0.0002 ***	-0.0003 ***	0.0004 **
δ_4	0.0081 **	-0.0060	0.0161 ***	-0.0121
δ_5	0.0559	-0.3102 ***	-0.1143 ***	-0.3169 ***
δ_6	-0.0109	0.0835 **	-0.0502	0.3124 ***
<i>Spatial component</i>				
d_1	24.5381	203.0630 ***	97.7920 **	75.7932
d_2	-182.5720	-240.8970 **	-169.2840 **	-945.6830 **
d_3	-1.9537	-14.6711 ***	-7.2783 **	-6.4794
d_4	34.2778	-21.6544	91.0388 ***	160.2830 *
d_5	25.2288	35.8469 ***	21.7540 *	131.0860 **
<i>Terrain component</i>				
c_1	-0.0027	-0.0476 ***	0.0032	0.0667
c_2	0.0072	-0.0154	-0.0070	-0.0303
c_3	0.0035	0.0121	0.0060	0.0983 **
c_4	0.2144 **	0.1727	0.0262	0.0651
c_5	0.1074	-0.0983	0.0657	0.0834
c_6	-0.0089	-0.1412	-0.2113 ***	-0.7993 *
c_7	-0.4002 ***	-0.0990	-0.0362	-0.4884
c_8	-0.2906 **	0.0995	-0.0471	-0.3213
c_9	0.0261	0.0699	0.2898 ***	1.0591 *
c_{10}	-1.4572	-4.2703 ***	-2.5270 ***	-2.2706
c_{11}	2.3799	2.4327	1.4191	-1.5049
R^2	0.17	0.16	0.14	0.12
n	6885	6011	12161	5014

Note:

-Dependent variable =mortality rate in y^{-1} .

- R^2 = McFadden's pseudo R-squared value.

- n = degrees of freedom.

-Level of significance: *: $P<0.10$; **: $P<0.05$; ***: $P<0.01$.

-The complete model was:

$$M_{ij}(\mathbf{s}) = \Phi(\delta_0 + \delta_1 D(\mathbf{s}) + \delta_2 D^2(\mathbf{s}) + \delta_3 D^3(\mathbf{s}) + \delta_4 B(\mathbf{s}) + \delta_5 P(\mathbf{s}) + \delta_6 H(\mathbf{s}) + (d_1 x + d_2 y + d_3 x^2 + d_4 y^2 + d_5 x \cdot y) + l(\mathbf{s})[c_1 + c_2 \cos(\alpha(\mathbf{s})) + c_3 \sin(\alpha(\mathbf{s}))] + \ln(z(\mathbf{s}) + 1) \cdot l(\mathbf{s})[c_4 + c_5 \cos(\alpha(\mathbf{s})) + c_6 \sin(\alpha(\mathbf{s}))] + z(\mathbf{s})^2 \cdot l(\mathbf{s})[c_7 + c_8 \cos(\alpha(\mathbf{s})) + c_9 \sin(\alpha(\mathbf{s}))] + c_{10} z(\mathbf{s}) + c_{21} z(\mathbf{s})^2$$

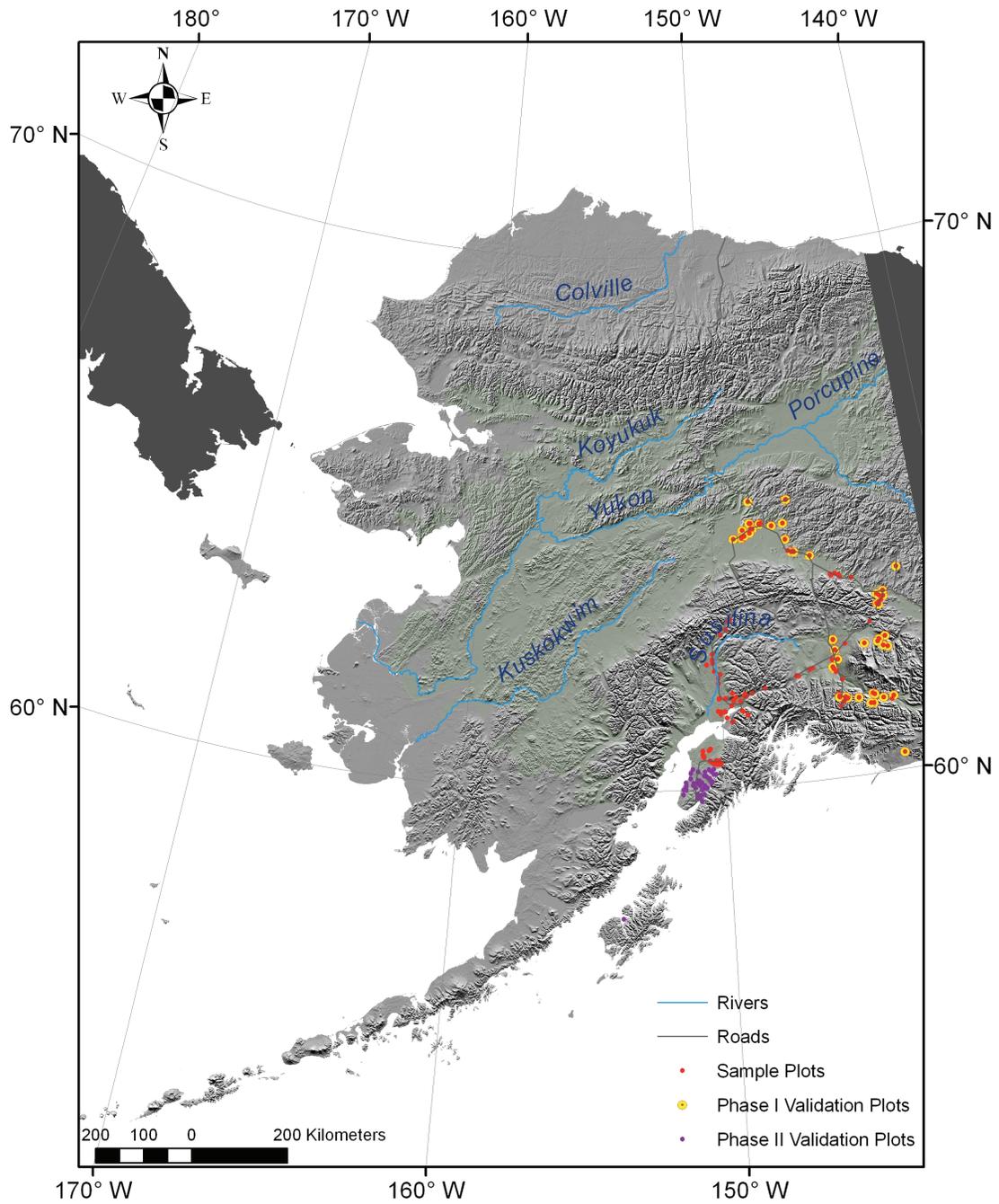


Figure 1—Geographic distribution of the sample and validation plots and their relative location in the Alaska boreal forest region (green area. Source: the U.S. Geological Survey Ecoregions Map of Alaska, <http://agdc.usgs.gov/data/projects/fhm/>). Albers equal area map projection with standard parallels.

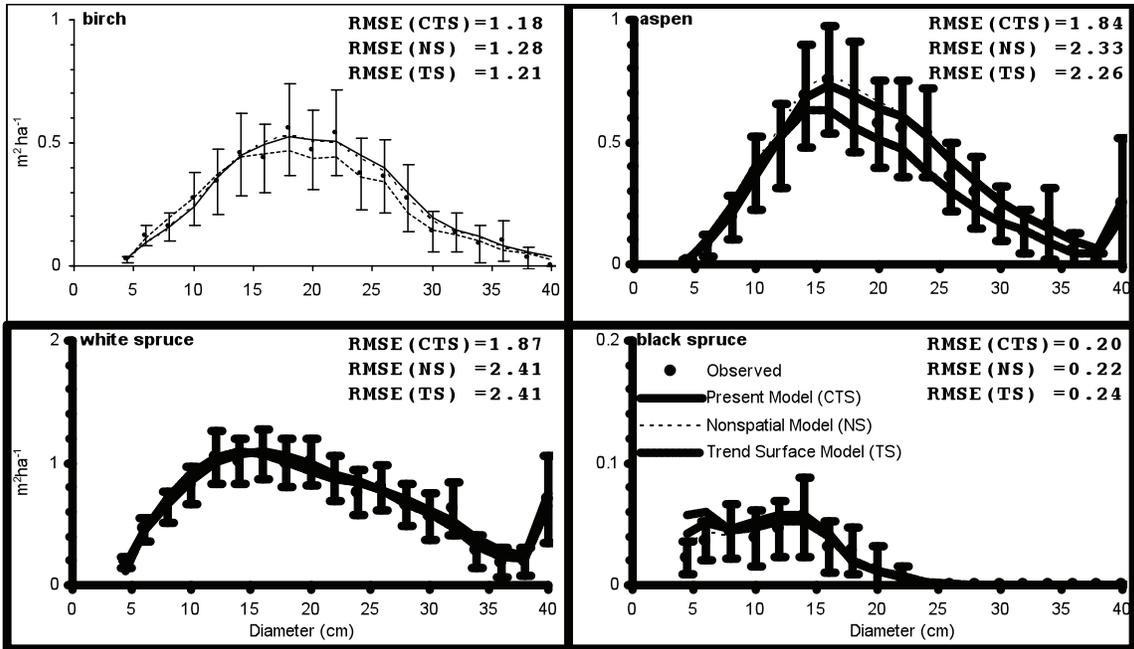


Figure 2—Average predicted and observed basal area by diameter class and species with 95 percent confidence interval over the 175 Phase I validation plots. Predictions were obtained with the present model (1), the nonspatial model (2), and the uncontrolled trend surface model (3). RMSE represents root mean squared errors calculated for that species by the three different models.

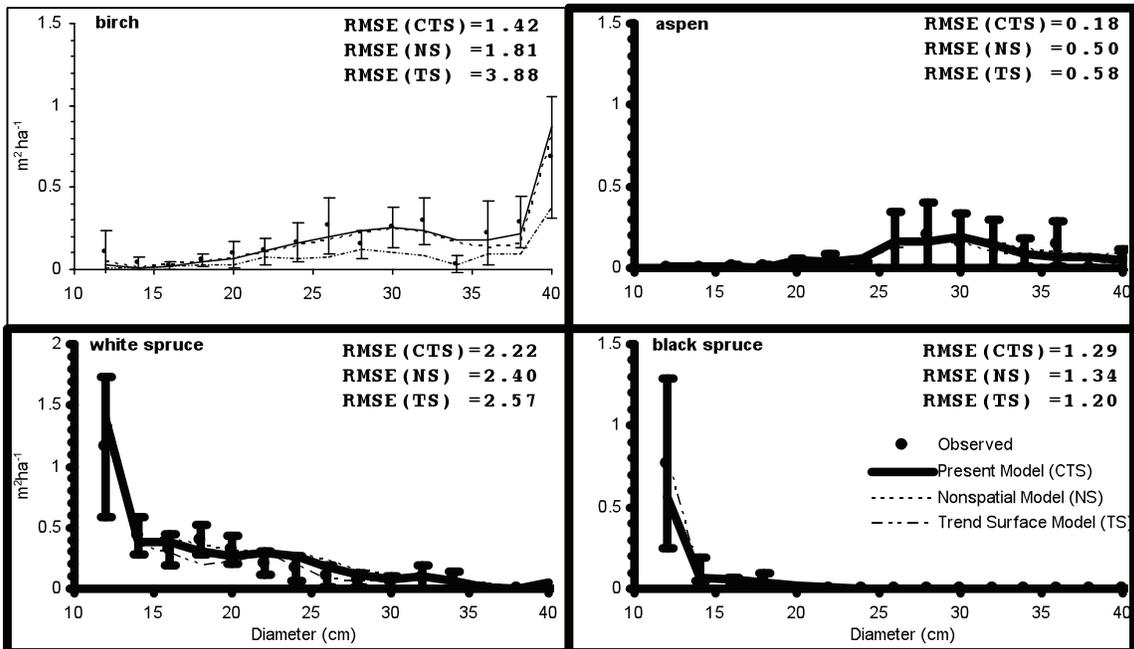


Figure 3—Average predicted and observed basal area by diameter class and species over the 40 Phase II validation plots. Vertical bars represented 90 instead of 95 percent confidence interval of observed values due to the small number of plots. Predictions were obtained with the present model (CTS), the nonspatial model (NS), and the conventional trend surface model (TS). RMSE represents root mean squared errors calculated for that species by the three different models.

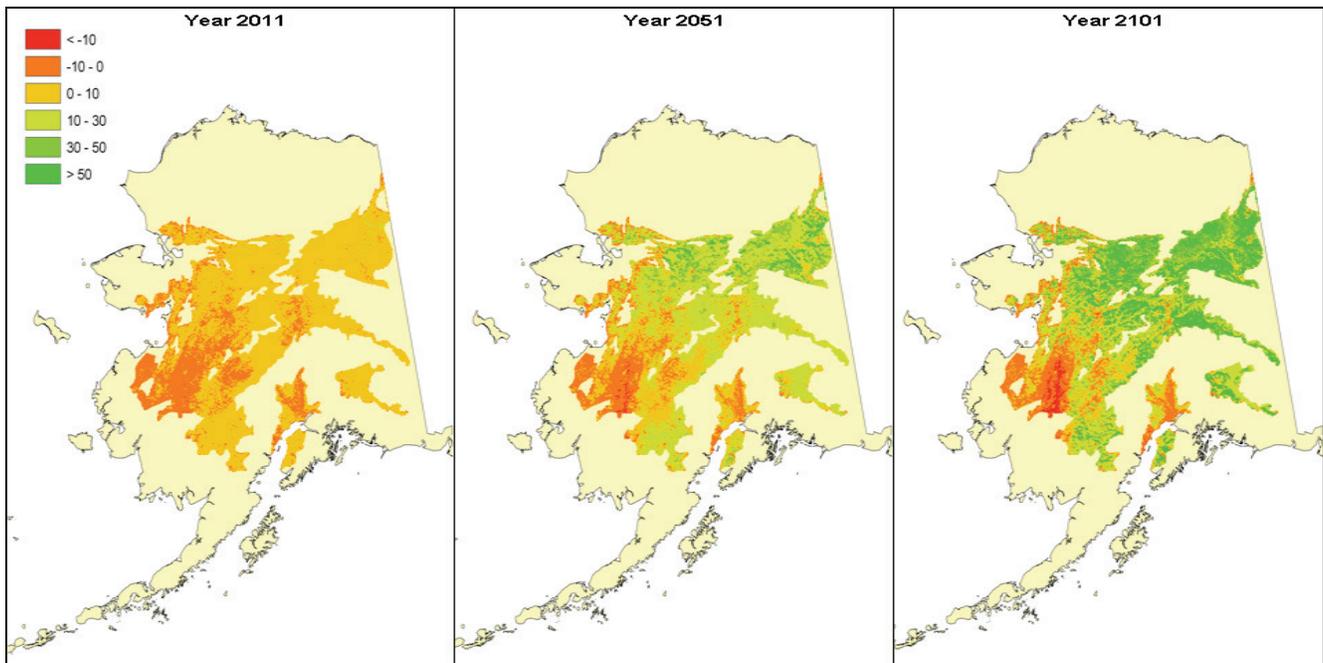


Figure 4—Predicted stand basal area change (m^2ha^{-1}) of the Alaska boreal forest in the year 2011, 2051, and 2101, assuming constant climate conditions and no major natural disturbances. The initial stand states were obtained from the 2001 NLCD Landsat remote sensing data.

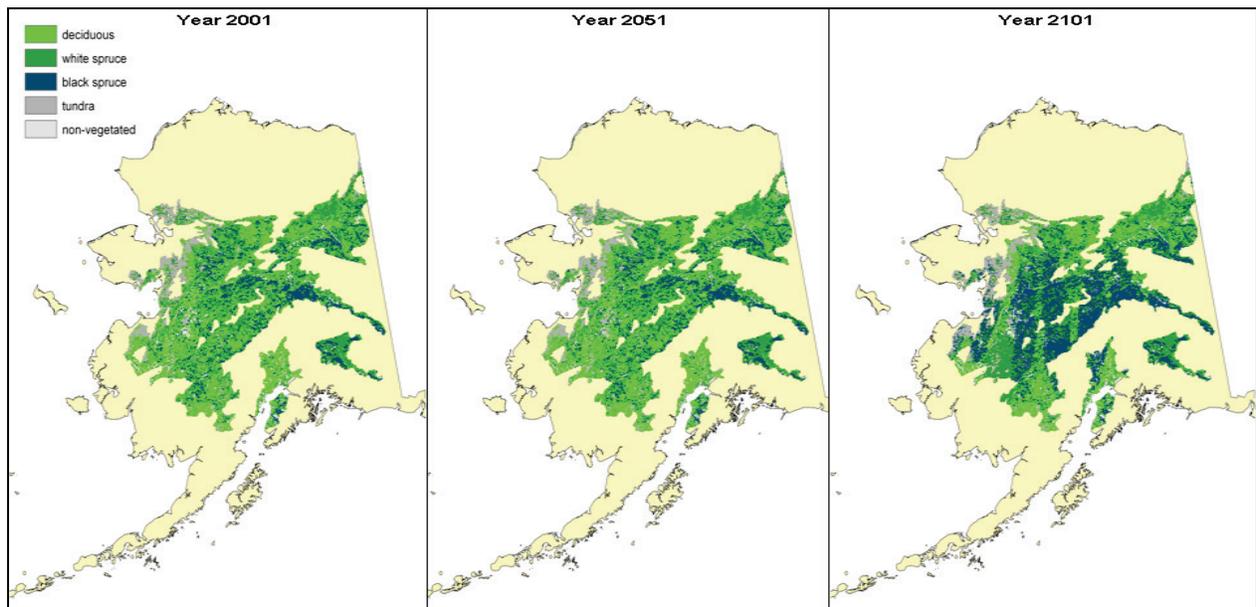


Figure 5—Observed (year 2001) and predicted (year 2051 and 2102) tree species coverage in the boreal forest region of Alaska, assuming constant climate conditions and no major natural disturbances. The initial stand states were obtained from the 2001 NLCD Landsat remote sensing data.

CURIOUS OR SPURIOUS CORRELATIONS WITHIN A NATIONAL-SCALE FOREST INVENTORY?

Christopher W. Woodall and James A. Westfall

ABSTRACT

Foresters are increasingly required to assess trends not only in traditional forest attributes (e.g., growing-stock volumes), but also across suites of forest health indicators and site/climate variables. Given the tenuous relationship between correlation and causality within extremely large datasets, the goal of this study was to use a nationwide annual forest inventory to determine levels of correlation among a wide array of database fields to aid foresters in separating correlation from causality in comprehensive forest resource assessments. In examining more than 15,000 individual correlations, we found the overwhelming majority (> 85 percent) of correlation coefficients were under 0.1. Site variables (e.g., elevation) had the highest mean correlations, while tree variables (e.g., live aboveground biomass) had the lowest mean correlations with all other variables. Nearly all the high correlations (>0.6) were between variables substantially autocorrelated (e.g., site class code and site index). Given that most correlations within a large-scale forest inventory dataset are very low with the remainder being nonsensical or autocorrelates, finding a highly correlated pair of variables with no apparent autocorrelation deserves further exploration.

INTRODUCTION

For most of the 20th century, forest resource assessments in the United States and abroad were often conducted purposively at small scales using spatially inconsistent sample techniques (i.e., relevé sampling such as stand exams) or conducted periodically at large scales using temporally inconsistent sample techniques (e.g., periodic forest inventory programs in the U.S., Frayer and Furnival 1999). In addition to the lack of spatially and temporally consistent forest inventories, the absence of computing resources available to forest professionals prevented complex forest inventory analyses and resource hypothesis testing. Until the 1990s, the analysis of large-scale forest resource datasets was severely limited to a few analysts with access to inconsistent datasets in computationally limited data management systems.

With the emergence of international agreements focused on the health of forest biomes (USDA 2004) and greenhouse gas accounting, nations have responded by developing nationally consistent forest inventories including numerous

variables complementary to traditional tree attributes (e.g., soils and downed dead wood, Perry and others 2009). In addition to field implementation of large-scale forest inventories, data management systems have been developed such that the multitude of data can be rapidly distributed to the public via well-documented web sites. Perhaps never before have forest professionals or the public had access to such large and extensive datasets for exploration of forest resource questions. For example, there are currently 1.1 and 15.0 million records within the plot and tree tables of the U.S. national inventory, respectively (Woudenberg and others 2011). Coupling the millions of inventory records with the hundreds of database fields provides the opportunity to explore numerous facets of forest ecosystems such as fire ecology (Woodall and Nagel 2007), climate change impacts (Woodall and others 2009), forest health (Huebner and others 2009), growth and mortality (Shaw and others 2005), and ownership patterns (Butler and Leatherberry 2005).

With the ability to rapidly assess forest resource attributes using extensive datasets comes the danger of inferring causality from possibly spurious correlations. Given that the U.S. national forest inventory data are publicly available for rapid download, most analyses will be conducted by users not affiliated with the actual data collection or management. Forest professionals have received little guidance on the frequency of high correlations within large-scale forest inventory datasets. Are strong correlations a common occurrence? Does autocorrelation confound many analyses? The goal of this study was to use a nationwide annual forest inventory to determine levels of correlation among a wide array of database fields to help foresters separate correlation from causality in comprehensive forest resource assessments.

METHODS

This study used data exclusively from the national inventory of all U.S. forests. The U.S. Department of Agriculture,

C.W. Woodall, Research Forester, U.S. Department of Agriculture, Forest Service, Northern Research Station, Forest Inventory and Analysis Program, St. Paul, MN 55108

J.A. Westfall, Research Forester, U.S. Department of Agriculture, Forest Service, Northern Research Station, Forest Inventory and Analysis Program, Newtown Square, PA 19073

Forest Service's Forest Inventory and Analysis (FIA) program is charged by Congress with providing an annual inventory of all forest lands. The FIA sampling framework is based on a systematic network of ground plots (Bechtold and Patterson 2005) obtained by dividing the U.S. into a series of 2,400-ha hexagons. Within each hexagon, FIA operates a multi-phase inventory. In phase 1 (P1), land area is stratified using aerial photography or classified satellite imagery to increase the precision of estimates using stratified estimation. In second phase (P2), permanent fixed-area plots are installed in each hexagon when field crews visit plot locations that have accessible forest land. Field crews collect data on more than 300 variables, including land ownership, forest type, tree species, tree size, tree condition, and other site attributes (e.g., slope, aspect, disturbance, land use) (USDA 2009). The plot design for FIA inventory plots consists of four 7.2-m fixed-radius subplots spaced 36.6 m apart in a triangular arrangement, with one subplot in the center. All trees with a diameter at breast height of at least 12.7 cm are inventoried within forested conditions. Within each subplot, a 2.07-m microplot offset 3.66 m from the subplot center is established where live tree seedlings and trees with a d.b.h. between 2.5 and 12.7 cm are inventoried. In addition to the trees measured on these plots, data are also gathered on the condition of the area in which the trees are located (e.g., stand-age class, ownership group, tree-density class). During the third phase of the inventory (P3), forest health indicators are measured on a 1/16th subset of the entire FIA ground plot network. The suite of forest health indicators includes tree crown condition, lichen communities, forest soils, vegetation diversity, down woody material, and ozone injury (Woodall and others In Press).

Using FIA's national database (FIADB version 4.0), we extracted forest inventory data for the most recent inventory in 49 states (currently no inventory available for Hawaii or interior Alaska). Given the multitude of database fields and tables examined in this study, FIA's documented nomenclature will be used in this study (Woudenberg and others 2011). The data extraction was limited to fields in the plot, condition, and tree tables, or variables calculated from those fields (e.g., total tree biomass on a plot): INVYR, STATECD, UNITCD, COUNTYCD, PLOT, PLOT_STATUS_CD, MEASYEAR, MEASMON, MEASDAY, REMPER, KINDCD, DESIGNCD, RDDISTCD, WATERCD, LAT, LON, ELEV, P2PANEL, CONGCD, MANUAL, EMAP_HEX, CYCLE, SUBCYCLE, CONDID, COND_STATUS_CD, RESERVCD, OWNCD, OWNGRPCD, FORTYPCD, FLDTYPCD, MAPDEN, STDAGE, STDSZCD, FLDSZCD, SITECLCD, SICOND, SIBASE, SISP, STDORPCD, CONDPROP_UNADJ, MICRPROP_UNADJ, SUBPPROP_UNADJ, SLOPE, ASPECT, PHYSCLCD, GSSTKCD, ALSTKCD,

DSTRBCD1, DSTRBYR1, TRTCD1, TRTYR1, BALIVE, FLDAGE, ALSTK, GSSTK, FORTYPCDCALC, SITETREE_TREE, SITECL_METHOD, CARBON_DOWN_DEAD, CARBON_LITTER, CARBON_SOIL_ORG, CARBON_STANDING_DEAD, CARBON_UNDERSTORY_AG, CARBON_UNDERSTORY_BG, CYCLE2, SUBCYCLE2, TREE, AZIMUTH, DIST, SPCD, SPGRPCD, DIA, DIAHTCD, HT, HTCD, ACTUALHT, TREECLCD, CR, CCLCD, TREEGRCD, CULL, DAMLOC1, DAMTYP1, DAMSEV1, STOCKING, VOLCFNET, VOLCFGRS, VOLBFNET, BOLBFGRS, VOLCFSND, DRYBIO_BOLE, DRYBIO_TOP, DRYBIO_STUMP, DRYBIO_SAPLING, DRYBIO_BG, CARBON_AG, CARBON_BG. All tree-level variables were summed to the plot and condition for live and standing dead trees. These calculated tree-level variables were delineated for live or dead by preceding each variable with a "L" or "D," respectively. Not all fields from the database tables were extracted for this study. Excluded were variables that were not alphanumeric or were a duplication of variables (e.g., secondary and tertiary tree damages). Finally, records were excluded when one or more fields were null. With these constraints, this study's data records totaled 42,617.

Correlations were calculated using SAS's CORR procedure with Pearson's correlation coefficients as the primary output. To assess the distribution of correlations from a large-scale forest inventory, the frequency of correlations from the correlation matrix of all this study's variables was determined. Correlations among the same variables were excluded from the matrix calculation (coefficient=1) for a total of 15,751 correlations. Mean absolute correlations were determined among broad categories of variables according to plot (i.e., plot selection information such as measurement year and county), site (i.e., physiographic information such as elevation and latitude), condition (i.e., stand condition information such as forest type and stand age), and tree (i.e., summed tree attributes such as height and volume). Actual individual correlations were examined when correlation coefficients exceeded 0.7.

RESULTS AND DISCUSSION

In examining of 15,625 individual absolute correlations, we found the overwhelming majority (> 85 percent) to be under 0.1 while less than 1 percent was above 0.5 (Fig. 1). Site variables (e.g., elevation and latitude) had the highest mean correlations (≈ 0.09), while tree variables (e.g., live aboveground biomass) had the lowest mean correlations (≈ 0.05) with all other study variables. Nearly all the high correlations (>0.7) were between variables substantially autocorrelated (e.g., algorithm calculated forest type and field estimated forest type) (Table 1). The remainder of high correlations could be attributed to spurious effects of

database manipulation (e.g., latitude and plot number) or possible curious ecological relationships (e.g., physiographic class and disturbance year).

If variables would be randomly chosen from a strategic-scale forest inventory dataset such as FIA's national inventory, it is extremely unlikely that any appreciable level of correlation would be found. If indeed the correlation exceeded 0.5, then these variables would stand a strong chance of being autocorrelated. Examples of autocorrelation in this study were measurement year and manual number, inventory cycle and kind code, and sum of live tree numbers and sum of distances to live trees. Most of these spurious correlations should be readily identified by even novice inventory analysts. Other spurious correlations, such as longitude and site index base, may take identification by experts in forest inventory databases and sampling designs. Only about a dozen correlations exceeded 0.6 and were ecologically interesting. Physiographic class was strongly correlated with the year of the most recent disturbance, soil organic carbon, understory aboveground biomass, and sum of live-tree board foot gross volume. Poor physiographic sites (i.e., ridge tops) may have shallow soils with little organic soil carbon and may be more prone to disturbances thus reducing their aboveground biomass. It appears as though approaching such large-scale datasets with readily testable ecological hypotheses may be the best method to derive meaningful relationships as opposed to the often spurious results of massive database computations using no *a priori* assumptions.

CONCLUSIONS

Given that most correlations within a large-scale forest inventory dataset are very low with most of the remainder being autocorrelated, finding a highly correlated pair of variables with no apparent autocorrelation is very unlikely. Because all correlations were assumed to be linear in this study, we suggest that non-linear correlations be examined in future studies. With the ever increasing availability of large datasets of ecosystem conditions (i.e., national forest inventories), a tenet can be forwarded: given the extreme rarity of finding highly correlated natural ecosystem variables lacking autocorrelation, when identified their further investigation is warranted.

LITERATURE CITED

- Bechtold, W.A.;** Patterson, P.L. eds. 2005. Forest Inventory and Analysis national sample design and estimation procedures. Gen. Tech. Rep. SRS-GTR-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p.
- Butler, B.J.;** Leatherberry, E.C. 2004. America's family forest owners. *Journal of Forestry*. 102: 4-14.
- Frazer, W.E.;** Furnival, G.M. 1999. Forest survey sampling designs, a history. *Journal of Forestry*. 97: 4-10.
- Huebner, C.D.;** Morin, R.S.; Zurbriggen, A.; White, R.L.; Moore, A.; Twardus, D. 2009. Patterns of exotic plant invasions in Pennsylvania's Allegheny National Forest using intensive Forest Inventory and Analysis plots. *Forest Ecology and Management*. 257: 258-270.
- Perry, C.H.;** Woodall, C.W.; Amacher, M.C.; O'Neill, K.P. 2009. An inventory of carbon storage in forest soil and down wood of the United States. In: McPherson, B.J.; Sundquist, E., eds. Carbon sequestration and its role in the global carbon cycle. AGU Special Monograph 183. Washington, DC: American Geophysical Union: 101-116.
- Shaw, J.D.;** Steed, B.E.; DeBlander, L.T. 2005. Forest Inventory and Analysis (FIA) annual inventory answers the question: What is happening to pinyon-Juniper woodlands? *Journal of Forestry*. 103: 280-285
- U.S. Department of Agriculture Forest Service.** 2004. National report on sustainable forests – 2003. FS-766. Washington, DC: U.S. Department of Agriculture, Forest Service. 139 p.
- U.S. Department of Agriculture Forest Service.** 2009. Forest Inventory and Analysis national core field guide (Phase 2 and 3), version 4.0. Washington, DC: U.S. Department of Agriculture Forest Service, Forest Inventory and Analysis. <http://www.fia.fs.fed.us/library/field-guides-methods-proc/> (accessed December, 2009).
- Woodall, C.W.;** Conkling, B.L.; Amacher, M.C. [and others]. 2010. The Forest Inventory and Analysis phase 3 indicators database 4.0: Description and users manual. Gen. Tech. Rep. NRS-xx. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station.
- Woodall, C.W.;** Nagel, L.M. 2007. Down woody fuel loadings dynamics of a large-scale blowdown in northern Minnesota. *Forest Ecology and Management*. 247: 194-199.
- Woodall, C.W.;** Oswalt, C.M.; Westfall, J.A.; Perry, C.H.; Nelson, M.D.; Finley, A.O. 2009. An indicator of tree migration in forests of the eastern United States. *Forest Ecology and Management*. 257: 1434-1444.
- Woudenberg, S.W.;** Conkling, B.L.; O'Connell, B.M. [and others]. 2011. The Forest Inventory and Analysis Database: database description and users manual version 4.0 for Phase 2. Gen. Tech. Rep. RMRS-GTR-245. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Table 1—Matrix of absolute correlation coefficients for all study correlations exceeding 0.7 (live and dead tree variables designated by L and D, respectively)

Variables	Correlations		
	0.7 – 0.8	0.8 – 0.9	0.9 – 1.0
Plot	Lon		
Measyear	Manual		
Lon	Plot		
Manual	Measyear		
Condid		Condprop_unadj, microprop_unadj, subprop_unadj	
Owncd			Owngprcd
Owngprcd			Owncd
Fortyped	Fldtyped		
Fldtyped	Fortyped		
Siteclcd	Second		
Sicnd	Siteclcd		
Conprop_unadj		Condid	Micprop_unadj, subprop_unadj
Micprop_unadj		Condid	Condprop_unadj, subprop_unadj
Subprop_unadj		Condid	Conprop_unadj, Micprop_unadj
Gsstkcd	Carbon_understory_ag		
Dstrbyr1		Carbon_soil_org	
Trtcd1		Carbon_standing_dead	
Trtyr1	L_carbon_bg		
Fldage		L_dist	
Carbon_soil_org		Dstrbyr1	
Carbon_standing_dead		Trtcd1	
Carbon_understory_bg	Gsstkcd		
L_tree	L_azimuth, l_sprgrpcd		
L_azimuth	L_tree	L_spced, l_sprgrpcd	
L_dist		Fldage	
L_spced		L_azimuth	L_sprgrpcd
L_sprgrpcd	L_tree	L_azimuth	L_spced
L_actualht		L_volbfnet	
L_volbfnet		L_actualht	
L_carbon_bg	Trtyr1		
D_tree			D_damtyp1
D_azimuth			D_damsev1
D_dist			D_decaycd
D_damtyp1			D_tree
D_damsev1			D_azimuth
D_decaycd			D_dist

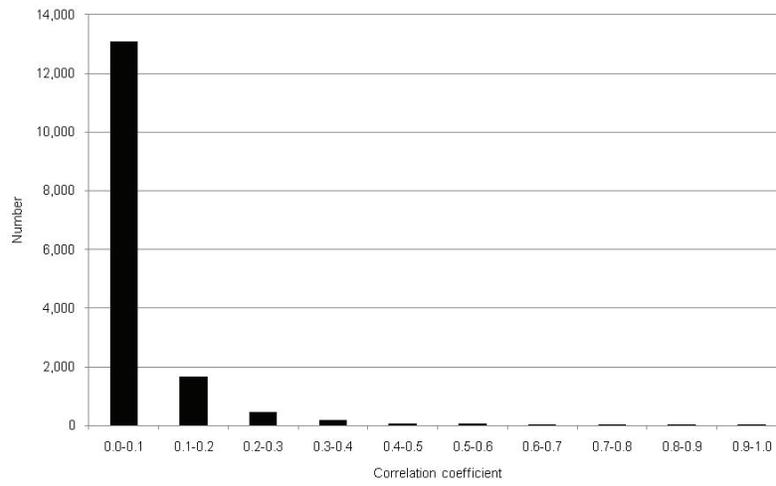


Figure 1—Frequency of absolute correlation coefficients among a multitude of variables sampled during an inventory of U.S. forests.

ESTIMATING TREE CROWN WIDTHS FOR THE PRIMARY ACADIAN SPECIES IN MAINE

Matthew B. Russell and Aaron R. Weiskittel

ABSTRACT

In this analysis, data for seven conifer and eight hardwood species were gathered from across the state of Maine for estimating tree crown widths. Maximum and largest crown width equations were developed using tree diameter at breast height as the primary predicting variable. Quantile regression techniques were used to estimate the maximum crown width and a constrained nonlinear equation was developed for estimating the largest crown width. We noted an improvement in predictions in 11 of 14 species compared to currently-used largest crown width equations in the region. The models performed well across the wide range of stand conditions present in the dataset and proved effective when used in the computation of alternative measures of stand density (crown competition factor and percent canopy cover). Results from this analysis can be used in examining tree crown dynamics and assessing alternative measures of stand density.

INTRODUCTION

Accurately determining individual tree crown width has a broad applicability to forestry and natural resource sciences, yet, the measurement of crown width is lacking in most forest inventories. Estimating crown width can be used to calculate stand canopy closure, which is important for assessing wildlife habitat suitability, fire risk, and understory light conditions for regeneration (Crookston and Stage 1999). Crown width variables have become integral in fields using airborne laser scanning technologies (e.g. Salas and others 2010). Tree diameter at breast height (dbh) generally accounts for much of the variability in predicting tree crown dimensions.

Open-grown trees are commonly selected to estimate maximum crown width (mcw). The mcw of a tree is generally defined as the potential crown width at a given diameter if the tree is open-grown. Methods for selecting trees in the field that display open-grown characteristics are available (Paine and Hann 1982), but making this determination can be laborious, often involves extensive traveling to subject trees, and may lead to subjectivity. Despite its importance, there are currently no regional mcw equations for most of the important commercial species in the northeastern U.S. Ek (1974) developed mcw equations for several species in the Lake States region but their applicability for use with trees found in Maine has not previously been assessed.

A forest-grown tree of a given species displays a horizontal crown extension that is less than that of an open-grown tree. This measure of crown width is termed the largest crown width (lcw). Hence, lcw equations differ from mcw equations in that they predict the crown widths of trees growing in forested settings. Bechtold (2003) developed lcw equations that covered a broad portion of the eastern U.S. The performance of the Bechtold (2003) equations to trees growing in Maine is unknown.

Crown width measurements are integral to estimating alternative measure of stand density used throughout forestry. Determining the mcw of trees is needed for estimating crown competition factor (CCF; Krajicek and others 1961), and lcw equations are needed to estimate percent canopy cover (PCC; Crookston and Stage 1999). CCF is an estimate of the area available to the average tree in the stand in relation to the maximum area it could use if it were open grown. PCC is defined as the percentage of the ground area that is covered by a vertical projection of tree crowns. Consequently, quantification of these stand density measures is an important component of many forest growth and yield models. As an example, the Forest Vegetation Simulator employs equations developed using open-grown trees to compute CCF, but uses separate equations developed with forest-grown trees for estimating PCC (Crookston and Dixon 2005).

The goal of this analysis was to employ data gathered from across Maine to determine tree crown width attributes for its primary species. The primary objectives were to: (1) develop mcw and lcw equations for seven conifer and eight hardwood tree species; (2) compare model predictions with existing equations used throughout the region; and (3) evaluate the crown width equations in the determination of measures of stand density.

METHODS

SPECIES

Maine is part of the Acadian forest, which is a transition zone between the conifer-dominant boreal forests of the

north and the mixed hardwood forests of the south (Braun 1950, Rowe 1972). Forests are typically established under natural regeneration and are comprised of mixed-species stands with even- or uneven-aged stand structures. Common conifer species include balsam fir [*Abies balsamea* (L.)], red spruce [*Picea rubens* (Sarg.)], white spruce [*Picea glauca* (Moench) Voss.], eastern white pine [*Pinus strobus* L.], eastern hemlock [*Tsuga canadensis* (L.) Carr.], black spruce [*Picea mariana* (Mill.) B.S.P.], and northern white-cedar [*Thuja occidentalis* L.]. Hardwoods commonly found include red maple [*Acer rubrum* L.], paper birch [*Betula papyrifera* Marsh.], gray birch [*Betula populifolia* Marsh.], yellow birch [*Betula alleghaniensis* Britt.], quaking aspen [*Populus tremuloides* Michx.], bigtooth aspen [*Populus grandidentata* Michx.], American beech [*Fagus grandifolia* Ehrh.], northern red oak [*Quercus rubra* L.], and sugar maple (*Acer saccharum* Marsh.). Red spruce, white spruce, and balsam fir dominate the relatively low-lying sites of poorer drainage, but the proportion of eastern hemlock and white pine increases as drainage improves.

DATA

The data for this analysis came from three primary sources at a range of locations throughout Maine: (1) USDA Forest Service Forest Health Monitoring program (FHM); (2) USDA Forest Service Northern Research Station, Penobscot Experimental Forest (PEF); and (3) University of Maine Cooperative Forestry Research Unit (CFRU).

FHM—The FHM program collected information on 123 0.07-ha plots throughout each of the 16 counties in Maine from 1991-1999. Some of the trees in these plots were remeasured during this period. The maximum horizontal diameter of the widest axis of the tree crown and the distance perpendicular to this axis were measured. Crown measurements were collected on trees with a dbh greater than 12.7 cm.

PEF—The PEF, located in the towns of Bradley and Eddington, ME, is a long-term experiment investigating impacts of even-, two-, and uneven-aged silvicultural systems in the Acadian forest (Sendak and others 2003). Tree crown and height measurements have been collected since 2000 on a subset of continuous forest inventory (CFI) plots at the PEF. Crown measurements in this analysis were obtained from individual trees on 81 plots (0.08-ha in size) across the PEF. In addition, 2,698 crown measurements made on 20 CFI plots across the PEF were used in this analysis (Saunders and Wagner 2008).

CFRU—The CFRU dataset came from an early investigation of thinning in spruce-fir forests (McCormack 1989). It consisted of four locations across Maine in the townships of

Lakeville Plantation, T5 R15 WELS, T11 R16 WELS, and T11 R13 WELS. Stand ages at time of establishment ranged from 17 to 70 years. Thirty-one plots of varying size were measured up to four times from 1978 to 1994 across these locations.

In the PEF and CFRU datasets, tree dbh was recorded and crown radii (r) were measured from the center of the bole of each tree to the edge of its crown in each of the cardinal directions (N, S, E, W). For the FHM data, the maximum horizontal diameter of the widest axis of the tree crown and its perpendicular distance were each divided by two and considered as crown radii measurements. Quadratic mean crown width was computed for all datasets to provide an unbiased estimation of crown area irrespective of crown shape (Gregoire and Valentine 1995). Crown width by dbh and dataset is presented in Figure 1. Depending on the minimum size dbh measured in the inventory and the relative distribution of species, crown width and dbh vary according to dataset.

Tree data according to species are summarized in Table 1. Observations that were coded as displaying a broken or dead top, or greater than 50 percent crown dieback were excluded in this analysis.

MODEL DEVELOPMENT

To develop mcw equations, open-grown trees are often used in model development. This can be problematical as the determination of “open-grownness” could lead to subjectivity, is not always specified in the field, and datasets of open-grown trees are often comprised of small sample sizes. One approach to overcome this is to employ quantile regression techniques to estimate a species-specific mcw for a given tree diameter. Least squares regression techniques estimate a response variable that is conditioned solely on the statistical mean, while quantile regression methods allow estimation of response variables for any quantile of the data (Koenker and Hallock 2001). Given that the data comprised a wide range of tree crown widths (both open- and forest-grown) and that the interest is in estimating the maximum potential crown width for a species at a given dbh, the 99th quantile was fit to represent the maximum crown width for open-grown trees. A nonlinear allometric equation of the following form was used:

$$mcw = a_i dbh^{a_2} \quad [1]$$

where dbh is tree diameter at breast height (cm) and a_i 's were coefficients estimated from the 99th percentile used to represent maximum crown width (m) for each species.

An lcw equation predicts the crown width for trees that are not able to reach their biological maximum due to competition in a stand. Recognizing the relationship

between open-grown and forest-grown trees, constrained predictions of *lcw* were made to be less than or equal to that of the *mcw* curve. A nonlinear equation was fit to the data:

$$lcw = \frac{mcw}{b_1 dbh^{b_2}} \quad [2]$$

where *mcw* is the predicted maximum crown width of the tree at its corresponding *dbh* (Eq. [1]) and b_1 's were estimated coefficients. The b_1 parameter was non-significant for four species. To evaluate the *lcw* equations, we computed fit index (FI) and mean absolute error (MAE):

$$FI = 1 - \left(\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right)$$

$$MAE = \sum_{i=1}^n |y_i - \hat{y}_i| / n$$

where y_i , \hat{y}_i , and \bar{y} are observed, predicted, and mean *lcw*, respectively, and n is number of trees in a species. The *mcw* and *lcw* model parameters were estimated in R using the nonlinear quantile regression (**nlrq**) and generalized nonlinear least squares (**gnls**) functions, respectively. Further analysis indicated that incorporating tree crown ratio as an additional predictor showed a minor improvement in fit index for some species.

To assess the crown width equations in terms of computing stand density measures, data from the Penobscot Experimental Forest (PEF) long-term silvicultural study were obtained. The permanent sample plots within these stands represent differing silvicultural practices with a variety of stand compositions and structures (Sendak and others 2003). The *mcw* equation (Eq. [1]) was used to compute maximum crown area for each individual tree, and crown competition factor (CCF) was calculated (Krajicek and others 1961). The *lcw* equation (Eq. [2]) was used to compute percent canopy cover (PCC) with a correction for crown overlap (Crookston and Stage 1999). CCF and PCC were also estimated for the PEF plots using the Ek (1974) and Bechtold (2003) equations, respectively. These estimates of CCF and PCC were compared with plot basal area ($m^2 ha^{-1}$).

RESULTS

All coefficients were positive in estimating *mcw*, indicating that *mcw* increases nonlinearly with *dbh* (Table 2). Differences between the *mcw* equations developed in this analysis and the equations of Ek (1974) were observed for several species (Figure 2). Generally, the predicted *mcw* using equations developed in this analysis were 45 percent

higher than *mcw* predicted from the equations of Ek (1974). For the *lcw* equations, increasing *dbh* resulted in a larger predicted *lcw* (Table 3). For conifers, fit index (FI) ranged from 0.25 for northern white-cedar to 0.56 for eastern hemlock and eastern white pine. For hardwoods, FI ranged from 0.12 for American beech to 0.59 for paper birch. Mean absolute error (MAE) for conifers ranged from 0.46 m for black spruce to 1.18 m for eastern white pine. For hardwoods, MAE ranged from 0.46 m for gray birch to 1.52 m for American beech. Compared to previously published equations (Bechtold 2003), reductions in MAE were observed for 6 of 7 conifer species, with the exception of eastern white pine. For hardwoods, we observed reductions in MAE when compared to Bechtold's (2003) equations for 5 of 7 species, with the exceptions of northern red oak and sugar maple.

CCF and PCC were found to be positively correlated with plot-level basal area (BA) for permanent sample plot data at the Penobscot Experimental Forest (Figure 3). The Pearson correlation coefficients of CCF-BA and PPC-BA were 0.56 and 0.70, respectively. The Pearson correlation coefficient for CCF-BA using the equations of Ek (1974) was 0.45. The correlation coefficient for PPC-BA using the equations of Bechtold (1974) was 0.61.

DISCUSSION

For the primary species occurring across Maine, *mcw* and *lcw* relationships appear to be adequately captured using tree *dbh*. Using quantile regression provides the ability to model the biological maximum of tree crown width, hence, determining whether or not trees are open-grown is not required in developing *mcw* equations. Adapting the *mcw* estimate using a nonlinear equation form results in an accurate and constrained estimate of *lcw*. For most species, improved predictions resulted when compared to previously published equations developed at more broad scales.

The large differences in the presented *mcw* equations with those of Ek (1974) likely arise from three primary sources. First, geographic differences that influence tree crown attributes are likely apparent, as the Ek (1974) equations were developed using trees grown in the Lake States. Second, the data used in this analysis included trees with a full range of crown widths and *dbh*, including measurements of trees with a minimum *dbh* of around 1.4 cm for most species. Lastly, employing quantile regression allowed for a quantitative estimation of the maximum potential crown width, whereas Ek (1974) used least squares procedures conditioned solely on the statistical mean of the data. Shade tolerant hardwoods like sugar maple and American beech likely had low FI's for *lcw* because a limited number of

trees were used for those species and a complete range of diameters was not available when compared to other species.

Predicted crown widths from the developed equations appear to adequately represent alternative measures of stand density. We observed a stronger correlation between CCF and basal area and PPC and basal area using the crown width equations developed herein when compared to CCF and PCC using crown width equations found in Ek (1974) and Bechtold (2003). This speaks to the importance of an accurate and reliable estimate of individual tree crown width in determining stand density measures. Stand basal area is extensively used in growth models because it is easy to measure and is highly correlated with volume, but it is often confounded with site quality and stand age. Alternatively, the mcw equations can be used in computing CCF, a measure that is assumed to be independent of site and age and can be applied to both even- and uneven-aged stands. CCF is an effective measure of stand density because the determination of crown width is species-specific. Hence, CCF is a metric that takes into account the contribution of individual species to stand density. PCC has implications not only in forestry but also in assessing wildlife habitat suitability and evaluating fire risk potential. The computation of PCC, however, does not taken into account the spatial distribution of trees and as a percentage is constrained to be between 0 and 100%, which may prove difficult in quantifying competition. Field measurements of canopy cover and/or derived canopy cover estimates from LiDAR could aid in further evaluating the performance of the crown width equations.

Methods developed for developing mcw equations do not require the discernment between open- and forest-grown trees. Provided that trees are sampled across a wide range of stand conditions, estimating maximum crown width using quantile regression performs well. Bounding a tree's crown width by its potential maximum is biologically logical and resulted in accurate estimates of largest crown width. The equations of Bechtold (2003) performed well when using data from Maine, however we did find improvements for most species examined. Results from this analysis can be used in exploring measurements of stand density, examining tree crown profiles, and investigating canopy dynamics for species common to these forests.

ACKNOWLEDGMENTS

Data were provided by the USDA Forest Service Forest Health Monitoring Program, USDA Forest Service Northern Research Station (Penobscot Experimental Forest), and the University of Maine Cooperative Forestry Research Unit. We acknowledge Max McCormack and Mike Saunders

for providing access to their data. We thank Mark Ducey, KaDonna Randolph, Chris Toney, and two anonymous reviewers who provided comments on this material. This project was supported by the National Science Foundation Center for Advanced Forestry Systems, research grant number 0855370.

LITERATURE CITED

- Bechtold, W.A.** 2003. Crown-diameter prediction models for 87 species of stand-grown trees in the eastern United States. *Southern Journal of Applied Forestry*. 27(4): 269-278.
- Braun, E.L.** 1950. *Deciduous forests of eastern North America*. Hafner, New York, NY. 596 p.
- Crookston, N.L.;** Dixon, G.E. 2005. The Forest Vegetation Simulator: a review of its applications, structure, and content. *Computers and Electronics in Agriculture*. 49: 60-80.
- Crookston, N.L.;** Stage, A.R. 1999. Percent canopy cover and stand structure statistics from the Forest Vegetation Simulator. Res. Pap. RMRS-GTR-24. Ogden, UT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 11 p.
- Ek, A.** 1974. Dimensional relationships of forest and open grown trees in Wisconsin. University of Wisconsin.
- Gregoire, T.G.;** Valentine, H.T. 1995. A sampling strategy to estimate the area and perimeter of irregularly shaped planar regions. *Forest Science*. 41(3): 470-476.
- Koenker, R.;** Hallock, K.F. 2001. Quantile regression. *The Journal of Economic Perspectives*. 15(4): 143-156.
- Krajicek, J.E.;** Brinkman, K.A.; Gingrich, S.F. 1961. Crown competition-a measure of stand density. *Forest Science*. 7: 35-42.
- McCormack, M.L.** 1989. Thinning spruce and spruce-fir stands. In: 1989 Annual report of the Cooperative Forestry Research Unit. Orono, ME: Maine Agricultural and Forestry Experiment Station: 9-10.
- Paine, D.P.;** Hann, D.W. 1982. Maximum crown width equations for southwestern Oregon tree species. Res. Pap. 46. Corvallis, OR: Oregon State University, Forest Research Laboratory. 20 p.
- Rowe, J.S.** 1972. *Forest regions of Canada*. Publication 1300-251. Ottawa, ON: Department of the Environment, Canadian Forest Service. 172 p.
- Salas, C.;** Ene, L.; Gregoire, T.G.; Naesset, E.; Gobakken, T. 2010. Modelling tree diameter from airborne laser scanning derived variables: A comparison of spatial statistical models. *Remote Sensing of Environment*. 114(6): 1277-1285.
- Saunders, M.R.;** Wagner, R.G. 2008. Height-diameter models with random coefficients and site variables for tree species of central Maine. *Annals of Forest Science*. 65(2): 203.
- Sendak, P.E.;** Brissette, J.C.; Frank, R.M. 2003. Silviculture affects composition, growth, and yield in mixed northern conifers: 40-year results from the Penobscot Experimental Forest. *Canadian Journal of Forest Research*. 33(11): 2116-2128.

Table 1—Summary statistics for data used in the development of crown width models for seven conifer and eight hardwood species in Maine

Species	Code	n	dbh (cm)				Quadratic mean crown width (m)			
			Mean	SD	Min	Max	Mean	SD	Min	Max
Conifers										
Balsam fir	BF	3605	12.2	7.0	1.1	40.1	2.8	1.3	0.3	12.5
Black spruce	BS	400	14.6	3.5	4.3	23.6	2.4	0.7	0.8	5.2
Eastern hemlock	EH	1127	23.1	11.1	1.3	57.4	5.3	2.0	1.2	12.1
Eastern white pine	WP	866	25.6	15.7	1.4	92.2	5.0	2.4	0.6	14.8
Northern white-cedar	WC	866	22.7	7.6	1.6	59.7	3.7	1.2	0.8	10.5
Red spruce	RS	2994	19.7	7.2	1.2	56.9	3.6	1.4	0.7	12.1
White spruce	WS	339	17.4	7.1	1.5	40.6	3.5	1.2	0.6	9.3
Hardwoods										
American beech	AB	325	21.3	6.7	11.9	43.4	6.0	2.0	1.4	12.2
Gray birch	GB	251	6.4	4.7	1.3	25.4	2.1	1.3	0.2	9.4
Northern red oak	RO	102	24.3	10.3	12.7	66.8	6.3	2.3	2.1	13.1
Paper birch	PB	576	16.4	8.4	1.3	37.3	4.3	2.1	0.2	14.0
Quaking aspen	QA	353	18.2	7.5	1.3	45.4	4.2	1.7	0.4	10.7
Red maple	RM	1785	18.2	8.6	1.3	54.9	4.9	2.1	0.2	12.2
Sugar maple	SM	355	24.4	9.6	12.7	73.9	6.2	2.0	2.4	14.3
Yellow birch	YB	388	23.6	8.9	1.5	54.1	6.3	2.2	1.9	13.0

Table 2—Parameter estimates (standard errors in parentheses) by species for predicting the maximum crown width (mcw; m) using tree diameter at breast height (dbh; cm) for seven conifer and eight hardwood species growing in Maine

Species	a ₁	a ₂
Conifers		
Balsam fir	1.37 (0.039)	0.572 (0.021)
Black spruce	0.535 (0.21)	0.742 (0.14)
Eastern hemlock	2.44 (0.42)	0.408 (0.055)
Eastern white pine	1.24 (0.49)	0.585 (0.10)
Northern white-cedar	1.63 (0.44)	0.436 (0.087)
Red spruce	1.80 (0.46)	0.461 (0.075)
White spruce	1.50 (0.46)	0.496 (0.10)
Hardwoods		
American beech	2.93 (0.65)	0.434 (0.077)
Gray birch	2.24 (1.8)	0.382 (0.28)
Northern red oak	4.08 (2.0)	0.310 (0.16)
Paper birch	1.48 (0.24)	0.623 (0.056)
Quaking aspen	1.31 (0.24)	0.586 (0.059)
Red maple	2.17 (0.19)	0.491 (0.030)
Sugar maple	3.31 (0.66)	0.356 (0.06)
Yellow birch	4.04 (0.79)	0.308 (0.062)

Table 3—Parameter estimates (standard errors in parentheses) by species with fit index (FI) and mean absolute error (MAE) for predicting the largest crown width (lcw; m) using tree diameter at breast height (dbh; cm) for seven conifer and eight hardwood species in Maine. Comparisons of MAE with those of Bechtold (2003) equation [4] is presented

Species	b ₁	b ₂	FI	MAE	
				This Study	Bechtold (2003) Eq. [4]
Conifers					
Balsam fir	1.49 (0.017)	0.105 (0.0050)	0.55	0.58	0.65
Black spruce	-	0.174 (0.0045)	0.35	0.46	0.47
Eastern hemlock	1.90 (0.058)	-0.057 (0.010)	0.56	1.01	1.06
Eastern white pine	-	0.147 (0.0033)	0.56	1.18	1.08
Northern white-cedar	2.19 (0.20)	-0.080 (0.029)	0.27	0.77	0.79
Red spruce	4.33 (0.21)	-0.264 (0.015)	0.43	0.77	0.96
White spruce	2.09 (0.16)	-0.069 (0.027)	0.51	0.59	0.64
Hardwoods					
American beech	-	0.194 (0.0058)	0.12	1.52	1.58
Gray birch	3.10 (0.27)	-0.214 (0.04)	0.49	0.62	-
Northern red oak	4.10 (0.89)	-0.272 (0.065)	0.43	1.31	1.29
Paper birch	2.10 (0.13)	-0.035 (0.021)	0.59	1.00	1.04
Quaking aspen	2.65 (0.26)	-0.157 (0.034)	0.57	0.87	0.88
Red maple	2.63 (0.11)	-0.132 (0.014)	0.56	1.05	1.15
Sugar maple	-	0.161 (0.0049)	0.17	1.44	1.38
Yellow birch	4.23 (0.51)	-0.294 (0.037)	0.41	1.33	1.58

$$\text{Model is: } lcw = mcw / (b_1 dbh^{b_2})$$

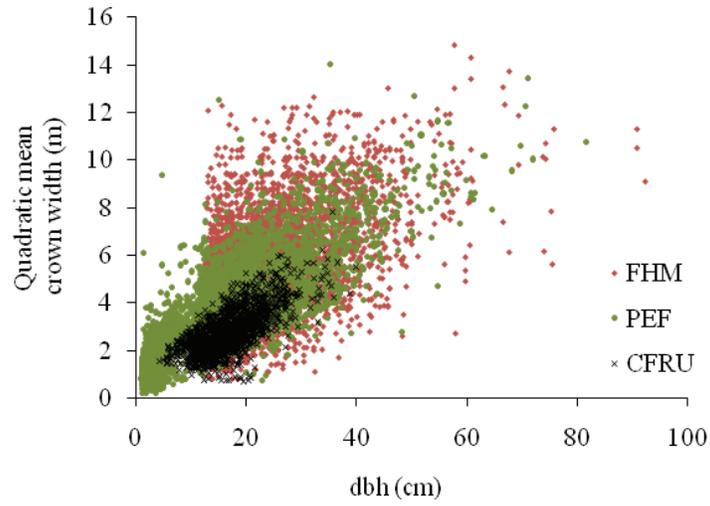


Figure 1—Quadratic mean crown width and tree diameter at breast height (dbh) for Penobscot Experimental Forest (PEF), Cooperative Forestry Research Unit (CFRU), and Forest Health Monitoring (FHM) data.

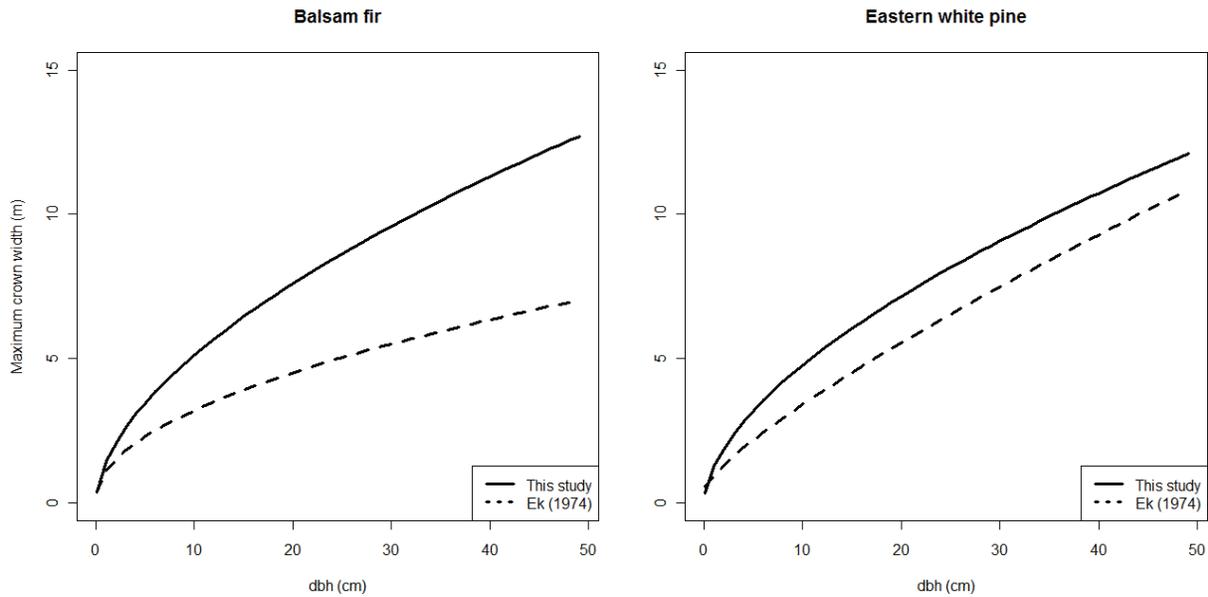


Figure 2—Predicted maximum crown width using tree diameter at breast height (dbh) for the equation developed in this study and that of Ek (1974) for balsam fir and eastern white pine.

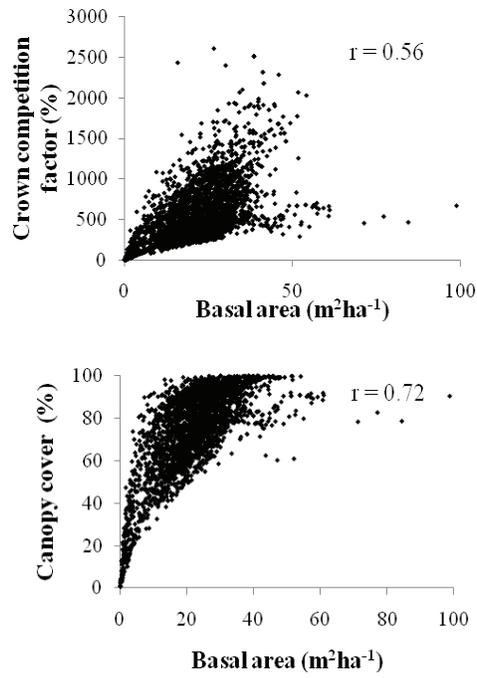


Figure 3—Crown competition factor and percent canopy cover related to basal area using permanent sample plots obtained from the Penobscot Experimental Forest, Bradley and Eddington, Maine.

AN EVALUATION OF THE PROPERTIES OF THE VARIANCE ESTIMATOR USED BY FIA

John P. Brown and James A. Westfall

ABSTRACT

The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service currently conducts inventories utilizing the protocols of the national enhanced FIA Program. Due to the permanent locations of the sample plots, the stratification of the population occurs after the selection of sample units, i.e., post-stratification. In situations where the population is of limited areal extent, this may result in small within-stratum sample sizes. The survey literature provides some guidance on post-stratified sample sizes, but does not specifically address the behavior of estimators when sample sizes are smaller than recommended. It is important for FIA to evaluate how estimators perform across a range of sample sizes, such that samples of sufficient size can be constructed to ensure accurate estimates. The variance estimator used by FIA accounts for a secondary source of variation (V_2) due to random within-strata sample sizes that is introduced beyond that obtained from standard proportional allocation of samples to a stratified sample (V_1). Thus, each estimate's variance is composed of two parts. This study utilizes a Monte Carlo simulation to examine the relative contributions of V_1 and V_2 to the total variance (V_{Total}) of the estimate. FIA plots from Pennsylvania were treated as a population from which samples of size n are repeatedly drawn and V_1 , V_2 , and V_{Total} calculated for forest area and cubic volume estimates. The sample size varied from 25 to 200 plots. With increasing sample size n , the V_1 variance stabilized at sample sizes greater than 60 plots, whereas the V_2 variance required sample sizes greater than 125. The ratio of the two variance components ($V_{\text{RAT}} = V_1/V_2$) was found to increase with increasing n , ranging from 6 to 32 plots for the area estimates and from 8 to 45 plots for the volume estimates.

INTRODUCTION

The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service currently inventories forested land across the United States using procedures detailed in Bechtold and Patterson (2005). During Phase 1, remotely sensed information is used to stratify the population to reduce the variance of estimates. This stratification varies by region but generally includes at a minimum forest and nonforest as strata (Bechtold and Patterson 2005). In Phase 2, permanent ground plots are visited and data on numerous attributes are collected at various levels of detail. Plots determined as clearly nonforest from aerial imagery are assessed remotely.

Weights for strata are determined during Phase 1. However, the Phase 2 sample determines strata sample sizes as plots are permanently located without respect to stratum boundaries. This sampling design is considered to be a post-stratified simple random sample (Cochran 1977, Schaeffer

and others 2006) and it has an added source of variation due to stratum sample sizes not being fixed in advance. The magnitude of this additional variation within a forest inventory has not been well studied. It is the goal of this study to examine how the use of this post-stratification estimate affects the variances of total area and total cubic foot volume. Specifically, the variance estimates for these values will be split into their components and examined both separately and jointly in order to better understand what role each plays in the total variance under several sample size situations. This is important to FIA to insure that sufficient sample sizes are available for accurate estimates. Sample strata weights will also be tested for agreement with population strata weights using χ^2 tests of agreement.

METHODS

Plot data is from a complete set of panels for Pennsylvania measured from 2003 to 2007. Phase I strata were developed by classifying the percent tree canopy cover from the NLCD 2001 map product (Homer et al. 2004) into five classes. There were 4,628 plots that were treated as the population from which samples of plots were drawn.

VARIANCE ESTIMATOR

The variance of the estimate (Bechtold and Patterson, 2005) is given by

$$v(\hat{Y}_d) = \frac{A_T^2}{n} \left[\underbrace{\sum_h W_h n_h v(\bar{Y}_{hd})}_{V_1} + \underbrace{\sum_h (1 - W_h) \frac{n_h}{n} v(\bar{Y}_{hd})}_{V_2} \right] \quad (1)$$

where

A_T = total area of the population.

$h \dots H$ = strata in the domain of interest.

n = sample size.

W_h = weight for stratum h within the population.

n_h = sample size for stratum h .

\bar{Y}_{hd} = mean of attribute of interest (plot proportion forest land or cubic-foot volume) for stratum h

d = domain of interest

$v(\bar{Y}_{hd})$ = variance of the mean for stratum h

The left side addend within the bracketed sum in (1) will be referred to as V_1 . This part of the variance results from the stratification of the population during Phase I. The right side addend within the bracketed sum in (1) will be referred to as V_2 . This second part of the variance is a consequence of stratum sizes being random within the strata determined in Phase I. The ratio V_1/V_2 will be defined as V_{RAT} and the total variance as V_{Total} .

MONTE CARLO SIMULATION

The first stage of the Monte Carlo (MC) simulation (Metropolis and Ulam 1949) was to determine how many sets of 50 plots would result in stable values for the V_1 , V_2 , V_{RAT} , and total variance for both the area and cubic-foot estimates. Each of the four values was calculated for 5,000 draws. Then, the variance of each value was calculated for the first three sets. Subsequently, an additional set was added and the variance recalculated for the specific value. Variances were plotted against the number of sets drawn and it was determined that 5,000 draws were sufficient to stabilize the several measures of interest. As computation length was not extensive, 10,000 draws were performed for each sample size.

Procedures for the second stage of the MC simulation were performed separately for the forest land area and cubic-foot volume estimates. Initially 10,000 sets of plots were selected, for each of a number of selected sample sizes. Plots in a set are drawn without replacement and sample sizes were: 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 125, 150, 175, 200. Every plot in the set had its stratum, proportion of forest, and cubic volume recorded. Each set of n plots had its strata sizes, strata weights, strata means, strata variances, and stratified mean and variance for the total forested area calculated. While each plot in a set is not replaced, at the next iteration (next set of plots) all plots are then replaced, thus all aggregate measures on the plot are considered to be sampled with replacement. The MC variance was then determined as the variance of the 10,000 stratified sample means. These procedures were repeated for cubic-foot volume.

Trends in variance behavior were visually analyzed with the use of boxplots. A boxplot was generated for each set of 10,000 plots at the selected sample sizes for V_1 , V_2 , V_{RAT} and V_{Total} . Boxplots used the first, second (median), and third quartiles for the lower, middle and upper horizontal lines of the boxes. Minimum and maximum values were represented by the lower and upper whiskers respectively of the boxplots. Means also were calculated and shown as points (triangles). Patterns were examined specifically for the means and medians.

Stratum sample weights were calculated for each set of plots drawn for all sample sizes. A χ^2 test was used to test agreement of stratum sample weights and population stratum weights ($H_0: n_h/n = W_h$) for all sets of plots at all sample sizes. The significance level was set at 0.05. If the test was not significant, the set was recorded to be in agreement with the population stratum weights. Frequencies of agreement for a fixed sample size were calculated for the 10,000 simulations. This agreement testing will be used to assess if deviations from population stratum weights exist and whether they may be influencing the variances of the estimates.

RESULTS

VARIANCES FOR THE AREA ESTIMATES

For the V_1 variance, median and mean values for a given n stabilize around a sample size of 60 plots (Figure 1). Median and mean values are less for a sample size of 50 and below. V_2 variances approximately stabilize for a sample sizes of 125 or greater (Figure 2) and are slightly higher for a sample size lower than 125. V_{RAT} values do not approach a stable point (Figure 3). When considering the median and mean values, V_{Total} approaches an asymptote after 125 samples as well (Figure 4), yet is still decreasing slightly for larger values.

VARIANCES FOR THE VOLUME

Again considering the median and mean values for a given n , V_1 values for the volume estimates also stabilize around a sample size of 60 (Figure 5). As was the case for the area estimates, values for a sample size less than 60 are smaller on average. V_2 values similarly stabilize for sample size 125 and greater (Figure 6). V_{RAT} values are increasing and range from 8 to 45 for sample size 25 and 200 respectively (Figure 7). V_{Total} values approach an asymptote for sample size 125 and greater (Figure 8) when focusing on the mean and median values.

MC VARIANCE ESTIMATES

The MC variance estimates for both area of forest land and volume follow a similar pattern, they decrease at a decreasing rate (Figures 9 and 10). As compared to the mean V_{Total} for identical sample sizes, the MC variance is in close agreement.

SAMPLE WEIGHTS

Sample weights were consistently in agreement 95 percent of the time or better for all sample sizes (Table 1). There were no apparent patterns related to sample size, as all agreement levels were either 95 or 96 percent in all cases.

DISCUSSION

Patterns for the V_1 , V_2 , and V_{Total} values were quite similar between the area and volume estimates. For both area and volume, V_1 increased to a stable value at sample size 60 and above, while V_2 and V_{Total} decreased to a stable value for sample size greater than 125. There were differences in ranges for V_{RAT} with V_{RAT} for area ranging between 6 and 32 over the given sample size range, while V_{RAT} for volume ranged between 8 and 45 therein.

The increasing values for V_{RAT} stem from minute changes in V_2 relative to V_1 (Figures 1, 2 and Figures 5, 6). While the V_2 values have what appears to be an asymptote, small changes downward are enough to continue inflating the value of V_{RAT} . In regard to V_{Total} however, the overall addition from V_2 is small, and V_{Total} stabilizes when V_2 stabilizes.

Three factors suggest that V_{Total} is biased for smaller samples. First, V_1 increased as it approached 60 samples then approached an asymptote (Figures 1 and 5). Second, V_{RAT} was continuously increasing as well, implying that V_1 dominated V_2 (Figures 3 and 7). Even though V_2 initially decreases, V_{RAT} shows that V_1 is still much greater than V_2 , therefore an increasing V_1 offsets the decreasing by V_2 . Third, V_{Total} shows a similar pattern as V_1 , increasing to an asymptote at 60 (Figures 4 and 8). These factors demonstrate then that for low sample sizes V_{Total} is underestimated. The main factor to this downward bias for appears to be V_1 , with minor offsetting by V_2 .

Stratum sample weights agreed with population stratum weights for all sample sizes (Table 1). Agreement percentages were 95 percent and above, which is where they should be given that the significance level for the χ^2 test was set at 95 percent as well. It was thought that perhaps the lower sample sizes might fail to generate similar sample stratum weights as compared to the population stratum weight as some of the class sizes were small, but this hypothesis was not supported. Approximately 5 percent of the samples deviated from the population weight and the other 95 percent were similar.

CONCLUSIONS

With increasing sample sizes, the penalty factor for post-stratification, V_2 , diminishes greatly compared to the

variance component stemming from stratified design (V_1). Cochran (1977) states that the effect of the V_2 variance will be small if the mean number of sampling units per stratum is reasonably large. For these data, asymptotes are approached for V_2 and V_{Total} at sample size of 125. The mean number of sampling units per stratum is therefore 25 here, which may provide some insight of minimum bound for 'reasonably large.' Cochran (1977, p.134) states also that stratum samples greater than twenty are 'reasonably large' and Schaeffer et al. (2006, p. 150) suggest that stratum samples sizes greater than 20 provide "...nearly as accurate sample sizes as stratified sampling with proportional allocation." This may be too conservative a rule of thumb for this data, as the smallest stratum sample weight was about 0.06, resulting in just seven samples on average in that stratum at an overall sample size of 125. Users of FIA data should be aware that stratifications which later have small sample sizes may result in an underestimate of the variance of the intended estimate. Further study may more accurately determine what within-stratum minimums are achievable.

What was not varied in this study was the state from which the plots were located, as this study was conducted using only one population with a specific stratification scheme. Weights for the five strata ranged from 0.06 to 0.38. Results from other populations having differing structures should be examined to determine if the results found in this study are more broadly applicable.

LITERATURE CITED

- Bechtold, W.A.;** Patterson, P.L., eds. 2005. The enhanced Forest Inventory and Analysis Program—National sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 85 p.
- Cochran, W. G.** 1977. Sampling techniques. 3rd ed. New York: John Wiley. 428 p.
- Homer, C.;** Huang, C; Yang, L. [and others]. 2004. Development of a 2001 national landcover database for the United States. Photogrammetric Engineering and Remote Sensing. 70:829-840.
- Metropolis, N.;** Ulam, S. 1949. The Monte Carlo method. Journal of American Statistical Association. 44: 335-341.
- Schaeffer, R. L.;** Mendenhall, W.; Ott, R. L. 2006. Elementary survey sampling. 6th ed. Belmont: Thomson Brooks/Cole. 464 p.

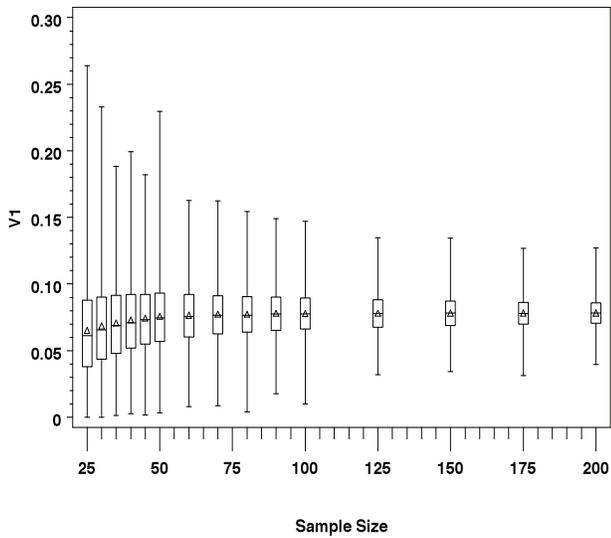


Figure 1—Boxplots for the 10,000 simulations of the area V_1 variance using the given sample sizes. Lower and upper whiskers represent minimum and maximum values. Lower and upper box edges represent 1st and 3rd quartiles, with the median represented by the line inside the box. Means are symbolized with triangles.

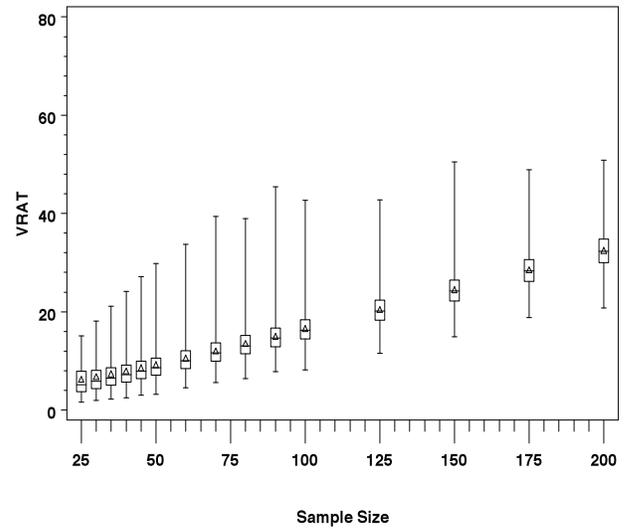


Figure 3—Boxplots for the 10,000 simulations comparing the ratio (V_{RAT}) of the V_2 and V_1 area variances using the given sample sizes. Lower and upper whiskers represent minimum and maximum values. Lower and upper box edges represent 1st and 3rd quartiles, with the median represented by the line inside the box. Means are symbolized with triangles.

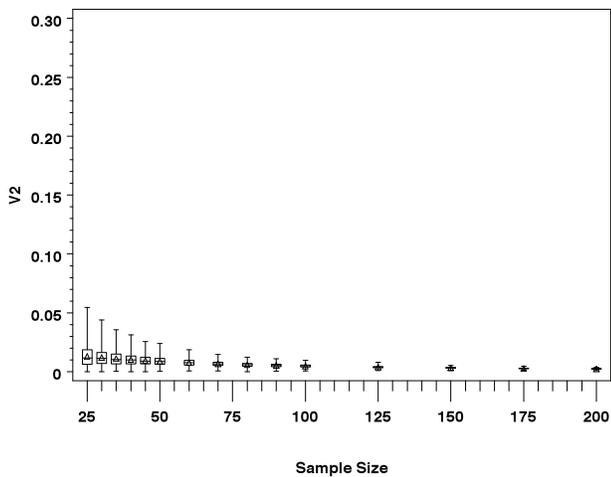


Figure 2—Boxplots for the 10,000 simulations of the area V_2 variance using the given sample sizes. Lower and upper whiskers represent minimum and maximum values. Lower and upper box edges represent 1st and 3rd quartiles, with the median represented by the line inside the box. Means are symbolized with triangles.

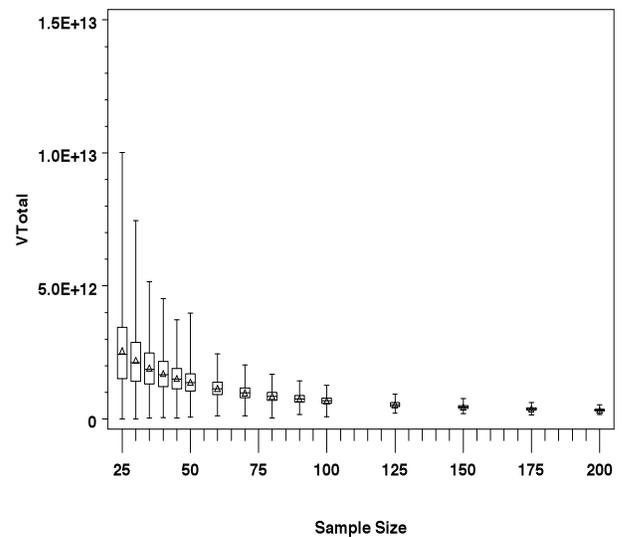


Figure 4—Boxplots for the 10,000 simulations of the total area variance (V_{Total}) using the given sample sizes. Lower and upper whiskers represent minimum and maximum values. Lower and upper box edges represent 1st and 3rd quartiles, with the median represented by the line inside the box. Means are symbolized with triangles.

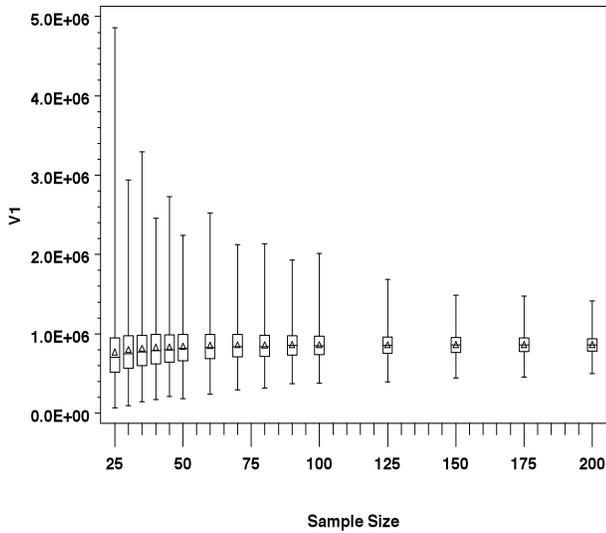


Figure 5—Boxplots for the 10,000 simulations of the cubic volume V_1 variance using the given sample sizes. Lower and upper whiskers represent minimum and maximum values. Lower and upper box edge represent 1st and 3rd quartiles, with the median represented by the line inside the box. Means are symbolized with triangles.

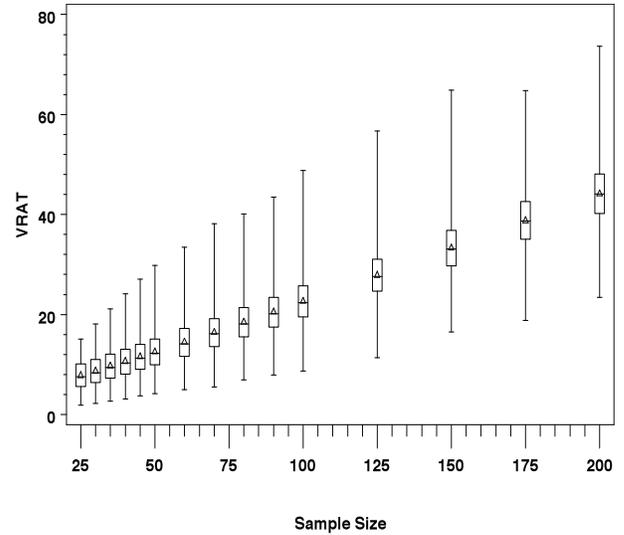


Figure 7—Boxplots for the 10,000 simulations comparing the ratio (V_{RAT}) of the V_2 and V_1 cubic volume variances using the given sample sizes. Lower and upper whiskers represent minimum and maximum values. Lower and upper box edges represent 1st and 3rd quartiles, with the median represented by the line inside the box. Means are symbolized with triangles.

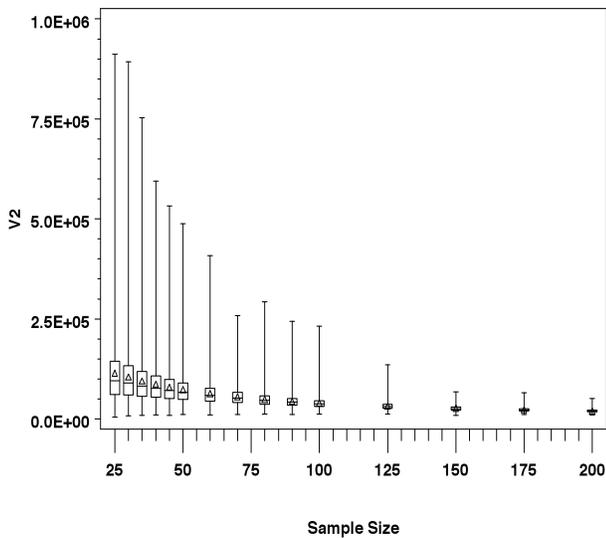


Figure 6—Boxplots for the 10,000 simulations of the cubic volume V_2 variance using the given sample sizes. Lower and upper whiskers represent minimum and maximum values. Lower and upper box edges represent 1st and 3rd quartiles, with the median represented by the line inside the box. Means are symbolized with triangles.

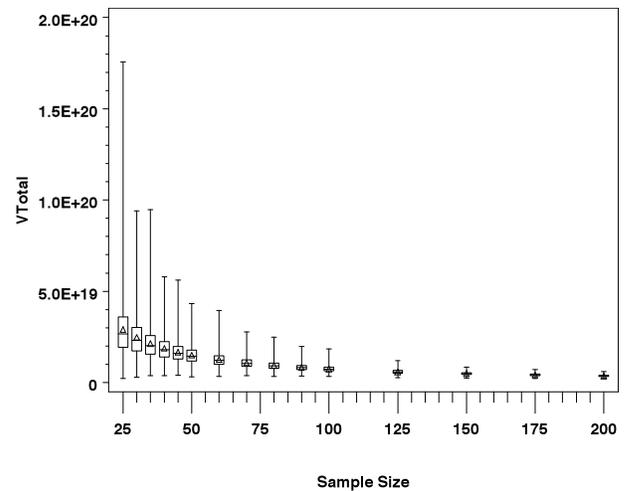


Figure 8—Boxplots for the 10,000 simulations of the total cubic volume variance (V_{Total}) using the given sample sizes. Lower and upper whiskers represent minimum and maximum values. Lower and upper box edges represent 1st and 3rd quartiles, with the median represented by the line inside the box. Means are symbolized with triangles.

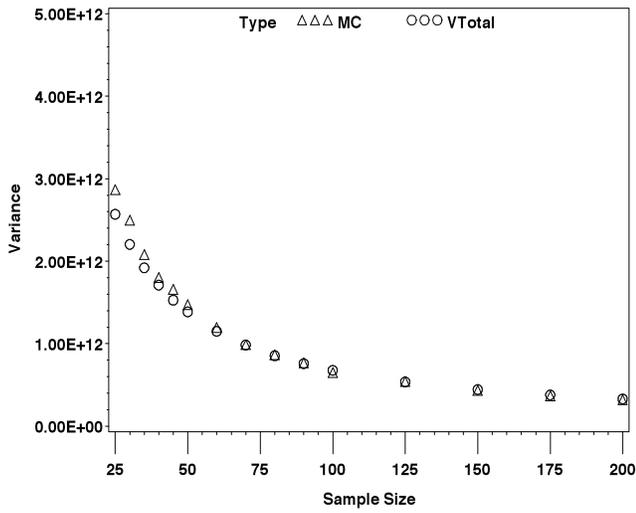


Figure 9—Comparison of the Monte Carlo variance for the 10,000 simulations of the mean total area and mean V_{Total} for a given sample size.

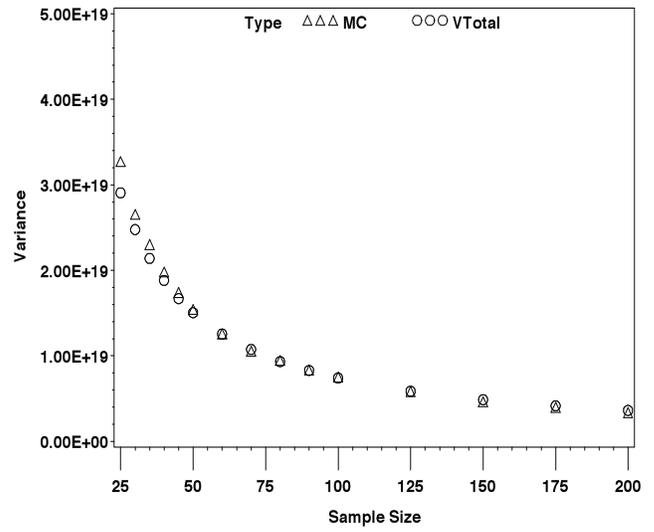


Figure 10—Comparison of the Monte Carlo variance for the 10,000 simulations of the mean total cubic volume and mean V_{Total} for a given sample size.

Table 1—Proportion of simulations where the sample weight agreed with the population weight as tested by a χ^2 goodness-of-fit test

Sample size	Area			Volume		
	Lower confidence limit	p	Upper confidence limit	Lower confidence limit	p	Upper confidence limit
25	0.9464	0.9508	0.9550	0.9463	0.9507	0.9549
30	0.9479	0.9523	0.9564	0.9453	0.9498	0.9540
35	0.9525	0.9567	0.9606	0.9482	0.9525	0.9566
40	0.9483	0.9526	0.9567	0.9488	0.9531	0.9572
45	0.9477	0.9521	0.9562	0.9474	0.9518	0.9559
50	0.9482	0.9525	0.9566	0.9460	0.9504	0.9546
60	0.9510	0.9552	0.9592	0.9469	0.9513	0.9554
70	0.9508	0.9550	0.9590	0.9504	0.9547	0.9587
80	0.9523	0.9565	0.9604	0.9462	0.9506	0.9548
90	0.9501	0.9544	0.9584	0.9520	0.9562	0.9601
100	0.9492	0.9535	0.9575	0.9488	0.9531	0.9572
125	0.9512	0.9554	0.9594	0.9533	0.9574	0.9613
150	0.9518	0.9560	0.9599	0.9557	0.9597	0.9635
175	0.9550	0.9591	0.9629	0.9565	0.9605	0.9642
200	0.9515	0.9557	0.9597	0.9545	0.9586	0.9624

A MULTIVARIATE MIXED MODEL SYSTEM FOR WOOD SPECIFIC GRAVITY AND MOISTURE CONTENT OF PLANTED LOBLOLLY PINE STANDS IN THE SOUTHERN UNITED STATES

Finto Antony,¹ Laurence R. Schimleck,¹ Alex Clark III,² and Richard F. Daniels¹

ABSTRACT

Specific gravity (SG) and moisture content (MC) both have a strong influence on the quantity and quality of wood fiber. We proposed a multivariate mixed model system to model the two properties simultaneously. Disk SG and MC at different height levels were measured from 3 trees in 135 stands across the natural range of loblolly pine and the stand level values were used for the modeling SG-MC system. Regional variation in mean trend of the properties was incorporated in the model. Contemporaneous correlation between the SG and MC was accounted by defining within stand error structure appropriately. Compared to univariate models, predictions based on the multivariate model were improved by 29 and 26 % in root mean square prediction error for disk SG and MC after taking account of the contemporaneous correlation.

INTRODUCTION

A forest is a complex dynamic system with inter-related individual components. Foresters commonly rely on simultaneous modeling systems to explain such inter-dependent systems. One familiar example of such a system to forest biometricians is simultaneous modeling of dominant height, basal area, trees per hectare and volume (Borders 1989; Fang et al. 2001; Hall and Clutter 2004). Two main reasons for the popularity of simultaneous modeling systems in forestry are: 1) compatibility requirement of individual components in the system (Clutter 1963); 2) contemporaneous correlation of error among individual components in the system.

Specific gravity (SG) and moisture content (MC) both have a strong influence on the quantity and quality of wood. SG describes the mass of woody material present in a given volume of wood. It is a unit-less measure and expressed as the ratio of wood basic density (oven dry weight divided by green volume) with the density of water at 4°C (Megraw 1985). SG is considered an important wood property because of its strong correlation with the strength of solid wood products, as well as the yield and quality of

pulp produced (Panshin and deZeeuw 1980). Generally the moisture content of wood is expressed as a percentage of the oven dry weight of wood. Moisture content influences the physical and mechanical properties of wood, resistance to biological deterioration and dimensional stability (Haygreen and Bowyer 1996).

SG and MC vary considerably within loblolly pine (*Pinus taeda* L.) trees. SG follows a decreasing trend with tree height (He 2004; Megraw 1985; Phillips 2002; Zobel and Blair 1976), while MC increases with height (Koch 1972; Phillips 2002). It has been reported that these two variables are highly negatively correlated with high SG associated with low MC and vice-versa (Koch 1972; Zobel and Blair 1976). The primary factor controlling the longitudinal variation in disk SG and MC in a loblolly pine tree is the proportion of juvenile wood (Zobel and Blair 1976; Zobel and vanBuijtenen 1989). In general, the proportion of juvenile wood is higher towards the top of a tree than at the base and juvenile wood has lower SG and higher MC than mature wood.

The objective of this study was to model the longitudinal variation in disk SG and MC as a simultaneous multivariate mixed model system. We will show how contemporaneous correlation between these two variables (disk SG and MC) can be potentially utilized to improve the prediction of disk SG or MC for loblolly pine at any height.

DATA

The Wood Quality Consortium at the University of Georgia and the United States Department of Agriculture (USDA) Forest Service Southern Research station sampled planted loblolly pine across its natural range to study the longitudinal variation in wood SG and MC. Trees were sampled from

¹Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA-30602

²(Retired) USDA Forest Service, Athens, GA-30602

135 stands from six physiographic regions across the southeastern United States. Regions sampled included: 1- southern Atlantic Coastal Plain (**R1**), 2- northern Atlantic Coastal Plain (**R2**), 3- Upper Coastal Plain (**R3**), 4- Piedmont (**R4**), 5- Gulf Coastal Plain (**R5**) and 6- Hilly Coastal Plain (**R6**). A minimum of 12 plantations from each of the six physiographic regions were sampled. The stands selected for sampling included 20- to 25-year-old loblolly pine plantations planted at 1250 or more trees per hectare and having 625 trees per hectare or more after thinning. Only stands that were conventionally managed with no fertilization (except phosphorus at planting on phosphorus deficient sites) and no competition control were sampled. Three trees from each stand were felled and cross sectional disks of 3.8 cm thickness were collected from 0.15, 1.37 m and then 1.52 m intervals along the stem up to a diameter of 50 mm outside bark. The disks were sealed in plastic bags and shipped to the USDA Forest Service laboratory for physical property analysis. Disk SG (based on green volume and oven-dry weight) and disk MC (based on green and oven-dry weights) were determined for each sampling height. Stand averages (at each height) for disk SG and MC were calculated using the three trees sampled per stand. A summary of average stand characteristics for each region is presented in Table 3.1. Plots of stand average disk SG and MC with relative height are presented in Figures 3.1 and 3.2.

MODEL DEVELOPMENT

Two response components are considered in this simultaneous model system, disk SG and MC measured at the same heights for 3 trees in a stand. The basic models adopted for these two components are

$$SG = f_1(x, \beta) = \beta_{0,1} + \beta_{1,1}x + \beta_{2,1}x^2 + \beta_{3,1}(\alpha_{1,1} - x)_+^2 + \beta_{4,1}(\alpha_{2,1} - x)_+^2 + \varepsilon_{SG} \quad [1]$$

$$MC = f_2(x, \beta) = \beta_{0,2} + \beta_{1,2}x + \beta_{2,2}(\alpha_{1,2} - x)_+^2 + \varepsilon_{MC} \quad [2]$$

where SG = disk SG; MC = disk MC; x = relative height h/H , h is the average height above ground and H is the average total height of the stand calculated from the three sampled trees;

$$[\beta_{0,1} \quad \beta_{1,1} \quad \beta_{2,1} \quad \beta_{3,1} \quad \beta_{4,1} \quad \beta_{0,2} \quad \beta_{1,2} \quad \beta_{2,2} \quad \alpha_{1,1} \quad \alpha_{2,1} \quad \alpha_{1,2}]^T$$

are parameters to be estimated, with knot parameters $[1 > \alpha_{1,1} > \alpha_{2,1} > 0]$ and $[1 > \alpha_{1,2} > 0]$; ε_{SG} and ε_{MC}

are error terms for disk SG and MC respectively.

The $(\alpha_j - x)_+^2$ terms indicates the positive part of the function $\alpha_j - x$ where “+” sets it to zero for those values

of x where $\alpha_j - x$ is negative (here $x > \alpha_j$). The basic model form for disk SG is equivalent to the standard form of the taper model proposed by Max and Burkhart (1976), which is not constrained to have a value of zero at the tip of the tree.

In order to account for stand-to-stand variability in the data, we used a nonlinear mixed effect model (NLMM). Let y_{ijk} represent the k^{th} response ($k = 1, 2$) variable measured at j^{th} relative height from i^{th} stand; the univariate nonlinear mixed model for each property can be represented as

$$y_{ij1} = \theta_{0,i1} + \theta_{1,i1}x_{ij} + \theta_{2,i1}x_{ij}^2 + \theta_{3,i1}(\alpha_{1,i1} - x_{ij})_+^2 + \theta_{4,i1}(\alpha_{2,i1} - x_{ij})_+^2 + \varepsilon_{ij1} \quad [3]$$

$$y_{ij2} = \theta_{0,i2} + \theta_{1,i2}x_{ij} + \theta_{2,i2}(\alpha_{1,i2} - x_{ij})_+^2 + \varepsilon_{ij2} \quad [4]$$

The mixed effect parameter θ_{ijk} in the above models takes the form

$$\theta_{ik} = A_{ik}\beta_k + B_{ik}b_{i,k} \quad [5]$$

where $b_{i,k}$ is the i^{th} stand level random effect vector specific to the k^{th} response variable with $b_{i,k} \sim N(0, \Psi_k)$; B_{ik} is the associated random effect design matrix; A_{ik} is the fixed effect design matrix and β_k is the fixed effect parameter vector specific to the k^{th} response variable.

In order to develop the bivariate model, we first fitted the univariate stand level NLMM’s model for disk SG (Eq. 3) and MC (Eq. 4) separately. Initially we assumed all the parameters in the univariate models were mixed. Final specification of mixed effect parameters in the univariate models were decided based on Akaike’s Information Criteria (AIC), a model selection criterion used for NLMM’s. Parameters $\beta_{0,i1}$, $\beta_{1,i1}$, $\beta_{2,i1}$, $\beta_{0,i2}$ and $\beta_{1,i2}$ were selected as mixed, with random stand level intercepts in these parameters. The regional variation in mean trend for both properties was incorporated by appropriate fixed effect specification (fixed effect design matrix) for all parameters, except the knot parameters, in both univariate modes. The knot parameters were assumed as common for all regions for both properties. Since we had six distinct physiographical regions in the study, we assumed different fixed effect parameters for each region with the southern Atlantic Coastal Plain as the reference region with all other regions having their own parameters which are

deviations from the reference region. The final fixed effect specifications for each parameter were identified using univariate models for each property and likelihood ratio test between full model and reduced model. The fixed effect specifications corresponds to all parameters used in the bivariate model are presented in Table 3.2.

The variance-covariance structure for $\text{var}(b_{i,1})$ and $\text{var}(b_{i,2})$

in the univariate models were selected based on the model selection criteria (AIC and Bayesian information criterion (BIC)). We selected a general positive definite form of variance-covariance structure for disk SG and a diagonal form of variance-covariance structure for disk MC. The model information criteria and log likelihood values for the final selected univariate models, called **SG1** and **MC1** respectively for each response, are presented in Table 3.3.

For fitting the bivariate model, the univariate model equations for two responses were stacked together and can be represented as

$$y_{ij} = f(\mathbf{x}_{ij}, \boldsymbol{\theta}_i) + \varepsilon_{ij} \quad [6]$$

where $\mathbf{y}_{ij} = (y_{ij1}, y_{ij2})$. To take account of the correlation between responses measured from the same stand at the same height level, we assumed the within stand variance-covariance matrix as

$$\boldsymbol{\varepsilon}_{ij} \stackrel{i.i.d}{\sim} N(\mathbf{0}, \sigma^2 \boldsymbol{\Lambda})$$

where $\boldsymbol{\varepsilon}_{ij} = (\varepsilon_{ij1}, \varepsilon_{ij2})$. Following Eq. 5, after stacking the fixed effect and random effect vectors and design matrices for two response variables, we can write $\boldsymbol{\theta}_i = (\theta_{i1}, \theta_{i2})$ as

$$\boldsymbol{\theta}_i = A_i \boldsymbol{\beta} + B_i b_i \quad [7]$$

where $A_i = \text{diag}(A_{i1}, A_{i2})$; $B_i = \text{diag}(B_{i1}, B_{i2})$;

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T)^T; b_i = (b_{i1}^T, b_{i2}^T)^T$$

and we assumed that $b_i \stackrel{i.i.d}{\sim} N(\mathbf{0}, \boldsymbol{\Psi})$.

All the models were fitted using the nlme package in R, version 2.9.1 (Pinheiro et al. 2009). Initially the two univariate models (Eq. [3] and [4]) were simultaneously fitted, referred to as **SGMC1**, with a positive definite form of variance-covariance structure for disk SG, a diagonal form of variance-covariance structure for disk MC and unique variance parameter estimate for each response variable. Here, a block-diagonal form was used to define the random effect structure of two responses as follows

$$\boldsymbol{\Psi} = \begin{pmatrix} \text{var}(b_{i,1}) & \mathbf{0} \\ \mathbf{0} & \text{var}(b_{i,2}) \end{pmatrix}$$

The advantage of multivariate fitting over univariate fitting is that we can incorporate correlation among errors and random effects associated with different response variables in the model by specifying different forms of $\boldsymbol{\Lambda}$ and $\boldsymbol{\Psi}$ (Fang et al. 2001; Hall and Clutter 2004). The contemporaneous correlation between responses was incorporated by relaxing the form of $\mathbf{\ddot{E}}$ from an identity matrix to a symmetric positive definite matrix (referred to as **SGMC2**). We also allowed for correlation among random effects associated with the two models. The final best fitted model (referred to as **SGMC3**) is represented as follows

$$y_{ij1} = (\beta_{0\ell,1} + b_{0i,1}) + (\beta_{1\ell,1} + b_{1i,1})x + (\beta_{2\ell,1} + b_{2i,1})x^2 + \beta_{3\ell,1}(\alpha_{1,1} - x)_+^2 + \beta_{4\ell,1}(\alpha_{2,1} - x)_+^2 + \varepsilon_{ij1}$$

$$y_{ij2} = (\beta_{0\ell,2} + b_{0i,2}) + (\beta_{1\ell,2} + b_{1i,2})x + \beta_{2\ell,2}(\alpha_{1,2} - x)_+^2 + \varepsilon_{ij2}$$

$$\begin{pmatrix} b_{0i,1} \\ b_{1i,1} \\ b_{2i,1} \\ b_{0i,2} \\ b_{1i,2} \end{pmatrix} \sim N(\mathbf{0}, \boldsymbol{\Psi}) \quad \text{where} \quad \boldsymbol{\Psi} = \begin{pmatrix} \varphi_{00,1} & \varphi_{01,1} & \varphi_{02,1} & \varphi_{00,12} & 0 \\ & \varphi_{11,1} & \varphi_{12,1} & \varphi_{10,12} & 0 \\ & & \varphi_{22,1} & \varphi_{20,12} & 0 \\ & & & \varphi_{00,2} & 0 \\ & & & & \varphi_{11,2} \end{pmatrix} \quad [8]$$

$$(\varepsilon_{ij1}, \varepsilon_{ij2})^T | \mathbf{b}_i \sim N(\mathbf{0}, \mathbf{R}_i),$$

where

$$\mathbf{R}_i = \sigma^2 \mathbf{G}_i^{1/2}(\delta) \boldsymbol{\Gamma}(\rho) \mathbf{G}_i^{1/2}(\delta)$$

$$\mathbf{G}_i(\delta) = \text{diag}(\mathbf{I}_1, \delta^2 \mathbf{I}_2)$$

$$\boldsymbol{\Gamma}(\rho) = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

In [8] the fixed effect $\beta_{\ell,k}$, $k(k=1,2)$ indicates parameter β specific to ℓ^{th} region specified in Table 3.2 for response variable SG ($k=1$) and for response variable MC ($k=2$).

The random effect $b_{i,k}$, $k(k=1,2)$ indicates the random effect parameter specific to the i^{th} stand for response variable SG ($k=1$) and for response variable MC ($k=2$).

The model information criteria (AIC and BIC) and log likelihood values from simultaneous fitting of the models (**SGMC1**, **SGMC2** and **SGMC3**) are presented in Table 3.3. The log likelihood and information criteria from **SGMC1** were equal to the sum of log likelihood and information

criteria from univariate fitting SG1 and MC1. Incorporation of contemporaneous correlation into the model (SGMC2) significantly improved the model fitting criteria. The final model SGMC3 found to have a significant improvement in model information criteria over SGMC2. The estimated fixed effect parameter from the final simultaneous model is presented in Table 3.4. The estimated random effect variance-covariance matrix Ψ is

$$\Psi = \begin{pmatrix} 0.00074 & -0.00100 & 0.00058 & -0.23368 & 0 \\ & 0.00335 & -0.00257 & 0.32946 & 0 \\ & & 0.00241 & -0.20459 & 0 \\ & & & 121.92 & 0 \\ & & & & 284.19 \end{pmatrix}$$

and the within stand residual parameters are $\delta = 638.68$ and $\rho = -0.779$.

PREDICTION

Our primary objective of developing a simultaneous system is to make predictions. The reported advantage of using a multivariate method over univariate method is its improvement in predictive performance (Fang et al. 2001; Hall and Clutter 2004). The information on contemporaneous correlation among response variables can be potentially utilized to improve the prediction of a variable at a particular measurement occasion (here at a particular stand height level) given that the observed value of other response variables at the specified measurement occasion. For example in the proposed multivariate system, information of disk SG at any specific height can be utilized to improve the prediction of disk MC at that height. Similarly, observed disk MC at any specific stand height can be utilized to improve the prediction of disk SG at that height.

There are several situations where we can utilize a multivariate model to make predictions. Fang et al. (2001) dealt with several such prediction scenarios based on their height-basal area-volume simultaneous mixed model system. In the present study, we are primarily interested in prediction from a multivariate model system where observations on one of the correlated response variables are available. For example, we may want to predict disk MC for a stand at different heights when measurements of disk SG are available. To this extent, we can utilize a predictor proposed by Hall and Clutter (2004) for NLMM's which is based on a linear mixed model (LMM) approximation

of NLMM. The proposed predictor is analogous to the empirical best linear unbiased predictor (BLUP) of LMM. It is supposed to perform better than the plug-in-predictor proposed for NLMM by Pinheiro and Bates (2000). The following on the derivation of a predictor was extracted from Hall and Clutter (2004). Generically a NLMM can be represent as

$$\hat{o} = f(\beta, b, \mathbf{A}, \mathbf{B}) + \varepsilon \tag{9}$$

where β is p x 1 vector of fixed effect parameters and \mathbf{A} is a corresponding fixed effect design matrix; b is q x 1 vector of random effect parameters and \mathbf{B} is a corresponding random effect design matrix; and ε is N x 1 vector of error term with $\varepsilon \stackrel{i.i.d.}{\sim} N(0, \sigma^2 \Lambda)$.

Taking first-order Taylor series linearization of Eq. [9] around the estimates of $(\beta, b) = (\hat{\beta}, \hat{b})$ gives

$$\hat{o} \approx f(\hat{\beta}, \hat{b}, \mathbf{A}, \mathbf{B}) + \tilde{\mathbf{A}}(\beta - \hat{\beta}) + \tilde{\mathbf{B}}(b - \hat{b}) + \varepsilon \tag{10}$$

where $\tilde{\mathbf{A}} = \left. \frac{\partial f(\beta, b, \mathbf{A}, \mathbf{B})}{\partial \beta} \right|_{\beta = \hat{\beta}, b = \hat{b}}$, $\tilde{\mathbf{B}} = \left. \frac{\partial f(\beta, b, \mathbf{A}, \mathbf{B})}{\partial b} \right|_{\beta = \hat{\beta}, b = \hat{b}}$

Now the Eq. 10 can be represented as a LMM on $\mathbf{z} = \mathbf{y} - \mathbf{f}(\hat{\beta}, \hat{b}, \mathbf{A}, \mathbf{B}) + \tilde{\mathbf{A}}\hat{\beta} + \tilde{\mathbf{B}}\hat{b}$ as follows

$$\mathbf{z} = \tilde{\mathbf{A}}\beta + \tilde{\mathbf{B}}b + \varepsilon \tag{11}$$

Let us decompose the response vector $\hat{o} = (\hat{o}_s^T, \hat{o}_h^T)$, where \hat{o} represents the observed component and \hat{o}_h represents the unobserved component. Accordingly, all other model quantities can be divided as

$$\mathbf{z} = \begin{pmatrix} \mathbf{z}_s \\ \mathbf{z}_h \end{pmatrix}; \quad \tilde{\mathbf{A}} = \begin{pmatrix} \tilde{\mathbf{A}}_s \\ \tilde{\mathbf{A}}_h \end{pmatrix}; \quad \tilde{\mathbf{B}} = \begin{pmatrix} \tilde{\mathbf{B}}_s \\ \tilde{\mathbf{B}}_h \end{pmatrix};$$

$$\mathbf{f}(\hat{\beta}, \hat{b}, \mathbf{A}, \mathbf{B}) = \begin{pmatrix} \mathbf{f}_s(\hat{\beta}, \hat{b}, \mathbf{A}_s, \mathbf{B}_s) \\ \mathbf{f}_h(\hat{\beta}, \hat{b}, \mathbf{A}_h, \mathbf{B}_h) \end{pmatrix}$$

Then based on LMM [11], the empirical BLUP of \mathbf{z}_h based on \mathbf{z}_s is given as

$$\mathbf{z}_h = \tilde{\mathbf{A}}_h \hat{\beta} + \tilde{\mathbf{V}}_{hs} \tilde{\mathbf{V}}_{ss}^{-1} (\mathbf{z}_s - \tilde{\mathbf{A}}_s \hat{\beta}) \tag{12}$$

where $\tilde{\mathbf{V}} = \tilde{\mathbf{B}} \hat{\text{var}}(b) \tilde{\mathbf{B}}^T + \hat{\text{var}}(\varepsilon)$, the variance-covariance matrix of \mathbf{z} based on LMM approximation [11], which can be decomposed into

$$\tilde{\mathbf{V}} = \begin{pmatrix} \tilde{\mathbf{V}}_{ss} & \tilde{\mathbf{V}}_{sh} \\ \tilde{\mathbf{V}}_{hs} & \tilde{\mathbf{V}}_{hh} \end{pmatrix}$$

By rearranging [12] using the relation between \mathbf{z} and $\hat{\boldsymbol{\theta}}_h$, we will get our predictor for $\hat{\boldsymbol{\theta}}_h$ as

$$\hat{\mathbf{y}}_h = \mathbf{f}_h(\hat{\boldsymbol{\beta}}, \hat{\mathbf{b}}, \mathbf{A}_h, \mathbf{B}_h) - \tilde{\mathbf{B}}_h \hat{\mathbf{b}} + \tilde{\mathbf{V}}_{hs} \tilde{\mathbf{V}}_{ss}^{-1} \{ \mathbf{y}_s - \mathbf{f}_s(\hat{\boldsymbol{\beta}}, \hat{\mathbf{b}}, \mathbf{A}_s, \mathbf{B}_s) + \tilde{\mathbf{B}}_s \hat{\mathbf{b}} \} \quad [13]$$

When $\text{COV}(\boldsymbol{\varepsilon}_s, \boldsymbol{\varepsilon}_h) \neq \mathbf{0}$, the predictor specified in Eq. [13] takes account of this dependence through $\tilde{\mathbf{V}}_{hs}$. However when $\text{COV}(\boldsymbol{\varepsilon}_s, \boldsymbol{\varepsilon}_h) = \mathbf{0}$, $\hat{\boldsymbol{\theta}}_h$ and $\hat{\boldsymbol{\theta}}_s$ are correlated only through the shared random effects and is best approximated by the plug-in-predictor $\hat{\mathbf{y}}_h = \mathbf{f}_h(\hat{\boldsymbol{\beta}}, \hat{\mathbf{b}}, \mathbf{A}_h, \mathbf{B}_h)$. Since we are interested in predicting the value of one response variable using data where another response variable is available or measured at the same height from the same stand, we expect that the predictor [13] performs better than the plug-in-predictor.

In order to evaluate the predictive performance of the fitted multivariate model, we randomly selected data from 25 stands. We created a new data set with data from the 25 selected stands excluded (apart from data measured at relative heights equivalent to heights of 1.37 m and 13.7 m to get the estimate of random effect while fitting) and refitted the final model SGMC3 to this new data. We made predictions based on [13] for both disk SG and MC for the selected 25 stands that were not used for model fitting. Disk SG was predicted for the 25 excluded stands assuming that disk MC measurements were available for all heights and stands. The same assumption was made for disk SG when disk MC was predicted for the excluded stands.

Plots showing the univariate plug-in-prediction, multivariate plug-in-prediction and multivariate improved prediction (based on Eq. [13]) of disk SG and MC for 5 stands randomly selected from the excluded 25 are presented in Figure 3.3 and 3.4. We can see from the figures that additional information for one response variable significantly improved the prediction of the other response variable using Eq. [13] compared to the plug-in-predictors. The curves are closer to their observed values for both disk SG and MC using the Eq. [13] predictor. Table 3.5, presents the root mean square prediction error (RMSPE) for the three prediction methods based on predictions of SG and MC for trees from the 25 excluded stands. Prediction from multivariate approaches, both plug-in-predictor and Eq. [13], was considerably better than those of the univariate approach. Prediction based in Eq. [13] were improved by 29 (SG) and 26 % (MC) (Table 3.5).

DISCUSSION

Nonlinear mixed models are an important tool for modeling and predicting growth and wood quality attributes in

forestry (Fang 1999; Hall and Bailey 2001; Jordan et al. 2008; Jordan et al. 2006). Univariate mixed models were commonly used in forestry to model different growth and wood properties. Compared to conventional methods univariate mixed models provide improved predictions because of their ability to capture different levels of variability within the data, e.g. variability from stand-to-stand, plot-to-plot and tree-to-tree (Fang et al. 2001) through random effects in the models. In addition to variability observed at different levels of the data, individual components (properties) measured from a forest are usually inter-dependent. The simultaneous modeling technique can take account of the inter-dependency in a system through random effects and the inter-dependency among different components in the system through contemporaneous correlation.

In this article, we proposed a multivariate simultaneous mixed model for stand average disk SG and MC at different tree heights. We observed a high correlation (-0.78) between two components in our system. The inverse relation between SG and MC was identified by Koch (1972), Zobel and Blair (1976) and Zobel and van Buijtenen (1989). Various explanations have been proposed for the inverse relation between SG and MC within trees such as the amount of heartwood, the presence of extractives and the proportion of juvenile wood. According to Zobel and Blair (1976), the dominant factor controlling SG and MC variation within a loblolly pine tree is the proportion of juvenile wood and the proportion of juvenile wood increases longitudinally from stump-to-tip of loblolly pine trees.

The advantage of multivariate simultaneous systems is their improvement in prediction in one component given the other components in the system (Fang et al. 2001; Hall and Clutter 2004). Based on this study, we found a significant improvement in prediction for both properties, approximately 29 and 26 percent reduction in RMSPE for both disk SG and MC respectively, based on the simultaneous system after taking account of the contemporaneous correlation between the components. The multivariate plug-in-predictor improved by 5 and 11 percent in RMSPE compared to univariate approach for both disk SG and MC respectively. This clearly indicates the potential of multivariate model fitting over univariate approach. Operationally, the proposed system can be used to improve the prediction of stand disk SG at different height levels using the measured disk MC using non-destructive sampling methods.

ACKNOWLEDGEMENTS

Thanks are due to the Wood Quality Consortium, University of Georgia for providing the data for this study. Thanks also

to the USDA Forest Service Southern Research Station, Athens, GA for preparing the samples and providing the lab facilities.

LITERATURE CITED

- Borders, B.E.** 1989. System of equation in forest stand modeling. *Forest Science* 35:548-556.
- Clutter, J.L.** 1963. Compatible growth and yield models for loblolly pine. *Forest Science* 9(3):354-370.
- Fang, Z.** 1999. A simultaneous system of linear and nonlinear mixed effects models for forest growth and yield prediction (Ph.D dissertation), University of Georgia, Athens. 177 p.
- Fang, Z., R.L. Bailey, and B.D. Shiver.** 2001. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. *Forest Science* 47(4):550-562.
- Hall, D.B., and R.L. Bailey.** 2001. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *Forest Science* 47(3):311-321.
- Hall, D.B., and M. Clutter.** 2004. Multivariate multilevel nonlinear mixed effects models for timber yield predictions. *Biometrics* 60(1):16-24.
- Haygreen, J.G., and J.L. Bowyer.** 1996. *Forest Products and Wood Science: An Introduction. Third Edition.* Ames: Iowa State University Press, Iowa. pp. 80-102.
- He, R.** 2004. Mixed effects modeling of wood properties of loblolly pine in the southeastern United States (Ph.D dissertation), University of Georgia, Athens. 213 p.
- Jordan, L., A. Clark, L.R. Schimleck, D.B. Hall, and R.F. Daniels.** 2008. Regional variation in wood specific gravity of planted loblolly pine in the United States. *Canadian Journal of Forest Research* 38:698-710.
- Jordan, L., R. He, D.B. Hall, A. Clark, and R.F. Daniels.** 2006. Variation in loblolly pine cross-sectional microfibril angle with tree height and physiographic region *Wood and Fiber Science* 38(3):390-398.
- Koch, P.** 1972. *Utilization of the Southern Pines.* USDA Forest Service, Southern Forest Experiment Station, Agriculture Handbook No. 420. 235-260.
- Max, T.A., and H.E. Burkhart.** 1976. Segmented polynomial regression applied to taper equations. *Forest Science* 22:283-289.
- Megraw, R.A.** 1985. Wood quality factors in loblolly pine. TAPPI Press, Atlanta, GA.
- Panshin, A.J., and C. deZeeuw.** 1980. *Textbook of Wood Technology.* McGraw-Hill Book Co., New York. 705 p.
- Phillips, K.M.** 2002. Modeling within tree changes in wood specific gravity and moisture content for loblolly pine in Georgia (M.S. Thesis), University of Georgia, Athens. 95 p.
- Pinheiro, C.J., and D.M. Bates.** 2000. Mixed effects models in S and S-plus. Springer, New York. 523 p.
- Pinheiro, C.J., M.D. Bates, S. DebRoy, D. Sarkar, and R.C. team.** 2009. nlme: Linear and Nonlinear Mixed Effects Models. R package version 3. 1-92.
- Zobel, B.J., and R. Blair.** 1976. Wood and pulp properties of juvenile wood and topwood of the southern pines. *Journal of Applied Polymer Science* 28:421-433.
- Zobel, B.J., and J.P. vanBuijtenen.** 1989. *Wood Variation: Its Causes and Control.* Springer-Verlag, Berlin. 363 p.

Forest Ecosystems

MODELING FOREST ECOSYSTEM CHANGES RESULTING FROM SURFACE COAL MINING IN WEST VIRGINIA John Brown, Andrew J. Lister, Mary Ann Fajvan, Bonnie Ruefenacht, and Christine Mazzarella.....	67
A PRELIMINARY TEST OF AN ECOLOGICAL CLASSIFICATION SYSTEM FOR THE OCONEE NATIONAL FOREST USING FOREST INVENTORY AND ANALYSIS DATA W. Henry McNab, Ronald B. Stephens, Richard D. Rightmyer, Erika M. Mavity, and Samuel G. Lambert.....	77
CHANGES IN EARLY-SUCCESSIONAL HARDWOOD FOREST AREA IN FOUR BIRD CONSERVATION REGIONS ACROSS FOUR DECADES Sonja N. Oswalt, Kathleen E. Franzreb, and David A. Buehler.....	87
RELATIONSHIPS BETWEEN HARVEST OF AMERICAN GINSENG AND HARDWOOD TIMBER PRODUCTION Stephen P. Prisley, James Chamberlain, and Michael McGuffin.....	95

MODELING FOREST ECOSYSTEM CHANGES RESULTING FROM SURFACE COAL MINING IN WEST VIRGINIA

John Brown, Andrew J. Lister, Mary Ann Fajvan, Bonnie Ruefenacht, Christine Mazzarella

ABSTRACT

The objective of this project is to assess the effects of surface coal mining on forest ecosystem disturbance and restoration in the Coal River Subbasin in southern West Virginia. Our approach is to develop disturbance impact models for this subbasin that will serve as a case study for testing the feasibility of integrating currently available GIS data layers, remote sensing, and existing Forest Inventory and Analysis program (FIA) data.

Using a set of 30-m-pixel based GIS-based predictor layers (topography, soils and imagery), we developed models that predict total forest carbon for each pixel in the study area. By combining the vegetation change tracker (VCT) year of disturbance outputs with an annual biomass map derived from modeling the FIA data, we will be able to determine biomass losses from mining and estimate potential forest regrowth.

INTRODUCTION

The challenge of mitigating greenhouse gases has resulted in considerable focus being placed on the carbon storage capacities of forests. Trees and other plants naturally remove carbon dioxide (CO₂) from the atmosphere and temporarily convert (sequester) carbon in wood, roots, leaves and the soil. In the Appalachian region of Kentucky, Virginia, Tennessee, and West Virginia, mountaintop removal mining has been prevalent since 1985 (US EPA 2005). This mining technique requires the removal (flattening) of mountain peaks to access the coal layers below. The waste material that is removed is pushed into adjacent valleys (valley fills), burying many headwater streams. Utilization of this mining technique increased with the 1990 amendments to the Clean Air Act, when mining and electric companies focused on extraction of low-sulfur coal to meet the new standards (Fox 1999). At about the same time, larger and more efficient machinery became available for excavation and removal (Szwilski and others, 2001). Between 1985 and 2001, 6,697 valley fills were approved by agencies in these States, and these fills would eventually cover 339 square kilometers (US EPA 2005).

In 2006, 43 percent of all coal extracted from West Virginia came from surface mining, 70 percent of which was mined using mountaintop removal methods (Britton 2007). Not only is forest directly lost, but recent studies have demonstrated that the integrity of the residual forest is significantly altered due to fragmentation and the introduction of edge (Wickham and others, 2007). Conversion of large tracts of interior forest to edge results in a host of ecological changes, both aquatic and terrestrial (SAMAB 1996).

Prior to the 1977 Surface Mining Control and Reclamation Act (SMCRA), most mined land in the Appalachian region was planted with trees. The composition and productivity of the resulting forests are highly variable and spatially irregular due to the physical and chemical properties of the residual mine spoil material (Rodrigue and Burger 2002). SMCRA was enacted to reduce problems with severe erosion, sedimentation, landslides and mass instability caused by pre-SMCRA surface mining (Angel and others 2005). SMCRA regulations require mining companies to post a bond that is sufficient to cover the cost of reclaiming a surface mined site. Because of the 5-year timeframe required to demonstrate successful soil stabilization and vegetation reclamation, many surface mined soils are severely and purposely compacted by machinery and converted to grasslands and shrubs. Native forests have not been successfully restored due to several soil factors: poor aeration, high alkalinity, and reduced water infiltration, in addition to severe compaction (Ashby and others, 1984, Andrews and others, 1998). As a result, millions of hectares of grassland and scrubland, in various successional stages, fragment the otherwise forested mountains and reduce the forest's potential to produce timber and sequester carbon (Burger and Maxey 1998).

The Forestry Reclamation Approach (FRA) is a new approach being tested as a method for reclaiming surface-

John Brown, Research Forester, USDA Forest Service, NRS, Princeton, WV
Andrew J. Lister, Research Forester, USDA Forest Service, NRS, Newtown Sq., PA 19073
Mary Ann Fajvan, Research Forester, USDA Forest Service, NRS, Morgantown, WV 26505
Bonnie Ruefenacht, Remote Sensing Specialist, Red Castle Resources, Salt Lake City, UT
Christine Mazzarella, US EPA Region 3, Philadelphia, PA 19103

coal-mined land to forest within the guidelines imposed by SMCRA (Burger and others 2005). FRA recommendations are founded on restoring mine-soil quality to increase potential carbon sequestration. Restoration guidelines include the creation of deep soil rooting medium, suitable for planting native ground covers and tree species, to improve ecological values. Post-mining forest restoration is slowly gaining acceptance; about 30 million trees have been planted since 2005 (Personal communication, Patrick Angel, forester/soil scientist, USDI Office of Surface Mining Reclamation and Enforcement, 421 West Highway 80, London, Kentucky 40741). These forests are very young, hence, the future productivity, value, and carbon sequestration potential of these restored forests is still unknown.

The objective of this project is to assess the effects of surface coal mining on forest ecosystem disturbance and simulated restoration in the Coal River Mountain watershed in southern West Virginia. This watershed already has active surface mining. Three new and proposed mountaintop removal mines are projected to produce more than 47 million tons of coal from 2009 through 2025 (WV DEP 2008). Our approach develops disturbance impact models for a sub-watershed that will serve as a case study for testing the feasibility of integrating currently available GIS data layers, remote sensing, and existing data from the USDA Forest Inventory and Analysis (FIA) program. Specifically, we will 1) identify specific areas and ecosystems that have been depleted of carbon stocks; and 2) calculate the reduction relative to a previous condition. This paper presents the methods used to accomplish these two tasks and presents initial results of our biomass modeling efforts. Our ultimate goal is to model the change in carbon stocks from anticipated forest restoration activities using FRA guidelines and make comparisons with the previous condition to determine the long-term effects of the proposed mining on the watershed.

MATERIALS AND METHODS

To identify the year and spatial extent of forest disturbance due to surface mining and to generate maps to estimate the pre- and post- disturbance carbon stocks in these areas, a regression tree predictive modeling approach was employed using Cubist software (www.rulequest.com), which is based on a process created by Quinlan (1992). While the algorithm that Cubist employs is proprietary, generally speaking, regression trees work by using classification trees to classify instances into groups based on values of a set of independent variables and a dependent variable, and then developing regression models that describe the relationship between the dependent and independent variables using the instances contained in each of the classification tree's terminal nodes.

For our regression tree, we used several GIS-based predictor layers as the independent (predictor) variables, and we used total aboveground carbon estimates generated from forest inventory plots as the dependent variables.

INDEPENDENT VARIABLES

Landsat image data were obtained from the US Geological Survey (USGS) Global Visualization Viewer (GLOVIS) data distribution system (<http://glovis.usgs.gov>), and consisted of a set of annual Landsat 5 scenes collected over path/row 18/34 during the growing season. Image dates (month/day/year) included the following days: 9/17/1984, 9/20/1985, 7/5/1986, 6/6/1987, 6/8/1988, 8/17/1990, 9/21/1991, 6/3/1992, 8/25/1993, 10/15/1994, 8/31/1995, 10/4/1996, 9/5/1997, 8/7/1998, 6/23/1999, 6/9/2000, 10/2/2001, 8/2/2002, 6/2/2003, 6/20/2004, 9/11/2005, 8/13/2006, 9/17/2007, 7/17/2008, and 6/2/2009; suitable data were unavailable for 1989. These scenes were 30-m pixel size and processed by the USGS to Level 1T (terrain corrected) using the Level 1 Product Generation System (USGS 2011) and were further processed using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) software (Masek and others, 2006). LEDAPS software produces atmospherically-corrected, surface reflectance-calibrated imagery that can be used to assess environmental and land cover change (Masek and others, 2006). From the scenes that were available for each year within the growing season, bands 1-5 and 7 of the scene with the greatest cloud-free area were selected.

Other data used for this study are listed in table 1 and included a 10-m elevation dataset obtained from a subset of the National Elevation Dataset (NED) (Gesch and others, 2002), raster elevation derivative datasets created using the NED data, and data from the Soil Survey Geographic (SSURGO) database (NRCS 2011). Also, for each Landsat scene, the disturbance magnitude of the difference Normalized Burn Ratio (dNBR) was created using vegetation change tracker (VCT) software (Huang and others, 2010).

DEPENDENT VARIABLE

Estimates of total aboveground carbon (TAG) were obtained using allometric equations that were applied to data collected by the FIA on the 69 inventory plots found in the portion of the Coal River watershed found within Landsat path/row 18/34 (fig. 1). TAG is calculated as described in Woudenberg and others (2011) and includes the carbon mass of the aboveground portion of live trees with a diameter of 2.5 cm or larger and dead trees with a diameter of 12.7 cm or larger. The FIA data were collected between 2004 and 2008 and consisted of plots with pure stands or hardwoods or conifers.

MODEL DEVELOPMENT

The latitude and longitude of the FIA plots were used to intersect them with the set of predictor data using a GIS, and values for each independent variable were assigned to the TAG value associated with each plot to create the training data for the regression tree modeling. The elevation, elevation derivatives, and SSURGO data were assumed to be temporally constant and these and LEDAPS-calibrated Landsat and VCT Landsat derivatives (dNBR and NDVI) from 2007 were used to build the initial model. Model results were assessed using cross validation (10 percent holdout) statistics: mean absolute error (MAE); relative error (RE), or the ratio of the MAE to the error magnitude that would result from always predicting the mean value; and the correlation coefficient (r) that describes the strength of the relationship between each set of predictions and carbon values from the holdout data. Using a combination of these metrics, correlation matrices, and experience from prior modeling, data reduction was performed automatically and heuristically until a set of independent variables was chosen to produce the final model for 2007 imagery.

Because the Landsat imagery was calibrated using LEDAPS, we, like Powell and others, (2010), made the assumption that variations in pixel values between corresponding surface reflectance-calibrated images were due to changes in the reflective characteristics of the landscape and not due to differences in the atmosphere or sensor position. We thus applied the regression tree model developed for the 2007 Landsat and ancillary data to the corresponding data for each year of Landsat data between 1984 through 2009 to produce a set of 25 (yearly between 1984 and 2009) maps of carbon estimates for the watershed.

RESULTS AND DISCUSSION

The nonlinear portion of the regression tree process does not have many of the assumptions of linear modeling and is generally effective at choosing the best attributes to use in decision rules from among several potentially collinear variables. However, through a combination of examining cross validation (10 percent holdout) results from Cubist and arbitrary decisions, only 35 of the original variables were used to produce the final model.

The Cubist model output is shown in figure 2. Cubist used 13 exploratory variables. Five variables were important to the classification portion of the Cubist analysis: dNBR, landform, X, Y, and profile curvature. Of these, profile curvature was present in five of the six rules developed, while the remaining four were present in at least half of the rules. Two variables, landform and Y, were only used in the decision process (table 2). Each of the remaining 11 variables was involved infrequently with the linear models for each rule. Only one variable, heatload, was

present in half the rules (3 of 6) while the remaining variables were present for only one or two of the six rules generated. In general, coefficients calculated for specific variables during the linear model steps were consistent in sign from rule to rule, i.e., if a coefficient was positive for a variable in one rule it was positive as well in other rules. The actual values plotted against the predicted values have a reasonably linear relationship (fig. 3). The correlation coefficient was 0.89 ($r^2 = 0.79$).

The Cubist model rules (fig.2) were then applied to the aforementioned LEDAPS processed Landsat scenes resulting in TAG estimates maps for nearly all years from 1984-2009. Four of these maps are illustrated in figure 4, where an 8-year interval was used to demonstrate applicability of the model. Rivers and streams clearly appear as white lines within the maps, and irregular patches correspond with areas of disturbance, some of which is already identified as surface coal mining activity. The distinct boundaries that appear in the final map are due to the use of the Easting and Northing in the decision rules. While the existence of these lines creates a visual anomaly, the use of the map is a geospatial dataset that will provide pixel value summaries that serve as estimates. It is recognized that the presence of these discontinuities indicates that additional effort is needed to further refine the predictive models.

CONCLUSIONS

Methodology developed to date demonstrates the feasibility of utilizing a set of GIS predictor layers to generate temporal maps of total aboveground carbon for a watershed containing surface mining activity in West Virginia. This is an important step in the ultimate goal of assessing the amount of carbon stock removed in disturbance events, specifically surface coal mining. Subsequent steps will compare output from the VCT disturbance maps and the predicted TAG maps which will enable temporal removals of carbon stock for the period 1984-2009. Additionally, it is hoped that these later results will have broader applicability to other watersheds containing surface mining activity.

DISCLAIMER—The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency. The U.S. government has the right to retain a nonexclusive royalty-free license in and to any copyright covering this article.

LITERATURE CITED

Andrews, J.A.; Johnson, J.; Torbert, J. [and others]. 1998. Minesoil and site properties associated with early height growth of eastern white pine. *Journal of Environmental Quality*. 27:192-199.

- Angel, P.;** Davis, V.; Burger, J.A. [and others]. 2005. The Appalachian regional reforestation initiative. US Office of Surface Mining. Forest Reclamation Advisory No. 1. 2 p. <http://arri.osmre.gov/>. [Date accessed: March 8, 2009].
- Ashby, W.C.;** Vogel, W.G.; Kolar, C.A.; Philo, G.R. 1984. Productivity of stony soils on strip mines. In: Nichols, J. D.; Brown, P.L., eds. Erosion and productivity of soils containing rock fragments; Proceedings of a symposium. Special Publ. No. 13. Madison, WI: Soil Science Society of America: 31-44.
- Bolstad, P.V.;** Lillesand, T.M. 1992. Improved classification of forest vegetation in Northern Wisconsin through a rule-based combination of soils, terrain, and Landsat TM data. *Forest Science*. 38(1): 5-20.
- Britton, J.Q.;** Blake, B. M. Jr.; McColloch, G. H. 2007. West Virginia. *Mining Engineering* 59(5):123-125
- Burger, J.A.;** Maxey, W.R. 1998. Maximizing the value of forests on reclaimed mined land. *Green Lands*. 28:37-46.
- Burger, J.;** Graves, D.; Angel, P. [and others]. 2005. The forestry reclamation approach. Forest Reclamation Advisory No. 2. US Office of Surface Mining. 4 p. Available at <http://arri.osmre.gov/>. [Date accessed: March 8, 2009].
- ESRI.** 2011a. How hillshade works. Redlands, CA: ESRI Development Network. Available at http://edndoc.esri.com/arcobjects/9.2/net/shared/geoprocessing/spatial_analyst_tools/how_hillshade_works.htm. [Date accessed: February 1, 2011].
- ESRI.** 2011b. How curvature works. Redlands, CA: ESRI Development Network. Available at http://edndoc.esri.com/arcobjects/9.2/NET/shared/geoprocessing/spatial_analyst_tools/how_curvature_works.htm. [Date accessed: February 1, 2011].
- Fox, J.** 1999. Mountaintop removal in West Virginia. *Organic Environment*. 12:163-183.
- Gesch, D.;** Oimoen, M.; Greenlee, S. [and others]. 2002. The national elevation dataset. *Photogrammetric Engineering & Remote Sensing*. 68(1):5-32.
- Huang, C.;** Goward, S.N.; Masek, J.G. [and others]. 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*. 114:183-198.
- Masek, J.G.;** Vermote, E.F.; Saleous, N. [and others]. 2006. A Landsat surface reflectance data set for North America, 1990-2000. *Geoscience and Remote Sensing Letters*. 3:68-72.
- McCune, B.;** Keon, D. 2002. Equations for potential annual direct incident radiation and heat load. *Journal of Vegetation Science*. 13:603-606.
- McNab, H.W.** 1989. Terrain shape index: quantifying effect of minor landforms on tree height. *Forest Science*. 35(1):91-104.
- NRCS (Natural Resources Conservation Service).** 2011. Soil survey geographic (SSURGO) database for West Virginia. Fort Worth, TX: U.S. Department of Agriculture, Natural Resources Conservation Service. Available <http://soildatamart.nrcs.usda.gov>. Date accessed: February 1, 2011].
- Parker, A.J.** 1982. The topographic relative moisture index: an approach to soil-moisture assessment in mountain terrain. *Physische Geographie*. 3:160-168.
- Powell, S.L.;** Cohen, W.B.; Healey, S.P. [and others]. 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sensing of Environment*. 114(5): 1053-1068.
- Quinlan, J.R.** 1992. Learning with continuous classes. In: Adams, A.; Sterling, L., eds. Proceedings, 5th Australian joint conference on artificial intelligence. Singapore:World Scientific: 343-348.
- Riley, S.J.;** DeGloria, S.D.; Elliot, R. 1999. A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences*. 5:1-4.
- Roberts, D.W.;** Cooper, S.V. 1989. Concepts and techniques of vegetation mapping. In: Land classifications based on vegetation: Applications for resource management. Gen. Tech. Rep. INT-257. Ogden, UT: U.S. Department of Agriculture Forest Service, Intermountain Forest and Range Experiment Station: 90-96.
- Rodrigue, J.A.;** Burger, J.A. 2002. Forest productivity and commercial value of pre-law reclaimed mined land in the eastern United States. *Northern Journal of Applied Forestry*. 19(3): 106-114.
- SAMAB (Southern Appalachian Man and the Biosphere).** 1996. The Southern Appalachian assessment terrestrial resources technical report. Report 5 of 5. Atlanta, GA: U.S. Department of Agriculture, Forest Service, Southern Region. Available at http://samab.org/saa/saa_reports.html [Date accessed: April 5, 2011].
- Stage, A.R.** 1976. An expression of the effects of aspect, slope, and habitat type on tree growth. *Forest Science*. 22:457-460.
- Szwilski, T.B.;** Dulin, B.E.; Hooper, J.W. 2001. An innovative approach to managing the impacts of mountaintop coal mining in West Virginia. *International Journal of Surface Mining, Reclamation and Environment*. 15:73-85.
- US EPA (U.S. Environmental Protection Agency).** 2005. Mountaintop mining/valley fills in Appalachia: Final programmatic environmental impact statement. EPA 9-03-R-05002.
- U.S Environmental Protection Agency.** Philadelphia, PA: Region 3. Available at http://www.epa.gov/region3/mnttop/pdf/mtm-vf_fpeis_full-document.pdf [Date accessed: April 5, 2011].
- USGS (United States Geologic Survey).** 2011. Landsat 7 Science Data Users Handbook, Chapter 10: Level 1 Processing. <http://landsathandbook.gsfc.nasa.gov/level/> [Date accessed: April 5, 2011]
- WV DEP (West Virginia Department of Environmental Protection).** 2008. Mining permit search. Charlestown, WV: West Virginia Department of Environmental Protection. Available at: <http://www.dep.wv.gov/insidedep/Pages/miningpermitsearch.aspx>. [Date accessed: April 5, 2011].
- Wickham, J.D.;** Riitters, K.H.; Wade, T.G. [and others]. 2007. The effect of Appalachian mountaintop mining on interior forest. *Landscape Ecology*. 22:179-187.
- Woudenberg, S.W.;** Conkling, B.L.; O'Connell, B.M. [and others]. 2011. The Forest Inventory and Analysis database: database description and user's manual version 4.0 for Phase 2. Gen. Tech. Rep. RMRS-GTR-245. Fort Collins, CO: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station.

Table 1—List of datasets assessed for inclusion in Cubist regression tree modeling procedure

Dataset Name	Dataset Description	Source
Forest productivity of yellow poplar	Index of forest productivity	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Forest productivity of red oak	Index of forest productivity	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Forest productivity of white oak	Index of forest productivity	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Site index northern red oak	Index of forest productivity	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Site index white oak	Index of forest productivity	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Site index yellow poplar	Index of forest productivity	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Seedling mortality index	Index of seedling mortality likelihood	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Depth to fragipan layer	Depth to a fragipan restrictive soil layer	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Depth to lithic bedrock	Depth to a lithic bedrock restrictive soil layer	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Depth to paralithic bedrock	Depth to a paralithic restrictive soil layer	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Depth to restrictive layer	Depth to any restrictive layer	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Depth to water table	Depth to the water table	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Liquid limit	Index related to the range of water contents over which a soil exhibits liquidity	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Plasticity index	Index related to range of water content over which a soil exhibits solidity	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Soil organic matter percent	Percent soil organic matter in the top soil layer	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Clay percent	Percent clay content of the surface soil layer	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Sand percent	Percent sand content of the surface soil layer	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Silt percent	Percent silt content of the surface soil layer	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)

Table 1—(Continued) List of datasets assessed for inclusion in Cubist regression tree modeling procedure

Dataset Name	Dataset Description	Source
Rock type	Categorical value representing different bedrock types	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Soil pH in water	pH of soil mixed in water	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Cation-exchange capacity (CEC-7)	Cation exchange capacity of the surface soil layer	Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture (2011)
Elevation	Elevation	Gesch <i>et al.</i> (2002)
Filtered elevation range	Elevation range within a 90-m square buffer centered on each pixel	Gesch <i>et al.</i> (2002)
Filtered mean elevation	Mean elevation within 90-m square buffer centered on each pixel	Gesch <i>et al.</i> (2002)
Filtered mean-minimum elevation range	Mean elevation - minimum elevation within 90-m square buffer centered on each pixel	Gesch <i>et al.</i> (2002)
Transformed aspect	Linear transformation of aspect	Roberts and Cooper (1989)
Cosine-transformed aspect-slope	Cos(aspect) X percent slope	Stage (1976)
Sine-transformed aspect-slope	Sin(aspect) X percent slope	Stage (1976)
Relative moisture index	Index of relative amount of moisture available at a site	Parker (1982)
Modified relative moisture index	Variation of relative moisture index	Parker (1982)
Heatload	An index of the relative amount of solar radiation that a site receives	McCune and Keon (2002)
Hillshade	An index of solar radiation a site receives, incorporating shadows and illumination angle	ESRI (2011a)
Bolstad's landform	A landform index	Bolstad and Lillesand (1992)
McNab's landform	A landform index	McNab (1989)
Planform curvature	An index of curvature of the land surface	ESRI (2011b)
Slope curvature	An index of curvature of the land surface	ESRI (2011b)
Profile curvature	An index of curvature of the land surface	ESRI (2011b)
Relative slope position	An index of slope position between valley bottom and ridge top	Unknown; based on ESRI topographic functions
Slope position	Position of the pixel as a percentage between the valley floor and ridgetop.	Unknown; based on ESRI topographic functions
Landform type	A categorical variable representing landform shape and position	Parker (1982)
Surface area : ground area ratio	An index of topographic complexity	Unknown; based on ESRI topographic functions
Topographic roughness index	An index of topographic complexity	Riley <i>et al.</i> (1999)
Easting	The value of geographic coordinate in UTM meters	Native ESRI functionality (xmap and ymap environment variables)
Northing	The value of geographic coordinate in UTM meters	Native ESRI functionality (xmap and ymap environment variables)
Easting X Northing	Easting X Northing	Native ESRI functionality (xmap and ymap environment variables)

Table 2— Frequencies of occurrence and general coefficient patterns for important variables in Cubist rules

Variable	Decision process	Regression models	Coefficient positive	Coefficient negative
	<i>No. of rules used in</i>	<i>No. of rules used in</i>	<i>Number</i>	<i>Number</i>
dNBR	3	2	2	0
Landform	4	0	-	-
X	4	1	1	0
Y*	4	0	-	-
Profile curvature	5	1	0	1
Slope*	0	2	2	0
COS(Aspect) transformation*	0	2	2	0
Relative slope position	0	1	0	1
Landsat band 6	0	2	0	2
Landsat band 4	0	2	2	0
Transformed aspect	0	2	0	2
Heatload	0	3	3	0
Slope position	0	1	0	1



Figure 1— The study site in southern West Virginia, comprised of the portion of the Coal River watershed found within the boundary of Landsat scene 18/34.

Rule 1: [7 cases, mean 1759.924, range 0 to 12319.47, est err 4022.683]

if
 dnbr <= 133
 then
 Total Above Ground Carbon = 1759.924

Rule 2: [6 cases, mean 27173.246, range 11515.6 to 42868.64, est err 15844.396]

if
 dnbr > 133
 landform in {4, 7, 8}
 profile curvature <= -0.02502192
 x * y > 1.846332e+012
 then
 Total Above Ground Carbon = -241151.452 + 1865 dnbr + 29039 slope * COS(aspect) transformation

Rule 3: [32 cases, mean 49472.605, range 4402.104 to 98225.22, est err 17879.621]

if
 dnbr > 133
 profile curvature > -0.02502192
 then
 Total Above Ground Carbon = -254641.664 + 0.72 x - 534 relative slope position - 8 landsat band 6 + 2 landsat band 4

Rule 4: [8 cases, mean 65291.813, range 42630.82 to 80570.83, est err 16716.725]

if
 landform in {3, 6, 9, 10}
 profile curvature <= -0.02502192
 x * y > 1.846332e+012
 then
 Total Above Ground Carbon = -129803.498 + 1086 dnbr - 32917 transformed aspect + 1.9 heatload - 8 landsat band 6
 + 6841 slope * COS(aspect) transformation + 2 landsat band 4

Rule 5: [12 cases, mean 78180.602, range 59397.14 to 121845.6, est err 12838.607]

if
 landform in {6, 7, 8, 10}
 profile curvature <= -0.02502192
 x * y <= 1.846332e+012
 then
 Total Above Ground Carbon = -15215.85 - 29792 profile curvature - 508 slope position + 3.8 heatload

Rule 6: [4 cases, mean 122093.297, range 100926.2 to 153119.2, est err 11975.873]

if
 landform in {3, 5, 9}
 profile curvature <= -0.02502192
 x * y <= 1.846332e+012
 then
 Total Above Ground Carbon = -323169.019 - 99972 transformed aspect + 2398 dnbr + 4.7 heatload

Average lerrorl 10856.566

Relative lerrorl 0.42

Correlation coefficient 0.89

Figure 2—Cubist output modeling total aboveground carbon.

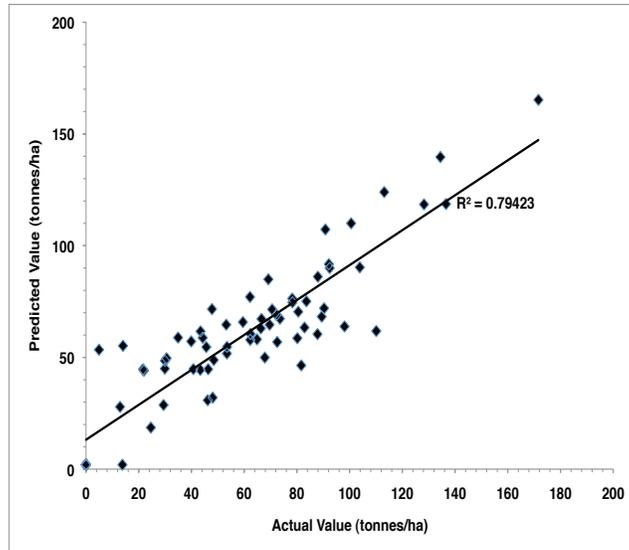


Figure 3—Cubist total aboveground carbon predictive values vs. the actual total aboveground carbon values.

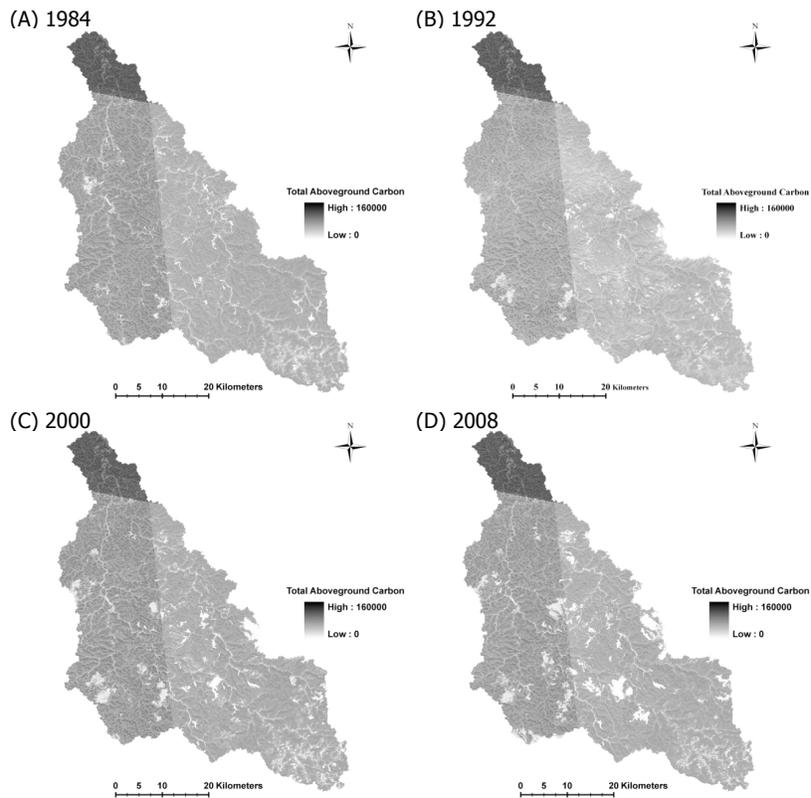


Figure 4—Prediction maps for total aboveground carbon. Selected maps were produced at 8-year intervals.

A PRELIMINARY TEST OF AN ECOLOGICAL CLASSIFICATION SYSTEM FOR THE OCONEE NATIONAL FOREST USING FOREST INVENTORY AND ANALYSIS DATA

W. Henry McNab, Ronald B. Stephens, Richard D. Rightmyer, Erika M. Mavity, Samuel G. Lambert

ABSTRACT

An ecological classification system (ECS) has been developed for use in evaluating management, conservation and restoration options for forest and wildlife resources on the Oconee National Forest. Our study was the initial evaluation of the ECS to determine if the units at each level differed in potential productivity. We used loblolly pine (*Pinus taeda*) site index from field plots inventoried by the forest inventory and analysis group of the Forest Service as a measure of productivity at each hierarchical level. The classification system performed best at the landtype level where it identified significant differences in site index between exposed slopes (82 feet) and sheltered slopes (94 feet). Results were less conclusive at the landtype association level, where no clear differences in site index were found among seven units. Results of this preliminary test suggest the ECS will be useful as a guide for diversifying forest cover composition by identifying land units that differ in environmental properties associated with productivity.

INTRODUCTION

The USDA Forest Service adopted a policy in 1992 of using an ecological approach for management of natural resources on national forests and grasslands. To assist managers implement that policy consistently at all administrative levels throughout the agency, an eight-level hierarchical framework concept of ecological units was developed for application from national to local scales (Cleland and others 1997). Ecosystems of national and regional extent have been identified and delineated using a “top-down” method of successive stratification of large regions into subregions that represent smaller ecosystems of increasing uniformity (Cleland and others 2007). For identification of the smallest ecosystems, at landscape and local scales, however, a “bottom-up” method is commonly used where data representing environmental components and associated

vegetation are analyzed and grouped into units of similar ecological potential, productivity, and predictable response to disturbance (VanKley 1993, Hix and Pearcy 1997). Where field data are not initially available to develop a bottom-up ECS, however, a survey method based on existing knowledge of environmental relationships, especially as modeled and analyzed with a geographic information system, can be used for the initial subdivision of large areas to form smaller, tentative ecological units. Testing and validation of a survey-based ECS is highly desirable to identify units that require refinement and to gain confidence from users who did not participate in its development (Rowe and Sheard 1981, Barnes and others 1982).

The Oconee National Forest (ONF) used the survey method to develop an ECS consistent with the national ecological framework to form the basis for a large-scale assessment of opportunities for management of forest resources¹. An interdisciplinary team of resource specialists used expert knowledge of environmental gradients on selected areas of the ONF represented by a range of combinations of bedrock formation, topography, and soils to identify and classify land areas with similar ecological characteristics at a range of scales². The classification system was then applied to the entire ONF using a geographic information system to delineate polygons of similar ecological potential, each of which is hypothesized to enclose an area that differs from its neighbors. The purpose of this study was to begin the process of testing and evaluating the validity of the ONF classification. Our specific objective was to use data from an independent source to determine if the ECS identified land areas that differed in biological response. Our study is considered preliminary because it utilized a small set of existing data to test the classification for only one

¹ Oconee Large Scale Assessment. Unpublished report on file. Chattahoochee-Oconee National Forests, Gainesville, Ga.

² Technical report: Process used in mapping ecological classification system units on the Oconee National Forest of the Georgia Piedmont, July 2009. Unpublished report on file. Chattahoochee-Oconee National Forests. Gainesville, Ga. 9 p.

W. Henry McNab, Research Forester, USDA-Forest Service, Southern Research Station, Asheville, NC 28806
Ronald B. Stephens, Forest Silviculturist (Retired), Richard D. Rightmyer, Forest Soils Scientist, Erika M. Mavity, GIS Specialist, USDA-Forest Service, Chattahoochee-Oconee National Forest, Gainesville, GA 30501
Samuel G. Lambert, Forester, USDA-Forest Service, Southern Research Station, Forest Inventory and Analysis, Knoxville, TN 37919

environmental response, forest productivity, which was not the main goal of the large scale assessment on the ONF. This study is the first part of an ongoing project to test and refine an ECS for the ONF for integration of ecological concepts with natural resource management (Sharitz and others 1992) to evaluate, for example, the effects of forest restoration on water yields in the southern Piedmont (Trimble and others 1987).

STUDY AREA AND HIERARCHICAL ECOLOGICAL UNITS

The ONF is in the Midland Plateau-Central Uplands Subsection, one of ten ecological units that stratify the Southern Appalachian Piedmont Section into smaller areas of more uniform environments associated with climate and surficial geologic materials (Cleland and others 2007). Extending from central Alabama northeast to South Carolina, this large (11,884 miles²) subsection is a region of highly weathered metamorphic gneisses and schists that includes most of the north-central portion of Georgia (Figure 1). Much of this subsection is an exotic terrain that was accreted to the continent during formation of the Appalachian Mountains, which now forms an extensive shield-like plateau underlain by a complex of granitic gneisses and schists that vary in resistance to weathering and associated soils. Two major river systems draining the region have cut deeply into parts of the plateau surface, forming extensive areas of highly dissected topography that extend more than 50 miles north from the boundary of this subsection with the coastal plains. Harper (1930) subdivided the Piedmont physiographic province into upper and lower parts based on the amount of landscape-scale dissection associated with the major river basins. The ONF is in the highly dissected lower part of the Piedmont, where little of the original plateau surface remains. Almost all of the ONF lies off of the plateau surface, on the broad and highly eroded sides of the extensive drainage basins of the Oconee and Ocmulgee Rivers and their tributaries.

Physiography of the subsection varies, but can be generally characterized as a slightly southerly sloping, moderately to strongly dissected peneplain with few surface features. Occasional granite monadnocks are present in the northern part of the subsection and areas of strongly dissected landforms increase to the south, particularly along the east-west Fall Line transition to the coastal plains (Fenneman 1931, Burbanck and Platt 1964). Staheli (1976) found the dendritic drainage pattern of this region was consistent throughout, but differed markedly from the trellis pattern of the Schist Plains Subsection of the Piedmont farther north. Pehl and Brim (1985) show no noteworthy variation of forest habitats in the region they delineate as the Midland Plateau Region of the Piedmont and which they describe

as "...topography gently to steeply undulating, with forest vegetation associated extensively with steeper topography." Wharton (1989) identified a midland subprovince within the Piedmont (which is similar to the Midland Plateau-Central Uplands Subsection) without further subdivision, and described 14 plant communities associated with topographic and soil moisture regimes ranging from hydric river swamps to xeric bluffs.

Quantitative relationships among environmental gradients and vegetation in the lower Piedmont of Georgia are limited to studies by Cowell (1993, 1998). On a landscape scale, he found vegetative communities could be subdivided into two groups: upland and bottomland forests. Cowell found soil fertility (in the upper 4 inches) was more important than moisture (expressed by topographic position and aspect) when accounting for variation in the distribution of tree species on upland sites. Elsewhere in the broader Appalachian Piedmont region, Golden (1979) reported that composition of forest tree and shrub communities in the highly disturbed landscapes of central Alabama was associated with macroscale landscape position ranging from xeric ridgetops to subhydric stream bottoms. Working in South Carolina, Jones (1988) associated composition of old-growth forest vegetation with a moisture gradient, which he suggested was related to landform, aspect, and soil properties. Brender and Davis (1959) concluded that the effects of topography (as it affects site moisture relations) was more important than soil types in determining the rate of hardwood encroachment into pine stands in the lower Piedmont of Georgia. Considerable study, however, has been made of the unusual flora occurring on soils weathered from materials associated with two intrusive geologic formations: granite (Burbanck and Phillips 1983) and gabbro (Schmidt and Barnwell 2002). Wharton (1989) comments that effects of over 200 years of disturbance to soil and vegetation related to European settlement have largely obscured many ecological relationships but historical accounts suggest that arborescent vegetation was associated with "red land" and "gray land" soil types weathered from different types of bedrock. Nelson (1957) provides a county-level map of the "gray lands" that were usually occupied by a pine-hardwood mixture, "granitic lands" (generally near Elberton, Ga.) that were dominated consistently by pine forests, and "red lands" that supported hardwood stands before European settlement. Following almost two centuries of intensive disturbance, hardwood stands are currently found on about 18 percent of the Piedmont landscape, equally distributed between bottomlands and lower slopes of coves (Nelson and others 1957).

For our study, the Midland Plateau-Central Uplands subsection was subdivided into landtype associations (LTA) following the hierarchical structure of the national ecological framework. Seven recurring LTAs, based

primarily on composition of mapped geologic formations occurring within the proclamation boundary of the ONF, were delineated as closed polygons enclosing an area of about 1.4 million acres (Figure 1). Bedrock in this subsection is predominately a mixture of highly weathered northeast-trending bands of metamorphic granitic gneisses and schists that have formed soils that vary mostly in depth and degree of erosion. Upland soils, which make up over 90 percent of the study area, are primarily Ultisols that have a thermic temperature regime, a udic moisture regime, are well drained, highly acidic, and low in fertility. Slope gradients of upland soils range from 2 to 35 percent. Most upland soils are classified as eroded, resulting from a long period of intensive cultivation. Climate of this area is a combination of maritime and continental influences that varies little throughout the subsection. Average monthly temperatures range from 44°F in January to 80°F in July. Almost all precipitation occurs as rain, which averages 48 inches annually. The wettest month is March (5.5 inches) and the driest is October (2.8 inches). Soil moisture deficits usually occur annually during the late growing season as a result of high temperature and low precipitation and often are cumulative during successive years of below average rainfall. Elevation averages about 510 feet (range 321-711 feet) for the study area.

Six landtype (LT) units of the ECS, which occurred within all LTAs, were recognized within the three separated land areas forming the proclamation boundary of the ONF (Table 1). All LTs except one (glade) identify segments of the landscape that define a perceived moisture gradient associated with topography, ranging from dry ridges to wet stream banks. Ridges were separated into three classes: (1) Piedmont plain, (low hills atop the plateau), which occurred only slightly (65 acres) within the LT analysis area, (2) broad ridges, and (3) narrow ridges. (major land divides between tributary streams within the river basins). Slopes were stratified in two groups based on the relative amounts of solar radiation received: (1) exposed (aspects between 158° - 292°) or (2) sheltered (aspects from 293° - 157°). Riparian LTs occurred in bottomlands on sites with moisture regimes ranging from supermesic or subhydic on high floodplains to hydric beside streams. Glades are small (1 - 2 acres) "island-like" LTs occurring on nearly flat uplands underlain by gabbro rock formations that have weathered to form soils with clay B-horizons that are highly impervious to water movement (Schroeder and others 2000). These areas are typically flooded during winter and early spring, but usually experience drought during late summer when precipitation declines. Glades are sites with a unique moisture regime that varies seasonally from xeric to hydric

(Schmidt and Barnwell 2002). On the ONF glades occur as two large areas of about 4,000 acres.

Landtypes were further subdivided into landtype phases (LTP), the lowest and most homogenous level of the ECS. Thirteen units (including water) were identified, 5 of which were associated with upland sites and the others with bottomlands (Table 1). The broad ridge LT was subdivided into two LTPs: (1) broad ridge or (2) narrow ridge. Broad ridges were generally those along the ridge divides of 5th level hydrologic units, termed watersheds in the USGS classification scheme, which generally range in area from 40,000 to 250,000 acres. Narrow ridges typically followed 6th level hydrologic units (sub-watersheds) that range from 10,000 to 40,000 acres. Two LTPs associated with slopes were identified using criteria similar to that for LTs: (1) exposed and (2) sheltered. The LTP designated as upland flat was restricted to the glade LT.

Finally, LTPs were modified (LTPm) to account for the biological effects of differential soil erosion. Each LTP was assigned a code representing one of seven mapped or perceived classes of soil erosion, ranging from slight to severe, resulting in a total of 84 potential classification units. When all national forest lands were classified at the LTPm level, however, only 34 ecological units were identified. Most of the riparian LTPs were represented by a single level of erosion, such as forested wetland-slight erosion or sand levee-slight erosion. Each LTPm represents an ecological unit of varying size with sufficiently uniform physical and chemical properties that combine to form environmental conditions suitable for establishment and maintenance of a characteristic vegetative community.

METHODS

Field data used for testing the classification were obtained from FIA through a standard data service request. Sample plots were restricted to those occurring on sites classified as forest land³. Site index (50 years) of loblolly pine on each sample plot was used as the biological response variable. Site index, a timber-related measure of site quality, was not an ideal choice of response variable considering the ecological objectives of the study, but was the best of those available in the FIA data set. Where site index had been determined for a species other than loblolly pine, it was converted to an equivalent value for loblolly pine using relationships reported by Olson and Della-Bianca (1959), Harrington (1987), and other sources.

³ Forest land is an area >1 acre with at least 10 percent cover by live trees of any size or species, as defined in the Forest Inventory and Analysis Database: Database Description and Users Manual Version 4.0 for Phase2. Draft revision 3. USDA Forest Service. Forest Inventory and Analysis Program. 368 p.

A fixed-effects model analysis of variance was used to determine if the biological response variable (site index) was affected by treatments for each of the four ECS levels. A treatment consisted of the various randomly occurring combinations of environmental variables represented by the classification units within each ECS level. At the LTA level of the ECS, for example, we hypothesized that environmental variation of the Piedmont landscape affecting site index of loblolly pine would be reduced if the underlying geologic formations (see Fig. 1) were taken into account. The seven fixed categories of geology were considered as natural treatments in a completely randomized experimental design. Our null hypothesis, therefore, was that mean site index of loblolly pine did not differ among treatments (i.e. geologic groups or LTAs). Rejection of the null hypothesis resulted in non-rejection of the alternate hypothesis, which stated that mean site index differed among the LTA ecological units. Sample field data to test the hypothesis came from the LTA ecological units in which FIA plots had been placed. Similarly, the six LT ecological units were assumed to be moisture regime treatments that were sampled with randomly located FIA plots to determine if site index differed among them. Each of the four ECS levels was a separate experiment with a different set of treatments.

Sample sizes varied among treatments for each ECS level and depended on criteria used by FIA for establishing field plots and extent of the geographic area being investigated. At the large LTA level (Fig. 1), ECS units sampled with ≥ 4 plots were judged as adequate replication for meaningful analysis. For the LT level and below, where the study area was restricted to the smaller area of the ONF, ECS units with ≥ 3 field sample plots were included in the analysis. Although the minimal replication used in our study would likely result in an analysis with little power to detect real differences among treatments (Zar 1996), it was justified on the basis of increasing knowledge about the function and application of the ECS.

Bartlett's test was used to determine the homogeneity of variances of site index among units at each ECS level. A square root transformation of site index was used where necessary to achieve homogeneity of variance (Zar 1996). When the analysis of variance indicated significant differences were present among mean site index of the ECS units, Tukey's test was used to determine differences among treatments (Zar 1996). All tests of statistical significance were made at the $P = 0.1$ level of probability. We used the increased type I error rate (probability of falsely detecting an effect) of $P = 0.1$, instead of the traditional $P = 0.05$, because of the small-size and the exploratory nature of our study.

RESULTS

A total of 241 FIA plots were present on forest land in the study area surrounding the ONF, which was defined by delineation of the large geologic based LTAs, as shown in Figure 1. However, 63 plots were discarded because site index was missing (i.e. stand was too young for its determination) or it had been determined for a tree species that could not be converted to an equivalent value for loblolly pine, leaving 178 plots potentially available for analysis. Four of the 178 plots had been installed on sites classified as hydric bottomlands, which were discarded because of the low representation of this group of plots in the data set. The remaining 174 plots were located on sites classified as mesic uplands and were available for analysis at the LTA level of the ECS.

Analysis at the LT level of the ECS and below was restricted to the area where those smaller and more detailed ecological units had been delineated, which was only within the boundary of the ONF (Fig 1). Only 18 FIA plots had been installed in the ONF and therefore could be used for analysis of data at the LT level of the ECS and below. Classification groupings of the 18 FIA field plots were identical for analysis at the LT and LTP levels. For example, upland units at the LT and LTP levels differed only by ridge type: broad versus narrow. Because the three FIA plots were all on narrow ridges, the LTP analysis would have been identical to that for the LT; therefore it was omitted. Finally, to obtain sample sizes adequate for analysis ($n \geq 3$) at the LTPm level, the 18 plots were grouped into three broad classes (low, medium, and high) of erosion instead of the seven detailed categories recognized in the ECS.

Forest type of the large study area was predominately pine (59 percent) but it varied considerably among LTAs, from 46 percent in LTA4 to 75 percent in LTA7 (Table 2). The pine type was primarily loblolly (96 percent); the oak-hickory type was classified mostly as white oak/red oak/hickory (30 percent) or mixed upland hardwoods (27 percent). Most of the oak/pine forest type (74 percent) occurred in LTA3 and LTA4, and almost half of LTA5 was classified as oak/hickory type. Among all sample plots site index was highest for four plots associated with the oak/gum and elm/ash forest types. Although those plots had been classified in the FIA data as having a mesic moisture regime, they were likely located on drier parts of very mesic and fertile floodplains.

LANDTYPE ASSOCIATION ECOLOGICAL UNITS

Mean loblolly pine site index for the entire study area averaged 88.1 feet and ranged from an average of 83.6 feet (LTA4) to 103.1 feet (LTA7) (Table 3). Excluding LTA7,

represented by only 4 plots, variation of plot site index was wide for the other LTAs (44 - 71 feet) and particularly for three LTAs (2, 4, and 6) each of which likely (based on forest type) included a plot associated with a floodplain (Table 2). The analysis indicated significant differences ($p < 0.02$) of site index were present among some or all of the seven LTAs. Site index differences were present between two groups of LTAs (Fig. 2). The Tukey Test indicated that average site index did not differ among LTA 2, LTA6, or LTA7, but it was statistically higher for LTA7 than for LTA1, LTA3, LTA4, or LTA5.

LANDTYPE ECOLOGICAL UNITS

Only 18 of the 174 FIA plots were available for the analysis of site index for LTs delineated within the boundary of the ONF, an area of about 115,300 acres. However, an adequate number of plots for analysis ($n \geq 3$) were available for only three LTs: ridges ($n=3$), exposed slopes ($n=5$) and sheltered slopes ($n=10$). Analysis of data for the 18 plots revealed mean site index of exposed slopes (80.7 feet) was lower ($P=0.02$) compared to sheltered slopes (97.1 feet) (Fig. 3). Neither of those ECS units differed in site index compared to the ridge LT, which was intermediate (93.0 feet) between exposed and sheltered slopes.

LANDTYPE PHASE ECOLOGICAL UNITS

As explained previously, the analysis for LTPs would be identical to that for LTs, and therefore is not presented.

LANDTYPE PHASE - MODIFIED ECOLOGICAL UNITS

The LTPm level of the classification grouped LTPs based on the severity of soil erosion. Analysis of data from 13 plots located on the three classes of soil erosion revealed no significant difference of site index of loblolly pine for sheltered slopes with high erosion (98.0 feet) compared with moderate erosion (97.6 feet) (Fig 4). Although average site index was lowest on exposed slope with moderate erosion (88.8 feet), it was not statistically different from that measured on plots located on the two sheltered slope units.

DISCUSSION

The results of our analysis suggest that the land units delineated using the ECS define areas of differing site quality, and perhaps ecological potential, over a range of scales, from large LTAs to small LTPs. The strongest findings of the study occurred at the LT level of the ECS, where we found clear differences in site index between exposed (80.7 feet) and sheltered units (97.1 feet). We could not detect real differences in site index among ecological units at the LTPm level of the ECS, which was a measure

of soil erosion. Because soil erosion clearly affects site quality in the Georgia Piedmont (Harrington 1991) the small number of FIA plots (13) available for our analysis at the LTPm level was likely a contributing factor in our inability to demonstrate a difference in site index.

A recognized limitation of our study was use of site index, not composition of vegetation, as the biological response variable. Composition is generally used to evaluate hypothesized ecological units (Rowe and Sheard 1981). We used site index for several reasons primarily because it was available in the FIA data set and also because it is a vegetative variable that indirectly integrates physical components of ecosystems including long-term climate and soil characteristics (Spurr and Barnes 1973). Harrington (1991) in an extensive study of loblolly pine site index found the species was sensitive to many environmental variables, including those considered important to differentiate ecological units, such as climate, geology, and soil. In comparison with other Piedmont tree species, particularly hardwoods, loblolly pine is less responsive to variation in site quality (Nelson and Beaufait 1956). Our study is perhaps noteworthy because we found no references from other studies where site index of southern pines had been used to test for differences among ecoregion units. In a highly replicated, large-scale study of ponderosa pine (*P. ponderosa*) site quality in Arizona and New Mexico, Mathiasen and others (1987) found site index did not vary among seven habitat types.

Results of our preliminary study suggest the possible need for refinement of the ECS at the LTA level, which is currently based on types of bedrock. Loblolly pine site index varied little among LTAs when compared across the seven groups. Except for the exposed granitic domes and localized areas of gabbro, the mostly buried geology of the Piedmont resembles an extensive shield of gneisses and schists that have weathered differentially to form a coarse mosaic of soils with slightly varying moisture and nutrient characteristics. Unlike LTA7, which is associated with an unusual type of rock, environmental conditions associated with the other six LTAs did not result in identification of ecological units associated with detectable differences of site index for loblolly pine.

CONCLUSIONS

In conclusion, our preliminary evaluation of the ECS developed for the ONF using a small FIA data set demonstrated a promising relationship between ecological units and environmental gradients expressed by site index of loblolly pine. An analysis using a larger data set, with

vegetation as the biological response variable, is needed to clarify and strengthen the ecological relationships at the lower levels of the ECS. Such an analysis will likely indicate the need for revision of classification units at the LTA level. This region of the Georgia Piedmont is particularly challenging for ecological classification due to lack of topographic relief and its long history of intensive past disturbance resulting in variable soil erosion. As Rowe and Sheard (1981) make clear, ecosystem classification is done not only to reduce environmental variation by stratification of land units for management planning, but also gain a better understanding of the underlying interactions among the important physical components that combine to make the ecosystems unique, which was one of the objectives for developing an ECS for the Oconee National Forest.

ACKNOWLEDGMENTS

We thank Brian D. Jackson and Jeff McDonald, Chattahoochee-Oconee National Forests, and Sonja N. Oswalt, Southern Research Station, for their helpful review comments on an earlier version of this paper. Stanley J. Zarnoch, Southern Research Station, provided particularly insightful comments on the statistical analysis.

LITERATURE CITED

- Barnes, B.V.;** Pregitzer, K.S.; Spies, T.A.; Spooner, V.H. 1982. Ecological forest site classification. *Journal of Forestry*. 80:493-498.
- Breder, E.V.;** Davis, L.S. 1959. Influence of topography on the future composition of the lower Georgia Piedmont. *Journal of Forestry*. 57:33-34.
- Burbanck, M.P.;** Platt, R.B. 1964. Granite outcrop communities of the Piedmont plateau in Georgia. *Ecology*. 45:292-306.
- Burbanck, M.P.;** Phillips, D.L. 1983. Evidence of plant succession on granite outcrops of the Georgia Piedmont. *American Midland Naturalist*. 109:94-104.
- Cleland, D.T.** and others 1997. National hierarchical framework of ecological units. In: Boyce, M.S.; Haney, A., eds. *Ecosystem management: applications for sustainable forest and wildlife resources*. New Haven, CT: Yale University Press. 181-200.
- Cleland, D.T.** and others 2007. *Ecological Subregions: Sections and Subsections of the Conterminous United States*. Washington, DC: U.S. Department of Agriculture Forest Service. Scale 1:3,500,000.
- Cowell, C.M.** 1993. Environmental gradients in secondary forests of the Georgia Piedmont, U.S.A. *Journal of Biogeography*. 20:199-207.
- Cowell, C.M.** 1998. Historical change in vegetation and disturbance on the Georgia Piedmont. *American Midland Naturalist*. 140:78-89.
- Fenneman, N.M.** 1931. *Physiography of eastern United States*. New York: McGraw-Hill Book Company. 714 p.
- Golden, M.S.** 1979. Forest vegetation of the lower Alabama Piedmont. *Ecology*. 60:770-782.
- Harper, R.M.** 1930. The natural resources of Georgia. School of Commerce, Bureau of Business. Research Study No. 2 Bulletin University Georgia. 39(3):xi-105.
- Harrington, C.A.** 1987. Site index comparisons for naturally seeded shortleaf and loblolly pines. *Southern Journal of Applied Forestry*. 11:86-91.
- Harrington, C.A.** 1991. PTSITE-A new method of site evaluation for loblolly pine: model development and user's guide. General Technical Report SO-81. New Orleans, LA: U.S. Department of Agriculture Forest Service, Southern Forest Experiment Station. 45 p.
- Hix, D.M.;** Percy, J.N. 1997. Forest ecosystems of the Marietta Unit, Wayne National Forest, Southeastern Ohio: multifactor classification and analysis. *Canadian Journal Forest Research*. 27:1117-1131.
- Jones, S.M.** 1988. Old-growth forests within the Piedmont of South Carolina. *Natural Areas Journal*. 8:31-37.
- Mathiasen, R.L.;** Blake, E.A.; Edminster, C.B. 1987. Estimates of site potential for ponderosa pine based on site index for several southwestern habitat types. *Western North America Naturalist*. 47:467-473.
- Nelson, T.C.** 1957. The original forests of the Georgia Piedmont. *Ecology*. 38:390-397.
- Nelson, T.C.;** Ross, R.D.; Walker, G.D. 1957. The extent of moist sites in the Georgia Piedmont and their forest associations. Research Note 102. Asheville, NC: U.S. Department of Agriculture Forest Service, Southeastern Forest Experiment Station. 2 p.
- Nelson, T.C.;** Beaufait, W.R. 1956. Studies in site evaluation for southern hardwoods. *Proceedings Society American Foresters*. p. 67-70.
- Olson, Jr., D.F.;** Della-Bianca, L. 1959. Site index comparisons for several tree species in the Piedmont of Virginia - Carolina Piedmont. Station Paper 105. Asheville, NC: U.S. Department of Agriculture Forest Service, Southeastern Forest Experiment Station. 9 p.
- Pehl, C.E.;** Brim, R.L. 1985. Forest habitat regions of Georgia, Landsat IV imagery. Special Publication 31. Athens, GA: Georgia Agricultural Experiment Station. Map supplement. Scale 1:1,000,000.
- Rowe, J.S.;** Sheard, J.W. 1981. Ecological land classification: a survey approach. *Environmental Management*. 5:451-464.
- Schmidt, J.M.;** Barnwell, J.A. 2002. A flora of the Rock Hill Blackjacks Heritage Preserve, York County, South Carolina. *Castanea*. 67:247-279.
- Schroeder, P.A.;** Melear, N.D.; West, L.T.; Hamilton, D.A. 2002. Meta-gabbro weathering in the Georgia Piedmont: implications for global silicate weathering rates. *Chemical Geology*. 103:235-245.
- Sharitz, R.R.;** Boring, L.R.; VanLear, D.H.; Pinder, III, J.E. 1992. Integrating ecological concepts with natural resource management of southern forests. *Ecological Applications* 2:226-237.
- Spurr, S.H.;** Barnes, B.V. 1973. *Forest Ecology*. The Ronald Press Company. New York. 571 p.
- Staheli, A.C.** 1976. Topographic expression of superimposed drainage on the Georgia Piedmont. *Geological Society of America Bulletin*. 87:450-452.

Trimble, S.W.; Weirich, F.H.; Hoag, B.L. 1987. Reforestation and the reduction of water yield on the Southern Piedmont since circa 1940. *Water Resources Research* 23:425-437.

Van Kley, J.E. 1993. An ecological classification system for the central hardwood region: The Hoosier National Forest. West Lafayette, IN: Purdue University. 436 p. PhD dissertation.

Wharton, C.H. 1989. The natural environments of Georgia. Bulletin 114. Atlanta, Ga: Department of Natural Resources, Environmental Protection Division, Georgia Geologic Survey. 227 p.

Zar, J.H. 1996. *Biostatistical Analysis*. Upper Saddle River, NJ: Prentice Hall. 662 p.

Table 1—Preliminary non-hierarchical units occurring at the landtype, landtype phase, and landtype phase-modified levels of the ecological classification system for the Oconee National Forest

ECS levels ¹	Description
Landtype	
Piedmont plain ²	Largely undissected land surface of the "original" penneplain or plateau
Broad ridge	Ridges along watershed divides of 5th level hydrologic units
Narrow ridge	Ridges along watershed divides of 6th level and smaller hydrologic units
Exposed slope	Linear part of a slope below the ridge with an aspect from 158°-292°
Sheltered slope	Linear part of a slope below the ridge with an aspect from 293°-157°
Riparian	Concave land surface enclosing large streams and rivers
Glade	Flat area of a ridge associated with gabbro rock formations
Landtype phase	
Broad ridge	Ridges along watershed divides of 5th level hydrologic units
Narrow ridge	Ridges along watershed divides of 6th level and smaller hydrologic units
Exposed slope	Linear part of a slope below the ridge with an aspect from 158°-292°
Sheltered slope	Linear part of a slope below the ridge with an aspect from 293°-157°
Upland flat	Flat area of a ridge associated with gabbro rock formations
Others ³	Concave land surfaces associated with subhydric to hydric riparian sites
Landtype phase-modified⁴	
Slight	Little or no erosion
Moderate	From 25 - 50 percent of surface horizon lost
Severe	Over 50 percent of surface lost, often "gullied"

¹ These units are not hierarchical. They are common to all and may occur in any of the seven landtype associations within the larger Midland Plateau-Central Uplands Subsection.

² Present in a very small area (65 acres) on the ONF; it was combined with broad ridge at the LTP level.

³ Land units associated with wetter parts of the landscape (forested wetland, open wetland, riparian, river floodplain, stream terrace, sand levee, upland flat, and water).

⁴ The seven categories of erosion in the ECS (slight, slight-moderate, moderate-slight, moderate, moderate-severe, severe-moderate, and severe) were grouped into three classes for this study.

Table 2—Distribution of FIA plots by forest type and landtype association in the Midland Plateau-Central Uplands Subsection where site index was determined for loblolly pine on sites classified as upland mesic

Forest type	Landtype association							Total all	Per-cent	Site index
	1	2	3	4	5	6	7			
	-----Number of plots-----									
Pine	15	20	27	18	9	11	3	103	59.2	90.7
Oak/pine	1	2	7	10	-	3	-	23	13.2	84.7
Oak/hickory	5	10	8	10	8	3	-	44	25.3	82.2
Oak/gum	-	1	-	-	-	-	1	2	1.1	97.0
Elm/ash	-	-	-	1	-	1	-	2	1.1	105.5
Total	21	33	42	39	17	18	4	174	100.0	88.1

Table 3—Characteristics of ecological units classified by landtype association from FIA sample plots within the Midland Plateau-Central Uplands Subsection study area of the Oconee National Forest

Item	Landtype association ¹						
	1	2	3	4	5	6	7
Site Index (feet)	85.2	94.3	87.3	83.6	97.9	90.0	103.1
Basal area (feet ² /acre)	124	110	100	109	98	112	129
Elevation (feet)	232	350	306	374	369	251	415
Aspect (degrees)	202	262	197	160	132	87	91
Gradient (percent)	12.0	9.8	8.0	12.0	8.4	10.2	7.0

¹ An 8th geologic group, aluminous schist, occurred in the subsection but was not present within the area delineated by LTAs in the proclamation boundary of the Oconee National Forest.

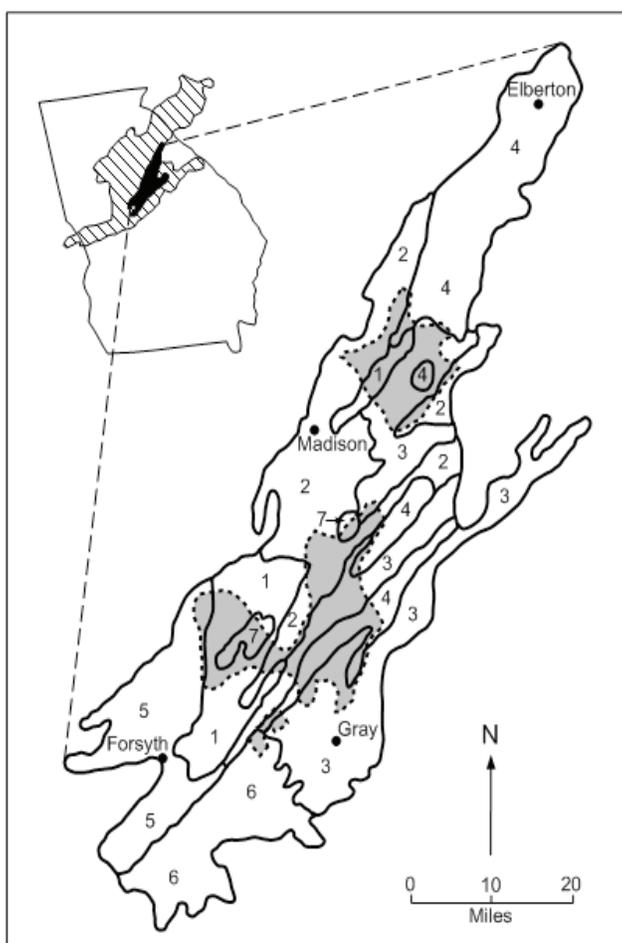


Figure 1—Extent of the Midland Plateau—Central Uplands Subsection (hatched area in small inset map) in Alabama, Georgia, and South Carolina. The study area (black overlay in Georgia) was defined by the closed polygons of seven proposed landtype associations (LTAs) that occur within the proclamation boundary of the Oconee National Forest (three gray areas in the enlarged LTA area). The LTAs (identified by a number in each polygon) represent the predominate geologic bedrock formations: 1, intermediate gneiss; 2, granitic gneiss; 3, mica schist; 4, granite; 5, biotite gneiss; 6, metamorphosed mafic; 7, mafic and ultramafic (gabbro).

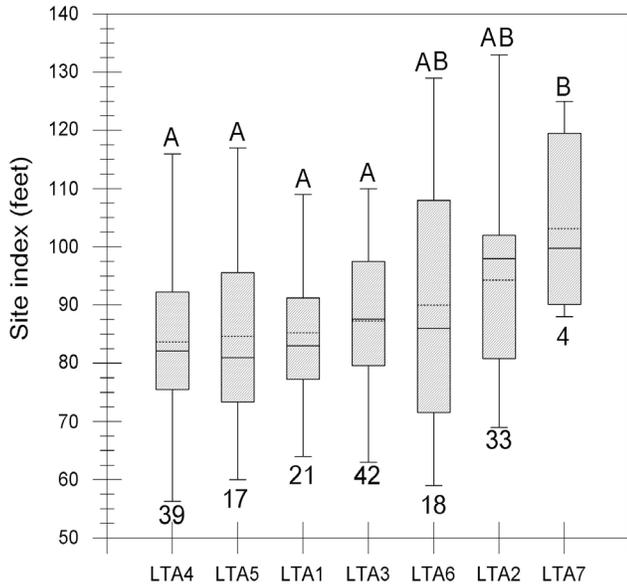


Figure 2—Box plot for loblolly pine site index by landtype association (LTA). The bottom and top of the box represent the 25th and 75th percentiles; respectively; the mean is represented by the horizontal dashed bar and the median by the solid bar in each box. The cross bars below and above each box indicate the range of site index. LTAs with the same letters are not significantly different at the 0.1 level of probability. The number of plots present in each LTA is shown below each box.

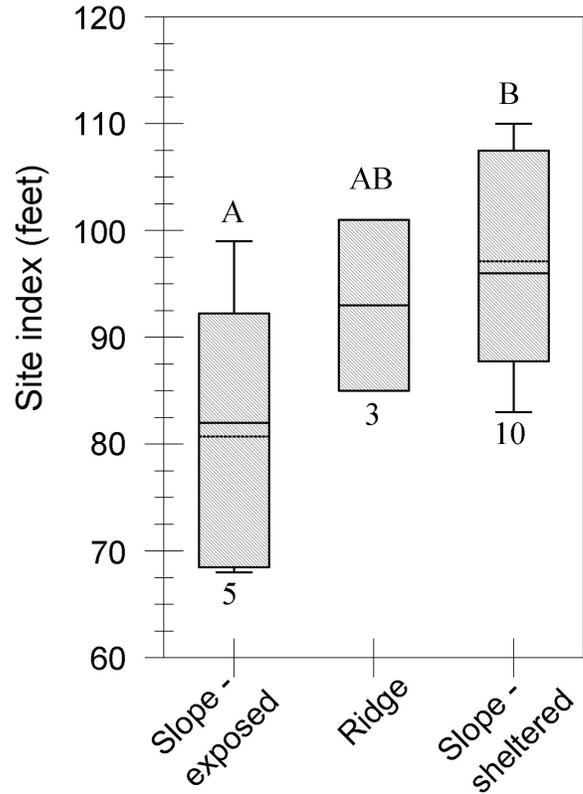


Figure 3—Box plot for loblolly pine site index by landtype and landtype phase levels of the ecological classification system (ECS) for the Oconee National Forest. The bottom and top of the box represent the 25th and 75th percentiles, respectively; the mean is represented by the horizontal dashed bar and the median by the solid bar in each box. The cross bars below and above each box indicate the range of site index. Bars with the same letters are not different at the 0.1 level of probability. Below each bar is the number of plots present in that unit of the ECS.

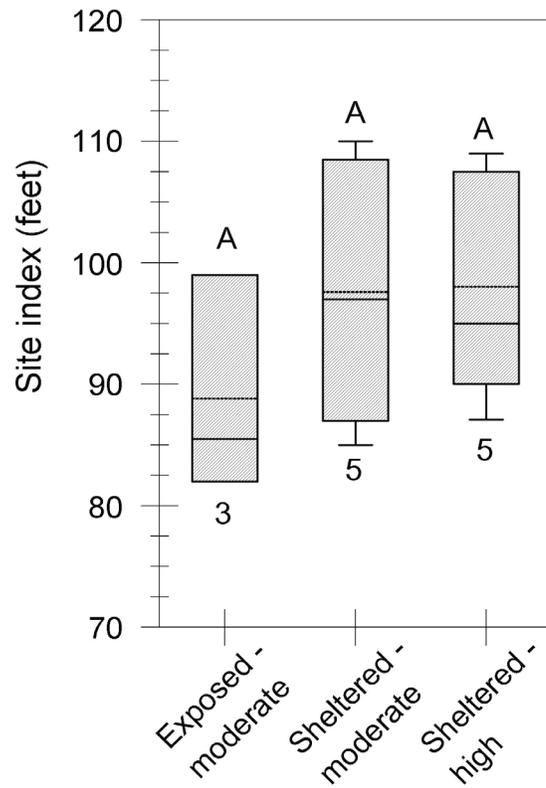


Figure 4—Box plot for loblolly pine site index by landtype phase-modified level of the ecological classification system (ECS) for the Oconee National Forest. The bottom and top of the box represent the 25th and 75th percentiles, respectively; the mean is represented by the horizontal dashed bar and the median by the solid bar in each box. The cross bars below and above each box indicate the range of site index. Bars with the same letters are not different at the 0.1 level of probability. Below each bar is the number of plots present in that unit of the ECS.

CHANGES IN EARLY-SUCCESSIONAL HARDWOOD FOREST AREA IN FOUR BIRD CONSERVATION REGIONS ACROSS FOUR DECADES

Sonja N. Oswalt,¹ Kathleen E. Franzreb,² David A. Buehler³

ABSTRACT

Early successional hardwood forests constitute important breeding habitat for many migratory songbirds. Declines in populations of these species suggest changes in habitat availability either on the species' wintering grounds or on their early successional breeding grounds. We used Forest Inventory and Analysis data from 11 states across four decades to examine changes in early successional (small-diameter) hardwood forests in four Bird Conservation Regions (BCRs) where migratory songbirds of interest have exhibited population declines: Appalachian Mountains, Central Hardwoods, Piedmont, and Southeastern Coastal Plains. We hypothesized that 1) proportional to the amount of timberland on the landscape, hardwood area in the four BCRs of interest has remained stable across the four decades studied and 2) proportional to the total amount of hardwood timberland on the landscape, the area of small-diameter hardwood forest in the four BCRs of interest has declined across the four decades studied. In the Central Hardwood BCR, proportional hardwood area declined slightly ($P=0.0033$), while in the Southeastern Coastal Plain, proportional hardwood area remained stable (0.2705). The Appalachian Mountains and Piedmont experienced increases ($P<0.0001$). Total timberland area and proportional area of early successional forests across the entire sample of interest remained stable from the 1970s through the 1980s, experienced an increase in the 1990s, then declined in the 2000s ($P<0.0001$)—a pattern reflected in the individual BCRs. Implications of our findings are discussed.

INTRODUCTION

Early-successional, or small-diameter, hardwood forests constitute an important habitat component for many wildlife species, including numerous migratory songbird and game animal species. Historically, natural and anthropogenic disturbances like fire, insects and disease, domestic and wild animal grazing, and storms helped to create and maintain early successional habitat in the central hardwoods, Appalachian, and Piedmont regions (Lorimer 2001). Lorimer and White (2003) estimate that in the pre-settlement hardwood forests of the northeast the average proportion of the landscape occupied by early successional habitat was between 1-3 percent, with some coastal pine/oak forests

exhibiting proportions of 10 percent or higher. Following European settlement, land clearing for agriculture, development and commercial timber management replaced fire as primary disturbances in the hardwood forests of Eastern North America, resulting in widespread areas of early-successional habitat reaching proportions of as much as 75 percent of the forested landscape by the late 19th and early 20th centuries (Lorimer and White 2003).

More recently, some studies suggest that forests throughout the central and northern hardwood regions are maturing, resulting in a reduction in the amount of early successional habitat on the landscape (Trani 2001, Brooks 2003, Oswalt and Turner 2009). Lorimer and White (2003) and Brooks (2003) suggest that, for the northeastern United States, the proportion of forest that is early-successional may be nearing pre-settlement levels following the widespread clearing that occurred during settlement and expansion.

Early successional hardwood forests constitute important breeding habitat for many migratory songbirds of concern like the golden-winged warbler (*Vermivora chrysoptera*), prairie warbler (*Dendroica discolor*), chestnut-sided warbler (*D. pensylvanica*), and Bewick's wren (*Thryomanes bewickii*), among others. Changes in the availability of early successional habitat are of interest to wildlife managers and ornithologists who are concerned with declines in disturbance-dependent avian species (Hunter and others 2001, DeGraaf and Yamasaki 2003). Mitchel and others (2001) found that birds associated with early successional habitat respond to changes in habitat availability at a landscape scale, and inferred that the extent of contiguous habitat may be limiting for those populations. Declines in populations of these species suggest changes in habitat availability either on the species' wintering grounds or on their early successional breeding grounds. Regional patterns

¹Forest Resource Analyst, USDA Forest Service Southern Research Station, Forest Inventory and Analysis, 4700 Old Kingston Pike, Knoxville, TN 37919

²Research Wildlife Biologist and Adjunct Associate Professor, USDA Forest Service Southern Research Station, University of Tennessee Department of Forestry Wildlife and Fisheries, 274 Ellington Plant Sciences, Knoxville, TN 37996

³Professor, University of Tennessee Department of Forestry Wildlife and Fisheries, 246 Ellington Plant Sciences, Knoxville, TN 37996

of change in early successional habitat are, therefore, important for understanding the role that declining small-diameter forest area may play in changing populations of breeding songbirds.

We used Forest Inventory and Analysis (FIA) data from 11 states across four decades to examine changes in early successional (small-diameter) hardwood forests in portions of four bird conservation regions where migratory songbirds of interest have exhibited population declines: Appalachian Mountains, Central Hardwoods, Piedmont, and Southeastern Coastal Plains. We hypothesized that 1) proportional to the amount of timberland on the landscape, total hardwood area in the four BCRs of interest has remained stable across the four decades studied and 2) proportional to the total amount of hardwood timberland on the landscape, the area of small-diameter hardwood forest in the four BCRs of interest has declined across the four decades studied.

METHODS

Data from the USDA Forest Service national FIA Database (FIADB) were compiled and analyzed to examine the status and trends of small diameter hardwood forests among four decadal time periods (1970s, 1980s, 1990s, and 2000s) within four Bird Conservation Regions of interest. The sample population was defined by intersecting the outline of Bird Conservation Regions (BCRs) of interest with FIA plot locations in 11 states using ESRI ArcGIS (figure 1). Four BCRs were of interest in this study: Central Hardwoods, Southeastern Coastal Plain, Appalachian Mountains, and Piedmont. FIA plots were located on the map using actual coordinates collected in the field, with the exception of plot locations in Missouri and West Virginia, where FIA “perturbed and swapped” locations were used (see Bechtold and Patterson 2005 for detailed documentation of FIA inventory methods, and LaPointe 2005 for an explanation of fuzzed and swapped coordinates). Not all states were available for all years, and survey years varied among states. States, survey periods, and numbers of plots used in this analysis are given in table 1.

Data were aggregated to the county level for analysis, and counties were used as the sample unit (Fei and Steiner 2007, Oswalt and Turner 2009). The total timberland area in hectares (TTA), total hardwood timberland area (THA), and total small-diameter hardwood timberland area (TSD) were calculated for each Decade-State-BCR-County combination. Sample area and size differed through time; therefore, area estimates were normalized for analysis by converting raw numbers to proportions, yielding the proportion of total timberland area that was hardwood (PTTA), the proportion of total timberland area that was small-diameter hardwood (PTSD), and the proportion of total hardwood timberland

that was small-diameter (PTHA). Concerns that the use of proportions might produce erroneous results with regards to changes in avian habitat if raw TTA and raw TSD both experienced declines but PTSD remained stable were relieved by Smith and others (2009), who showed that in the regions encompassing the BCRs of interest, timberland area has remained stable or increased since the mid-1970s. Hardwood stands were identified as those falling within a pre-selected set of FIA forest-type groups containing primarily hardwood species (table 2). Small-diameter (seedling/sapling) stands were identified using the FIA variable STNDSZCD, which defines small diameter stands as: Stands with an all live stocking value of at least 10 (base 100) on which at least 50 percent of the stocking is trees less than 12.7 cm in diameter (U. S. Forest Service 2009). Analyses of variance were used to determine changes in PTTA, PTSD, and PTHA over time across the whole study area and by BCR. Proportions were arcsin-transformed to improve normality. Means were back-transformed for reporting purposes. Generalized least square means were compared among decades for each ANOVA.

RESULTS

HYPOTHESIS 1

Proportional to the amount of timberland on the landscape, hardwood area in the four BCRs of interest has remained stable across the four decades studied. Hardwood area trends, as a proportion of total timberland, varied by BCR and time. In the Appalachian Mountain BCR, PTTA increased between the 1970s and 1990s, and then increased again between the 1990s and 2000s ($p < 0.0001$; figure 2). The Central Hardwoods experienced a moderate increase in PTTA from 1970 to 1980 (88.0 ± 1.9 and 89.3 ± 0.8), followed by a gradual decrease in 2000 to levels statistically lower than 1980, but comparable to 1970 (86.2 ± 0.9 ; $p = 0.0033$). The Southeastern Coastal Plains BCR PTTA remained stable across all four decades ($p = 0.2705$). The PTTA increased in the Piedmont BCR between the 1980s and 1990s ($p < 0.0001$). Timberland in the Appalachian Mountains and Central Hardwood BCRs was predominately hardwood, and contained the highest proportion of hardwood to softwood timberland in the study (91.2 ± 4.1 and 86.2 ± 0.9 percent in the 2000s, respectively). In comparison, the Piedmont BCR sample area was composed of approximately 60.8 ± 1.4 percent hardwood area, while the Southeastern Coastal Plain BCR was only 39.1 ± 1.1 percent hardwood area.

HYPOTHESIS 2

Proportional to the total amount of hardwood timberland on the landscape, the area of small-diameter hardwood forest in the four BCRs of interest has declined across the four decades studied.

Proportionally, the area of small-diameter hardwood timberland across the entire sample of interest remained stable from the 1970s to the 1980s (27.0 ± 0.7 and 26.8 ± 0.7 percent, respectively), increased in the 1990s to 32.3 ± 0.8 percent, then declined in the 2000s to 21.7 ± 0.6 percent ($p < 0.0001$; figure 3). In the Appalachian Mountains BCR, no differences occurred from the 1970s to the 1980s (18.0 ± 1.3 and 16.0 ± 0.9 percent, respectively), but small-diameter area increased in the 1990s to 19.6 ± 1.4 percent of hardwood timberland before declining precipitously to 11.7 ± 0.9 percent in the 2000s ($p < 0.0001$; figure 4). Small-diameter hardwood area was stable in the Central Hardwoods BCR from the 1970s through the 1990s (23.8 ± 2.2 , 21.5 ± 1.2 , and 21.8 ± 1.8 percents, respectively) but declined to 9.1 ± 0.6 percent of total hardwood timberland area in the 2000s ($p < 0.0001$). In the Piedmont BCR, small-diameter area experienced no significant changes ($p = 0.1329$). Small-diameter area increased in the Southeastern Coastal Plain between the 1970s and 1980s (34.7 ± 1.0 and 38.3 ± 1.1 percent of hardwood timberland area, respectively), reached a peak in the 1990s at 43.7 ± 1.0 percent, then declined back to pre-1990s levels in the 2000s (36.7 ± 1.1 percent, $p < 0.0001$).

DISCUSSION

In contrast to our original hypothesis that the hardwood proportion of timberland area remained stable from the 1970s to the 2000s in the BCRs studied, total hardwood area actually increased in the Appalachian Mountains and Piedmont BCRs and remained stable, overall, in the Central Hardwoods and Southeastern Coastal Plain BCRs. Because of the stability of the total timber resource, and the relative stability of the overall hardwood resource, we were able to focus on the proportion of that resource that was small-diameter habitat.

Declines in early successional stands as a proportion of the overall hardwood resource were most notable in the Central Hardwood and Appalachian Mountain BCRs with 15 and 6 percent declines from the 1970s to the 2000s, respectively. Current proportions of early-successional forest for the Central Hardwoods and Appalachian Mountains appear to be similar to presettlement levels for the upper Midwest and Northeast, but possibly much lower than presettlement levels for the central hardwoods region as reported by Lorimer (2001), though that study used different definitions of early successional forests, different regional boundaries, and included both softwood and hardwood forests, savannas, and prairies. Comparisons with presettlement landscapes are also confounded by overall changes in forest area that occurred with the onslaught of development. Oswalt and Turner (2009) studied the Appalachian Hardwood Region (similar to, but distinct from the Appalachian Mountains

BCR), and also note that total diameter distributions of hardwood trees shifted to larger diameter classes between the 1980s and 2000s (Oswalt and Turner 2009).

In contrast to the Central Hardwoods and Appalachian Mountains BCRs, while we noted proportional declines from the 1990s to the 2000s in the Southeastern Coastal Plain BCR, there was no net change from the 1970s and small-diameter stands still comprised between 34 and 36 percent of total hardwood timberland. The Piedmont and Southeastern Coastal Plain BCRs may experience more natural disturbance from hurricanes and associated fire and storms than the northern and central interior forests, or a larger proportion of timberland in the Piedmont and Southeastern Coastal Plain may be affected by commercial timber harvests, resulting in a larger proportion of small-diameter forests. However, overall hardwood forest area (and, subsequently, small-diameter hardwood forest area) is lowest in both of these predominately pine and mixed oak/pine regions than in the Central Hardwoods and Appalachian Mountains regions.

The loss of early successional hardwood forest habitat on the landscape is suggested as one potential reason for declining migratory songbird populations that typically rely on small-diameter forests for a portion of their lifecycle (Richardson and Brauning 1995, Nolan and others 1999, Gill and others 2001, Klaus and Buehler 2001). In a study examining bird population status in three of these BCRs, we found that most of the scrub-shrub birds as a group were declining significantly (Franzreb and others in press). Thus, it is particularly concerning that we found significant declines in small-diameter forests in the two BCRs that contained the largest proportion of hardwood timberland investigated in this study. However, factors beyond overall area loss may be playing a role in avian species declines. For example, although our study addresses declines in landscape-scale early successional hardwood forest area, it does not address shifts in tree, shrub, or herb species composition since the 1970s. Changes in the dominant vegetation occupying small-diameter stands may affect the structure of breeding habitat and available food sources, which may, in turn, impact populations (Lynch and Whigham 1984). This paper and other papers addressing landscape-level changes in small-diameter forest (e.g. Trani and others 2001) also fail to take into account the distribution of small-diameter forests in relation to the overall forest matrix, and in relation to surrounding land uses. Overall changes in the forest matrix, particularly patch size, may also play an important role in avian population dynamics (Lynch and Whigham 1984).

Early successional forests as defined in this paper may not adequately represent changes in habitat used by disturbance dependent birds on the landscape. For example, this study

does not assess changes in scrub-shrub habitat that would not meet the FIA definition of forestland. Additionally, some species that depend on early successional structure for breeding may be able to make use of relatively small canopy gaps or multi-storied forests that may not be captured within the definition of “small-diameter stand size” utilized in this paper.

The FIA program has undergone many changes since the 1970s, including switching from measuring plots using a variable-radius prism plot design to a fixed-radius annual remeasurement plot design, changing plot remeasurement cycles, fluctuating plot lists, and changes in definitions and estimation methods (Bechtold and Patterson 2005). These changes have accompanied the transition of FIA from a series of regional programs to a nationally consistent program that is comparable from state to state across regional boundary lines. Therefore, some changes noted in the paper may be due in part to changing FIA methodologies, though we anticipate that those influences are minimal.

CONCLUSIONS

Data from FIA suggest that early successional habitat in hardwood forests of the Central Hardwoods, Appalachian Mountains BCRs have declined since the 1970s, despite a stable or increasing hardwood timberland resource, and that Piedmont and Southeastern Coastal Plain BCR hardwood forests have declined since the 1990s, but are similar to areas noted in the 1970s. These declines are concerning with regards to disturbance-dependent migratory songbird populations that have been declining over the last several decades. However, multiple factors may also play a role in avian population declines, and changes in other types of early successional habitat that were not captured in this study, like scrub-shrub habitat, prairies, and small canopy gaps may be affecting populations.

LITERATURE CITED

- Bechtold, William A.**; Patterson, Paul L. Editors. 2005. The enhanced Forest Inventory and Analysis program—national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 85 p.
- Brooks, Robert T.** 2003. Abundance, distribution, trends, and ownership patterns of early-successional forests in the northeastern United States. *Forest Ecology and Management* 185(2003): 65-74.
- DeGraaf, Richard M.**; Yamasaki, Mariko. 2003. Options for managing early-successional forest and shrubland bird habitats in the northeastern United States
- Fei, Songlin**; Steiner, Kim C. 2007. Evidence for increasing red maple abundance in the Eastern United States. *Forest Science* 53(4) 2007:473-477.
- Gill, F. B.**; Canterbury, R.A.; Confer, J.L. 2001. Blue-winged Warbler (*Vermivora pinus*). *The Birds of North America Online* (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology; Retrieved from the Birds of North America Online: <http://bna.birds.cornell.edu.proxy.lib.utk.edu:90/bna/species/584doi:10.2173/bna.584>.
- Hunter, W.C.**; Buehler, D.A.; Canterbury, R.A. [and others]. 2001. Conservation of disturbance-dependent birds in Eastern North America. *Wildlife Society Bulletin* 29(2):440-455.
- Klaus, N. A.**, Buehler, D.A. 2001. Golden-winged Warbler breeding habitat characteristics and nest success in clearcuts in the southern Appalachian mountains. *Wilson Bulletin* 113: 297-301.
- LaPoint, E.** 2005. Access and use of FIA data through FIA spatial data services. In: McRoberts, R.E.; Reams, G.A.; VanDeusen, P.C.; McWilliams, W.H., eds. 2005. Proceedings of the fifth annual forest inventory and analysis symposium; 2003 November 18-20; New Orleans, LA. Gen. Tech. Rep. WO-69. Washington, D.C.: U.S. Department of Agriculture Forest Service. 222 p.
- Lorimer, Craig G.** 2001. Historical and ecological roles of disturbance in Eastern North American Forests: 9,000 years of change. *Wildlife Society Bulletin*, 29(2): 425-439.
- Lorimer, Craig G.**; White, Alan S. 2003. Scale and frequency of natural disturbances in the northeastern US: implications for early successional forest habitats and regional age distributions. *Forest Ecology and Management* 185(2003): 41-64.
- Lynch, James F.**; Whigham, Dennis F. 1984. Effects of forest fragmentation on breeding bird communities in Maryland, USA. *Biological Conservation* 28(1984): 287-324.
- Mitchell, Michael S.**; Lancia, Richard A.; Gerwin, John A. 2001. Using landscape-level data to predict the distribution of birds on a managed forest: effects of scale. *Ecological Applications* 11(6): 1692-1708.
- Nolan Jr., V.**; Ketterson, E.D.; Buerkle, C.A. 1999. Prairie Warbler (*Dendroica discolor*). *The Birds of North America Online* (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology; Retrieved from the Birds of North America Online: <http://bna.birds.cornell.edu.proxy.lib.utk.edu:90/bna/species/455doi:10.2173/bna.455>.
- Oswalt, Christopher M.**; Turner, Jeffery A. 2009. Status of hardwood forest resources in the Appalachian Region including estimates of growth and removals. Resour. Bull. SRS-142. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 16 p.
- Richardson, M.**; Brauning D.W. 1995. Chestnut-sided Warbler (*Dendroica pensylvanica*). *The Birds of North America Online* (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology; Retrieved from the Birds of North America Online: <http://bna.birds.cornell.edu.proxy.lib.utk.edu:90/bna/species/190doi:10.2173/bna.190>
- Smith, Brad W.** tech. coord.; Miles, Patrick D., data coord. Perry, Charles H., map coord.; Pugh, Scott A., Data CD coord. 2009. Forest Resources of the United States, 2007. Gen. Tech. Rep. WO-78. Washington, D.C.: U.S. Department of Agriculture Forest Service, Washington Office. 336 p.
- Trani, Margaret K.**; Brooks, Robert T.; Schmidt, Thomas L. [and others]. 2001. Patterns and trends of early successional forests in the eastern United States. *Wildlife Society Bulletin* 29(2): 413-424.

USDA Forest Service. 2009. The Forest Inventory and Analysis Database: Database Description and Users Manual Version 4.0 for Phase 2. Draft Revision 2. Forest Inventory and Analysis Program: U.S. Department of Agriculture Forest Service. 368 p. Available online at: <http://www.fia.fs.fed.us/tools-data/docs/default.asp> [Date accessed: March 17, 2010].

Table 1—States, years, and number of plots used for each decadal time period

Decade	1970s	1980s	1990s	2000s
State/Year	Alabama 1972	Alabama 1982	Alabama 1990	Alabama 2008
	Arkansas 1978	Arkansas 1988	Arkansas 1995	Arkansas 2007
	Georgia 1972	Georgia 1989	Georgia 1997	Georgia 2008
	Mississippi 1977	Kentucky 1988	Mississippi 1994	Kentucky 2007
	North Carolina 1974	Mississippi 1987	North Carolina 1990	Mississippi 2006
	South Carolina 1978	North Carolina 1984	South Carolina 1993	North Carolina 2007
	Tennessee 1980	South Carolina 1986	Tennessee 1999	South Carolina 2007
	Virginia 1977	Tennessee 1989	Texas 1992	Tennessee 2007
		Virginia 1985	Virginia 1992	Virginia 2008
		Missouri 1989		Missouri 2008
		West Virginia 1989		West Virginia 2006
Total Number of Plots	28,367	39,611	31,596	31,733

Table 2—Forest Inventory and Analysis forest type codes and definitions used for data selection

FIA Forest Type Code	Forest Type	FIA Forest Type Code	Forest Type
400	Oak/Pine group	510	Scarlet oak
401	Eastern white pine/northern red oak/white ash	511	Yellow-poplar
402	Eastern redcedar/ hardwood	512	Black walnut
403	Longleaf pine/oak	513	Black locust
404	Shortleaf pine/oak	514	Southern scrub oak
405	Virginia pine/southern red oak	515	Chestnut oak/black oak/scarlet oak
406	Loblolly pine/hardwood	516	Cherry/white ash/yellow-poplar
407	Slash pine/hardwood	517	Elm/ash/black locust
409	Other pine/hardwood	519	Red maple/oak
500	Oak/hickory group	520	Mixed upland hardwoods
501	Post oak/blackjack oak	800	Maple/beech/birch group
502	Chestnut oak	801	Sugar maple/beech/yellow birch
503	White oak/red oak/hickory	802	Black cherry
504	White oak/red oak/hickory	805	Hard maple/basswood
505	Northern red oak	809	Red maple/upland
506	Yellow-poplar/white oak/northern red oak	905	Pin cherry
507	Sassafras/persimmon	962	Other hardwoods
508	Sweetgum/yellow-poplar	971	Deciduous oak woodland
509	Bur oak	976	Miscellaneous woodland hardwoods

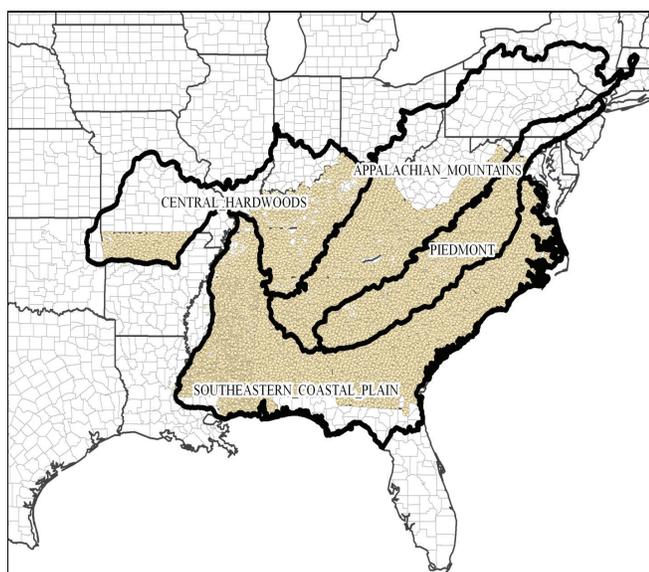


Figure 1—Bird Conservation Regions and plots (approximate locations) used in this study.

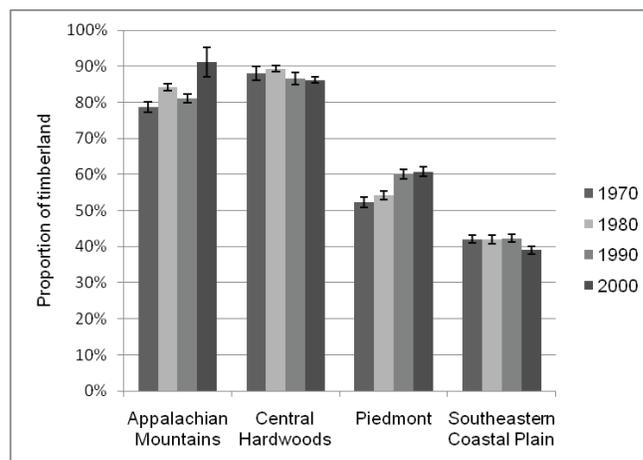


Figure 2—Proportion (± 1 s.e.) of timberland in selected hardwood forest types by BCR and time (all size classes).

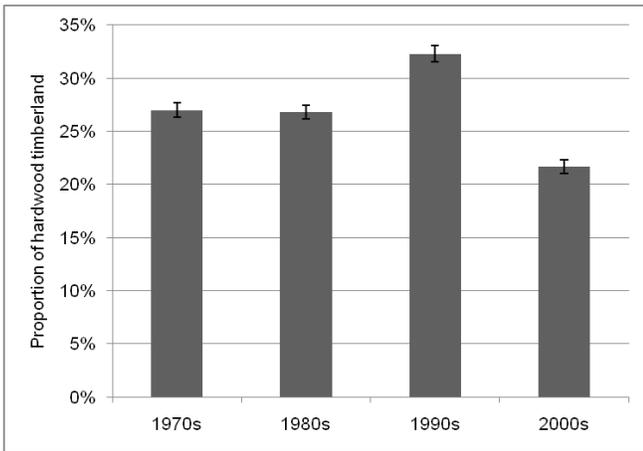


Figure 3—Proportion (± 1 s.e.) of all hardwood timberland that is small-diameter.

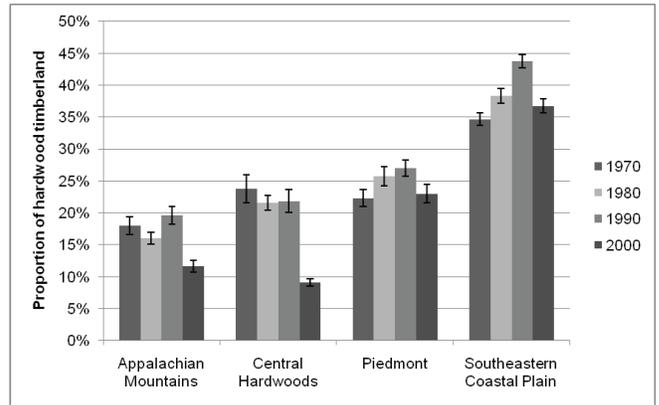


Figure 4—Proportion (± 1 s.e.) of hardwood timberland that is small diameter by BCR and time.

RELATIONSHIPS BETWEEN HARVEST OF AMERICAN GINSENG AND HARDWOOD TIMBER PRODUCTION

Stephen P. Prisley, James Chamberlain, and Michael McGuffin

ABSTRACT

The goal of this research was to quantify the relationship between American ginseng (*Panax quinquefolius*) and timber inventory and harvest. This was done through compilation and analysis of county-level data from public datasets: ginseng harvest data from U.S. Fish and Wildlife Service, US Forest Service (USFS) forest inventory and analysis (FIA) data, and roundwood production data from the USFS Timber Products Output (TPO) program. Data for the 18-state region from 2000 to 2007 were aggregated to the county level. Ginseng harvest was correlated with hardwood growing stock and hardwood forest area. No evidence of a relationship between timber harvest levels and ginseng harvest was observed. There is also no indication that ginseng harvest is higher in areas with more abundant public forestland. For the counties recording a ginseng harvest during the period, ginseng harvest was valued at \$25 million, while timber harvest value was estimated to be \$1 billion.

INTRODUCTION

Herbal medicines and other non-timber forest products have been a significant, yet underappreciated, part of the forest products industry in the United States for more than 300 years (Chamberlain and others 1998). Trade in American ginseng (*Panax quinquefolius*) between North America and China began in the early 1700s. By the mid-1700s, natural populations around Montreal had been depleted, and the plant was discovered in New England (Nash 1898). From the Revolutionary War until 1900, an estimated 20 million pounds of dried ginseng was exported to China from the US (Pritts 1995).

Since 1975, when American ginseng was put on Appendix II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), the United States has been tracking harvest and export of this important medicinal plant (Robbins 2000, U.S. Department of Interior 2009). Biannually, the U.S. Fish and Wildlife Service (FWS) must determine if export of wild-harvested ginseng will be detrimental to the species survival. For each of the years 2000-2010, the FWS determined that lawfully harvested ginseng could be exported from 19 states (Alabama, Arkansas, Georgia, Illinois, Indiana, Iowa, Kentucky, Maryland, Minnesota, Missouri, New York, North Carolina,

Ohio, Pennsylvania, Tennessee, Vermont, Virginia, West Virginia, and Wisconsin) without detriment to the survival of the species. Of the states approved to export American ginseng, most are found in the Appalachian region. The harvest data collected by FWS under CITES provides a county-level dataset of annual estimates that can inform analyses of the spatial distribution of ginseng harvest. Combined with data on forest conditions at the county level, there is an opportunity to examine relationships between ginseng harvest and forest inventory.

The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service (USFS) collects, analyzes and reports on the status and trends of America's forests: how much exists and where it is located, who owns it, and how it is changing, as well as the health and well-being of forest trees and other vegetation. It has been in continuous operation since 1930, with a mission to "make and keep current a comprehensive inventory and analysis of the present and prospective conditions of and requirements for the renewable resources of the forest and rangelands of the US" (Frayser and Furnival, 1999). FIA regularly reports on the status of forests in specific states. FIA also reports, through the Timber Products Output (TPO) program, production of roundwood through mills. Unlike FIA data, which are based on a sampling design involving plots on which trees are measured, the TPO dataset is based on surveys of mills, in which mill managers respond with estimates of production by wood product and source county. Thus, through TPO data we have an additional estimate of timber production at the county level (Johnson and others 2008).

The goal of this study was to improve the understanding of the relationship between American ginseng and hardwood timber harvests. We examine the relationship between standing timber volume, the amount of timber harvested and wild American ginseng harvest. There have been a few studies done to estimate amount of available ginseng habitat (Van Manen and others 2005), but no efforts have been made to quantify the relationship between timber and ginseng harvests.

METHODS

Data at the county level were compiled from two primary sources: ginseng harvest records from the FWS, and Forest Inventory and Analysis (FIA) data from the USFS. Data from each source were compiled from the states in the eastern US where recent (2000 – 2007) ginseng harvest data were available (Figure 1). Several states had missing ginseng harvest data for one or more years (Table 1), and Minnesota had no harvest data at the county level and was therefore omitted from this analysis. In all, data from 1,542 counties were compiled.

GINSENG HARVEST DATA

Ginseng harvest data provided by the FWS were entered manually into database tables. In some cases, dry weights were recorded in pounds and ounces and converted to decimal pounds. Also, for some states, green weights were recorded on data sheets provided by the FWS and were converted to dry weights using a factor of three pounds green weight per pound dry weight, a conversion ratio that is commonly used in the industry. Where the county of origin was not provided on data sheets (some records merely indicated “various” counties), the unassigned harvest numbers were allocated proportionally to counties where harvest was recorded. After entering all harvest data into the database and conducting error-checking for omitted or mis-entered data, average annual harvest across the time period was computed for each county.

FOREST INVENTORY DATA

FIA data are collected in all US states on an annual basis using a multiphase sampling scheme. Due to the transition from periodic to annual inventory, some states had incomplete inventories for the study period (Table 1). In such cases, however, state estimates are still available, but have larger variability than if complete data were available. The sampling intensity used in the FIA program results in estimates are not statistically reliable at the county level. The FIA program, therefore, recommends that totals for groups of counties called FIA units be used. We conducted this analysis at both the county and FIA unit level (Figure 1).

From the FIA data for each state, we compiled estimates of growing stock and removals (by softwood and hardwood), and forest area (by broad forest type and ownership) for each county and FIA unit. We anticipated that ginseng harvest may vary with forest type (hardwood versus softwood forests), and harvesters’ access may vary with land ownership (public versus private), so we summarized inventory and removals by forest type and ownership class for analysis. We included as public lands all federal, state,

and municipal forests except for military bases, in situations where we assumed ginseng harvesting would be restricted.

After compilation of FIA data and computation of relevant estimates, the FIA and ginseng harvest databases were merged by county identifier. This enabled creation of maps showing relevant variables as well as graphical and statistical analysis of relationships between ginseng harvest and forest inventory estimates. Both Pearson’s and Spearman’s correlation coefficients were calculated. These analyses were performed using ArcGIS software and JMP software (SAS Institute 2007).

PRODUCTION AND PRICE DATA

Data on sawtimber and pulpwood production from the states in the region were collected from the FIA Timber Product Output (TPO) dataset (Johnson and others 2008). Annual county figures from 2001 and 2006 were averaged to estimate annual wood product production for the period. Average stumpage price data for wood products were collected from Timber-Mart South (<http://www.tmart-south.com>) for the time period and applicable states in this study. Wood production and wood price data were used to compare economic value of ginseng and wood production for the individual states, averaged over the period 2000 – 2007.

Price data for ginseng is not as accessible or as readily available as for timber. Persons and Davis (2005) provide estimates of prices paid to ginseng harvesters for 1982 through 2005. Persons complemented this data with estimates for 2007 and 2008 through personal communications.

RESULTS AND DISCUSSION

GINSENG HARVEST

During the period of study, almost 500,000 pounds of American ginseng were harvested from the 18 states reported (Table 2). Kentucky accounted for more than 25 percent of the total, followed by Tennessee (13 percent), North Carolina (12 percent), West Virginia (9.5 percent), and Indiana (8.7 percent). These five states accounted for almost 70 percent of the total American ginseng harvest for the period 2000-2007. Maryland reported the lowest harvest of less than 600 pounds. The overall average annual ginseng harvest across the region during the period of study was 60,100 pounds. Annual harvest ranged from a high of 76,644 pounds in 2003 to a low of 42,085 in 2005.

Figure 2 illustrates the spatial distribution of American ginseng harvest. Counties reporting at least 600 pounds of annual harvest are located in eastern Kentucky, southern

West Virginia and western North Carolina. Fourteen states had counties with annual harvests greater than 90 pounds. The greatest majority of counties, though, reported less than 90 pounds of annual harvest. Clearly American ginseng harvest is concentrated in five states.

Across the 1,002 counties that reported some ginseng harvest between 2000 and 2007, the average annual harvest ranged from 0.008 to 1,113.3 pounds. The top 10 percent of producing counties reported at least 166 pounds per year and together accounted for approximately 34,718 pounds per year, or 60 percent of the overall harvest. The top 10 producing counties accounted for nearly 8,615 pounds or 15 percent of the overall average ginseng harvest. Five of these counties are in Kentucky, and four are in North Carolina.

RELATIONSHIPS WITH FOREST INVENTORY

Using correlation analysis we examined the relationship of a suite of variables with ginseng harvest. We did this for two subsets of the data. First, all counties with any reported ginseng harvest were used as a subset of the total dataset (which included 1,542 counties/cities, nearly a third of which had no reported harvest). Because many counties had only minimal harvest, we examined a second subset of only counties with an average annual harvest of at least 50 pounds. The first subset (all producing counties) consisted of 1,002 counties, and the second (producers of at least 50 pounds per year) consisted of 256 counties.

Table 3 summarizes the Pearson correlation coefficients at the county level for the suite of FIA variables examined. The number of counties included in the calculations may be less than the number of counties in a dataset because of missing observations (e.g., counties with no public land, no removals data, etc.).

Analysis of all counties with some harvest (first data subset) provides many statistically significant but low correlations. This dataset contains many counties that had very low harvest but might have large forest areas, growing stock volumes, etc. For example, many of these counties may be along the edges of the expected ginseng range, or may contain only small areas of forest that are conducive to ginseng growth and reproduction. Or, these counties may have a limited numbers of harvesters. Regardless, the strongest relationships were with hardwood growing-stock volume, total forest growing-stock volume, and hardwood forest area.

The analysis of counties producing at least 50 pounds annually (second data subset) presents a slightly different picture (Table 3). These counties, while numbering only a quarter of the total number of counties with any harvest, account for 84 percent of total ginseng harvest. Among these counties, we might expect to find more meaningful

relationships with forest inventory variables. Again, the strongest and most significant correlations are with growing stock volume and forest area. This is not surprising as it indicates more ginseng harvest in counties with more hardwood forest, and with more or larger hardwood trees.

Growing stock volume per acre is simply the total growing stock divided by number of forest acres, and represents relative timber density. This variable shows the one of the highest correlations among the variables in the second dataset. Figure 3 depicts the relationship between ginseng harvest and hardwood growing stock volume. This relationship had the strongest correlation for the 256 counties producing at least 50 pounds/year. However, there is tremendous variability, with some heavily forested counties (growing stock in excess of 800 million cubic feet) producing less than 200 pounds of ginseng annually, while some counties with much less forest volume (300 to 800 million cubic feet) are producing amounts of ginseng in excess of 600 pounds per year.

We also found, in the counties producing at least 50 pounds, positive (but non-significant) correlations with timber removals. This could be because counties with more timber removals also have more growing stock, which is positively correlated to ginseng harvest. Dividing timber removals by growing stock, therefore, gives us a variable that measures intensity of removals relative to standing inventory. For these, the correlations were negative, very low and not statistically significant, meaning the observed relationship could be based on chance alone.

As noted, FIA data are sparse within individual counties such that county-level estimates are not considered reliable as they have high variability. For some analyses, relevant patterns are clearer when data are aggregated to the FIA unit level. To test this effect, we examined correlation coefficients for total ginseng harvest within FIA unit aggregates (Table 4). At the FIA unit level, we see stronger and more significant correlations, due to the removal of county-to-county variability. Hardwood growing-stock volume and hardwood forest area are again significantly correlated with ginseng harvest. Correlations related to public land ownership are weaker or insignificant. The correlation of timber harvest (removals) to ginseng harvest is significant and positive, but lower than the correlations with growing-stock volume. Part of this effect could be due to the very strong and positive way in which removals are themselves correlated with growing-stock volume (0.72 correlation coefficient between hardwood growing stock and hardwood removals).

We found a negative but insignificant correlation with percent hardwood growing stock on public lands. The negative correlation (if significant) would suggest that

counties with a higher proportion of their hardwood forests under public ownership harvest less American ginseng than counties with less hardwood forest on public land. In fact, the FIA unit with peak ginseng harvest per hardwood forest area had only 8.5 percent of hardwood forest in public ownership, ranking 61st out of 76 FIA units. If public lands were a consistent, primary source of ginseng harvest, we would expect these correlations to be larger, positive, and significant.

At the aggregate level of FIA units, we looked at the harvest level of ginseng relative to hardwood forest area to get an indicator of production per unit area. It is impossible to extrapolate from this how much area might support ginseng harvest, because hardwood forest area alone does not account for all the factors relevant to ginseng growth, reproduction, survival, and harvest. But it is evident that ginseng harvest per hardwood acre varies widely, with the highest reported level being 2,615 pounds of ginseng produced per million acres of hardwood forest, in Eastern Kentucky (Figure 4). This eight-county area produced 27,375 pounds of ginseng in the six years for which we had data. An annual harvest of 4,562 pounds was derived from a hardwood forest area of 1.74 million acres. The top ten FIA units each produced over 1,000 pounds of ginseng per million acres of hardwood forest.

PRODUCTION AND ECONOMIC VALUE

While ginseng prices ranged from \$200/pound to an abnormal peak of \$1,150/pound (Persons and Davis 2005), we used a nominal average price of \$423.42/pound to obtain estimates of annual harvest value. Ginseng prices reflect the amount paid to harvesters for dried wild-harvested root.

For timber stumpage, we used averages of prices from southern states reported during the period: \$212 per thousand cubic feet for hardwood pulpwood and \$736.16 per thousand cubic feet for hardwood sawtimber. These prices may not reflect the entire study region, but are indicative of the active southern timber markets. Prices were for stumpage, the price paid to a landowner for standing timber before harvesting and transportation to a mill.

Timber product output data indicate that during the period 2000 – 2007, hardwood timber production in the 1,002 ginseng-producing counties averaged approximately 2.1 billion cubic feet per year, consisting of 0.982 billion cubic feet of pulpwood and 1.153 billion cubic feet of sawtimber.

While the average prices used may not reflect the variability over time and geographic region, they indicate the relative magnitude of the economic value of the timber and ginseng

crops. Annual hardwood timber revenue in the ginseng-producing counties was slightly more than \$1.0 billion, compared to approximately \$25 million for ginseng (Table 5). These numbers actually understate the difference in value, as the timber prices used are for stumpage (standing timber in the forest), and the ginseng prices are for dried ginseng delivered to a dealer.

CONCLUSIONS

Ginseng harvest in an area (county or FIA unit) is related to the amount of hardwood forests in the area, as well as other factors. There was a correlation between ginseng harvest and total hardwood forest area as well as hardwood growing stock. Also, there was a positive but statistically insignificant correlation between ginseng harvest and harvest of timber. Our findings suggest a slight negative relationship between ginseng harvest and amount of public lands. We also conclude that while the value of ginseng harvest may be significant to rural counties it is minor compared to hardwood timber values.

Further analysis of the relationship between ginseng harvest and forest conditions (including timber harvest) is possible. It is also reasonable to consider combining the spatial database of ginseng harvest (Figure 2) with other spatially-defined data that might help explain ginseng abundance. For example, temperature, precipitation, elevation, soil conditions, and other environmental parameters may be associated with ginseng distribution and abundance, and could be modeled with harvest data. Such analyses might provide further insights about factors explaining varying levels of ginseng harvest, and enhance the sustainable utilization of this valuable resource.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the American Herbal Products Foundation for supporting this research. We also wish to thank Huei-Jin Wang and Tiffany Wilson for assistance in data compilation.

LITERATURE CITED

- Chamberlain, J.;** Bush, R.; Hammett, A.L. 1998. Non-timber forest products: the other forest products. *Forest Products Journal*. 48(10):10-19.
- Frayser, W.E.;** Furnival, G.E. 1999. Forest survey sampling designs: a history. *Journal of Forestry* 97(12): 4-10.

Johnson, T.G.; Bentley, J.W.; Howell, M. 2008. The South’s timber industry—an assessment of timber product output and use, 2005. Resour. Bull. SRS-135. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 52 p.

Nash, G.V. 1898. American Ginseng: Its Commercial History, Protection and Cultivation. U.S. Department of Agriculture. Washington, D.C. Bulletin #16. 32p.

Persons, W.S.; Davis, J.M. 2005. Growing and Marketing Ginseng, Goldenseal and Other Woodland Medicinals. Bright Mountain Books, Inc. Fairview, NC. 466p.

Pritts, K.D. 1995. Ginseng: How to Find, Grow, and Use America’s Forest Gold. Stackpole Books. Mechanicsburg, PA. 148p.

Robbins, C.S. 2000. Comparative analysis of management regimes and medicinal plant trade monitoring mechanisms for American ginseng and goldenseal. Conservation Biology. 14(5): 1422-1434

SAS Institute. 2007. JMP Users Guide. Cary, NC: SAS Institute, Inc.

U.S. Department of the Interior. 2009. General advice for the export of wild and wild-simulated American ginseng (*Panax quinquefolius*) harvested in 2009 and 2010 from States with approved CITES Export Programs. 45p.

Van Manen, F.T.; Young, J.A.; Thatcher, C.A. [and others]. 2005. Habitat models to assist plant protection efforts in Shenandoah National Park, Virginia, USA. Natural Areas Journal 25(4): 339-350.

Table 1—Ginseng harvest and FIA data availability. Gray cells represent years in which county-level ginseng harvest data are missing. The FIA data years indicate the time period for the FIA data used for each State

State	2000	2001	2002	2003	2004	2005	2006	2007	Counties	FIA Data Years
Alabama									67	2001 - 2005
Arkansas									75	2000 - 2005
Georgia									159	1998 - 2004
Illinois									102	2002 - 2006
Indiana									92	2002 - 2006
Iowa									99	2002 - 2006
Kentucky									120	2000 - 2004
Maryland									24	2004 - 2006
Missouri									115	2002 - 2006
New York									62	2002 - 2006
North Carolina									100	2003 - 2006
Ohio									88	2001 - 2006
Pennsylvania									67	2002 - 2006
Tennessee									95	2000 - 2004
Vermont									14	2003 - 2006
Virginia									136	2002 - 2007
West Virginia									55	2004 - 2006
Wisconsin									72	2002 - 2006
Count	15	15	16	15	17	18	15	17	1542	

Table 2—Ginseng harvest by State and year (pounds dry weight). Where county-level data were unavailable, Statewide summary data were used and are shown in italics

State	2000	2001	2002	2003	2004	2005	2006	2007	Total
Alabama	256	<i>874</i>	457	1,011	649	221	761	340	4,569
Arkansas	519	927	2,075	2,633	1,717	496	863	990	10,220
Georgia	311	707	266	416	243	161	<i>167</i>	259	2,530
Illinois	2,781	2,884	1,748	2,844	2,682	1,234	2,000	2,082	18,255
Indiana	6,273	6,818	3,192	6,915	4,823	4,926	5,106	3,862	41,915
Iowa	940	783	798	554	286	230	609	1,014	5,215
Kentucky	<i>16,216</i>	<i>22,765</i>	12,149	22,572	16,672	9,393	13,713	11,332	124,813
Maryland	48	56	72	<i>109</i>	<i>160</i>	31	<i>62</i>	53	590
Missouri	1,477	1,703	1,907	2,452	1,358	2,093	1,722	1,097	13,809
New York	1,398	621	485	633	359	309	133	439	4,376
North Carolina	8,417	6,788	8,790	<i>6,548</i>	4,265	5,733	6,447	12,317	59,305
Ohio	3,492	3,254	3,135	4,559	3,978	3,311	2,265	3,126	27,120
Pennsylvania	<i>1,749</i>	1,370	<i>1,730</i>	920	1,025	930	1,355	<i>1,947</i>	11,025
Tennessee	<i>8,164</i>	<i>8,737</i>	<i>5,815</i>	<i>10,826</i>	8,204	5,034	<i>8,153</i>	8,730	63,663
Vermont	205	119	183	117	112	36	60	114	946
Virginia	5,723	3,821	3,810	4,675	3,435	1,569	2,798	3,050	28,881
West Virginia	8,602	5,409	5,206	7,170	5,882	4,785	4,561	4,150	45,765
Wisconsin	3,024	2,495	2,580	1,690	1,946	1,593	2,146	2,396	17,869
Totals	69,596	70,131	54,399	76,644	57,795	42,085	52,919	57,299	480,868

Table 3—Pearson correlation coefficients relating FIA-derived variables to average annual ginseng harvest at the county level. An asterisk indicates the correlations are statistically significant (at the 95 percent confidence level)

Variable	Counties with some production		Counties producing at least 50 pounds annually	
	Correlation Coefficient	Number of Counties	Correlation Coefficient	Number of Counties
Forest area	0.1629 *	1002	0.1584 *	256
Hardwood forest area	0.2177 *	1002	0.1844 *	256
Forest growing-stock volume	0.2340 *	1000	0.1978 *	256
Hardwood growing-stock volume	0.2884 *	1000	0.2189 *	256
Average growing stock per acre	0.2069 *	1000	0.2143 *	256
Hardwood growing stock on public lands	0.0822 *	678	-0.0297	196
Percent hardwood growing stock on public lands	-0.0418	678	-0.0886	196
Removals from all species	0.1175 *	783	0.1159	189
Removals of hardwood species	0.1746 *	782	0.1378	189
Removals as a percent of growing stock	-0.0163	781	-0.0314	189
Hardwood removals as percent of growing stock	0.0076	780	-0.0173	189

Table 4— Pearson correlation coefficients relating FIA-derived variables to average annual ginseng harvest at the FIA Unit level. An asterisk indicates the correlations are statistically significant (at the 95 percent confidence level)

Variable	Correlation Coefficient	Number of FIA Units
Forest area	0.3400 *	76
Hardwood forest area	0.3835 *	76
Forest growing-stock volume	0.4565 *	76
Hardwood growing-stock volume	0.4853 *	76
Average growing stock per acre	0.2897 *	76
Hardwood growing stock on public lands	0.2948 *	76
Percent of hardwood growing stock on public lands	-0.0620	58
Removals from all species	0.3389 *	58
Removals of hardwood species	0.4275 *	58
Removals as a percent of growing stock	-0.1121	58
Hardwood removals as a percent of growing stock	-0.0620	58

Table 5— Annual revenue from ginseng and hardwood timber harvest by State for 1,002 counties with recorded ginseng harvest during 2000-2007

State	Average Annual Ginseng Harvest (pounds)	Ginseng Revenue* (thousand \$)	Timber Revenue (thousand \$)
Alabama	571.1	\$242	\$32,996
Arkansas	1,277.5	\$541	\$27,464
Georgia	316.3	\$134	\$8,121
Illinois	2,281.9	\$966	\$35,404
Indiana	5,239.4	\$2,218	\$59,256
Iowa	651.9	\$276	\$11,315
Kentucky	15,601.6	\$6,606	\$118,108
Maryland	73.8	\$31	\$6,154
Missouri	1,726.1	\$731	\$78,073
New York	547.0	\$232	\$60,409
North Carolina	7,413.1	\$3,139	\$53,092
Ohio	3,390.0	\$1,435	\$38,006
Pennsylvania	1,378.1	\$584	\$109,602
Tennessee	7,957.9	\$3,370	\$127,923
Vermont	118.3	\$50	\$17,961
Virginia	3,610.1	\$1,529	\$78,640
West Virginia	5,720.6	\$2,422	\$93,249
Wisconsin	2,233.6	\$946	\$101,554
Total	60,108.3	\$25,451	\$1,057,327

* Based on \$423.42 per pound

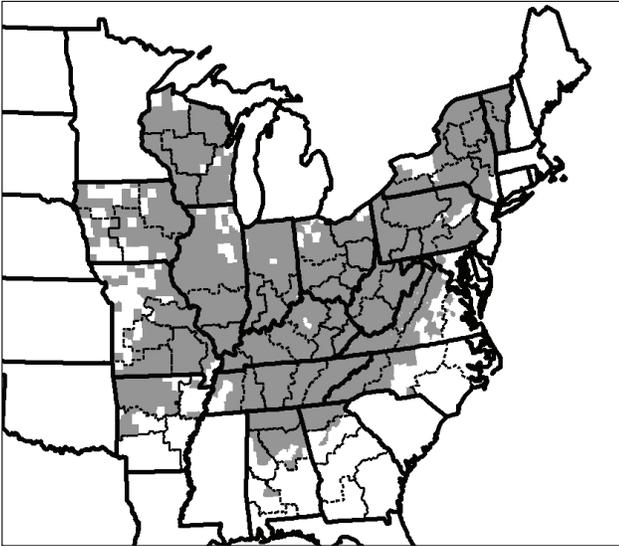


Figure 1—States for which county-level ginseng harvest data were available included Alabama, Arkansas, Georgia, Illinois, Indiana, Iowa, Kentucky, Maryland, Missouri, New York, North Carolina, Ohio, Pennsylvania, Tennessee, Vermont, Virginia, West Virginia, and Wisconsin. Counties shown in gray had at least one record of ginseng harvest during 2000-2007. Dashed lines within States indicate boundaries of FIA aggregation units.

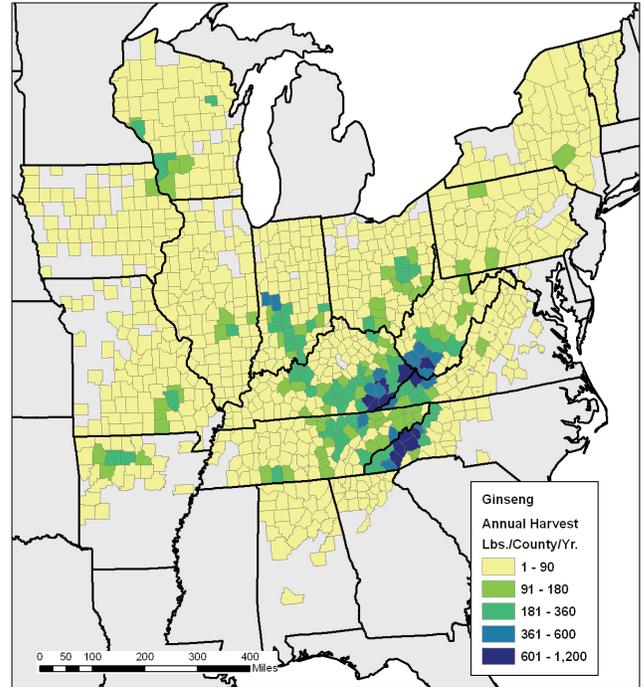


Figure 2—Map of average ginseng harvest.

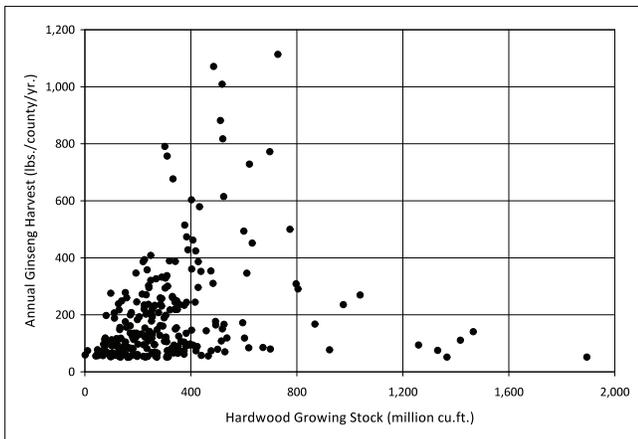


Figure 3—Scatterplot of hardwood growing stock volume and annual ginseng harvest for the 256 counties producing at least 50 pounds of ginseng per year.

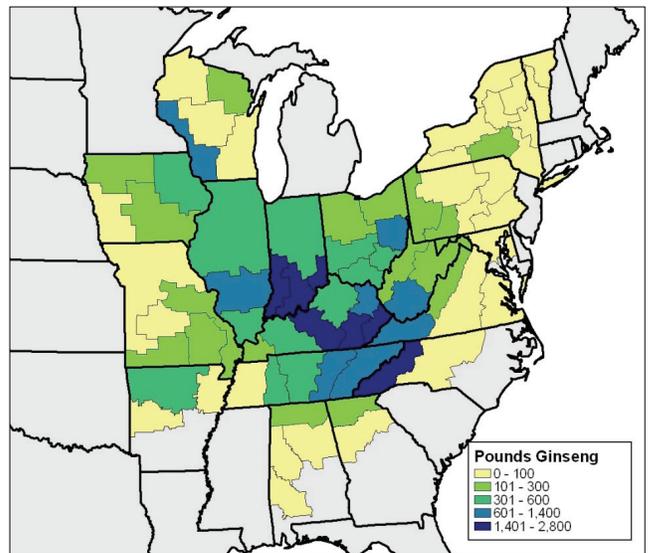


Figure 4—Ginseng harvest per million acres of hardwood forest, by FIA unit.

Forest Health

QUANTIFYING CHANGE IN RIPARIAN ASH FORESTS FOLLOWING THE INTRODUCTION OF EAB IN MICHIGAN AND INDIANA Susan J. Crocker and Dacia M. Meneguzzo.....	105
USING INVENTORY DATA TO DETERMINE THE IMPACT OF DROUGHT ON TREE MORTALITY Greg C. Liknes, Christopher W. Woodall, and Charles. H. Perry.....	109
POTENTIAL IMPACTS OF YEAR-ROUND SAMPLING ON MONITORING PRESENCE-ABSENCE OF INVASIVE FLORA IN THE SOUTHERN UNITED STATES Christopher M. Oswalt, Sonja N. Oswalt, W. Keith Moser.....	113
WHITE ASH (<i>Fraxinus americana</i>) HEALTH IN THE ALLEGHENY PLATEAU REGION, PENNSYLVANIA: EVALUATING THE RELATIONSHIP BETWEEN FIA PHASE 3 CROWN VARIABLES AND A CATEGORICAL RATING SYSTEM Alejandro A. Royo, Kathleen S. Knight, Jamie M. Himes, and Ashley N. Will.....	123
ESTIMATION OF INVASIVE PROBABILITY OF MULTIFLORA ROSA IN THE UPPER MIDWEST Weiming Yu, Zhaofei Fan, W. K. Moser, M. H. Hansen and M. D. Nelson.....	131

QUANTIFYING CHANGE IN RIPARIAN ASH FORESTS FOLLOWING THE INTRODUCTION OF EAB IN MICHIGAN AND INDIANA

Susan J. Crocker and Dacia M. Meneguzzo

ABSTRACT

The emerald ash borer (*Agrilus planipennis* Fairmaire; Coleoptera: Buprestidae; EAB) is an introduced beetle that kills ash (*Fraxinus* spp.) trees. While most EAB-related ash mortality has been documented in urban areas, the effects of EAB in forested settings, particularly in riparian forests, are not well known. This study utilizes forest inventory data to quantify changes in the composition and structure of riparian ash forests since the introduction of EAB to Michigan and Indiana. Estimates of the abundance, number of standing dead trees, mortality and regeneration of riparian ash were compared over time. The abundance of ash growing-stock significantly decreased across the study area between 2003/2004 and 2009. Mortality of riparian ash sharply increased in 2005. The preponderance of ash mortality was limited to riparian forests in the southeastern portion of Michigan.

INTRODUCTION

In recent years, the sustainability of the Nation's ash resource has been threatened by an exotic wood-boring beetle. Native to Asia, the emerald ash borer (*Agrilus planipennis* Fairmaire; Coleoptera: Buprestidae; EAB) was first detected in North America near Detroit, Michigan, in 2002 (Haack et al. 2002). Surveys conducted in the surrounding area soon revealed dead and dying ash (*Fraxinus* spp.) trees throughout southeastern Michigan. EAB was subsequently found in Indiana in 2004. Dendrochronological reconstruction by Siegert et al. (2009) has suggested establishment of EAB and initial mortality of ash originated in the Westland-Garden City area of Michigan around 1997-1998. Since tree mortality generally occurs 3 to 4 years after infestation, it could be concluded that EAB was introduced to southeastern Michigan during the early to mid-1990s, (Siegert et al. 2009) nearly 20 years ago.

While EAB poses a risk to ash in both urban and forested ecosystems, it represents a unique threat to riparian forests. Riparian forests tend to make up a small percentage of forested land area but they often contain a large proportion of ash. Data from the Forest Inventory and Analysis (FIA) Program of the USDA Forest Service show that riparian forests comprise 4.1 million acres, or 21

percent, of Michigan timberland, yet they contain nearly half (48 percent) of all ash trees in the State. Due to the predominance of ash in these areas, the composition and structure of riparian forests could be greatly altered by the activity of EAB.

The purpose of this investigation was to quantify the effects of EAB introduction and spread on the ash resource in riparian forests. To accomplish this goal, FIA data collected between 2003 and 2009 were analyzed to compare ash abundance, distribution of standing dead trees, mortality, and regeneration over time. Results provide an indication of how EAB presence has influenced riparian forest systems as well as insight into types of changes that may occur in other regions with EAB infestations.

METHODS

Annual inventory data from FIA, collected between 2003 and 2009, were used to analyze change in riparian ash composition in Michigan and Indiana. FIA began to collect data on an annual basis in 1999; under the annual inventory system, one-fifth of all plots (or one panel) in the State is measured each year. Once all five panels have been measured, each panel of plots will be remeasured on a 5-year cycle. For example, in Michigan, field plots measured in 2000 were remeasured in 2005. Subsequently, inventories are available for each year following the completion of the first annual inventory, using a 5-year rolling average. The first annual inventory was measured between 1999-2003 in Indiana and 2000-2004 in Michigan. For the sake of brevity, inventory periods are referred to using the last year of data collection.

Under the annual FIA plot design, all trees greater than 5 inches in diameter at breast height (d.b.h) are measured on four 24-foot radius subplots and saplings (d.b.h. between 1 and 4.9 inches) are measured on four 6.8-foot radius microplots (Bechtold and Patterson 2005). Tree variables and site attributes, including species, diameter, and physiographic class, are recorded on all subplots with

Susan J. Crocker, Research Forester, U.S. Department of Agriculture, Forest Service, Northern Research Station, 1992 Folwell Ave, St. Paul, MN 55108

Dacia M. Meneguzzo, Research Forester

a forested condition (for more information, see Bechtold and Patterson 2005). The presence and species of seedlings (d.b.h. less than 1 inch) are counted on the microplot, but detailed, individual measurements are not recorded.

Riparian ash forests were defined using only plots where ash was present and the physiographic class code was one of the following: narrow flood plains/bottomlands; broad floodplains/bottomlands; other mesic; swamps/bogs; small drains; bays and wet pocosins; beaver ponds; and other hydric. Reported estimates of abundance, number of standing dead trees, and mortality were limited to comparisons of growing-stock trees (trees 5 inches d.b.h. or larger) on timberland.

RESULTS

ASH ABUNDANCE

The number of ash trees in riparian forests significantly decreased in both Indiana and Michigan over the course of the study period. Between 2003 and 2009, the abundance of riparian ash trees in Indiana decreased by more than half, from 13 million to 6 million trees. While ash numbers declined across most of Indiana, change was concentrated in the northeastern and southeastern portions of the State. Michigan saw a 14 percent reduction in ash abundance over time, falling from 474 million trees in 2004 to 406 million in 2009. The sharpest declines in ash numbers occurred in the Lower Peninsula, particularly in counties surrounding Detroit.

STANDING DEAD TREES

In 2004, an estimated 5.6 million standing dead ash trees were recorded in riparian forests throughout Michigan. The majority of standing dead ash was located in the northern Lower Peninsula. Fewer, but a proportionally similar number of standing dead ash were recorded in Indiana in 2003—an estimated 1.2 million trees. Northeastern Indiana and the Indianapolis area had the highest numbers of standing dead ash. In 2009, the estimated number of standing dead riparian ash in Michigan increased to 6.3 million trees. In contrast, Indiana saw large decreases in the total number of standing dead riparian ash throughout the State.

ASH MORTALITY

Riparian ash mortality in Michigan was an estimated 3.3 million cubic feet per year in 2004, equal to 50 percent of total ash mortality (Fig. 1). By 2005, mortality sharply increased, more than doubling to nearly 8 million cubic feet per year; mortality remained high through 2009. In 2004, ash mortality was fairly evenly distributed throughout the state. However, by 2009, the majority of riparian ash mortality (66 percent) was located in the southern Lower Peninsula, predominately in the original six-county

quarantine area of the Detroit metro area (Oakland, Macomb, Washtenaw, Wayne and Monroe counties) (Haack et al. 2002).

A similar trend in riparian ash mortality was also seen in Indiana, where mortality began to increase in 2005 and had more than doubled by 2009 (Fig. 2). Increases in mortality were not evenly distributed across the State. For example, 100 percent of Indiana's riparian ash mortality in 2003 occurred in the southern half of the State; by 2009, 64 percent of riparian ash mortality was reported in the northern half of Indiana.

RIPARIAN REGENERATION

In terms of species composition, seedling regeneration in riparian forests in Michigan remained fairly similar between 2004 and 2009. In Indiana, however, changes between 2003 and 2009 included a decrease in black ash and green ash seedlings and an increase in boxelder, silver maple, sugar maple and white ash seedlings.

DISCUSSION

The orientation and nature of riparian forests makes them especially susceptible to insect invasion since they are small in area yet contain a large percentage of the ash resource (Crocker et al. 2009). Therefore, the aim of this investigation was to use FIA data to quantify the occurrence of change in riparian forests following the introduction of EAB.

Michigan, where EAB has been active the longest, showed the greatest transformation over time. The pattern of change appears to reflect the history of EAB introduction and its subsequent spread from the Detroit metropolitan area. The preponderance of standing dead ash trees in the northern half of the Lower Peninsula are indicative of later infestations of EAB and their distribution on the landscape. Detection of areas with increases in standing dead ash trees may be a mechanism for identifying future or currently undetected EAB infestations. Minimally, the location of these dead trees may highlight areas in which to allocate additional survey resources.

An increase in ash mortality in northern Indiana provides similar evidence of the pattern of spread. Trees killed around 2003 and 2004 began to be detected widely over the landscape in 2009. A continued look at riparian ash mortality over time will likely show more mortality in southern Indiana as EAB spreads further south. Seedling regeneration data from Indiana provides evidence of changes in future species composition within the State's riparian areas. Decreasing numbers of ash seedlings were accompanied by an increased number of maple seedlings.

Future work will include incorporating geospatial data to help refine estimates of riparian areas so that we can generate riparian estimates using FIA's periodic inventory, i.e., data collected prior to 1998. In doing so, we will be better able to create a picture of the riparian landscape prior to the introduction of EAB and construct a larger picture of landscape-level changes that may result from this insect.

ACKNOWLEDGMENTS

The authors thank Greg Liknes and Kathleen Knight of the Forest Service, Northern Research Station, for reviewing this manuscript.

LITERATURE CITED

Bechtold, W.A.; Patterson, P.L., eds. 2005. The enhanced Forest Inventory and Analysis Program: national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p.

Crocker, S.J.; McCullough, D.G.; Siegert, N.W. 2009. Predicting the ability to produce emerald ash borer: A comparison of riparian and upland ash forests in Southern Lower Michigan. In: McRoberts, R.E.; Reams, G.A.; Van Deusen P.C.; McWilliams, W.H., eds. 2009. Proceedings of the eighth annual forest inventory and analysis symposium; 2006 October 16-19; Monterey, CA. Gen. Tech. Report WO-79. Washington, DC: U.S. Department of Agriculture, Forest Service. 408 p.

Haack, R.A.; Jendek, E.; Liu, H. [and others]. 2002. The emerald ash borer: a new exotic pest in North America. Newsletter of the Michigan Entomological Society. 47: 1-5.

Siegert, N.A.; McCullough, D.G.; Leibold, A.M.; Telewski, F.W. 2009. Reconstruction of the establishment and spread of emerald ash borer through dendrochronological analysis. In: McManus, Katherine A.; Gottschalk, Kurt W., eds. Proceedings, 19th U.S. Department of Agriculture interagency research forum on invasive species 2008; 2008 January 8-11; Annapolis, MD. Gen. Tech. Rep. NRS-P-36. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station: 70.

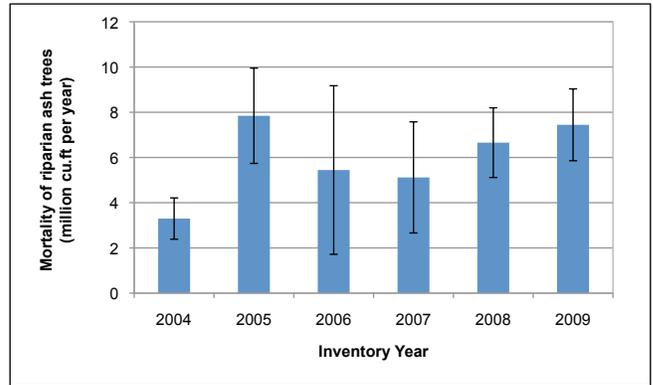


Figure 1—Mortality of growing-stock ash trees on riparian timberland, Michigan, 2004-2009.

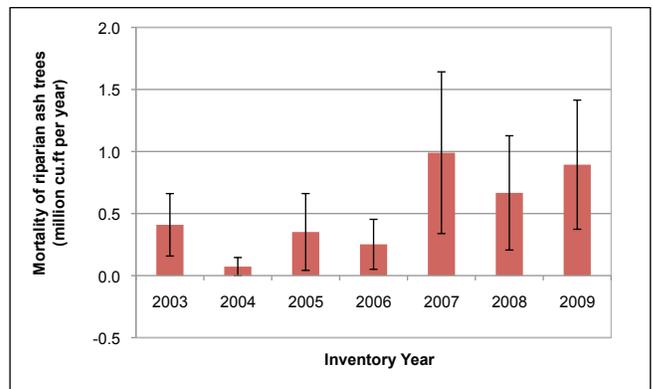


Figure 2—Mortality of growing-stock ash trees on riparian timberland, Indiana, 2003-2009.

USING INVENTORY DATA TO DETERMINE THE IMPACT OF DROUGHT ON TREE MORTALITY

Greg C. Liknes, Christopher W. Woodall, and Charles. H. Perry

ABSTRACT

Drought has been the subject of numerous recent studies that hint at an acceleration of tree mortality due to climate change. In particular, a recent global survey of tree mortality events implicates drought as the cause of quaking aspen mortality in Minnesota, USA in 2007. In this study, data from the Forest Inventory and Analysis program of the USDA Forest Service were analyzed for the period 2000-2009. The fate of individual trees was tracked on a 5-year return interval and the proportion of trees that died was examined in relation to the Palmer Drought Severity Index for the same time period. Quaking aspen mortality increased in northeastern Minnesota over the study period but was stable in north-central Minnesota. The rate of quaking aspen mortality was found to be significantly higher than the mortality rate of all tree species combined in northeastern Minnesota in recent remeasurement periods. Aspen mortality cannot be conclusively attributed to drought without further analysis of contributing factors. While anecdotal observations of small-scale mortality have been cited as evidence of climate-change-induced mortality in other studies, the results of this study suggest further exploration of statistical models for apportionment of inciting, predisposing, and contributing tree mortality factors.

INTRODUCTION

Coincident with the growing concern of climate change effects on ecosystems, the impact of drought on tree mortality has become a topic of interest in both popular and academic literature. Moisture deficiency and increased temperature were linked by van Mantgem et al. (2009) to substantial increases in tree mortality across the western United States, and the story was subsequently reported in popular media outlets (e.g., Boxall 2009). Rapid mortality of quaking aspen (*Populus tremuloides*) has been reported in the western United States (Worrall et al. 2008) and Canada (Hogg et al. 2008) with drought implicated as the cause.

While the aforementioned studies relied on direct observation of mortality at a small number of field sites, some alternative approaches have been pursued. Rehfeldt et al. (2009) used presence/absence data from the USDA Forest Service's Forest Inventory and Analysis (FIA) program to construct a bioclimatic model for the distribution of aspen. The model predicted the current range with only

4.5 percent error for 15,500 observations and allowed for predictions of range shifts (via mortality) under climate change scenarios. A self-identified first global survey of drought-induced tree mortality was developed by analyzing reports from around the world that link tree die-off to local drought conditions (Allen et al. 2010); the implication of this global drought survey is that mortality is increasing in response to global warming.

One of the mortality events attributed to drought by Allen et al. (2010) is the die-off of quaking aspen in northern Minnesota as reported by the Minnesota Department of Natural Resources (2007). It should be noted the report actually states the cause is "unknown," and later reports stated 30,000 acres of quaking aspen had perished from 2004 to 2009 indicating drought likely predisposed trees to attacks by secondary pests (Minnesota Department of Natural Resources 2009). The observations of mortality were derived partly from on-the-ground anecdotal observations and partly from aerial sketch mapping. Given emerging studies that suggest large-scale tree mortality events may be climate-change related, the purpose of this study was to objectively examine the evidence for drought-induced aspen mortality in northern Minnesota using FIA data.

METHODS

The area of investigation was limited to north-central (Climate Division 2) and northeastern (Climate Division 3) Minnesota in which severe drought occurred in 2006 and 2007. Climate Divisions in Minnesota are aggregations of counties with similar weather conditions (www.esrl.noaa.gov/psd/data/usclimate/map.html).

Individual tree data collected by FIA from 2000 to 2008 were analyzed. Because FIA revisited locations in Minnesota on a 5-year remeasurement interval, the status of individual trees was tracked over time (survived, died, or harvested). For the period analyzed, this approach resulted in six unique re-measurement intervals (2000-2004, 2001-

2005, 2002-2006, 2003-2007, 2004-2008, 2005-2009). The FIA database contained information for 1,764 quaking aspen trees and 3,494 total trees in the study area. In addition to the status of individual trees, other attributes were examined such as diameter at breast height (d.b.h.) and age.

Monthly averages of Palmer Drought Severity Index (PDSI) data were acquired from the National Oceanic Atmospheric Administration's National Climatic Data Center. Data are available from 1895 to the present and are reported by climate division. Hurst rescaling was applied to time series of PDSI using the method of Outcalt et al. (1997). The rescaled PDSI simplifies interpretation and provides additional information regarding long-term drought trends.

RESULTS AND DISCUSSION

For north-central Minnesota (Climate Division 2), no statistically significant difference was found in the proportion of quaking aspen trees that experienced mortality across the six overlapping time intervals (Figure 1). For northeastern Minnesota (Climate Division 3), quaking aspen mortality observed for the 2005-2009 time period was significantly higher than the mortality observed in the 2001-2005 time period (Figure 2). Quaking aspen mortality was not significantly different than the mortality of all tree species combined for five out of six time periods in Climate Division 2, but was higher for the four most recent time periods in Climate Division 3. The drought experienced in Climate Division 3 was more prolonged than Climate Division 2 since 2001 (Figure 3) and could have contributed to the increase in aspen mortality with time. Without more information, it is impossible to determine whether quaking aspen trees have been disproportionately affected by drought in Climate Division 3 or if the higher mortality relative to other species is a natural result of successional trajectories and stand age. Given the history of forest harvest activities across the Minnesota, one might expect numerous stands in northern Minnesota to be in latter stages of stand development and predisposed to density and age-related tree mortality.

This study represents a preliminary examination of drought-induced mortality using FIA data, and a variety of challenges in linking drought to mortality were uncovered. Due to the 5-year FIA re measurement interval, it will be difficult to attribute a single-year drought event to an increase in mortality. That is, for each dead tree, the exact year of mortality cannot be determined. There are also scale issues that must be overcome. Many of the mortality events

cataloged by Allen et al. (2010) were described as patchy, and the intensity of the FIA grid may not be well suited to such observations.

Perhaps the larger question is this: what is the correct way to link drought to tree mortality? There is a tendency in drought studies to use correspondence or correlation between the location of drought events and subsequent die off of trees as proof of causality. For example, if one examines Figure 2 and Figure 3 together, there appears to be a correspondence between increased aspen mortality and the prolonged drought of 2006-2007 in Climate Division 3. This correspondence does not prove causality, and it is generally accepted in the forestry community that drought is part of a complex that leads to mortality (as described by the decline spiral model, Manion 1991). At a minimum, an attempt should be made to eliminate other possible causes such as the age of the trees, other damage agents, or poor site quality. Future work should focus on development of statistical models to apportion the explanatory power of inciting, predisposing, and contributing factors that lead to mortality.

If establishing a causal link between a drought event and increased tree mortality requires more than correlation, the same standard should apply to attributing shifts in the drought/mortality cycle to climate change. Kampen (2010) elaborates on the shortfalls of correlational research and points out that models that verify or falsify can be misleading when feedback mechanisms exist and are not well understood.

CONCLUSIONS

Given extensive speculation that future climate change may result in widespread tree mortality, it may be ever more important to accurately establish causality between contemporary tree mortality and climatic events such as droughts. Erroneously attributing contemporary tree mortality to large-scale climatic events, rather than pursuing a better understanding of the host of factors involved, is likely to result in poor predictions for changing climatic conditions. This study found inconclusive evidence of increased quaking aspen tree mortality in Minnesota due to a long-term drought using large-scale inventory data, while anecdotal observations of small-scale mortality have been cited as evidence of climate-change-induced mortality. It is suggested that future studies explore statistical models for apportionment of inciting, predisposing, and contributing tree mortality factors.

LITERATURE CITED

Allen, C.D.; Macalady, A.K., Chenchouni, H.; Bachelet, D. [and others]. 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*. 259: 660-684.

Boxall, B. 2009. West’s trees dying faster as temperatures rise. *The Los Angeles Times*. January 23. Los Angeles, CA.

Hogg, E.H.; Brandt, J.P.; Michaelian, M. 2008. Impacts of a regional drought on the productivity, dieback, and biomass of western Canadian aspen forests. *Canadian Journal of Forest Research*. 38: 1373–1384.

Kampen, J.K. 2010. A methodological note on the making of causal statements in the debate on anthropogenic global warming. *Theoretical and Applied Climatology*. DOI 10.1007/s00704-010-0355-y.

Manion, P.D. 1991. *Tree disease concepts*. Englewood Cliffs, NJ: Prentice-Hall Inc. 399 p.

Minnesota Department of Natural Resources. 2007. Minnesota forest health annual report. St. Paul, MN: Minnesota Department of Natural Resources. 75 p. Available at http://files.dnr.state.mn.us/assistance/backyard/treecare/forest_health/annualreports/2009annualReport.pdf

Minnesota Department of Natural Resources. 2009. Minnesota forest health annual report. St. Paul, MN: Minnesota Department of Natural Resources. 74 p. Available at http://files.dnr.state.mn.us/assistance/backyard/treecare/forest_health/annualreports/2009annualReport.pdf

Outcalt, S.I.; Hinkel, K.M.; Meyer, E.; Brazel, A.J. 1997. Application of Hurst rescaling to geophysical serial data. *Geographical Analysis*. 29: 72-87.

Rehfeldt, G.E.; Ferguson, D.E.; Crookston, N.L. 2009. Aspen, climate, and sudden decline in western USA. *Forest Ecology and Management*. 258: 2353–2364.

van Mantgem, P.J.; Stephenson, N.L.; Byrne, J.C.; Daniels, L.D. [and others]. 2009. Widespread increase of tree mortality rates in the western United States. *Science*. 323(5913): 521-524.

Worrall, J.J.; Egeland, L.; Eager, T.; Mask, R.A. [and others]. 2008. Rapid mortality of *Populus tremuloides* in southwestern Colorado, USA. *Forest Ecology and Management*. 255: 686-696.

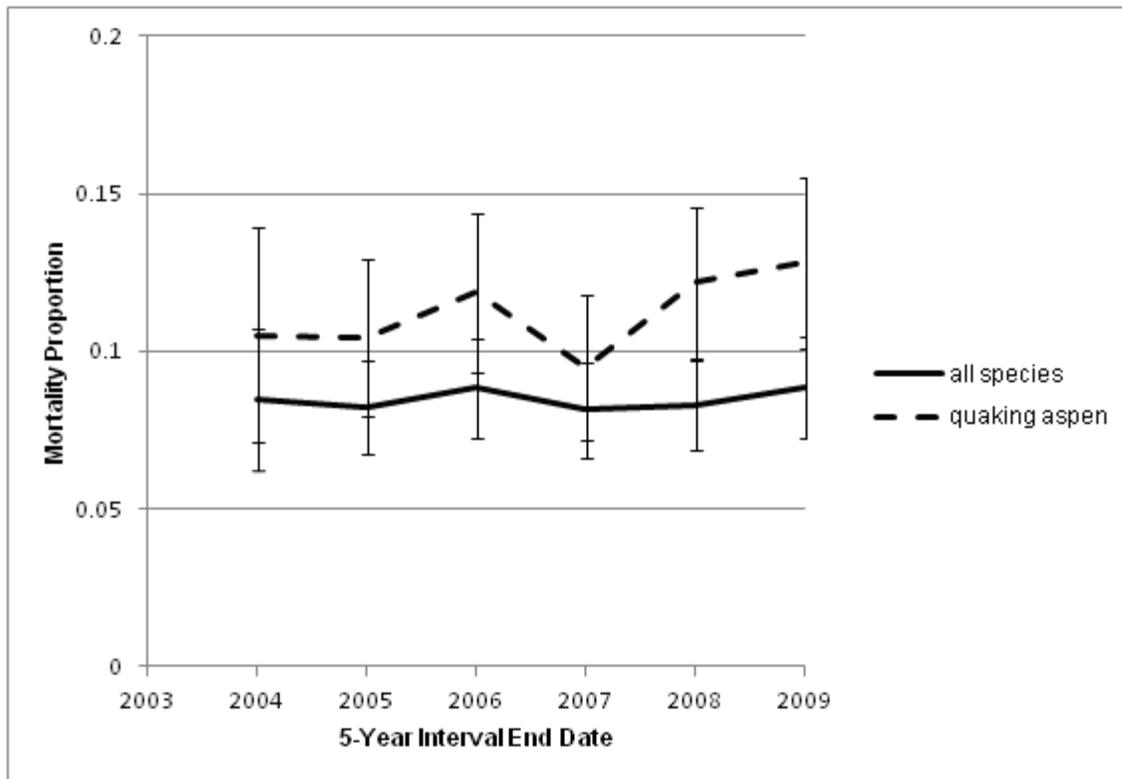


Figure 1—Proportion of trees that died in Climate Division 2 in Minnesota, USA. Trees were revisited on a 5-year interval, and the end year of the interval is depicted on the horizontal axis. Error bars represent the standard error for a sample proportion.

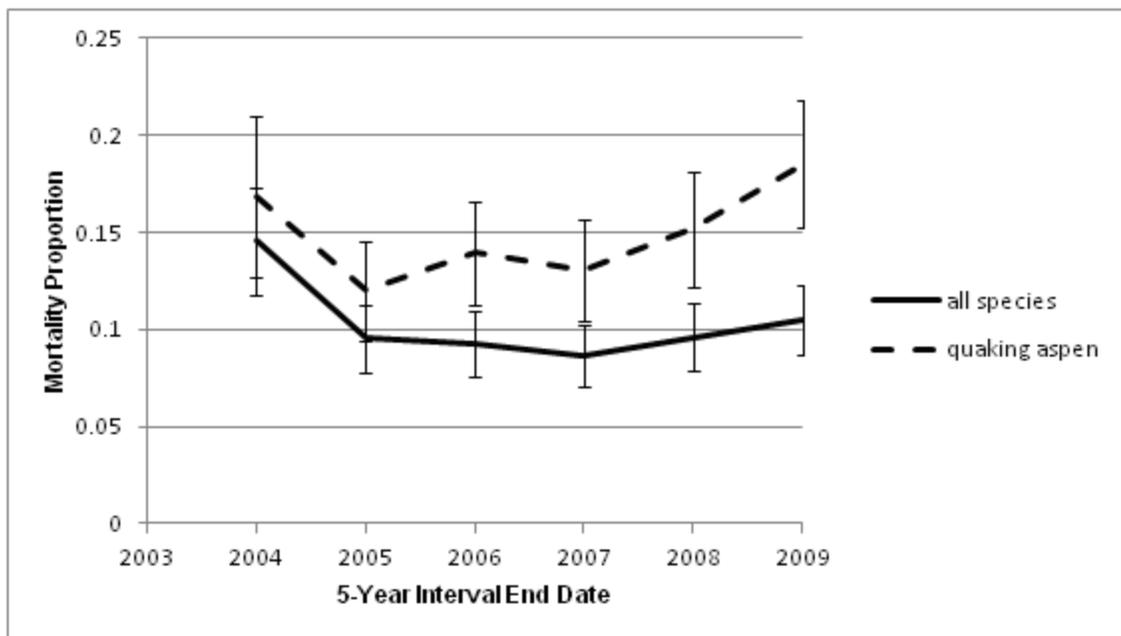


Figure 2—Proportion of trees that died in Climate Division 3 in Minnesota, USA. Trees were revisited on a 5-year interval, and the end year of the interval is depicted on the horizontal axis. Error bars represent the standard error for a sample proportion.

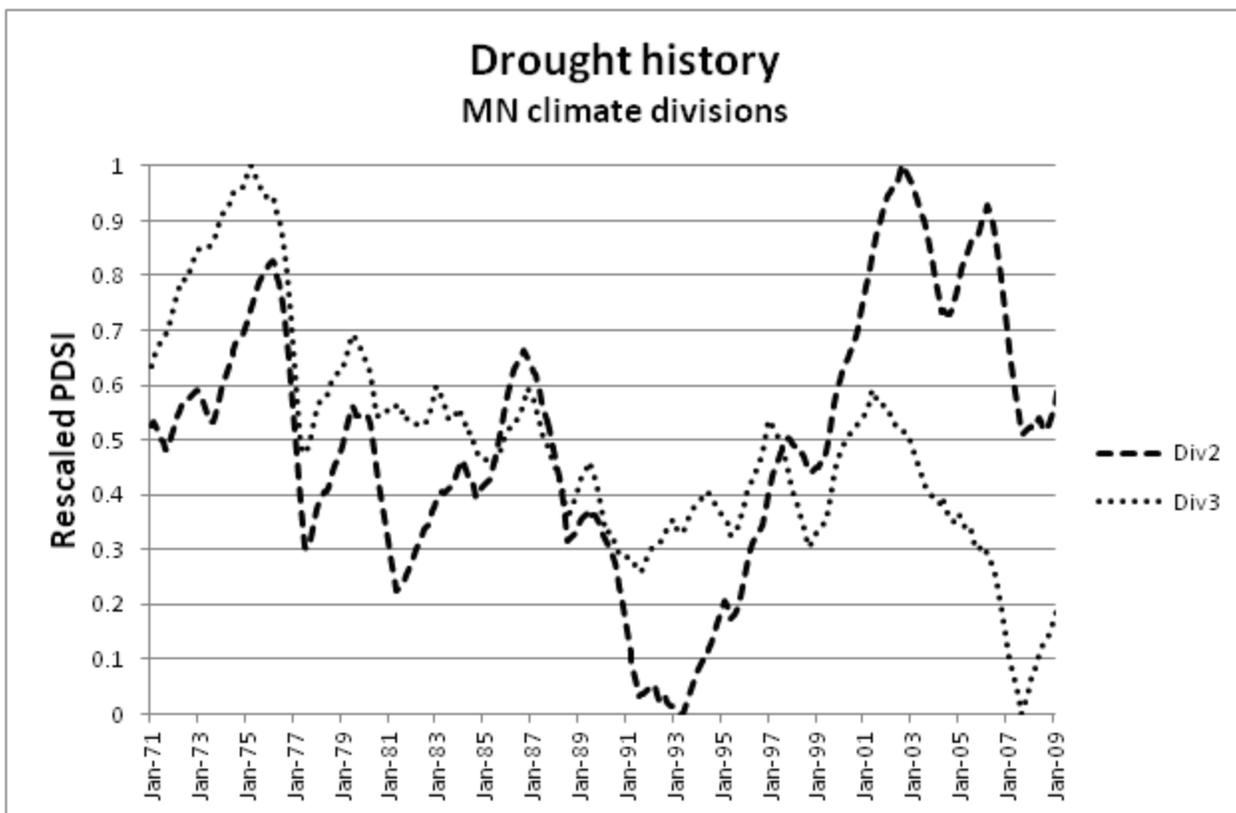


Figure 3—Drought as measured by the Palmer Drought Severity Index from 1971-2009 for Climate Divisions 2 and 3 in Minnesota, USA. Values have been transformed using Hurst rescaling such that positive slopes indicate a change from dryer to wetter conditions and negative slopes indicate a change from wetter to drier conditions. Climate Division 3 experienced its driest condition of the period in 2007 after a sustained shift toward drier conditions that began in 2001.

POTENTIAL IMPACTS OF YEAR-ROUND SAMPLING ON MONITORING PRESENCE-ABSENCE OF INVASIVE FLORA IN THE SOUTHERN UNITED STATES

Christopher M. Oswalt, Sonja N. Oswalt, W. Keith Moser

ABSTRACT

Studies suggest that the southern United States is an area of primary concern with regards to the spread of nonnative invasive plant species. Recent data show that species such as Japanese honeysuckle (*Lonicera japonica*) and Nepalese browntop (*Microstegium vimineum*) are invading forests and displacing native species throughout the southern United States. Monitoring on large spatial scales is among the most important mechanisms for the detection and prevention of the spread of nonnative species.

Accurate assessments of on-going biological invasions are a primary research priority in the Southeast. As one method for addressing this need, the US Forest Service Southern Research Station (SRS), in partnership with State forestry agencies across the South, initiated a southern region survey of 33 invasive plant taxa in 2001 on all forest ownerships as part of the SRS Forest Inventory and Analysis (FIA) program. Currently in the southern United States, presence-absence data is collected for select invasive flora throughout the calendar year. Little is known about the impacts of year-round sampling on the quality of invasive flora data collection. In this study we investigate the implications of year-round sampling on presence-absence data collected by the southern FIA program for states east of the Mississippi river. Chinese and European privets (*Ligustrum* spp) are observed on FIA plots most often between February and May, and least often between September and December. Exotic roses (*Rosa* spp) and Japanese honeysuckle follow a similar trend.

Nepalese browntop, however, is observed more often between August and October. Moreover, Nepalese browntop is observed more than four times as often during peak months than it is during the period between December and April. These results suggest that plant apparency may be impacting the quality of presence-absence data collected by the SRS-FIA program. While the systematic nature of the FIA sampling design minimizes the impact to population estimates of sampled invasive flora, year-round sampling may be impacting attempts to accurately portray the geographical distribution of a given plant.

Keywords: forest inventory, invasive species, exotic plants, sampling bias, plant survey

INTRODUCTION

Nonnative invasive plant species (NNIPS) are threats to southern forests through the displacement of native species (Mooney and Cleland 2001), the alteration of soil physical

and chemical properties (Bruce and others 1995, Jose and others 2002), and the disruption of successional pathways (Oswalt and others 2007) among other potential impacts (Gordon 1998, Jose and others 2002). Environmental impacts coupled with attempts to control and/or eradicate NNIPS are costly, as exemplified by the estimated \$3 to \$6 million spent annually by the State of Florida to manage the highly invasive Chinese tallowtree (*Melaleuca quinquenervia*; Pimentel and others 2005). Because of the environmental and ecological burdens posed by these species, NNIPS inventory and monitoring is considered a priority in the South.

Effective inventory and monitoring programs depend on reliable data. Monitoring vegetation on a large scale can be challenging, however. The potential effects of observer-bias in vegetation monitoring have been documented and include species misidentification and missed species occurrences, along with widely varying interpretations of area cover, all of which can result in inaccurate representations of species' abundances and spread (Archaux and others 2006, Leps and Hadincova 1992, Gotfryd and Hansell 1985, Hall and Okali 1978). Additionally, species diversity and abundance estimates are closely correlated with the seasonal sampling period, particularly as pertains to herbaceous ground flora, and single-season sampling may result in underestimates (Small and McCarthy 2002). Typically, vegetation surveys in the deciduous forests of the north and southeast are conducted during the growing season when flora are in leaf-on condition and/or are flowering. In some cases, however, overarching monitoring goals may result in sampling during the dormant season, as with the USDA Forest Service Southern Research Station Forest Inventory and Analysis (SRS-FIA) program.

Currently, the increase in the number of plots surveyed for invasive plants by using the FIA Phase 2 plots provides significant additional data over using the FIA Phase 3 plots.

Christopher M. Oswalt, Research Forester, USDA Southern Research Station Forest Inventory and Analysis, 4700 Old Kingston Pike, Knoxville, TN 37922

Sonja N. Oswalt, Forester, USDA Southern Research Station Forest Inventory and Analysis, 4700 Old Kingston Pike, Knoxville, TN 37922

W. Keith Moser, Research Forester, USDA Northern Research Station Forest Inventory and Analysis, 1992 Folwell Ave., St. Paul, MN 55108

While Phase 3 plots are only surveyed during the growing season months, the sample intensity is 1/16th of the Phase 2 plots. Moreover, the added logistical challenges of sampling a subset of the Phase 2 plots along with the fact that the invasive plant survey is a secondary goal render such an option untenable.

The SRS-FIA program began tracking forest health threats, including NNIPS, on forestland in 2001. The NNIPS-monitoring component provides a mechanism for monitoring the spread of common (known) invasive plants on both public and private land at a large scale by utilizing the existing FIA system of forest inventory plots. Data collection occurs year-round in all southern states, regardless of the expressed phenology of the vegetation (i.e., leaf-on, withered, brown, etc...).

The impacts of this year-round sampling on the quality of NNIPS flora data collection have not been quantified. In this study we investigate the implications of year-round sampling on presence-absence data collected by the southern FIA program for states east of the Mississippi river with the specific objective of quantifying the impact of year-round sampling on the SRS-FIA invasive plant data.

METHODS

DATA COLLECTION

The FIA program collects data on plots distributed in a random, systematic fashion on both private and public land across the United States. The plot design consists of four 1/6-acre fixed-radius subplots arranged in a “tri-areal” configuration (Bechtold and Patterson 2005). A number of environmental conditions, tree-level variables, and abiotic measures are recorded on each subplot. Detailed explanations of the FIA plot design and sampling phases are given in Bechtold and Patterson (2005).

Nonnative invasive plant species are among the variables sampled on all FIA plots in the Southern region. The NNIPS program was implemented in 2001 to meet the needs of State forestry agencies and other partners for tracking the emergence and spread of species known to cause ecological problems in southern forests. Observers are trained in the detection of 33 species classified into 6 life forms (trees, shrubs, herbs, grasses, vines, and ferns) from a pre-developed list of NNIPS (Table 1).

Additional NNIPS are recorded in Florida; however, for the purposes of this study the species analyzed were limited to those sampled in all southern states. Observers are instructed to note the presence/absence and percent cover (< 01 percent, 01-10 percent, 11-50 percent, 51-90 percent, and 91-100 percent) of (up to) the four most abundant species found on the forested portion (condition) of

each sampled subplot. Sampling occurs year-round, and observers are instructed to record an estimate of percent cover as though plants are in leaf-on condition when sampling occurs during the dormant season. Studies suggest that observer bias may be minimized by consistent and frequent calibration (training) and quality control procedures (Kercher, Frieswyk, and Zedler 2003). Standard FIA quality assurance procedures apply to the NNIPS program, and include randomly-selected plots subjected to checks by certified quality control personnel.

DATA ANALYSIS

We used 29,558 SRS-FIA plots from 9 southern States to examine the impact of year-round sampling (Figure 1). States included in the analysis were Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, and Virginia. We calculated the relative occurrence for each NNIPS at the plot level by the month in which the data were collected. Plots-by-month were grouped into season for analysis (Spring – March to May, Summer – June to August, Fall – September to November, and Winter – December to February) and subjected to a one-way analysis of variance (ANOVA) with Tukey mean separation to determine if differences in relative occurrence existed among sample seasons.

Additionally, we calculated the relative occurrence and estimate of cover for each NNIPS by subplot and month. We share those preliminary results graphically.

Results – Nonnative invasive plants from the predetermined list were detected on 15,720 (53 percent) of the sampled plots. Japanese honeysuckle was the most frequently observed species, while giant reed was least common (Table 2). Observers noted 3 or fewer species on most (92 percent) of the plots containing NNIPS, though 2 plots contained at least 8 species from the list (Table 3).

While inter-seasonal detection differences were found for 9 of the 33 species surveyed ($p < 0.05$; Table 4), four species appeared to exhibit the most notable differences (mimosa, Nepalese browntop, Chinese lespedeza, and shrubby lespedeza). The time of observation bias differed depending on the individual species and its phenotypic expression. For example, mimosa is a tree that presents an easily recognizable flower in the summer months and, indeed, detection rates were higher in the summer than in any other season (Figure 2a). Winter detections were least common for mimosa, when the plant is in leaf-off phase. Nepalese browntop was most frequently detected in summer and fall (Figure 2b), when foliage is most noticeable because of its height. Chinese and shrubby lespedeza exhibited similar patterns wherein relative occurrence was extremely low in the winter and highest in the summer and fall periods (Figure 2c and 2d, respectively).

Preliminary graphs of relative subplot occurrence by month indicate that detection differences appear to exist for species in each life form (Figure 3) with the exception of vines (Figure 3c) and possibly shrubs (Figure 3b). Mimosa detections occurred most often from April through July, the flowering period for the species, and Chinese tallowtree was most often detected from June through August (Figure 3a).

Shrubs and vines showed little monthly variation in relative occurrence, though ANOVA indicated that seasonal variation in detection did exist for privet species, which may be detected more frequently in the winter simply because it is an evergreen and perhaps more likely to be correctly identified during the months when other species are dormant. Two of the 6 grasses varied in their monthly detection rates (Figure 3d).

Microstegium detection rates peaked in July, August, and September, while tall fescue detection rates were highest from May through November—the typical growing season in the southern states. Ferns and other herbaceous species exhibited some monthly differences in detection rates (Figure 3e). Detection rates for both lespedeza species were highest from June through September, with peaks in August, while detection rates for Japanese climbing fern peaked in July.

Inter-seasonal differences in relative occurrence appeared strongest in the smaller cover classes. For example, the range in monthly relative occurrence of Chinese lespedeza was greatest for the <01 percent cover class, followed by the 01-10 percent, 11-50 percent, and 51-90 percent cover classes and was smallest for the >90 percent cover class (Figure 4a and 4b). Concomitantly, the overall relative occurrence was greatest for the smaller cover classes (Figure 4c). Moreover, the monthly mean relative occurrence deviated very little from the annual mean relative occurrence for the larger cover classes (Figure 4d). This pattern was similar for the majority of the species that exhibited a significant inter-seasonal bias.

DISCUSSION

Preliminary results from our study indicate that seasonal detection bias occurs for some species on the SRS-FIA NNIPS list. Species are most likely to be detected during the peak of the southern growing season (late spring, summer, early fall), or when some distinct characteristic (e.g. flowers, herb height, etc...) increases visibility in the forest understory. Seasonal detection bias may result in false negatives on FIA plots, thus underestimating invasion rates. Bias appears to be limited to the winter months, suggesting that sampling during the winter is less effective than sampling during the other three seasons for a limited number of species collected. In addition to underestimating invasion

rates, seasonal biases in estimates of percent cover add error to modeled representations of invasion threats on the ground. Underestimating invasion rates and/or extents may prevent managers from directing resources to appropriate areas for control and eradication efforts.

The most surprising result was that of seasonal differences in the relative occurrence of some tally tree species. For example, mimosa exhibited a five-fold increase in mean relative occurrence from winter to spring. Seasonal differences in relative occurrence among tally trees may suggest a need for increased winter identification training. Moreover, such results suggest that this analysis has potential for use within the quality control program of FIA in order to identify potential additional training needs.

The results of this study suggest that SRS-FIA may want to reconsider sampling during winter months, limiting sampling to the growing season. If sampling continues year-round, the study results suggest that additional measures are needed to train observers to recognize NNIPS during the dormant season, and that quality assurance personnel may need to pay extra attention to the NNIPS component of sampled variables during the winter months.

Data reliability is a key component of inventory and monitoring programs. The preliminary results of this study indicate that measures may need to be taken to ensure high quality NNIPS data are available year-round in the south. Further research is needed to quantify the impact of potential false-negatives to invasive plant distribution modeling using SRS-FIA invasive plant data. In addition, further research is needed to better understand this bias through multiple plots visits within a given year.

ACKNOWLEDGEMENTS

The authors thank Songlin Fei for discussions and Roger Conner for his review that helped improve this manuscript over the original version. The authors would also like to thank the many individuals at the Forest Inventory and Analysis Symposium in Knoxville, TN who contributed, through an open forum discussion, to ideas presented in this manuscript.

LITERATURE CITED

- Archaux, F.**; Gosselin, F.; Bergès, L.; Cevalier, R. 2006. Effects of sampling time, species richness, and observer on the exhaustiveness of plant censuses. *Journal of Vegetation Science* 17:299-306.
- Bechtold, W.A.**; Patterson, P.L. eds. 2005. The enhanced forest inventory and analysis program - national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 85 p.

- Bruce, K.A.;** Cameron, G.N.; Harcombe, P.A. 1995. Initiation of a new woodland type on the Texas coastal prairie by the Chinese tallow tree (*Sapium sebiferum* (L.) Roxb.). *Bulletin of the Torrey Botanical Club* 122(3):215-225.
- Gordon, D.R.** 1998. Effects of invasive, non-indigenous plant species on ecosystem processes: lessons from Florida. *Ecological Applications* 8:975-989.
- Gotfryd, A.;** Hansell, R.I.C. 1985. The impact of observer bias on multivariate analyses of vegetation structure. *Oikos* 45(2):223-234.
- Hall, J.B.;** Okali, D.U.U. 1978. Bias in a floristic survey of complex tropical vegetation. *Journal of Ecology* 66(1):241-249.
- Jose, S.;** Cox, J.; Merritt, S.; Miller, D.L.; Shilling, D.G. 2002. Alien plant invasions: the story of cogongrass in southeastern forests. *Journal of Forestry* 100:41-44.
- Kercher, S.M.;** Frieswyk, C.B.; Zedler, J.B. 2003. Effects of sampling teams and estimation methods on the assessment of plant cover. *Journal of Vegetation Science* 14:899-906.
- Leps, J.;** Hadincova, V. 1992. How reliable are our vegetation analyses? *Journal of Vegetation Science* 3:119-124.
- Mooney, H.A.;** Cleland, E.E. 2001. The evolutionary impact of invasive species. *PNAS* 98:5446-5451.
- Oswalt, C.M.;** Oswalt, S.N.; Clatterbuck, W.K. 2007. Effects of *Microstegium vimineum* (Trin.) A. Camus on native woody species density and diversity in a productive mixed-hardwood forest in Tennessee. *Forest Ecology and Management* 242(2007):727-732.
- Pimentel, D.;** Zuniga, R.; Morrison, D. 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics* 52(3):273-288.
- Small, C.J.;** McCarthy, B.C. 2002. Spatial and temporal variability of herbaceous vegetation in an eastern deciduous forest. *Plant Ecology* 164:37-48.

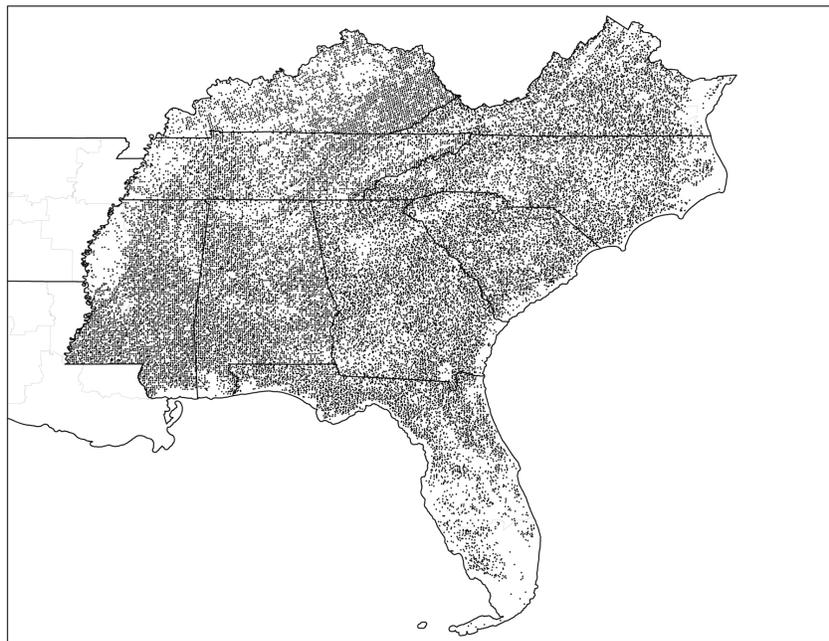


Figure 1—Approximate plot locations of forested plots where invasive plant data was collected by the Southern Research Station Forest Inventory and Analysis program between roughly 2002 and 2007.

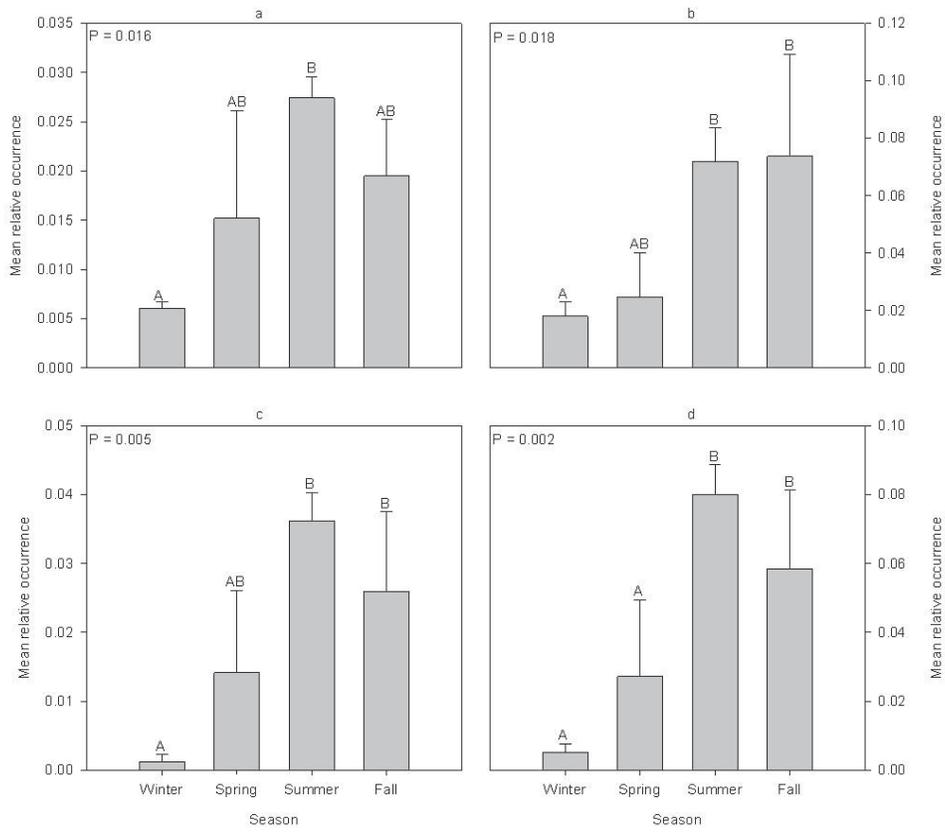


Figure 2—Mean relative occurrence of (a) mimosa, (b) microstegium, (c) Chinese lespedeza, and (d) shrubby lespedeza across four seasonal categories. Bars with different lettering indicates significant inter-seasonal differences (alpha 0.05).

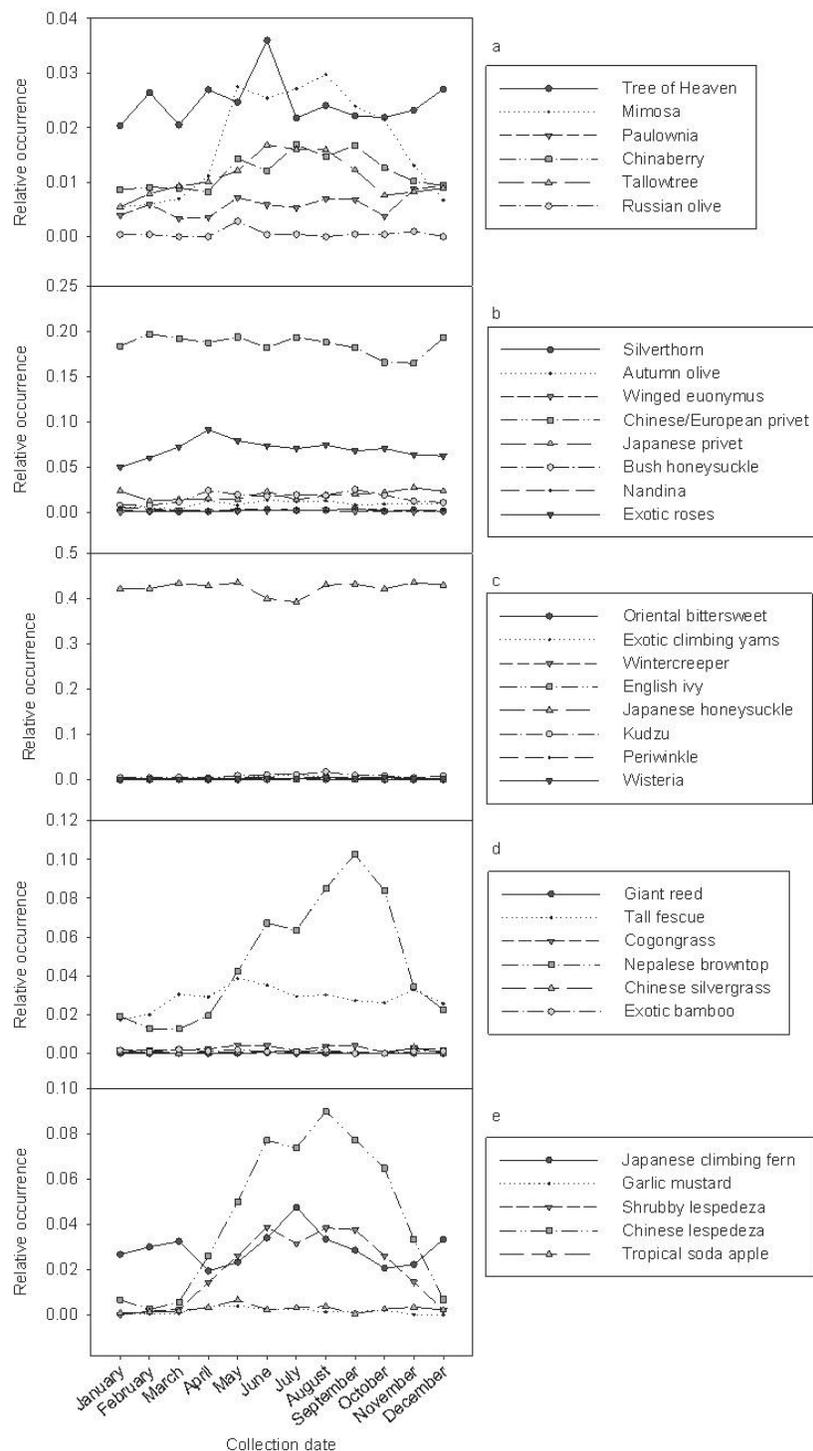


Figure 3—Monthly relative occurrence for invasive (a) trees, (b) shrubs, (c) vines, (d) grasses, and (e) ferns, forbs and other herbaceous plants collected by the Southern Research Station Forest Inventory and Analysis program between roughly 2002 and 2007.

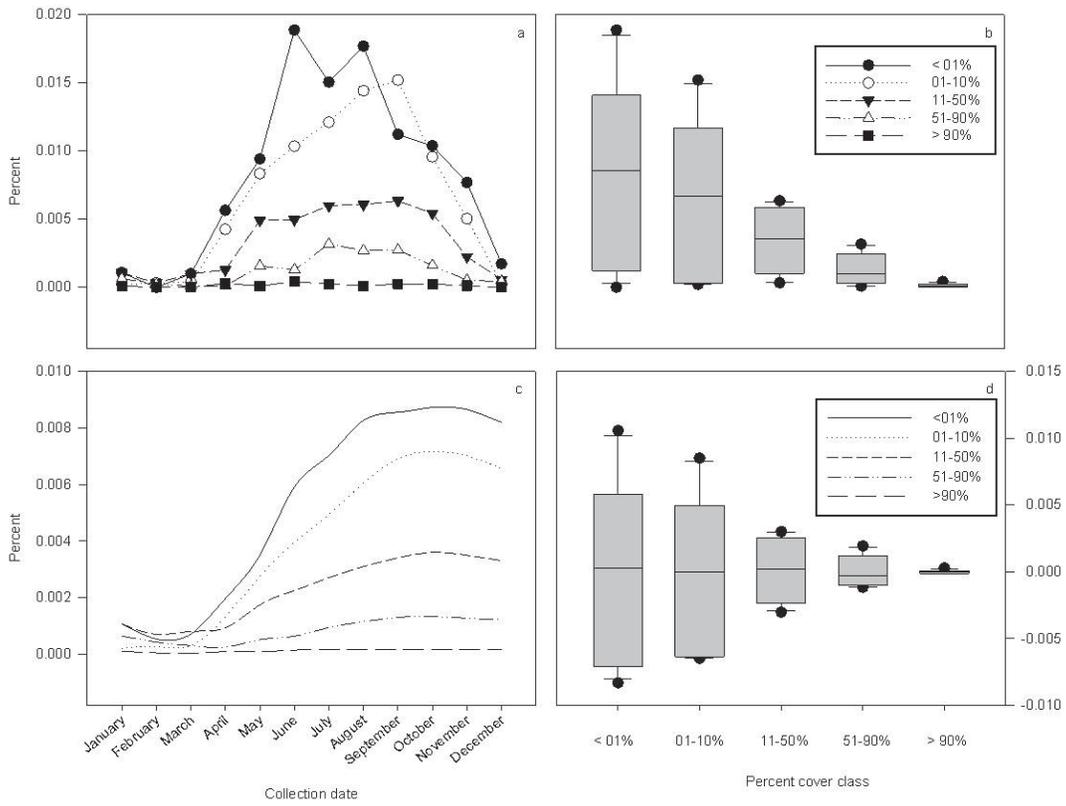


Figure 4—Estimates of (a) monthly relative occurrence, (b) boxplot for annual mean of monthly relative occurrence, (c) cumulative relative occurrence, and (d) boxplot of average monthly deviation from the annual mean for Chinese lespedeza across four cover classes.

Table 1—List of common and scientific names of each invasive plant collected in all states by the Southern Research Station Forest Inventory and Analysis program

Trees		Vines	
Tree of heaven	<i>Ailanthus altissima</i> (Mill.) Swingle	Oriental bittersweet	<i>Celastrus orbiculatus</i> Thunb.
Mimosa	<i>Albizia julibrissin</i> Durazz.	Nonnative-yams	<i>Dioscorea</i> Spp.
Paulownia	<i>Paulownia tomentosa</i> (Thunb.) Sieb. & Zucc. ex Steud.	Wintercreeper	<i>Euonymus fortunei</i> (Turcz.) Hand.-Maz.
Chinaberry	<i>Melia azedarach</i> L.	English Ivy	<i>Hedera helix</i> L.
Tallowtree	<i>Triadica sebifera</i> (L.) Small	Japanese honeysuckle	<i>Lonicera japonica</i> Thunb.
Russian-olive	<i>Elaeagnus angustifolia</i> L.	Kudzu	<i>Pueraria montana</i> var. <i>lobata</i> (Willd.)
Grasses		Nonnative vincas	<i>Vinca</i> Spp.
Giant reed	<i>Arundo donax</i> L.	Wisteria	<i>Wisteria</i> Spp.
Tall fescue	<i>Festuca arundinacea</i> Schreb.	Shrubs	
Cogongrass	<i>Imperata cylindrica</i> (L.) P. Beauv.	Silverthorn	<i>Elaeagnus pungens</i> Thunb.
Nepalese browntop	<i>Microstegium vimineum</i> (Trin.) A. Camus	Autumn olive	<i>Elaeagnus umbellata</i> Thunb.
Chinese silvergrass	<i>Miscanthus sinensis</i> Anders.	Winged euonymus	<i>Euonymus alatus</i> (Thunb.) Sieb.
Nonnative bamboos	<i>Bambusa</i> spp.	Chinese/European privet	<i>Ligustrum sinense</i> Lour.
Ferns, Forbs/Other Herbaceous			<i>Ligustrum vulgare</i> L.
Japanese climbing fern	<i>Lygodium japonicum</i> (Thunb. ex Murr.) Sw.	Japanese/Glossy privet	<i>Ligustrum japonicum</i> Thunb.
Garlic mustard	<i>Alliaria petiolata</i> (M. Bieb.) Cavara & Grande		<i>Ligustrum lucidum</i> W.T. Aiton
Shrubby lespedeza	<i>Lespedeza bicolor</i> Turcz.	Bush honeysuckle	<i>Lonicera</i> Spp.
Chinese lespedeza	<i>Lespedeza cuneata</i> (Dum.-Cours.) G. Don	Sacred-bamboo nandina	<i>Nandina domestica</i> Thunb.
Tropical soda apple	<i>Solanum viarum</i> Dunal	Nonnative roses	<i>Rosa</i> spp.

Table 2—Number of individual plots and percent of total forested plots sampled on which each invasive plant was observed by the Southern Research Station Forest Inventory and Analysis program during time period covering roughly 2002-2007

Life Form	Plant species	Plot Occurrences	Percent Occupied
Trees	Tree of Heaven	726	2.46
	Mimosa	501	1.69
	Paulownia	169	0.57
	Chinaberry	346	1.17
	Tallowtree	321	1.09
	Russian olive	16	0.05
Shrubs	Silverthorn	60	0.20
	Autumn olive	271	0.92
	Winged euonymus	49	0.17
	Chinese/European privet	5,484	18.55
	Japanese privet	553	1.87
	Bush honeysuckle	482	1.63
	Nandina	88	0.30
	Exotic roses	2,077	7.03
Vines	Oriental bittersweet	25	0.08
	Exotic climbing yams	123	0.42
	Wintercreeper	32	0.11
	English ivy	69	0.23
	Japanese honeysuckle	12,524	42.37
	Kudzu	255	0.86
	Periwinkle	99	0.33
	Wisteria	104	0.35
Grasses	Giant reed	1	0.00
	Tall fescue	845	2.86
	Cogongrass	73	0.25
	Nepalese browntop	1,356	4.59
	Chinese silvergrass	16	0.05
	Exotic bamboo	32	0.11
Ferns, Forbs/Other Herbaceous	Japanese climbing fern	859	2.91
	Garlic mustard	48	0.16
	Shrubby lespedeza	567	1.92
	Chinese lespedeza	1,242	4.20
	Tropical soda apple	77	0.26

Table 3—Number of plots by the number of unique invasive plant species that was observed by the Southern Research Station Forest Inventory and Analysis program during the time period covering roughly 2002-2007

Nonnative Plant Count	Plots (no.)
1	6,794
2	5,326
3	2,394
4	904
5	246
6	49
7	5
8	2

Table 4—F-test statistic and associated p-value for one way analysis of variance testing for inter-seasonal differences in relative occurrence of each invasive plant species collected by the Southern Research Station Forest Inventory and Analysis program during the time period covering roughly 2002-2007

Species	F	P
Autumn olive	2.409	0.142
Bush honeysuckle	3.539	0.068
Chinaberry	2.715	0.115
Chinese lespedeza	12.121	0.002
Chinese silvergrass	0.660	0.600
Chinese/European privet	6.205	0.018
Cogongrass	0.476	0.708
English ivy	2.004	0.192
Exotic bamboo	3.778	0.059
Exotic climbing yams	3.712	0.061
Exotic roses	7.629	0.010
Garlic mustard	2.789	0.109
Giant reed	1.000	0.441
Japanese climbing fern	3.831	0.057
Japanese honeysuckle	3.109	0.089
Japanese privet	1.936	0.202
Kudzu	4.655	0.036
Nepalese browntop	6.469	0.016
Mimosa	6.085	0.018
Nandina	0.770	0.543
Oriental bittersweet	2.185	0.168
Paulownia	0.427	0.739
Periwinkle	1.434	0.303
Russian olive	0.438	0.732
Shrubby lespedeza	9.353	0.005
Silverthorn	1.866	0.214
Tall fescue	4.962	0.031
Tallowtree	14.581	0.001
Tree of Heaven	0.586	0.641
Tropical soda apple	1.480	0.292
Winged euonymus	2.790	0.109
Wintercreeper	1.340	0.328
Wisteria	1.664	0.251

WHITE ASH (*FRAXINUS AMERICANA*) HEALTH IN THE ALLEGHENY PLATEAU REGION, PENNSYLVANIA: EVALUATING THE RELATIONSHIP BETWEEN FIA PHASE 3 CROWN VARIABLES AND A CATEGORICAL RATING SYSTEM

Alejandro A. Royo,^{1*} Kathleen S. Knight,² Jamie M. Himes,³ Ashley N. Will³

ABSTRACT

Following the detection of white ash (*Fraxinus americana*) decline in the Allegheny National Forest (ANF) of Pennsylvania, we established an intensified white ash monitoring network throughout the ANF. We rated crowns using both a categorical system as well as Forest Inventory and Analyses (FIA) Phase 3 measures of uncompact live crown ratio, density, dieback, and transparency. Across our plots, ash averaged 17.33 trees/acre and made up 19 percent of the total stand basal area. We found that trees on lower slopes were healthier than those on upper slopes. Categorical ratings correlated well with density and dieback and allowed us to develop and test conversion formulas to predict dieback and density using categorical scores. Our formulas proved robust in predicting density and dieback on an independent ash tree dataset. Erroneous predictions were generally linked to differences in what either measure defined as dieback. We suggest categorical assessments may provide a suitable alternative for rapid crown evaluations.

INTRODUCTION

Ash (*Fraxinus* spp.) is an important component of eastern deciduous forests and a valued timber species; an estimated 1.25 billion trees (≥ 5 inches DBH) grow throughout the 24 state region of the Upper Midwest and northeastern United States inventoried by the USDA Northern Research Station (USDA Forest Service 2010). Throughout much of this area, ash species, and in particular white ash (*Fraxinus americana*), have suffered from episodic periods of decline and dieback since at least the 1920s with more recent escalations in dieback and mortality heightening interest

in this problem (Hibben and Silverborg 1978, Sinclair and others 1990). Even more recently, the current and projected decimation of existing ash populations by the exotic emerald ash borer (*Agrilus planipennis*; EAB) beetle has focused attention on the current status and risk of ash regionally (Cappaert and others 2005). Indeed, baseline measures of ash health and decline status may be critical because susceptibility risk and mortality rates from EAB may be greater in stressed trees than in healthy trees during the early stages of EAB invasion (McCullough and others 2009).

While drought, fungal pathogens, and phytoplasmas all contribute to regional declines in ash health (Hibben and Silverborg 1978, Woodcock and others 1993), several lines of evidence suggest nutrient deficiencies may play a role in predisposing ash trees to decline. First, several studies indicate white ash is a base cation (e.g., Ca^{2+} , Mg^{2+}) demanding species that is consistently associated with soils with higher pH and greater base cation availability (Bigelow and Canham 2002, Finzi and others 1998, van Breemen and others 1997). Second, researchers have found strong relationships between exchangeable base cation availability and regional declines of sugar maple (*Acer saccharum*), another hardwood species with high cation requirements (Horsley and others 2000, Long and others 2009). Finally, Morin and colleagues (2006) found ash decline was primarily responsible for an observed 60 percent decrease in ash live basal area/acre in the 1990s in an intensified FIA/FHM monitoring plot network established in the Allegheny National Forest (ANF) and suggested mortality was concentrated on ridgetops and upper slopes. This latter finding is critical because within the unglaciated portion of the ANF, concentrations of exchangeable calcium and

¹ Research Ecologist, U.S. Department of Agriculture Forest Service, Northern Research Station, P.O. Box 335, Irvine, PA 16329. (814) 563-1040. aroyo@fs.fed.us

² Research Ecologist, U.S. Department of Agriculture Forest Service, Northern Research Station, 359 Main Road, Delaware, OH 43015.

³ Biological Science Technician, U.S. Department of Agriculture Forest Service, Northern Research Station, P.O. Box 335, Irvine, PA 16329.

magnesium are often greater on lower slopes than on upper slopes, often by as much as an order of magnitude (Bailey and others 2004).

Despite this provocative detection of ash decline on upper slope positions, the robustness of this finding is problematic because ash is greatly underrepresented with just 60 live trees distributed across 16 percent of the intensified FIA/FHM monitoring plots throughout the ANF. Furthermore, existing individuals are distributed unevenly across physiographic or topographic positions (e.g., 1 individual on dry, upper ridgetops vs. 40 in rolling uplands), making a well replicated analysis impossible. Given the small sample sizes and the highly unbalanced representation of ash, a definitive assessment of current white ash health status and its variability across topographic gradients (i.e., soil nutrition) requires enhancement of existing data and greater ash canopy sampling. Furthermore, given the impending invasion of emerald ash borer to the ANF, knowing what landscape positions are likely to contain stressed ash trees may lead to better risk models and mitigation strategies.

The objectives of this study were to assess white ash health status in the Allegheny Plateau region by establishing an expanded network of ash health monitoring plots throughout the ANF to complement existing FHM/FIA data. This enhanced sampling will allow us to explore, in detail, how topographic position and site characteristics (e.g., soil pH and nutrition) are related to ash decline and mortality patterns across the landscape. Additionally, we created a formula to relate FIA Phase 3 (P3) crown ratings (dieback, density, etc) to a user friendly categorical rating system. Finally, we tested the predicted FIA P3 crown ratings derived from our conversion formula on an independent dataset of ash crown health from Ohio.

METHODS

Throughout May – August 2009, we systematically surveyed areas across the entire ANF to establish new white ash monitoring plots (Figure 1). We superimposed a grid over the ANF ownership and, using existing stand information on ash basal area and landform classifications of both bottom/foot slope positions and shoulder/plateau top positions, we searched each 700 ha block to locate pairs of ash plots where one plot in the pair was established on a lower slope position and the other on an upper slope position. Because topographic position itself is not always a reliable indicator of site nutrition on the unglaciated portion of the ANF, we also conducted herbaceous plant surveys at each potential plot to detect species known to be reliable indicators of high base cation nutrition (Horsley and others 2008). At each plot, a focal ash tree was defined as plot center and the surrounding tree community was inventoried

in a variable radius plot using a 10 factor prism. Within each pair of upper and lower slope positions, care was taken to choose focal trees that were similar in diameter and crown class. Trees with two trunks were counted as separate trees if they divided below breast height (4.5 ft).

In summer 2009, we assessed the crown health status of all ash trees in the plots using a categorical rating system developed for managers for assessing ash decline due to emerald ash borer infestation; this rating system is itself a modification of protocols developed for bronze birch borer (Ball and Simmons 1980, Smith 2006). The rating scale is defined as follows:

1. Ash tree with a full, healthy canopy
2. Ash tree with a thinning canopy but no dieback
3. Ash tree with dieback, defined as dead twigs or branches near the top of the tree, exposed to sunlight. Dead branches that are low and shaded were not rated and considered a normal part of branch senescence
4. Ash tree with less than 50 percent of a full canopy, which could occur through a combination of dieback and thinning
5. Ash tree with a dead canopy, defined as no foliage in the canopy portion of the tree (The canopy is counted as dead even if live epicormic sprouts low on the trunk or stump sprouts are present.)

In summer 2010 we revisited each plot and assessed crown health using both the categorical measures and FIA P3 measures of uncompact live crown ratio (UCLR), crown density, dieback, and transparency. UCLR is estimated as the percentage of the actual tree length made up of live crown. Crown density estimates the amount of crown branches, foliage, and reproductive structures that block light visibility through the crown. Crown dieback is the percentage of the live crown area that exhibits signs of recent dieback, excluding snag branches and gaps in the canopy. Finally, transparency is calculated by estimating the percentage of skylight visible through the live, normally foliated portion of the crown (USDA Forest Service 2007).

We tested for differences in these canopy condition response variables using mixed model analysis of variance (PROC MIXED; SAS Institute Inc. 2005). Topographic position (upper/lower) was modeled as a fixed factor, and each 700 ha block was modeled as a random factor. We used Spearman rank correlations to examine the relationship between our categorical rating system and each of the FIA P3 crown measures and to identify the P3 measures that exhibited the strongest relationship. We then used linear and nonlinear regression analyses to derive the most parsimonious formula that best (e.g., high r^2 value) related our categorical measure with P3 metrics.

Finally, we tested the predictive ability of our Pennsylvania derived formulas on an independent dataset of ash canopy conditions from Ohio. This dataset contained 494 individuals of various native ash species (e.g., green ash, white ash) located throughout seven stands that varied in the severity of emerald ash borer infestation (Knight, unpub. data). We assessed the robustness of the predicted values by rounding the predicted value to the nearest 5 percent class and allowing a ± 10 percent tolerance (2 classes) as indicated in the FIA Phase 3 Field Guide (USDA Forest Service 2007).

RESULTS

Overall, we established 192 plots and assessed 538 white ash trees throughout the ANF. Ash basal area relative abundance, which is biased toward plots where ash was present, averaged 19.03 ± 1.1 percent of the basal area and ranged from 4.5 to 75 percent. Mean ash stem density was 17.33 ± 1.2 trees per acre and ranged from 2 to 111 trees per acre. Neither measure of ash abundance differed significantly between topographic positions:

Basal Area : $\bar{X}_{lower} = 17.9 \pm 1.4$ and $\bar{X}_{upper} = 20.8 \pm 1.6, P = 0.2$

Stem Density : $\bar{X}_{lower} = 16.3 \pm 1.6$ and $\bar{X}_{upper} = 19.1 \pm 1.8, P = 0.3$

Crown health assessments determined by the categorical method and FIA P3 variables of density, dieback, and transparency all indicated crowns were healthier on lower slope positions than on upper slopes (Figure 2). The P3 variable of uncompact live crown ratio did not differ significantly between topographic positions

($x = 31.6 \pm 0.7$ and $x = 31.2 \pm 0.7, P = 0.73$).

Correlation analyses revealed that all three FIA P3 ratings were significantly correlated with our categorical measure (Table 1); however only density and dieback exhibited moderately strong correlations (r values > 0.7 or < -0.7). We therefore further explored the relationship between the categorical rating and these two measures.

Regression analyses indicated our categorical rating was linearly related to Crown Density (adjusted $r^2 = 0.65$; Figure 3) and yielded the following formula:

$Y_{Density} = -10.59x + 59.114$; where $x = \text{Categorical Rating (1 - 5)}$

The best relationship between our categorical rating and Crown Dieback was obtained using an exponential model (adjusted $r^2 = 0.91$; Figure 4):

$Y_{Dieback} = 0.2626 \times e^{-1.18x}$; where $x = \text{Categorical Rating (1 - 5)}$

Predictions of crown density derived from the categorical ratings were accurate in 81.6 percent of the cases including

all trees and dropped to 71.2 percent excluding dead trees. The failure rates (i.e., predicted value fell outside actual value by > 10 percent) were 32, 20, 42, and 7 percent for condition classes 1, 2, 3, and 4, respectively. Crown dieback predictions derived from the categorical ratings were accurate in 96.8 percent of the cases including all trees and 92.8 percent excluding the dead trees. The prediction failure rate for classes 1 – 3 was negligible (< 1 percent). For condition class 4, the formula routinely overestimated dieback, resulting in a 92 percent failure rate.

DISCUSSION

LANDSCAPE LEVEL PATTERNS OF ASH HEALTH ON THE ALLEGHENY NATIONAL FOREST

We found that white ash was distributed throughout the ANF on both upper and lower slope positions. Nevertheless, we documented stark differences in crown health as measured by both a categorical condition class rating system and by the more quantitative FIA P3 measures of density, dieback, and transparency. White ash individuals located on upper slope positions throughout the ANF were generally rated more poorly, had less dense crowns, exhibited greater dieback, and appeared more transparent than similarly sized individuals at nearby lower slope positions. Although all four measures differed statistically between slope positions, we suggest that only the categorical condition rating and the FIA P3 measure of dieback differed enough (~ 1 condition class or 10 percent dieback) to operationally evaluate differences in crown health throughout the plateau.

We hypothesize that these stark differences in canopy health may be related to known site quality differences across topographic positions, and more specifically, the low concentrations of extractable pools of calcium and magnesium. Although we do not directly confirm differences in soil nutrition between our slope positions (e.g., soil analyses) several lines of evidence suggest these differences exist. First, research in the region has consistently documented that the highly weathered parent materials available to tree roots on the plateau and shoulder slopes in the unglaciated Allegheny Plateau region contain low concentrations of these ions. In contrast, lower slope positions tend to have higher concentrations due to delivery of these ions via water flow paths that percolate through deeper, mineral containing layers and enrich the soil as the water flows back out to the soil, sometimes in the form of seeps (Bailey and others 1999, 2004). Second, 50 percent of our selected lower slope plots had ≥ 1 herbaceous plant species known to be robust indicators of soils rich in base cations, and several other sites were characterized by the presence of seeps, which often transport cation supply to the soil complex (Horsley and others 2008). In contrast, only 7 percent of upper slope positions contained any of these indicator species. Finally, foliar nutrition analyses

from a subset of our focal trees found that both calcium and magnesium foliar concentrations were 39 and 29 percent greater, respectively, on trees found on lower slopes than on trees sampled on upper slopes (Royo, unpub. data).

ON THE UTILITY OF AN EASY CATEGORICAL SYSTEM

Categorical rating systems have a rich history in forest ecology and management as a method to rapidly assess tree health and position (Ball and Simmons 1980, Fajvan and Wood 1996, Gottschalk and MacFarlane 1993, Mader and Thompson 1969, Millers and others 1991). These protocols generally divide a relevant continuous variable into a few categories that are simple to recognize, biologically relevant, and easy to assess. Often, such protocols can be reliably used in the field by new users after just a few days of practice (e.g., Meadows and others 2001). Such rating systems allow forest managers to improve stand management via rapid assessments of current stand conditions. The FIA Phase 3 measures, in contrast, offer a far richer and detailed assessment of crown conditions, but are concurrently more complex involving as many as 14 categories in 5 percent classes and may require a week long training session to competently assess these in a reliable and repeatable manner. Finally, the rapid assessments possible using the categorical system may allow repeated sampling of an intensified plot network to monitor the rapid canopy decline and mortality that occur when outbreaking or invasive pests invade an area (e.g., EAB); at existing FIA implementation rates of 5 to 10 years (Bechtold and others 2008), regional or state level monitoring may miss the decline and only capture the end result.

Overall, our conversion formula proved fairly robust in correctly predicting crown density within tolerance in nearly three quarters of the cases in which it was tested. This finding is remarkable given that the conversion formula was generated on only white ash trees, 89 percent of which were in dominant or codominant crown positions, in an area without any signs of EAB. In contrast, the Ohio data contained five species of ash, 68 percent of which were in dominant or codominant crown positions, in stands spanning the full range of EAB infestation. Our dieback formula was highly reliable for classes 1 – 3, but had virtually no power in predicting crown dieback when trees were rated as a category 4. We believe the primary explanation for this inconsistency lies in the characterization of what is measured as dieback. Phase 3 dieback refers only to the severity of “recent stresses on a tree” on the upper and outer branches with fine twigs of the live crown and excludes snag branches (USDA Forest Service 2007). In contrast, our

categorical ratings were developed to assess the progressive decline and dieback of trees, and thus a categorical value of 4 includes any tree crown that contains 50 percent or greater dieback, including snag branches. Thus, trees may rate a 4 under scenarios of high recent dieback detected only in twigs or the cumulative dieback over a longer period of time that is evidenced by presence of dead twigs and snag branches. Indeed, examination of the range of dieback values of our trees in either Pennsylvania or Ohio reveals that P3 dieback assessments in our category 4 rating range from 1 to 80 percent. Additionally, our category 4 rating may encompass too great a range of conditions because any tree that is not completely dead but has >50 percent dieback is placed in category 4. Dividing this category into two distinct categories (e.g., 50 to 74 percent dieback and > 75 percent dieback to dead) may improve model reliability without complicating the rating system excessively. Indeed, other crown health and vigor rating systems often employ slightly more partitioned rating systems where dead trees are a sixth category (e.g., Fajvan and Wood 1996, Mader and Thompson 1969).

In summary, our survey of white ash throughout the ANF confirms Morin and colleague’s (2006) reports that white ash decline appears most prevalent on upper slope positions. Furthermore, we provided evidence that a simple, categorical rating system correlates well with FIA P3 crown measures of density and dieback. The advantage of this system is that it can readily and rapidly be assessed by managers and practicing foresters. With the use of our conversion formulas, these categorical ratings can estimate FIA P3 density and dieback with a fair amount of confidence. Finally, the establishment and assessment of ash monitoring plots in stands varying in ash abundance and crown health conditions may now be used as the basis of an EAB risk assessment and monitoring network.

ACKNOWLEDGMENTS

We thank Andre Hille, Lois Demarco, and Robert White of the Allegheny National Forest for assisting in project design and for graciously sharing their stand inventory data to identify potential sites. We would also like to thank the individuals who helped with field sampling including Eric Baxter, Kyle Costilow, Joshua Hanson, Lawrence Long, Marc Macdonald, Gregory Sanford, Stephanie Smith, and Ernie Wiltsie. The Forest Health Monitoring Base Evaluation Monitoring Grant Program provided funding for this project.

LITERATURE CITED

- Bailey, S.W.;** Horsley, S.B.; Long, R.P.; Hallett, R.A. 1999. Influence of geologic and pedologic factors on health of sugar maple on the Allegheny Plateau. In: Horsley, S.B., Long, R.P., editors. Sugar maple ecology and health: Proceedings of an international symposium. Radnor, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station: 63-65.
- Bailey, S.W.;** Horsley, S.B.; Long, R.P.; Hallett, R.A. 2004. Influence of edaphic factors on sugar maple nutrition and health on the Allegheny Plateau. *Soil Science Society of America Journal*. 68(1): 243-252.
- Ball, J.;** Simmons, G. 1980. The relationship between bronze birch borer and birch dieback. *Journal of Arboriculture*. 6: 309-314.
- Bechtold, W.A.;** Randolph, K.C.; Zarnoch, S.J. 2008. The power of FIA Phase 3 crown indicator variables to detect change. In: McWilliams, W., Moisen, G., Czaplowski, R., editors. 2008 Forest Inventory and Analysis (FIA) Symposium: RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Bigelow, S.W.;** Canham, C.D. 2002. Community organization of tree species along soil gradients in a north-eastern USA forest. *Journal of Ecology*. 90: 188-200.
- Cappaert, D.;** McCullough, D.G.; Poland, T.M.; Siegert, N.W. 2005. Emerald ash borer in North America: a research and regulatory challenge. *American Entomologist*. 51(3): 152-165.
- Fajvan, M.A.;** Wood, J.M. 1996. Stand structure and development after gypsy moth defoliation in the Appalachian Plateau. *Forest Ecology and Management*. 89(1-3): 79-88.
- Finzi, A.C.;** Canham, C.D.; Van Breemen, N. 1998. Canopy tree-soil interactions within temperate forests: species effects on pH and cations. *Ecological Applications*. 8(2): 447-454.
- Gottschalk, K.W.;** MacFarlane, W.R. 1993. Photographic guide to crown condition of oaks: Use for gypsy moth silvicultural treatments. Gen. Tech. Rep. NE-168. Radnor, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. 8 p.
- Hibben, R.;** Silverborg, S.B. 1978. Severity and causes of ash dieback. *Journal of Arboriculture*. 4: 274-279.
- Horsley, S.B.;** Bailey, S.W.; Ristau, T.E.; Long, R.P. [and others]. 2008. Linking environmental gradients, species composition, and vegetation indicators of sugar maple health in the northeastern United States. *Canadian Journal of Forest Research*. 38(7): 1761-1774.
- Horsley, S.B.;** Long, R.P.; Bailey, S.W.; Hallett, R.A. [and others]. 2000. Factors associated with the decline disease of sugar maple on the Allegheny Plateau. *Canadian Journal of Forest Research*. 30(9): 1365-1378.
- Long, R.P.;** Horsley, S.B.; Hallett, R.A.; Bailey, S.W. 2009. Sugar maple growth in relation to nutrition and stress in the northeastern United States. *Ecological Applications*. 19(6): 1454-1466.
- Mader, D.L.;** Thompson, B.W. 1969. Foliar and soil nutrients in relation to sugar maple decline. *Soil Science Society of America Proceedings*. 33(5): 794-800.
- McCullough, D.G.;** Poland, T.M.; Cappaert, D. 2009. Attraction of the emerald ash borer to ash trees stressed by girdling, herbicide treatment, or wounding. *Canadian Journal of Forest Research*. 39(7): 1331-1345.
- Meadows, J.S.;** Burkhardt, E.C.; Johnson, R.L.; Hodges, J.D. 2001. A numerical rating system for crown classes of southern hardwoods. *Southern Journal of Applied Forestry*. 25(4): 154-158.
- Millers, I.;** Lachance, D.; Burkman, W.G.; Allen, D.C. 1991. North American sugar maple decline project: Organization and field methods. Gen. Tech. Rep. NE-154. Radnor, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 26 p.
- Morin, R.S.;** Liebhold, A.M.; Gottschalk, K.W.; Woodall, C.W. [and others]. 2006. Analysis of forest health monitoring surveys on the Allegheny National Forest (1998-2001). Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station, Northeastern Area State and Private Forestry and Allegheny National Forest. 102 p.
- SAS Institute Inc.** 2005. SAS System for Windows. Version 9.1. Cary, NC: SAS Institute Inc.
- Sinclair, W.A.;** Iuloi, R.J.; Dyer, A.T.; Marshall, P.T. [and others]. 1990. Ash yellows: geographic range and association with decline of white ash. *Plant Disease*. 74(8): 604-607.
- Smith, A.** 2006. Effects of community structure on forest susceptibility and response to the Emerald Ash Borer invasion of the Huron River watershed in southeast Michigan. Columbus, OH: The Ohio State University. 122 p. M.S. Thesis.
- USDA Forest Service.** 2007. Forest Inventory and Analysis Phase 3 Crowns: Measurement and Sampling, Version 4.0. Washington, DC: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis. 22 p.
- USDA Forest Service.** 2010. Forest Inventory Data Online [FIDO] Version 2.0. Washington, DC: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis. <http://fiatools.fs.fed.us/fido/index.html>. [Date Accessed: October 22, 2010].
- van Breemen, N.;** Finzi, A.C.; Canham, C.D. 1997. Canopy tree - soil interactions within temperate forests: Effects of soil elemental composition and texture on species distributions. *Canadian Journal of Forest Research*. 27: 1110-1116.
- Woodcock, H.;** Patterson III, W.A.; Davies, K.M. 1993. The relationship between site factors and white ash (*Fraxinus americana* L.) decline in Massachusetts. *Forest Ecology and Management*. 60(3-4): 271-290.

Table 1—Spearman rank correlation matrix among FIA Phase 3 crown condition variables of uncompact crown live ratio (UCLR), crown density, dieback, and transparency, and a user friendly categorical rating of health conditions where 1=healthy, 5=dead, and 2 through 4 represent increasing stages of crown decline and dieback

	Condition	Density	Dieback	Transparency
Condition	1			
UCLR	-0.223 <0.0001			
Density	-0.723 <0.0001	1		
Dieback	0.719 <0.0001	-0.551 <0.0001	1	
Transparency	0.562 <0.0001	-0.539 <0.0001	0.434 <0.0001	1

For each cell, the upper value is the Spearman correlation coefficient (ρ) and the lower value is the significance (P value) of relationship.

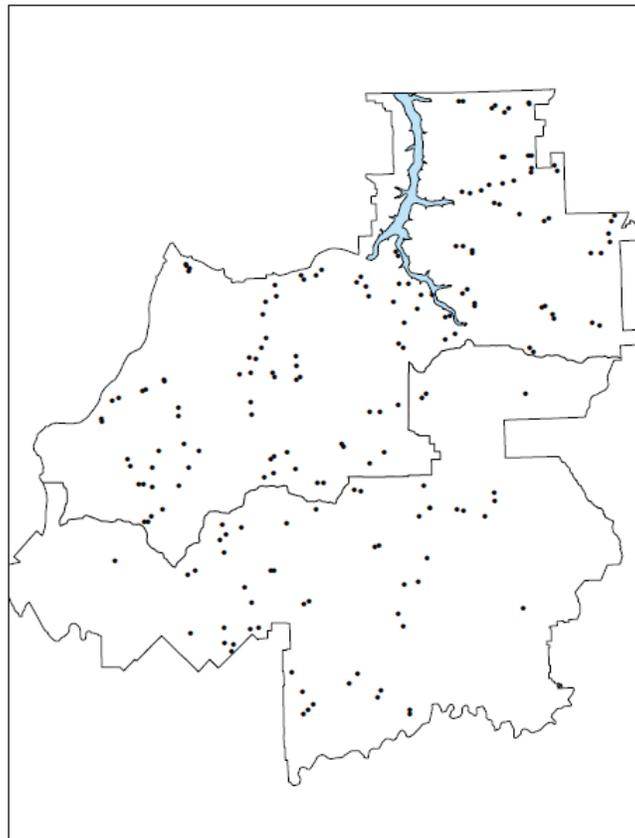


Figure 1—Approximate locations of the expanded network of ash health monitoring plots throughout the Allegheny National Forest proclamation boundary.

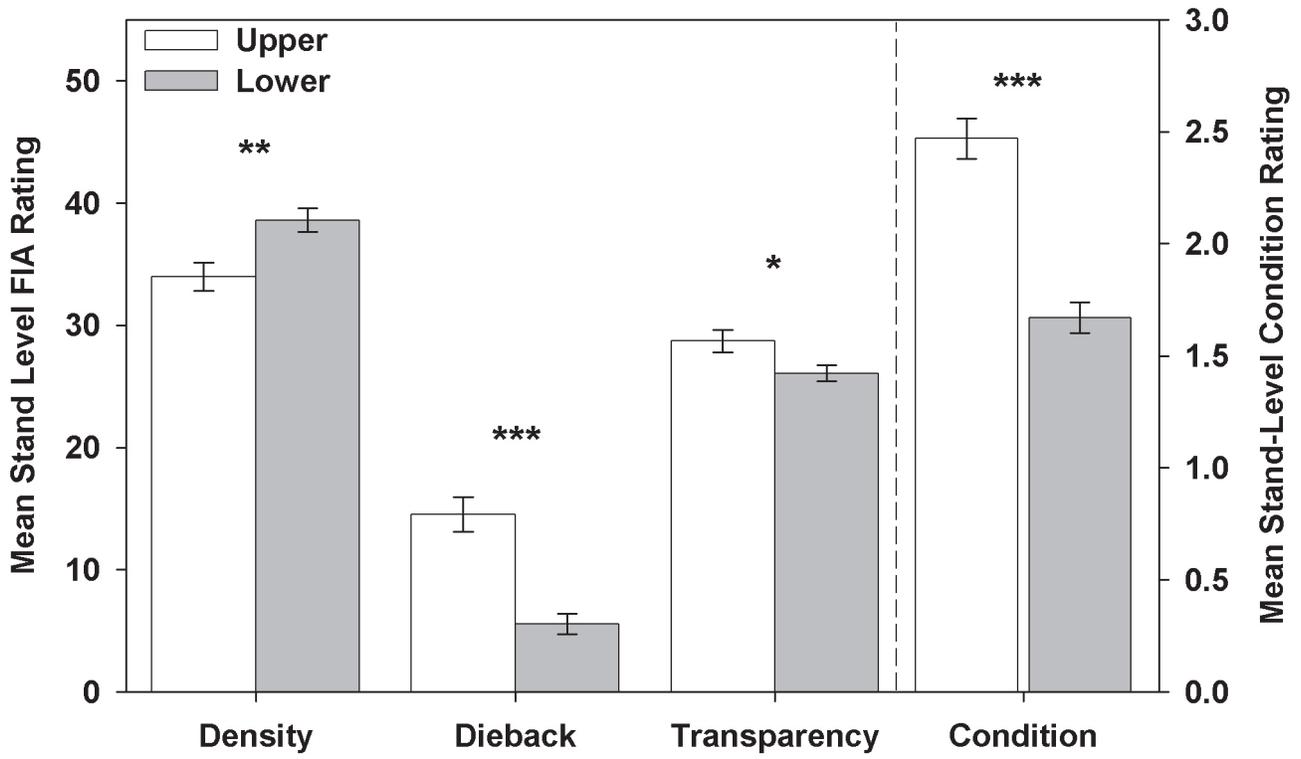


Figure 2—Mean crown health values (± 1 SE) of FIA Phase 3 measures of density, dieback, and transparency as well as a categorical rating (1=healthy, 5=dead, 2-4 increasing dieback) of white ash trees on stands on upper and lower slope positions. Asterisks denote significant differences from analysis of variance (* = < 0.05, ** = < 0.01, *** = < 0.001).

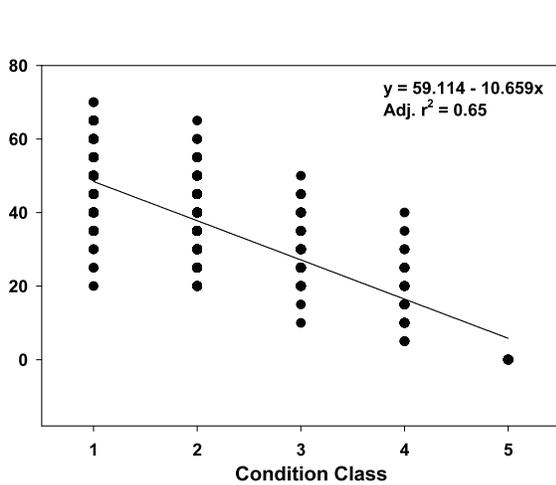


Figure 3—Linear regression between categorical crown rating system and FIA Phase 3 measure of crown density.

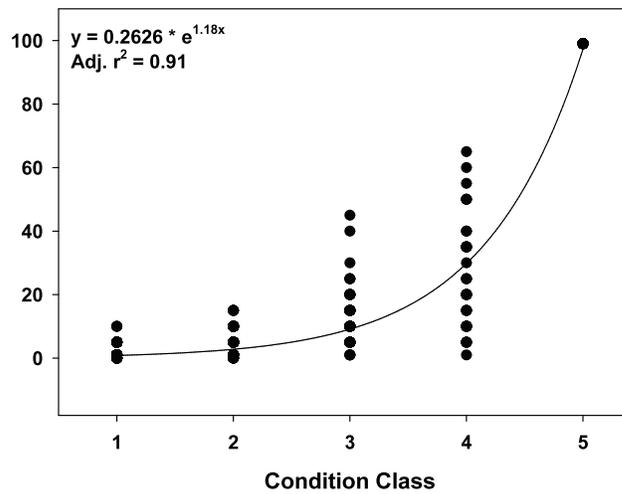


Figure 4—Exponential regression between categorical crown rating system and FIA Phase 3 measure of crown dieback

ESTIMATION OF INVASIVE PROBABILITY OF MULTIFLORA ROSA IN THE UPPER MIDWEST

Weiming Yu, Zhaofei Fan, W. K. Moser, M. H. Hansen and M. D. Nelson

ABSTRACT

Multiflora rose (*Rosa Multiflora* Thunb.) (MFR) is widely spreading across the United States, with up to 38 states in the contiguous United States reporting the presence of this species. In this study, U.S. Forest Service, Northern Research Station Forest Inventory and Analysis (FIA) data from the Upper Midwest states for the period of 2005-2006 were used to calculate MFR presence probability at the county level. The classification and regression tree technique is used to classify the counties into invasive stages based on the estimated presence probability and a map of invasive stages is obtained, which is helpful for forest managers to optimally allocate resources. A simultaneous autoregressive model (SAR) was used to identify the driving factors of the spread of MFR. The contiguous invasive stages indicated a strong invasive pattern in all directions, particularly southward and eastward. MFR presence shows a positive spatial autocorrelation and is negatively associated with latitude and county forest cover percentage. Our results suggest that the distribution of MFR is likely limited by its intolerance to extreme cold temperatures and anthropogenic disturbance (forest fragmentation and deforestation) plays an important role in the spread of MFR.

INTRODUCTION

Non-native invasive plants (NNIPs) are defined as those plants that: 1) are not indigenous to the ecosystem and 2) have a competitive advantage that causes deleterious impacts on structure, composition, and growth in forested ecosystems (Moser et al., 2009). Due to their competitive advantages compared with the native plants, some NNIPs can expel the native plants, alter the local ecosystem, threaten native biodiversity, and lower value of local ecosystem (Macdonald 1994). In the USA, the estimated loss due to NNIPs is more than \$33 billion per year (Pimentel et al., 2005).

The invasive process is classified into four stages: introduction, colonization, establishment and spread (Theoharides and Dukes 2007). The factors that affect the spread of NNIP include temperature, site quality, stand size, forest fragmentation, distance to road, county percent forest etc (Moser et al 2009). A factor may play a role in one or more stages. The management of NNIPs depends on the stages of invasion since the cost dramatically increases as

the population of NNIPs expands (Hobbs and Humphries 1995). It is prohibitively expensive to remove most of the NNIPs from the invaded region after they are well established. Webster et al. (2006) pointed out that the efforts to control invasive plants should focus on the establishment or earlier phases. Thus, the early detection of the NNIPs is very important.

In this study, we focus on multiflora rose (*Rosa multiflora* Thunb.) (MFR), which is widely established across United States. In fact, 38 states in the contiguous United States report the existence of MFR. MFR can exclude native ground flora and suppress tree regeneration and has been designated as a noxious weed in many states (Munger 2002, Denight et al., 2008). MFR was introduced to the North America in 1866 as rootstock for ornamental roses. Later, it was used for erosion control and “living fences” to confine livestock from the 1930s to the 1950s (Doll 2006). MFR is extremely prolific and its seeds are dispersed by birds and may remain viable in the soil for many years.

In this study, based on the estimated presence probability at the county level, we investigated the distribution of MFR, classified each county into four invasive stages using Classification and Regression Tree (CART) technique and identified the driving factors that affect MFR establishment and spread.

DATA AND METHODOLOGY

STUDY AREA

The Upper Midwest study region is comprised of seven states: Indiana, Illinois, Iowa, Missouri, Michigan, Wisconsin, and Minnesota. At the nexus of several ecoregions, this area is characterized by diverse vegetation compositions and structures. Northern Minnesota, and Wisconsin, northern Michigan, and southern Missouri are the most heavily forested areas. The middle of this region, most of it prairie during pre-European settlement times, is currently a mosaic of agricultural lands, with embedded

urban areas. Extensive human activities and fertile soil in this region can favor the establishment of NNIPs.

DATA

In this study, U.S. Forest Service, Northern Research Station Forest Inventory and Analysis (FIA) data from the Upper Midwest states for the period 2005-2006 were used. Phase 2 (tree inventory) data were collected on the standard FIA plot grid (1 plot per 2400 ha). Each Phase 2 plot consists of 4 subplots with radius 7.3 m. Associated with this overstory inventory data was sampling for 25 invasive plants species of interest. These 25 species were categorized as either grasses, vines, herbaceous, or woody species, such as multiflora rose. In total, 8663 phase 2 forested plots were assessed for MFR, where 1320 plots (15.3 percent), were invaded by MFR. Related factors including county forest percentage, inter-state highway density, forest type, and fragmentation were measured or calculated by using auxiliary GIS layers. The coverage of MFR is classified into 7 categories (Moser et al. 2009) and the midpoint values of each cover class are used.

METHODOLOGY

THE PRESENCE PROBABILITY OF MFR—We are interested in the presence probability of MFR at the county level. Therefore, the plot data were aggregated to the county level in order to obtain the number of the MFR-presence and total plots at each of the 649 counties. Mathematically, the presence probability is estimated by the ratio between the number of the MFR-presence and the total plots in each county. However, this estimation is severely biased due to the sample size of each county. For example, some counties only have a couple of plots and others may have more than 100 plots. In order to correct the bias due to the sample size in each county, we redefined the presence probability:

$$P_i = \frac{\sum_{j \in \eta_i} s_j}{\sum_{j \in \eta_i} n_j}$$

where: s_j is the number of the MFR-presence plots in the county j , n_j is the total plots in the county j , η_i is the set of neighbor for the county i , including the county i .

CLASSIFICATION AND REGRESSION TREE (CART)—Then we used the CART (Breiman et al. 1984) to classify each county, based on the estimated presence probability, into different stages: introduction, colonization,

establishment, and spread (Theoharides and Dukes 2007). The `rpart` package in R are used to implement CART (John, M. and John, B. 2003). Based on the classification of the estimated presence probability, we plot the map of invasive stages for the studied area.

MORAN'S I, AND GEARY'S C (WALLER AND GOTWAY 2004)—To investigate the spatial correlation of the presence probability of MFR, Moran's I and Geary's C were calculated via,

$$Moran's\ I = \frac{N * \sum_{i=1}^N \sum_{j=1}^N w_{ij} (Y_i - \bar{Y})(Y_j - \bar{Y})}{\sum_{i=1}^N \sum_{j=1}^N w_{ij} * \sum_{i=1}^N (Y_i - \bar{Y})^2}$$

and

$$Geary's\ C = \frac{(N - 1) * \sum_{i=1}^N \sum_{j=1}^N w_{ij} (Y_i - Y_j)^2}{2 \sum_{i=1}^N (Y_i - \bar{Y})^2 * \sum_{i=1}^N \sum_{j=1}^N w_{ij}}$$

where N is the number of counties; Y is the presence probability at each county; w_{ij} is a matrix of spatial weights.

The range of Moran's I is (-1, 1). The negative (positive) value of Moran's I indicates negative (positive) spatial autocorrelation. The value of Geary's C lies between 0 and 2. If Geary's C is less than 1, then it indicates positive spatial autocorrelation; otherwise, it indicates a negative spatial autocorrelation. Geary's C is inversely related to Moran's I and is more sensitive to local spatial autocorrelation.

SIMULTANEOUS AUTOREGRESSIVE MODEL (SAR) (WALLER AND GOTWAY 2004)—We used the simultaneous autoregressive model (SAR) to identify the driving factors of the spreading of MFR with the estimated presence probability (PP) as the dependent variable. The independent variables are: *Longitude (Lon)*, *Latitude (Lat)*, *forest fragmentation (Frag)*, *ecoregion (Eco)*, *road density (Rdens)*, and *county-level forest cover percentage (CFP)*. Though we had tested other driving factors such as forest type, but they were not significant and were not included in the model.

The SAR model is expressed as:

$$PP = \beta_0 + \beta_1 * Lon + \beta_2 * Lat + \beta_3 * Frag + \beta_4 * Eco + \beta_5 * Rdens + \beta_6 * CFP + \epsilon$$

$$\epsilon = \rho W \epsilon + v$$

where: β_i s are the parameters that need to be estimated, v is the independent error vector, ρ is the simultaneous autoregressive error coefficient, W is the spatial weight matrix.

RESULTS AND DISCUSSION

MFR is established in Missouri, Illinois, Indiana, Iowa, and the south of Wisconsin and Michigan and prevalent in northern Missouri, southeast Iowa, central and northern Illinois and Indiana. The forests of the northern of Minnesota, Wisconsin, and Michigan are not severely infected by MFR (Fig. 1). From this central core, MFR is spreading northward into the forested lands in southern Wisconsin and Michigan, and southward into the Ozark Highlands in southeast Missouri (Fig 3). There is a positive spatial autocorrelation among the MFR presence (Moral's $I = 0.93$, Geary's $C = 0.08$). The contiguous introduction, establishment and spread stages indicated a strong invasive pattern in all directions, but particularly southward and eastward.

The plot of X relative error vs. the tree size (Figure 2, left) suggests that tree size should be 4. And Figure 2 (right) suggested that each county should be classified into one of the four invasive stages: Introduction, <29.55%; Colonization, 29.55%-51.05%; Establishment, 51.05%-66.65%; and Spread, >66.65%. For each stage, we calculate the proportion of the plots for each cover class for each stage. Then the plot of proportion of plots of each class vs. the midpoint value of cover class (Figure 3, left) shows that this classification is reasonable. Finally, the map of invasive stages (Figure 3, right) is obtained based on the above classification.

SAR (Table 1) shows that county forest cover percentage, latitude and proportion of Lower Mississippi Riverine Forest are significantly negatively associated with the presence probability of MFR, while longitude and proportion of Eastern Broadleaf Forest and Prairie parkland are positively associated with the presence probability of MFR. Forest fragmentation and inter-state road density are not significantly related to the presence probability of MFR.

Denight et al (2008) observed that the northern distribution of MFR is likely limited by its intolerance to extreme cold temperatures as suggested by the negative association with latitude—a surrogate for the winter temperature gradient. The negative association between county forest cover percentage and MFR presence indicated the significance of human disturbances in the spread of MFR. Actually, MFR was widely planted for erosion control in many less forested counties in northern Illinois, southeastern Iowa and northern Missouri between 1930s and 1960s (Doll 2006). As reported, MFR frequently colonizes roadsides, old fields, pastures, prairies, open woodlands, and forest edges but is not found in standing water or in extremely dry areas (Doll 2006), which conforms to the findings of its distribution in different ecoregions/forest types in our study. The positive relationship with longitude also reflects a strong eastward

spreading pattern. Today, the species has become widely distributed in the eastern United States (Doll 2006). The non-significant relationship with forest fragmentation and road density seems to be a “scale” effect, for at the plot level; forest fragmentation and distance to roads have been suggested to be important influences of the establishment and spread of MFR.

SUMMARY

The strong spatial autocorrelation of MFR presence and its aggressive spreading southward and eastward reflects the fact that human disturbances and climatic factors, as well as forest conditions play an important role. Even though MFR's spread to the northern states seems to be curbed currently due to cold temperatures, climate change (warming) and the increasing number of anthropogenically-derived deforestation disturbances such as land clearing and urban development may increase the probability of infestation in the forests of the northern part of our study region. To prevent MFR from spreading into the highly forested areas in the northern states, a practical strategy is to eliminate MFR in the counties in the introduction stage and reduce MFR presence in the counties in the establishment stage. And the map of invasive stages can help forest managers to identify the counties in the introduction or establishment stage. Thus it provides a tool for forest manager to optimally allocate resources.

LITERATURE CITED

- Breiman, L.;** Friedman, J.H.; Olshen, R.A. [and others]. 1984. Classification and regression trees. Wadsworth and Brooks, Monterey, Calif.
- Denight, M.L.;** Guertin, P.J.; Gebhart, D.L. [and others]. 2008. Invasive Species Biology, Control, and Research, Part 2: Multiflora Rose (*Rosa Multiflora*). http://www.cecer.army.mil/techreports/ERDC_TR-08-11/ERDC_TR-08-11.pdf [Date Accessed: September 29, 2010]
- Doll, J.D.** 2006. Biology of Multiflora Rose. North Central Weed Science Society Proceedings 61:239. <http://ncwss.org/proceed/2006/abstracts/239.pdf> [Date Accessed: September 29, 2010]
- Hobbs, R.J.;** Humphries, S.E. 1995. An integrated approach to the ecology and management of plant invasions. *Conservation Biology*. 9(4):761–770
- John, M.;** John, B. 2003. Data Analysis and Graphics Using R: An Example-based Approach. Cambridge University Press.
- Macdonald, I.A.W.** 1994. Global Change and Alien Invasions: Implications for Biodiversity and Protected Area Management, In: Biodiversity and global change, by O. T.Solbrig, P. G. van Emden, and W. J. van Oordt. Wallingford-Oxon, UK: CAB International

Moser, W. K.; Hansen, M. H.; Nelson, M.D. [and others]. 2009. Relationship of invasive groundcover plant presence to evidence of disturbance in the forests of the upper Midwest of the United States. *Invasive Plants and forest ecosystems*. 29-58. CRC Press.

Munger, G. T. 2002. *Rosa multiflora*. In: Fire Effects Information System. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory. <http://www.fs.fed.us/database/feis/> [Date Accessed: September 29, 2010]

Pimentel, D.; Zuniga, R.; Morrison, D. 2005. Update on the environmental and economic costs Associated with alien-invasive species in the United States. *Ecological Economics*. 52(3): 273–288

Theoharides, K.A.; and Dukes, J.S. 2007. Plant invasion across space and time: factors affecting nonindigenous species success during four stages of invasion. *New Phytologist*.176(2): 256 – 273

Waller, L. A.; Gotway, C. A. 2004. *Applied Spatial Statistics for Public Health Data*. Published by John Wiley & Sons, Inc.

Webster, C.R.; Jenkins, M.A.; Jose, S. 2006. Woody invaders and the challenges they pose to forest ecosystems in the eastern United States. *Journal of Forestry*. 104(7): 366-374

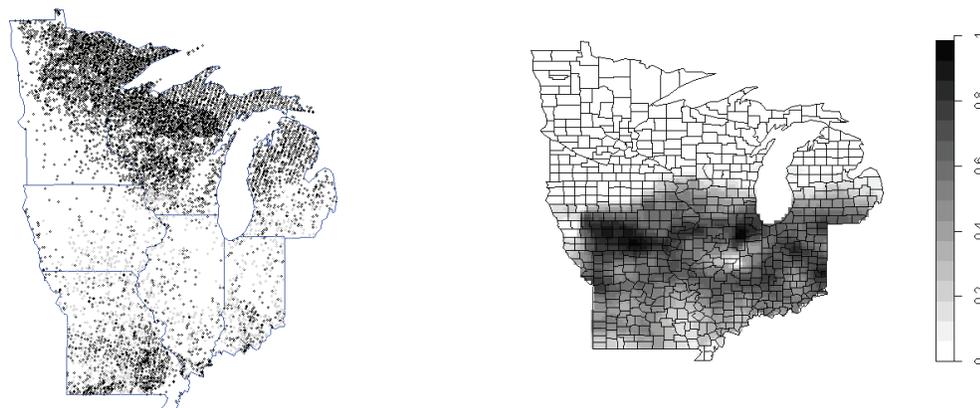


Figure 1—(Left) The distribution of FIA Phase 2 Plots without (black) and with the presence of Multiflora Rose (gray); and (right) the smoothed presence probability of MFR in the Upper Midwest, 2005–2006.

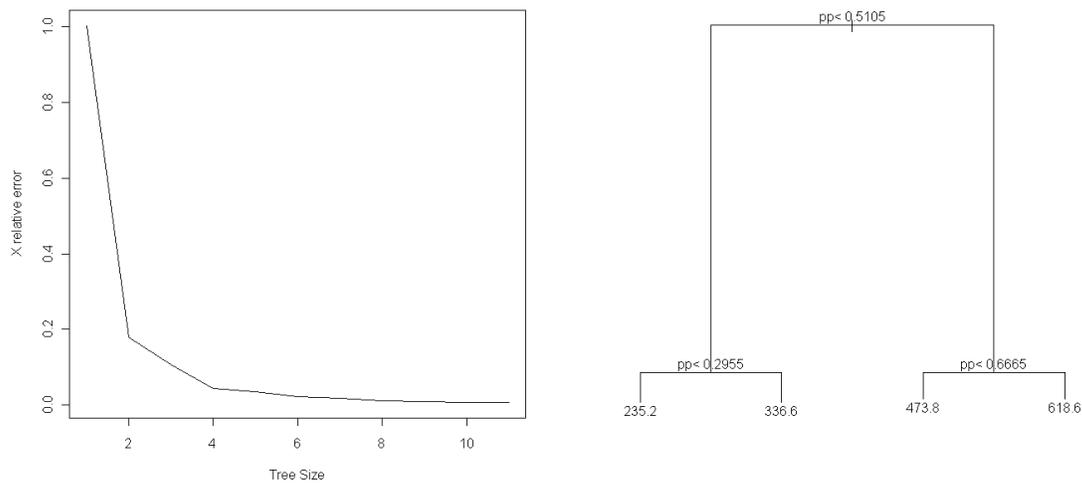


Figure 2—(Left) The plot of X relative error vs. the classification tree size; and (right) the Classification and regression tree partition of the estimated presence probability of MFR in the Upper Midwest, 2005–2006.

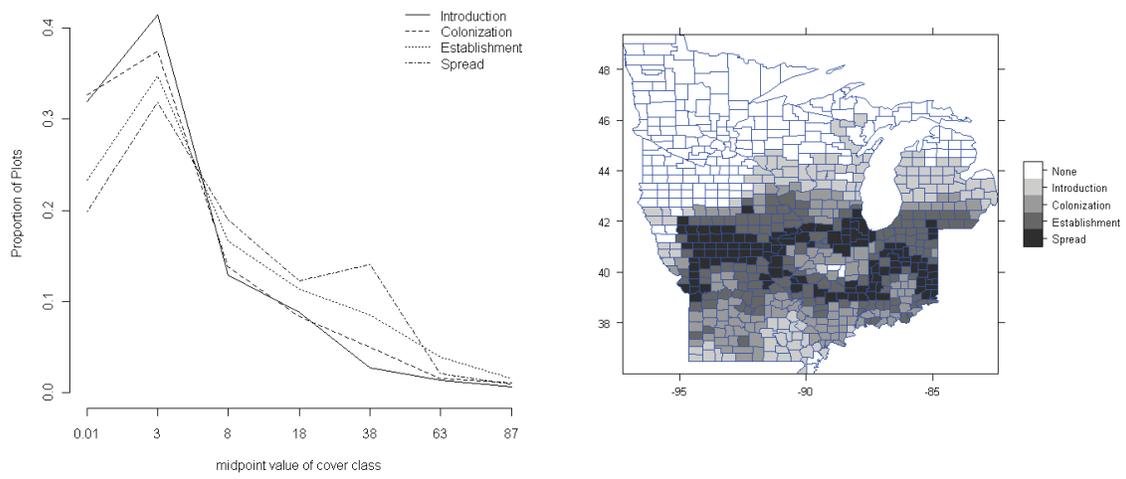


Figure 3—(Left) Proportion of plots vs. the midpoint value of cover class (see Moser and others, 2009 for the cover class categories); and (right) the maps of invasive stages based on the estimated presence probability of MFR in the Upper Midwest, 2005–2006.

Data Integrity

A NATIONAL ANALYTICAL QUALITY ASSURANCE PROGRAM: DEVELOPING GUIDELINES AND ANALYTICAL TOOLS FOR THE FOREST INVENTORY AND ANALYSIS PROGRAM Phyllis C. Adams and Glenn A. Christensen.....	139
RESOLVING THE PULPWOOD CANVASS WITH INVENTORY HARVEST INFORMATION Joseph M. McCollum and Tony G. Johnson.....	147
ALGORITHMIC DECISION RULES FOR ESTIMATING GROWTH, REMOVALS, AND MORTALITY WITHIN A NATIONAL-SCALE FOREST INVENTORY (USA) William H. McWilliams, Carol L. Alerich, William A. Bechtold, Mark Hansen, Christopher M. Oswalt, Mike Thompson, and Jeff Turner.....	155
APPLICATION OF AN ASSESSMENT PROTOCOL TO EXTENSIVE SPECIES AND TOTAL BASAL AREA PER ACRE DATASETS FOR THE EASTERN COTERMINOUS UNITED STATES Rachel Riemann, Ty Wilson, and Andrew Lister.....	161
A DATABASE STRATEGY FOR NEW VARIABLES B. Tyler Wilson, Ali Conner, Glenn Christensen, John Shaw, Jason Meade, and Larry Royer.....	175

A NATIONAL ANALYTICAL QUALITY ASSURANCE PROGRAM: DEVELOPING GUIDELINES AND ANALYTICAL TOOLS FOR THE FOREST INVENTORY AND ANALYSIS PROGRAM

Phyllis C. Adams and Glenn A. Christensen

ABSTRACT

A rigorous quality assurance (QA) process assures that the data and information provided by the Forest Inventory and Analysis (FIA) program meet the highest possible standards of precision, completeness, representativeness, comparability, and accuracy. FIA relies on its analysts to check the final data quality prior to release of a State's data to the national FIA database (FIADB). The analytical portion of the QA process varies considerably from region to region, aided by a variety of tools developed within each region. Recognizing that a QA process consisting of national standards would result in greater consistency and transparency, the national analytical QA task team is developing guidelines for analysts and oversees the development of QA Tools, a desktop charting and graphing application for checking compiled data prior to release. QA Tools includes a rigorous estimation engine based on national standard FIA estimation procedures. Batch processing is utilized to save time and reduce the cost of extensive QA. National consistency results in fewer resources allocated to maintaining multiple processes, thus allowing analysts to spend more time on error resolution.

INTRODUCTION

The U.S. Department of Agriculture Forest Service, Forest Inventory and Analysis (FIA) program is charged with assessing the status and trends in the Nation's forests. The FIA quality assurance (QA) program assures that the inventory produces complete, accurate, and unbiased forest information of known quality. FIA analysts play a critical role in the QA process, with primary responsibility for the quality of data released for publication. This paper outlines the flow of FIA data from the initial field planning stage through the collection, editing, storage, and compilation of data to analytical QA and publication. We highlight the role of QA within each step, with emphasis on the analytical QA stages in which analysts review the data prior to publication. Finally, we introduce QA Tools, a recently developed QA analysis tool.

Quality assurance, the overall system of management activities designed to assure that quality data are generated,

is composed of two components: quality control and quality assessment. Quality control includes the operational techniques used to reduce random and systematic errors, such as identifying and adopting standards for producing quality products, training, data collection field checks, developing efficient data flow procedures, data error checking, and assuring consistency through well documented procedures guides. Quality assessment evaluates how well our established standards are met by the program, telling us whether the FIA quality control system is satisfactory. The assessment procedure compares field production data with an independent "blind" measurement to evaluate the relative uncertainty associated with FIA field collected data.

FIA DATA FLOW AND QUALITY ASSURANCE

The FIA QA program includes comprehensive quality control and quality assessment steps from prefield preparation and field data acquisition to processing within the National Information Management System (NIMS) and post compilation analytical QA procedures that prepare the data for posting in public databases. Prior to presenting details about analytical QA, the primary focus of this paper, we describe the overall FIA data flow process. The flow of FIA data and the integral QA components are described through seven stages: 1) Prefield, 2) Data collection, 3) Precompilation data editing, 4) Stratification, 5) Data compilation in NIMS, 6) Analytical QA, and 7) Data delivery (fig. 1). As the FIA program evolved into a national program with the advent of the annual inventory design, procedures became more consistent. For example, a nationally consistent data collection protocol was defined and documented in national FIA field manuals and NIMS was implemented to store and process FIA data. National

consistency improves operating efficiency and the ability to deliver higher quality data. Furthermore, inconsistent protocols and QA procedures make assessments of the accuracy and repeatability of each regional program difficult to compare. Process documentation that increases consistency across FIA regions is described for each stage and diagrammed in figure 1.

PREFIELD PLANNING

The FIA data flow process begins with the prefield stage. This stage includes activities that occur in advance of field visits to FIA plot locations, such as land-use interpretation based on high-resolution imagery, assembling hard copy materials in support of the field staff, and verifying plot coordinates. Prefield programs have developed independently at each of the four FIA regions, leading to parallel, yet different, procedures, QA processes, tools, and database items. For example, multiple tools have been created to solve similar problems, and overlapping software programs have been developed. FIA land use definitions are not necessarily applied consistently across regions. Regional QA protocols generally include prefield determination of plot visit/nonvisit quality assessment. National prefield protocols and QA procedures are under development and will be documented in a prefield manual, adding consistency and transparency to the process.

DATA COLLECTION

Data are collected on FIA plots using national protocols. Field procedures and protocols are documented in the national FIA Field Manual. QA procedures in this stage have been formalized and are well documented in national and regional FIA QA plans. The QA effort is primarily focused on error control during the field measurement and data collection process. This is accomplished through extensive crew training, documentation of protocols and procedures used in the inventory, and field checks.

PRECOMPILED DATA EDITING

Raw field collected data undergo three different checks prior to database compilation: personal data recorder (PDR) logic checks, postfield data cleaning edits by field or office staff, and NIMS load error checks. Field data are collected using field data recorders that are programmed with multiple logic checks applied during data entry in the field. In 2007, a national task team was formed to develop the FIA Mobile Integrated Data Acquisition System (MIDAS), a data entry program that is used by all regions to collect both national and regional data. The MIDAS system is being integrated into the overall FIA information management structure. Postfield procedures that include various algorithms are used for checking and editing raw data. These procedures vary from region to region. Finally, prior to populating the NIMS database, additional NIMS load checks are applied to the data by the NIMS system.

STRATIFICATION

Stratification is applied to the inventory data prior to compilation to reduce the variance of estimates associated with forest land area and volume. Field data are grouped into broad strata using the stratification approach outlined in Bechtold and Patterson (2005). The number and type of strata used in the stratification process vary with regional requirements. Strata may include forest/nonforest, public/private ownership, or boundaries such as state, county, and ecological unit. High-resolution imagery and ownership maps are used to classify sampled plots into each stratum. Quality control and assessment of the stratification process is completed regionally, including checking plot locations, matching total area estimates to those provided by the Census Bureau, and double-checking boundary locations on GIS layers.

NIMS DATABASE COMPILATION

NIMS compilation includes calculating a variety of derived variables using estimation equations defined by Bechtold and Patterson (2005). NIMS design and implementation procedures as well as the compilation process are described in various unpublished documents produced by national implementation teams. Data checks of tables are performed informally on the database prior to analyst review. Data checks include an evolving list of data queries to identify errors. These include checks for missing values and checks for consistency between populated values of related variables, such as site index, stand age, and stand size class.

ANALYTICAL QA

After the data are compiled and summarized, FIA resource analysts conduct regional postcompilation checks and audits to ensure that core FIA data are complete and as accurate as possible before publication in FIADB (Miles and others 2001). There is presently no formal review process for all FIA analysts to follow, although each region's QA process generally includes a variety of standard checks and additional analyses that may be needed for unique resource issues. Documentation to define a rigorous nationally consistent process is being drafted. Additionally, tools have been and/or are being developed within each region to help analysts with QA.

DATA DELIVERY

Analyses and reports about status, trends, and location by land use class are generated as annual and 5-year reports and special reports. Data are posted to FIADB, FIA's public Web based database, with its suite of reporting tools. National FIADB tables are populated with regional data and displayed in standardized output tables. Detailed documentation is available in a Database Description and Users Manual (Miles 2008). Additional reporting includes peer-reviewed journal articles, articles in trade journals, extension publications, and other user-oriented outlets.

ANALYTICAL QUALITY ASSURANCE

The primary task of FIA analysts is to interpret and extract meaning from compiled inventory data and to communicate these findings to the public. Each State is assigned to an analyst who checks the State's compiled data prior to release to the public as FIADB tables, as annual and 5-year reports, or as peer-reviewed publications. Analysts must be experts on forestry and ecological issues in their assigned States and regions. They conduct research that adds value to the basic inventory products, and they participate on national FIA task teams.

FIA analysts play a critical role in the QA process; they are responsible for checking compiled data to assure the quality of data released to the public and the accuracy FIADB and FIA's annual and 5-year State reports. This is not a trivial task; in spite of extensive precompilation error checking; some errors are impossible to detect until analysts have attempted to interpret the numbers. After analysts conduct annual QA checks of information, they submit core tables to State contacts for annual review.

This analytical QA process varies considerably from region to region. Typically, FIA analysts use many different approaches, ranging from reviewing summary tables for suspicious values and checking map products carefully to a variety of in depth investigation methods. As the FIA program has expanded and moved to a fully annualized inventory, the need for a more thorough and efficient approach to data QA has increased dramatically. Thus, the national FIA analytical QA task team has proposed a nationally consistent process with standard minimum requirements for checking data prior to release for publication. The team's objective is to increase efficiency with well documented quality control standards and a consistent analytical QA process. Guidelines and checklists for analysts will outline routine QA steps required prior to releasing data to the public.

Analysts' toolboxes contain a variety of tools for data checking, including algorithms developed within each region. For example, FIA analysts may use a statistical application such as SAS (2002) to perform specific data queries and summaries. The same analyst may then use another application such as Microsoft (MS) Access that performs different summaries with the same data. Experienced analysts know what sorts of QA checks to apply to these data, and they generally have extensive knowledge of field collection methods to detect possible errors. But, what happens if the analyst is inexperienced with FIA protocols and the necessary understanding of how to query and summarize these data is not fully understood? Data errors may be missed and summary estimates may be incorrect. Each analyst within the four FIA regions of

the United States shares these analytical QA challenges. Similar analytical QA checks are repeated each year on that year's data. In this manner, analyst amasses a collection of queries that can summarize estimates and provide sampling errors from tables within NIMS or FIADB. They must evaluate both measured data collected in the field and compiled estimates (such as volume and biomass estimates). Attempting to streamline this process, the analytical QA team has overseen the development of QA Tools, a charting and graphing application for checking compiled data prior to release. QA Tools is a software application that addresses these issues and moves the FIA program toward a more thoughtful and comprehensive approach to analytical QA.

QA TOOLS: TOWARD A COMPREHENSIVE APPROACH TO ANALYTICAL QA

QA Tools is a desktop application developed as a joint effort by FIA analysts, programmers, and statisticians working toward a single integrated solution that provides a rigorous estimation engine based on national FIA statistical estimation procedures as described in Bechtold and Patterson (2005). The application allows an analyst to methodically, consistently, and efficiently perform QA checks and document the QA checks that were performed. The application can be quickly modified to reflect any changes in data storage, collection, and processing procedures within FIA. Batch mode processing improves efficiency by producing a predefined set of output reports, such as tables of estimates and graphics, during a single session. In setting up a specific QA query, the user can customize any input or output tables and charts. FIA data are brought into the tool directly from NIMS Oracle tables, as FIADB formatted tables, or as custom MS Access tables. User 'wizards' are utilized for quick access and ease of use in defining specific queries and outputs (fig. 2). The current version of the application works with FIADB 4.0 (Woudenberg and others 2010) structure, however it is not currently designed for general public use, but rather accesses prereleased data to perform analytical QA procedures. Any FIA analyst can load standard (or customized) tables and perform the same set of QA checks on data from any State in the country.

ESTIMATION ENGINE

A key strength of QA Tools is the built-in estimation engine that uses estimation procedures as described by Bechtold and Patterson (2005). Prior to QA tools, an analyst who wanted to calculate summary estimates not included as part of standard output tables had to understand the NIMS table structure and estimation procedures well enough to make this calculation. This is not an easy task for new analysts. In addition, a further level of statistical understanding is required to produce error estimates. Each analyst develops their own method to perform these calculations, all with

varying degrees of complexity and ease of customizing for specific inquiries. QA Tools provides the user with easy access to a complex database and a simplified approach for obtaining estimates with sampling errors. Interface ‘wizards’ guide the user through a choice of either basic estimation or more advanced reporting. Advanced estimation calculations include adding custom tables, user defined variables, and custom sub-groupings for reporting. With the estimation engine in QA Tools an analyst can also more easily evaluate FIA data as a ratio of means (e.g., estimates on a per acre basis).

BATCH PROCESSING

In addition to having the built-in estimation engine, another key strength of QA Tools is the ability for batch processing. The batch process approach allows each analyst to quickly produce tables and charts for error checking and QA documentation. QA Tools includes a ‘wizard’ to define and simplify a batch process session. Output tables and graphics are defined in report definition files (RDF) that can be saved for subsequent batch processing. These RDF files are loaded into a batch session and processed at one time. Each RDF file is simply a set of Extensible Markup Language (XML) code that tells QA Tools which input data to use, which statistical estimates or summaries are required, any filters applied to the data, format of the output table, and how to build each output graphic requested. The user can select an output table and/or graph for each report (fig. 3), and can designate where all input and output files are stored. Tables are output as MS Excel files and graphics are saved as either Joint Photographic Experts Group (JPEG) or Portable Document Format (PDF) files. Current and future releases of QA Tools include an initial set of RDF files that produce some of the nationally defined FIA tables.

With batch processing in QA Tools, analysts can now easily perform many QA checks by producing output graphics and tables, such as scatter and bar charts, for review. Scatter charts are useful for rapidly evaluating many individual measurement relationships at the tree level, such as diameter and height (fig. 4). Scatter charts also allow the analyst to quickly spot outliers or other unusual patterns in the data. Bar charts are another output graphic type available in QA Tools. Annual estimates of attributes such as forest area, tree volume, and tree biomass and their statistical reliability, can be depicted in bar charts if year by year regional stratification files were produced (fig. 5). A chart ‘wizard’ within QA Tools is available to help walk the user through the bar chart setup and evaluation process. Bar charts are useful as a QA check for errors or unusual changes over time within summary groups or series of estimates. Each batch process session can be saved as a Batch Definition File (BDF) which is essentially a set of RDF files that are processed at one time. Any BDF can be later loaded into QA Tools and reprocessed with the same dataset,

applied to another dataset such as an added panel of field measurements or a dataset from another State.

CONCLUSION

The objective of the analytical QA task team is to promote the collection and dissemination of high quality FIA data through the use of a nationally derived QA process. The task team is developing guidelines to facilitate this process and is promoting the development and use of analytical tools that support this objective. A nationally derived approach to QA promotes a consistent process across all FIA regions leading to the delivery of quality data products with known precision, completeness, representativeness, and comparability. Efficiencies are gained from national consistency because fewer resources go into maintaining separate processes, leaving more resources available to look at the data in greater detail and from many directions.

Consistency facilitates a nationally cohesive program through thorough documentation of FIA processes and methodologies that benefits quality control and leads to greater transparency. Examples include a draft guide to prefield procedures, data collection protocols outlined in a national FIA field manual, a national PDR program (MIDAS) with accompanying documentation and error checking logic, and published FIADB documentation and user guide. In addition, changes and updates in NIMS procedures are applied more consistently when accompanied by documentation. Data delivery procedures and tools, such as standard tables and FIA Web tools have been developed to assist clients in obtaining information.

Consistency is also achieved through the development of QA Tools, a desktop application available to FIA analysts for performing analytical QA of unprocessed data and compiled estimates. The estimation engine within QA Tools provides a quick method to calculate summary estimates and sampling errors. Efficiencies are realized by using QA Tools, especially with the batch processing capability, enabling a large set of predefined analytical QA checks to be completed expeditiously each year.

ACKNOWLEDGMENTS

In addition to the authors, other members of the analytical QA team contributed information for our description of the data flow and analytical QA processes. They include Randy Morin (Northern Research Station), Roger Conner (Southern Research Station), and Mike Thompson (Rocky Mountain Research Station). Olaf Kuegler provided extensive statistical expertise in developing the estimation engine. We especially want to recognize the work of Larry

Potts, the QA Tools programmer who devoted many hours to programming and modifying the application to meet the needs of the analysts.

LITERATURE CITED

Bechtold, W. A.; Patterson, P. L.; eds. 2005. The enhanced forest inventory and analysis program - national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 85 p.

Miles, P. D.; Brand, G. J.; Alerich, C. L.; Bednar, L. F.; Woudenberg, S. W.; Glover, J. F.; Ezzell, E. N. 2001. The forest inventory and analysis database: database description and users manual version 1.0. Gen. Tech. Rep. NC-218. St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Research Station. 130 p.

Miles, P. D. 2008. A simplified forest inventory and analysis database: FIADB-Lite. Gen. Tech. Rep. NRS-30. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 42 p.

SAS Institute. 2002. The SAS system: SAS OnlineDoc®, Version 9.2, HTML format [[CD-ROM]. Cary, NC.

U. S. Department of Agriculture Forest Service, Research and Development. 1998. Quality assurance implementation plan. Washington, DC: U.S. Department of Agriculture Forest Service. 24p.

Woudenberg, Sharon W.; Conkling, Barbara L.; O’Connell, Barbara M. [and others]. 2010. The Forest Inventory and Analysis Database Version 4.0: Database Description and Users Manual for Phase 2. Gen. Tech. Rep. RMRS-245. Fort Collins, CO: U.S. Department of Agriculture, Forest Service. Rocky Mountain Research Station. 355 p

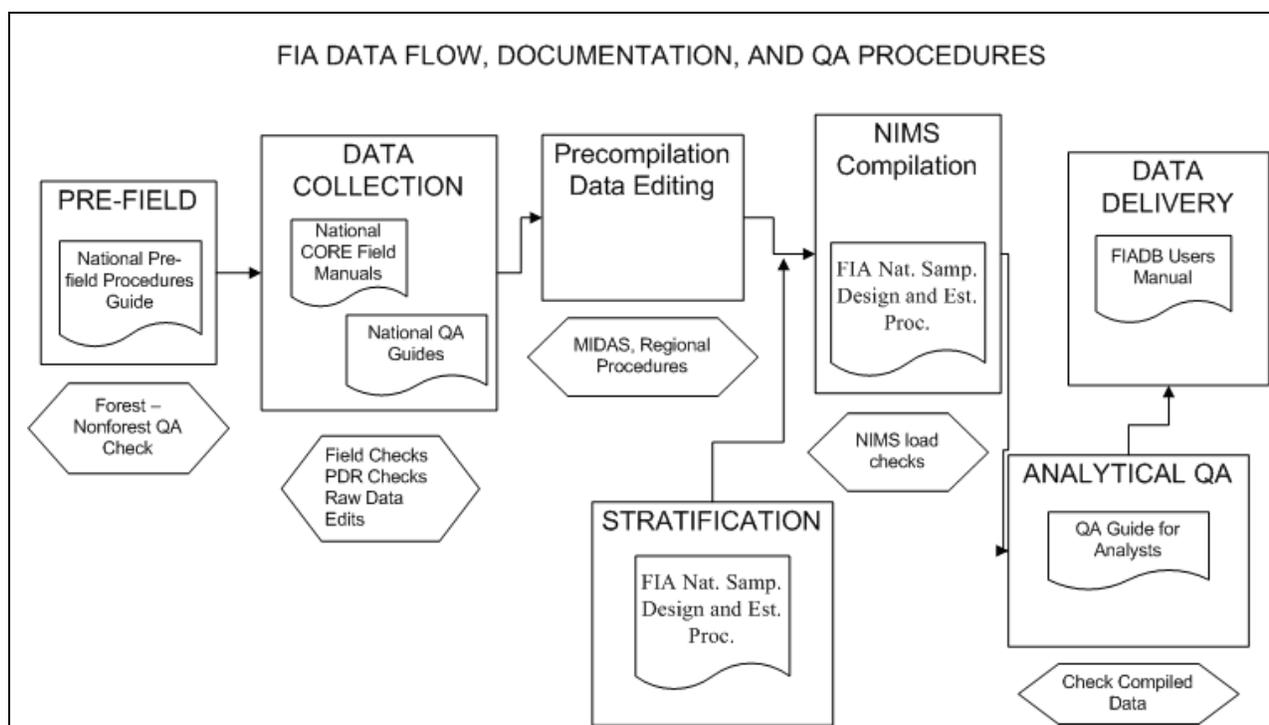


Figure 1—FIA data flow process showing the seven basic stages with examples of QA processes and documentation.

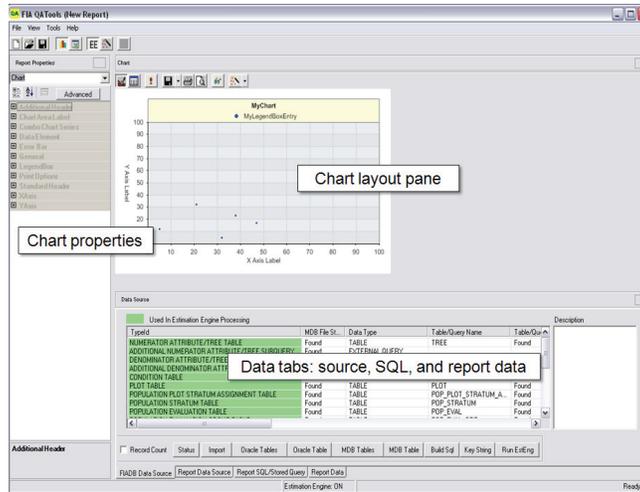


Figure 2—QA Tools “New Report” layout shows the user interface when defining a new analytical QA check.

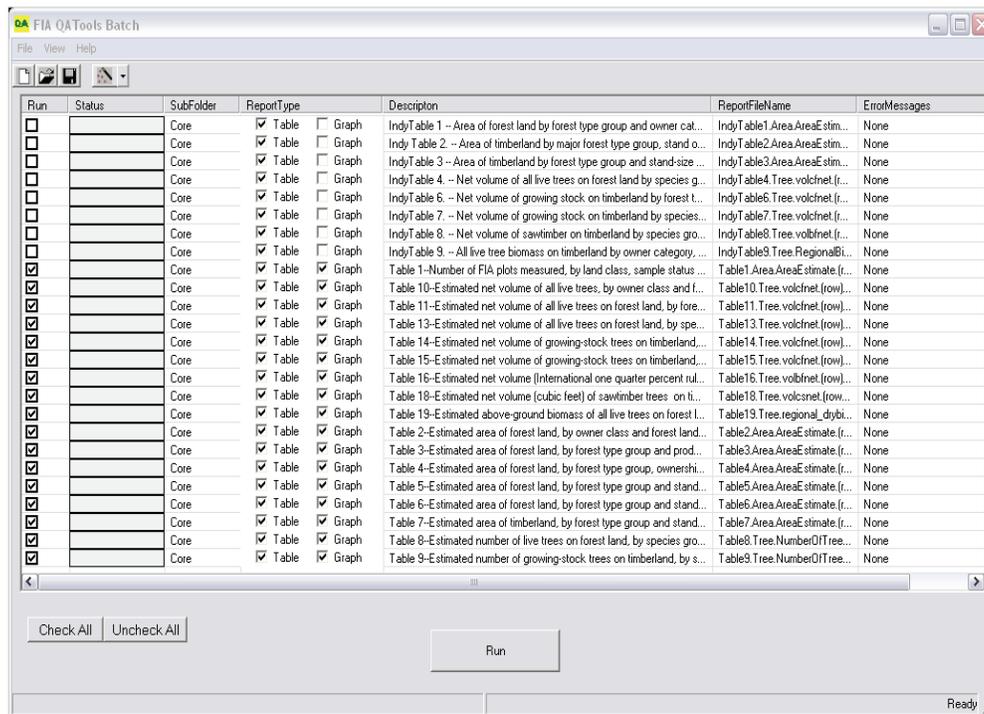


Figure 3—QA Tools “batch process” set-up window shows a group of selected, pre-defined QA checks that will be output as both tables and charts during this session.

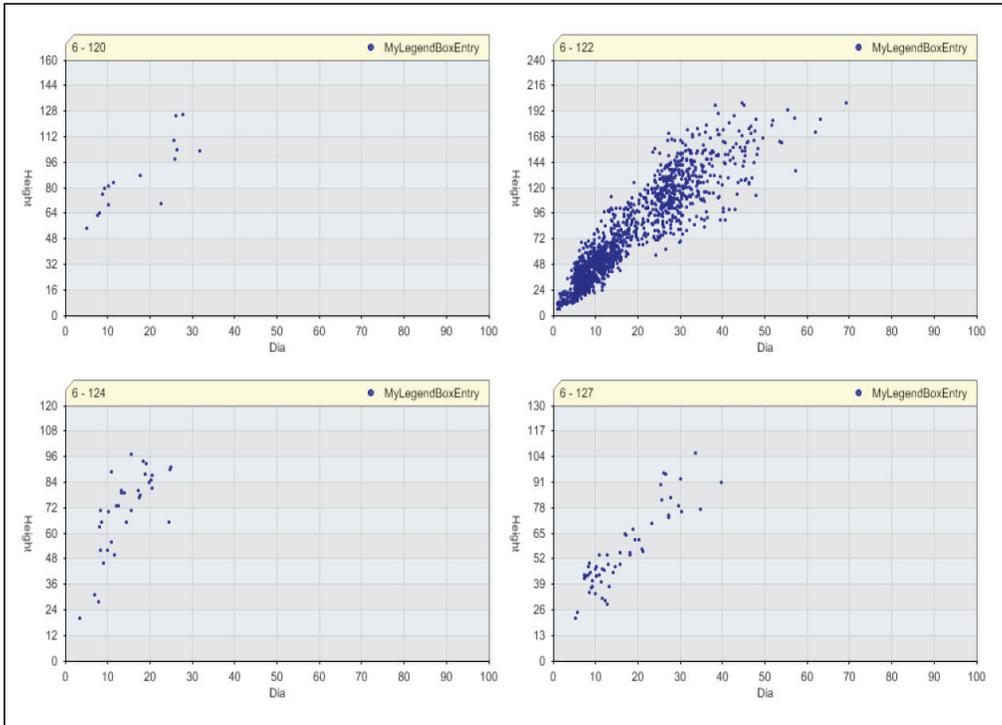


Figure 4—QA Tools scatter charts allow analysts to quickly evaluate many measurements for obvious data errors.

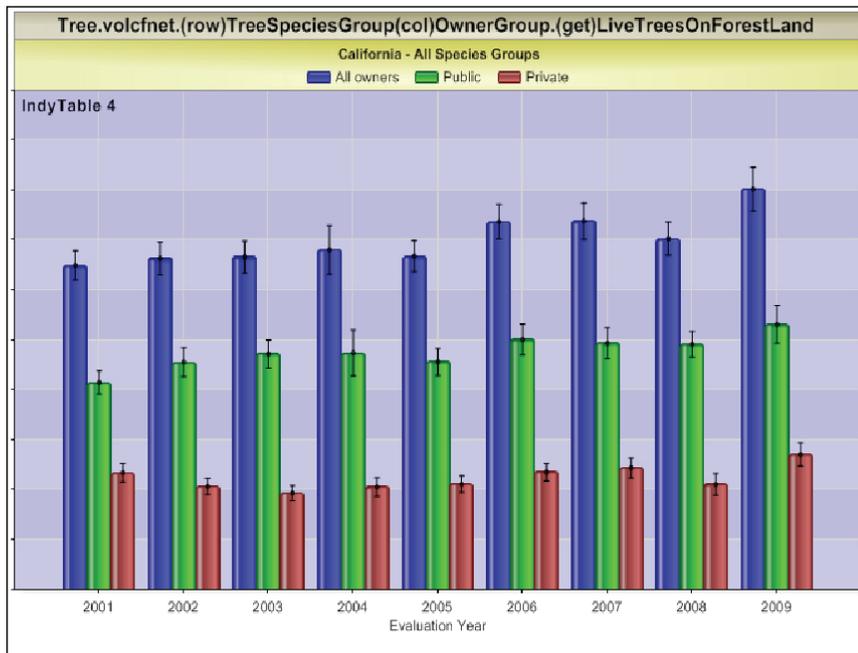


Figure 5— Bar charts from QA Tools enable an analyst to evaluate a current estimate against previously collected data, providing a quick method to spot potential errors.

RESOLVING THE PULPWOOD CANVASS WITH INVENTORY HARVEST INFORMATION

Joseph M. McCollum¹ and Tony G. Johnson²

ABSTRACT

The Resource Use section of the Forest Inventory and Analysis (FIA) Program has done a canvass of wood processing mills for timber product output (TPO) throughout the southern United States. Pulpmills in the South are canvassed on an annual basis, while all other mills (e.g., sawmills, veneer mills, etc.) are canvassed every two years. Attempts have been made to graph and map the amount of pulpwood harvested compared to the acres of forest treated (cut) in order to provide more information on harvesting rate or intensity (i.e., volume harvested per acre). It appears that one county's worth of plots is not enough data to accurately estimate the cutting rate in a county. The authors advocate smoothing the apparent cutting rate, with one suggested model being logistic regression based on forest density and ownership patterns in a county.

Keywords: Timber Products Output (TPO), Forest Inventory and Analysis (FIA), logistic regression, pulpwood harvest, harvesting rate.

INTRODUCTION

Each year since at least 1953, the Timber Products Output (TPO) section of the Forest Inventory and Analysis (FIA) unit, or their predecessor programs, has conducted a canvass of pulpwood mills in the Southern United States. The data have been reported tabularly, by county, and by a variety of maps. Maps have included dot density maps, starting with Cruikshank (1954) to Bertelson (1972). Bertelson (1975) experimented with some type of contouring for the 1974 report. Dennis May (1986) used raw choropleth maps (polygon maps with varying patterns to indicate the level of the response variable) for the 1984 report then experimented (1988) with a choropleth map based on acres of timberland. Johnson et al. (1997) used a choropleth based on acres of land, then used a choropleth based on acres of timberland (2010).

FIA field crews collect data on survey plots—including whether trees were harvested or not. Another variable that field crews collect is: TRTCD_x, where x = 1, 2, or 3, as appropriate. This is the “treatment code” that field crews believe a plot has received. For example, TRTCD_x = 10 indicates some type of cutting. Theoretically, a plot may be

cut twice, or even three times, in a cycle. In practice it will rarely happen more than once. In the Southern Research Station, if a condition of a plot receives TRTCD_x = 10, field crews will further classify that code into clearcut harvest, partial harvest, shelterwood harvest, commercial thinning, or timber stand improvement. In this study, all field calls of TRTCD_x = 10 were used and weighted equally.

Pulpwood production represents about 43 percent of overall harvesting (Johnson et al. 2009). However it would still be useful to see how the pulpwood canvass matches the Phase 2 survey.

The traditional calculation of acres treated per year is this one:

$$A = \sum_{i=1}^c \frac{EXPCURR_i \times ADJ_EXPCURR_i \times CONDPROP_i}{REMPER_i} (I_1 + I_2 + I_3) \quad [1]$$

where

EXPCURR_i is the current expansion factor for condition *i*,
ADJ_EXPCURR_i is the adjustment factor for condition *i*,
CONDPROP_i is the condition proportion of the plot in forest,
I_x is an indicator function for condition *i*; if TRTCD_x = 10, then I_x = 1, and 0 otherwise,

REMPER_i is the remeasurement period for condition *i*.

ADJ_EXPCURR is an adjustment factor that compensates for inaccessible portions of otherwise accessible plots.
REMPER is the time since the previous plot visit. In the case of a new plot, REMPER is assumed to be 5 years in the eastern United States and 10 years in the western United States (in the South, this restriction includes only western Texas and Western Oklahoma). See Rudis et al. (2008) and Harper (2010) for boundaries. Further details on the FIA database may be found in USDA (2007).

In standard pulpwood reports, softwood and hardwood are broken out separately, however in the interest of brevity they have been combined for this analysis. Figure 1 shows a map of pulpwood harvest per treated acre.

¹Information Technology Specialist, Forest Inventory and Analysis Unit, Southern Research Station, 4700 Old Kingston Pike, Knoxville, TN 37919.
e-mail: jmcollum@fs.fed.us

²Forester, Forest Inventory and Analysis Unit, Southern Research Station, 200 WT Weaver Blvd., Asheville, NC 28804.
e-mail: tjohnson09@fs.fed.us

A CLOSER LOOK AT PHASE 1

There are 98 counties (parishes in Louisiana, but often called counties hereinafter) with pulpwood harvest data but no treated acres identified by the P2 FIA plots. 25 of these counties have less than 100 cords of production, but another 25 counties show more than 10,000 cords of production. According to the FIA Database, there were about 357 billion cubic feet of wood on forestland on 257 million acres in the south. The average yield, then, is 19 cords per forested acre. Among 1227 forested counties, the median value is 20.6 cords per forested acre, with the 95th percentile being 35.4 cords per acre. The highest is Hampton City, VA at 127 cords per acre, based on one plot, followed by Carlisle County, KY with 4 plots at 66 cords per acre and Fayette County, KY and then several counties at 52 cords per acre. So, some of the 15 counties in the 32 – 569 cords per acre range should be considered suspect as well.

West Texas has 4 of 10 panels processed; West Oklahoma has no data processed at all. All other states have a full cycle of data. A cycle is a full set of plots. In general, there is one plot for every 5937 acres. Plots are divided into 5 to 7 panels in the Eastern United States, and one panel is done in roughly one year. In the Western United States, the plot list is divided into 10 panels, with one panel done roughly each year. At the end of one cycle, the next cycle begins.

Only three counties stood out in the 2008 canvass in terms of infinite production per acre of timberland: Johnston County, OK is in west Oklahoma. Dallas County, TX shows a small amount of production but no timberland according to the Phase 2 survey. Orleans Parish, LA shows no forestland, let alone any timberland; it too shows a small amount of production.

Potentially small amounts could be due to landowners or utility companies cutting down trees and sending them to a pulpwood mill. However, FIA's understanding of urban forestry is too weak to model this sort of activity.

A rigorous examination of the Phase 1 data shows that there just might be some forestland present in Orleans Parish. Phase 1 is the first phase of forest inventory. It involves taking a classified satellite scene, totaling the pixels, and overlaying the plots on the scene. *EXPCURR* is equal to the stratum size within a survey unit divided by the number of plots in the stratum. Results for the South Delta of Louisiana are shown in Table 1.

Survey unit lines were established because FIA believed contiguous counties to be ecologically similar. If we apply these correction factors to the results specifically for Orleans Parish, we get the result shown in Table 2.

The phase 1 map of Orleans Parish is shown in Figure 2.

The water layer was not used for plot stratification but it is included to give the reader context. One can see there are areas of likely forest in New Orleans East, between Interstate 510 and Bayou Sauvage National Wildlife Refuge, as well as on the West Bank of the Mississippi River near English Turn, and then a small amount along the intracoastal waterway.

While the method described in Table 2 gives the same number of acres in each survey unit as the current method, it distributes acres among counties differently. However there are no forested plots in Orleans Parish among which to distribute these acres. The current problem requires only calculation of forested acres rather than volume or biomass, but there are still no treated acres for Orleans or any of the other 97 county-equivalents with no observed cutting.

A CLOSER LOOK AT TRTCDx

It would appear then, that one county's worth of plots is not enough to accurately determine the cutting rate. The data need to be smoothed. There are many ways to accomplish this task.

Reams and McCollum (1999) found that important factors in probability of harvest were geographic region and ownership. Other factors were trees per acre and stand diameter.

It occurs to the authors that percent forest is a relevant factor as well, at least for privately owned forests. In most states the data appear to bear that out. For instance, in Louisiana, Figure 3 shows a graph of percent forest in a county to percent cut.

The regression line is a logistic regression, weighted by the number of plots, determined by the equation

$$Y = \text{logit}(b_0 + b_1X)$$

where Y is the percent of privately held acres treated and X is the proportion of acres in private forest, as a percent of total area of the county. The one severe outlier is St. Bernard Parish, based on two forested plots. Possibly there is particularly heavy cutting in this parish, but the cost of increasing the estimate in parishes with no observed cutting is reducing the high estimate of the high outliers as well.

The results for public lands in Louisiana are shown in Figure 4. These results are quite a bit noisier. Grant Parish, home to the Kisatchie National Forest, is the one over 40 percent publicly forested. The parishes with the highest

observed rates of public cutting are: Sabine (Ft. Polk), West Feliciana (Cat Island National Wildlife Refuge), Caddo (County/Municipal lands), and St. Landry (Other federally owned lands).

Most other states have fairly similar graphs for private land. One notable exception is Kentucky; the regression line is nearly flat. Georgia's and Mississippi's graphs are somewhat less steep as well. Table 3 has coefficients for all the southern states. One can see that management patterns on public lands are relatively diverse. There may be some accessibility issues in Kentucky. There would appear to be too much cutting in lightly forested counties for all to be flukes.

Another special case is Texas. Ordinarily, FIA processes Texas as though it were two separate states. From Figure 6, one can see that there is valid reason for doing that. Open circles represent counties in East Texas, open squares represent counties in West Texas that did not appear on the TPO canvass, and filled circles represent counties in West Texas that appeared on the TPO canvass. Only the counties that appeared on the canvass were used to compute the regression.

Public land in Texas has an extremely low cutting rate, as shown in Figure 7. The one outlier is Upshur County (other public land).

RESULTS

The next step is to fit these coefficients with the Phase 1 (not the raw Phase 2) derived area estimates, and divide the estimated treated acres into production. The combined result for softwood and hardwood is shown in Figure 6.

There are three counties with more than 32 cords per treated acre in the 2008 canvass, with Clay County, FL leading the way at 39 cords per treated acre.

There were five counties with more than 32 cords per acre in the 2009 canvass, with Washington County, TN leading the way.

CONCLUSIONS

The plan for the 2009 Southern Pulpwood Report, to be published in Fiscal Year 2011, is to use a method similar to this one. It marks a drastic shift in the way pulpwood data and FIA Phase 2 data has been processed and visualized. The authors welcome suggestions for improving the model.

ACKNOWLEDGEMENTS

The authors thank Todd Morgan, CF, of the University of Montana as well as Dr. Margaret Devall and Dennis Jacobs of the Southern Research Station and Mike Thompson of the Rocky Mountain Research Station for helpful review comments.

LITERATURE CITED

- Bertelson, D.F.** 1975. Southern Pulpwood Production, 1974. Resour. Bull. SO-54. New Orleans, Louisiana: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 24 p.
- Bertelson, D.F.** 1973. Southern Pulpwood Production, 1972. Resour. Bull. SO-41. New Orleans, Louisiana: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 25 p.
- Cruikshank, J.W.** 1954. 1953 Pulpwood Production In The South. Forest Survey Release No. 43. Asheville, NC: U.S. Department of Agriculture, Southeastern Forest Experiment Station. 34 p.
- Harper, R.** 2010. East Oklahoma, 2008 Forest Inventory and Analysis factsheet. Science Update SRS-025. Asheville, NC: USDA-Forest Service, Southern Research Station. 2p.
- Johnson, T.G.;** Steppleton, C.D.; Bentley, J.W. 2010. Southern Pulpwood Production, 2008. Resour. Bull. SRS-165. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 42 p.
- Johnson, T.G.;** Bentley, J.W.; Howell, M. 2009. The South's Timber Industry—an assessment of timber product output and use, 2007. Resour. Bull. SRS-164. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 52 p.
- Johnson, T.G.;** Steppleton, C.D. Southern Pulpwood Production, 1997. Resour. Bull. SO-138. New Orleans, Louisiana: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 45 p.
- Johnson, T.G.;** Steppleton, C.D. Southern Pulpwood Production, 1997. Resour. Bull. SO-138. New Orleans, Louisiana: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 45 p.
- May, D.M.** 1988. Southern Pulpwood Production, 1986. Resour. Bull. SO-138. New Orleans, Louisiana: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 45 p.
- May, D.M.** 1986. Southern Pulpwood Production, 1984. Resour. Bull. SO-111. New Orleans, Louisiana: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 30 p.
- Reams, G.A.;** McCollum, J.M. 1999. Predicting the Probability of Stand Disturbance. In: Haywood, J.D., ed. Proceedings of the tenth biennial southern silvicultural research conference. Gen. Tech. Rep. SRS-30. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 632 p.
- Rudis, V. A.;** Carraway, B.; Sheffield, R. M. [and others]. 2008. East Texas Forests, 2003. Resour. Bull. SRS-137, Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 145 p.

U.S. Department of Agriculture. 2007. The Forest Inventory and Analysis Database: Database Description and Users Guide Version 2.1. National Forest Inventory and Analysis Program. 196 p. http://fia.fs.fed.us/library/database-documentation/historic/ver2/FIADB_v2-1_03_09_07.pdf, last accessed 1 October 2010.

Table 1—Stratum statistics for the South Delta of Louisiana

Stratum	Acres	Plots	EXPCURR	Percent Forest
Nonforest	10062566	1681	5985.955	0.53
Nonforest Edge	669614	99	6763.374	13.54
Forest	315246	46	6852.933	66.71
Forest Edge	370913	76	4880.275	27.56
Bottomland Hardwoods	2336656	333	7016.816	81.20
Total	13754994	2235		

Table 2—Stratum statistics for Orleans Parish, Louisiana

Stratum	Acres	Plots	Percent Forest	Forested acres
Nonforest	201061	31	0.53	1070
Nonforest Edge	5381	1	13.54	729
Forest	3965	1	66.71	2645
Forest Edge	3243	0	27.56	894
Bottomland Hardwoods	10487	0	81.20	8516
Total	224137	33		13853

Table 3—Regression coefficients for southern states

State	Year	Private		Public	
		b ₀	b ₁	b ₀	b ₁
Alabama	2008	-4.1648	1.3859	-3.7556	-9.6544
Arkansas	2008	-4.1318	1.7803	-4.6951	1.5249
Arkansas	2009	-4.1513	1.7245	-4.5311	1.1438
Florida	2007	-4.1368	1.0346	-3.9876	-3.0698
Georgia	2008	-3.5806	0.4201	-4.0208	-1.8821
Kentucky	2007	-3.6491	0.0664	-3.9594	-3.8556
Louisiana	2005	-3.7813	1.1720	-4.4667	2.3616
Mississippi	2006	-3.6625	0.4711	-4.317	0.6019
North Carolina	2007	-3.9513	0.7585	-5.1279	-0.4179
Oklahoma	2008	-5.2298	2.7946	-5.6821	3.5580
South Carolina	2007	-3.9947	1.1854	-4.2465	0.2546
Tennessee	2007	-4.7584	1.5121	-5.3849	-1.0712
Texas	2007	-3.9199	1.1900	-5.7037	7.3037
Virginia	2008	-4.6398	1.5253	-3.9773	-5.3726

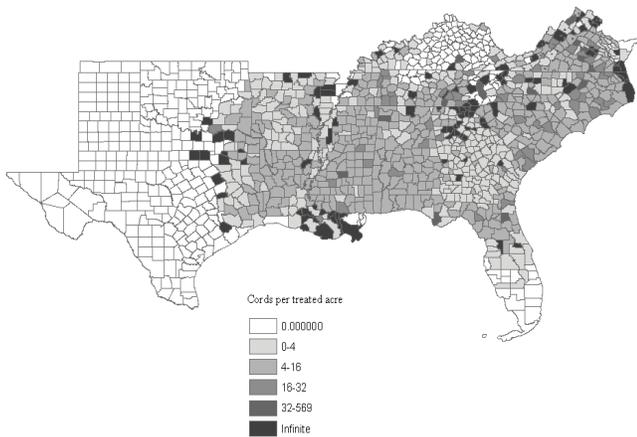


Figure 1—Choropleth map of pulpwood harvest per treated acre, 2008.

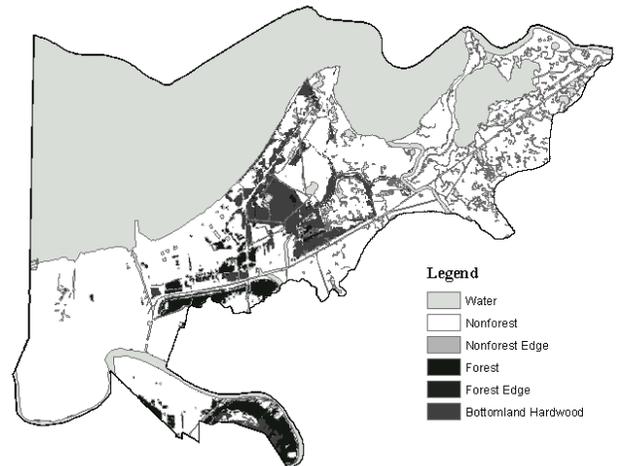


Figure 2—Phase 1 map of Orleans Parish, Louisiana.

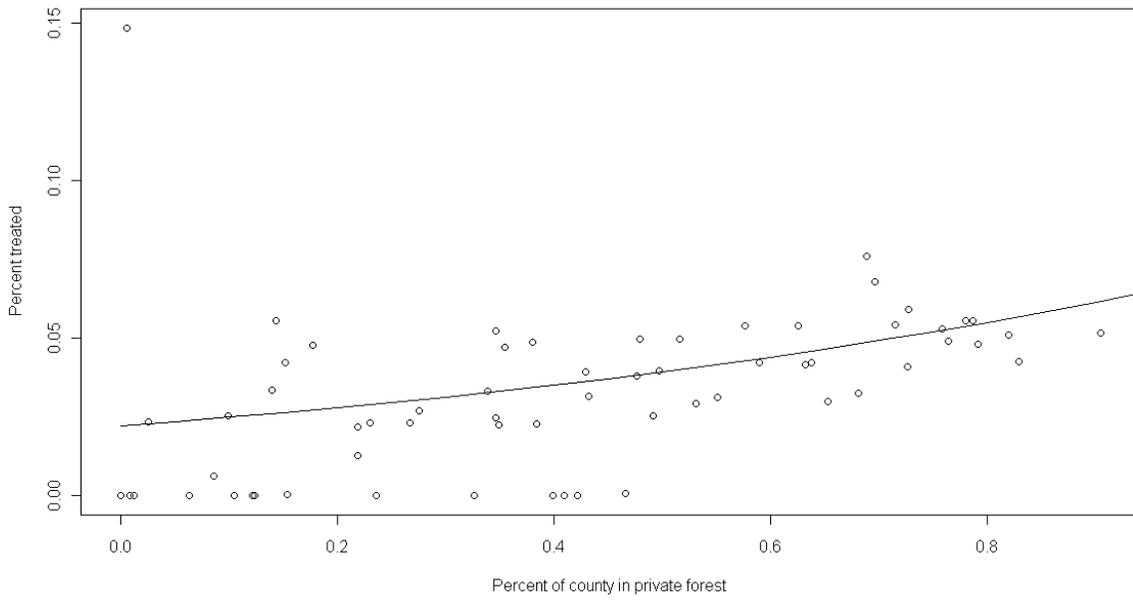


Figure 3—Percent of county in private forest versus percent treated, LA 2005.

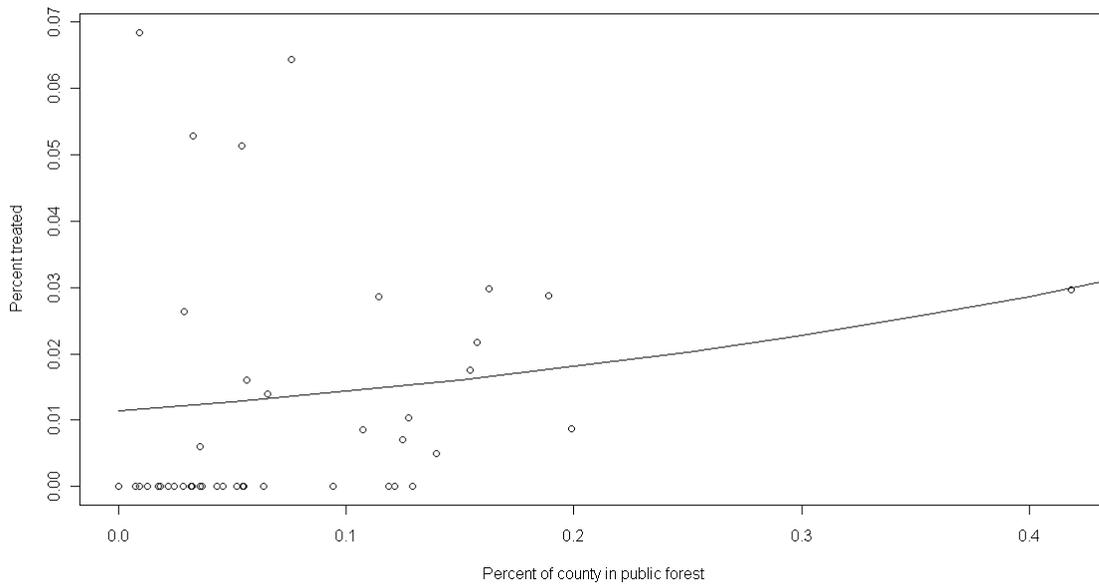


Figure 4—Percent of county in public forest versus percent treated, Louisiana 2005.

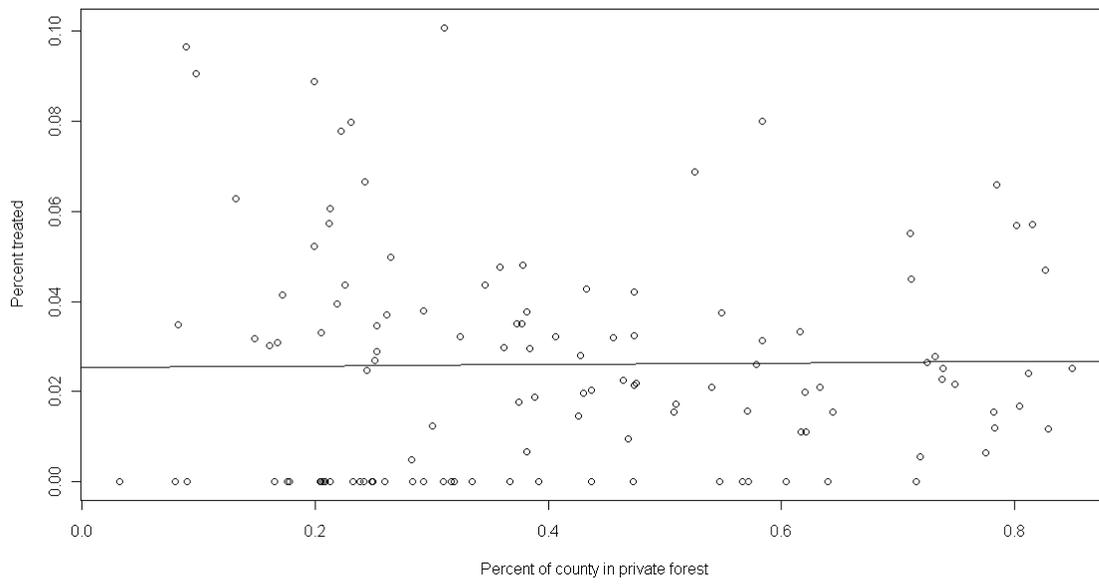


Figure 5—Percent of county in private forest versus percent treated, Kentucky 2007.

ALGORITHMIC DECISION RULES FOR ESTIMATING GROWTH, REMOVALS, AND MORTALITY WITHIN A NATIONAL-SCALE FOREST INVENTORY (USA)

William H. McWilliams, Carol L. Alerich, William A. Bechtold, Mark Hansen, Christopher M. Oswalt, Mike Thompson, and Jeff Turner

ABSTRACT

The U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) program maintains the National Information Management System (NIMS) that provides the computational framework for the annual forest inventory of the United States. Questions regarding the impact of key elements of programming logic, processing criteria, and estimation procedures were raised by national FIA Information Management Band (IMB) the review of the estimation procedures for forest “inventory” change, i.e. growth, removals, and mortality (GRM). The goal of the study was to review these questions and develop Decision rules to provide transparency to otherwise undocumented algorithmic pathways. These questions are important because they are the basis of FIA’s timber volume, tree biomass, and carbon stock GRM estimates for the United States. This study highlighted an incremental benefit of the annual system—the benefits of consistent rules, transparent methods, and reliable trend estimates for tracking forests in time and space accrue with each new panel.

INTRODUCTION

The U.S. Department of Agriculture Forest Service, Forest Inventory and Analysis (FIA) program maintains the National Information Management System (NIMS) that provides the computational framework for the annual inventory of the United States. (NIMS version 4.0 was used for this study.) The NIMS algorithms process all FIA field and other base measurements and provide all classified and estimated data for the national web service (Woudenberg and others 2011). NIMS is continually being revised to incorporate new field protocols and variable definitions, eliminate regional differences, and improve estimation procedures. The algorithms contained in NIMS follow the work of Patterson (2005). Although NIMS is a national system, it still uses regional procedures for estimating volumetric variables: timber volume, tree biomass, and

carbon. Some key stem measurements are diameter, height, and rotten or missing portion. Questions concerning the estimation procedures for forest “inventory” change; or growth, removals, and mortality (GRM) from the national FIA Information Management Band (IMB) have resulted questions regarding the impact of key elements of the programming logic, processing criteria, as well as estimation procedures. The IMB submitted these questions to the Techniques and Remote Sensing Band (TRSB) for decision rules. The TRSB convened the national Review Team. This paper represents the Review Team’s findings.

The goal of the study was to review these questions and develop Decision rules to provide transparency to otherwise undocumented algorithmic paths. More specific objectives include: compare of regional approaches, evaluate the impact of implementation options on the estimates, recommend implementation options for questions that significantly impact the results, and provide procedural recommendations for processing current and past GRM estimates. These questions are important because they are the basis of timber volume, tree biomass, and carbon GRM estimates for the United States. A national team was chartered to address these needs and to report findings.

METHODS

This national Review Team was comprised of representatives of all FIA Bands with direct experience with repeated FIA measurements, specifically field protocols, information technology, estimation, and quantitative analysis.

William H. McWilliams, Research Forester, U.S. Department of Agriculture, Forest Service, Northern Research Station, Forest Inventory and Analysis program, Newtown Square, PA 19073, USA.

Carol L. Alerich, Forester, NRS.

William A. Bechtold, Research Forester, Southern Research Station, Asheville, NC, 28804, USA, Retired.

Mark Hansen, Research Forester, NRS, Retired.

Christopher M. Oswalt, Research Forester, Southern Research Station, Forest Inventory and Analysis program, Knoxville, NC, 37919, USA.

Michael Thompson, Forester, Rocky Mountain Research Station, Forest Inventory and Analysis program, Ogden, UT, 84401, USA.

Jeff Turner, Research Forester, SRS, Asheville, NC, 28804, USA.

A series of criteria were used for sensitivity analysis of the impact that decision rules have on NIMS GRM estimates. The decision criteria included the level and degree of impact on estimates, and the degree of concurrence among the reviewers. The level of impact includes frequency and scale of occurrence (state, regional, or national) and whether the algorithm is for periodic inventories or annual inventories. The degree of impact was estimated by running simulations under various scenarios.

The NIMS algorithms contain many “pathways” for tree variables as they flow from input to output as volumetric estimates of average annual change. The complete NIMS system contains hundreds of thousands of program code and continues to expand. The pathways for “growing-stock” trees are particularly intricate. NIMS includes major pathways for periodic-to-annual (P2A) and annual-to-annual (A2A) change components. The P2A pathways are much more complicated because of changes in measurement protocols, methods, and procedures both in time and space. The A2A algorithms center on measurements that have been common since their inception around 2000. The team focused on A2A for context in providing transparent decision rules for current and future needs.

RESULTS

The national Review Team addressed nine questions. Some of these questions have been discussed for decades in various forms within FIA regions and national teams. The questions are listed below along with the Review Team’s final decision rules.

Question One: What should be done with trees that were recorded at T1 but missing at T2 with no explanation, or “non-reconciled” trees?

Decision rule: The occurrence of this situation is low and the degree of impact on estimates is very low. All trees loaded into the national database at T1 (Woudenberg and others 2011) must be reconciled and corrected at the regional level.

Regional estimation methods for reconciling parameters are a common theme in many of the questions. Each of four regional systems was built with specific models to predict volume for species or species groups. Independent variables typically include diameter, rotten portion, and other related variables, e.g. tree class. Approaches use modeled and measured independent variables and some use combinations of both. To illustrate one approach, NRS-FIA computes gross volume (cubic and board foot) using equations by

Scott (1979 and 1981). The independent variables are diameter and merchantable length. Length is estimated using a taper model (Westfall and Scott In Press). Percent rotten cull is also modeled and subtracted to provide “net” volume. It is important to note that precision requirements are the same for regions.

Tree biomass, and hence carbon, is estimated using a national approach termed Component Ratio Method (Heath and others 2009), which provides national harmonization for these variables.

The national annual inventory GRM algorithm has introduced many improvements to existing systems. For example, a method for “growing” trees to the midpoint of the inventory cycle (2.5 years for a five-year cycle) replaced previously disparate methods. The overall net change is equivalent under the new and previous theories, but individual components of change will vary. Not all regions had fully developed approaches for these calculations when the approach was implemented.

Question Two: What should be done about trees that are measured at T2 and found to be too large to be considered ingrowth, e.g., missed at T1?

Decision rule: The degree of impact is low. It is recommended to follow current NIMS procedures with some added details. First, it is recommended to assign T1 tree status to “live.” Then use regional-scale growth estimators to calculate T1 tree diameter. This Rule is subject to two sub-rules: if a tree is alive at T2, set T1 and T2 tree status to live; if tree status is “dead” at T2, set T1 tree class to “rotten cull.” The Team determined it was not possible for a tree to be considered alive at T1 then missed and removed at T2 due to data recorder edit procedures.

Question Three: What should be done about species mismatches between T1 and T2?

Decision rule: The impact is low and occurrence is relatively rare. The NIMS procedures for current estimates do not consider species code at T1. No suggested change is recommended. The Team did note that this will introduce some minor discrepancies for “net change” estimates. This is because NIMS estimates volumetric variables in both the GRM and the inventory estimation modules. Within the GRM context, net change is defined as gross growth minus mortality and removals. Net change can also be computed using inventory estimates by subtracting volume at T1 from T2. With no discrepancies, the GRM and inventory modules would provide identical estimates of net change. Discrepancies should not be significant for population totals,

such as the total volume of loblolly pine in Georgia. The discrepancy would be apparent for data cells or summaries with small sample size, e.g., loblolly pine in southern Georgia, on National Forest Land for growing-stock trees 18 inches and larger only.

Question Four: What should be done about trees with very large positive values or negative growth values?

Decision rule: The impact on estimates for trees with growth that is outside a normal range is considered low. Currently, if the diameter measurements at T1 and T2 are not taken at the same height, NIMS recalculates T1 diameter and associated T1 attributes. It was decided that other valid negative values for change in these attributes should be allowed to flow normally, as is the case with height discrepancies.

Question Five: What should be done about trees that change tree class between T1 and T2 (growing stock to rough or rotten and rough or rotten to growing stock)?

Decision rule: This question required simulations of the various combinations of tree class for T1 and T2, as well as related implications for all the various sub-components of GRM, e.g. growth on mortality of growing-stock trees on land that changed from forest to non-forest.

After evaluation of simulated results of the pathways that such a growing-stock tree could follow, the decision rules included in Table 1 are recommended.

The purpose of changing T1 attributes is to ensure variables used for volumetric estimates are as consistent as possible across the Nation and that the approach is transparent. The reason for changing T1 variables when a tree class changes from rough or rotten to growing stock is that these are often due to inconsistent methods and have significant impact on GRM estimates for growing-stock trees. Changes from growing stock to rough or rotten occur naturally and often, and hence, should not be changed. Any changes are made at T2 to integrate with estimates that are calculated separately in the inventory and GRM algorithms.

Monitoring real change in tree class is challenging because of objectivity in the classification. Suggestions for improved control include asking field staff to verify T1 tree class for all re-measured trees. Procedures used at the Southern Research Station provide a model for national implementation. Some trees may need to be added to the national database prior to implementing the rules. For these trees, it is recommended that if the tree is alive at T2, tree class at T1 should be set equal to tree class at T2; if the tree is dead at T2, set tree class at T1 to "rotten cull." In

the highly unlikely event that a tree was live and missed at T1 and removed at T2, set tree class to "growing stock." It should be noted that pathways for estimating growing-stock GRM's could be the most complicated programming component of NIMS.

Question Six: What should be done with trees that were classified as dead at T1 and found to be alive at T2?

Decision rule: The impact on estimates and frequency are both very low. This question addresses the same phenomena as Question Two and so, the decision rule for missing trees should be followed, i.e., compute missing variables using regional approaches.

Question Seven: How should NIMS accommodate storage of adjusted tree-level variables?

Decision rule: A fundamental paradigm of FIA has been that continuous improvement processes should address systematic differences in any field measurement or algorithm that is used to assign or estimate variable values, e.g., tree class, height, or rotten portion. This issue is very important because of the very large size and temporal nature of FIA data sets. It was decided that corrections to erroneous data should be replaced with corrected data; however, it is imperative that all original data is permanently archived and documented. The NIMS structure allows for this kind of archival and has more than one option that can be used.

Question Eight: Should trees with diameter measurements taken at different locations at T1 and T2 be used in GRM computations?

Decision rule: The degree of impact was considered low and the resolution follows other rules. These trees are used in the NIMS GRM calculations, so they are given T1 diameters and other variables needed to re-calculate volumes. As before, this assures consistent trend information to the extent possible.

Question Nine: Are denied/hazardous (DH) plots included in GRM calculations?

Decision rule: Access to the FIA samples can be denied by the landowner or have conditions too hazardous to conduct measurements (DH). The impact is considered to be non-existent because Bechtold and Patterson (2005) reviewed this issue in forming the theoretical constructs for the national FIA program. The existing rules are:

- if a sample is DH at T1 and visited at T2, do not include in GRM calculations,

- if visited at T1 and DH at T2, do not include in GRM calculations,
- if a part of the sample is DH at T1 and fully measured at T2, the portion included at both occasions is used in the GRM calculations,
- if fully measured at T1 and partially DH at T2, the portion included at both occasions is used in the GRM calculations,
- if partially DH at T1 and T2, the portion included at both occasions is used in the GRM calculation.

DISCUSSION AND CONCLUSIONS

As part of the review process, the Team developed a set of tenets to guide the sensitivity analyses, e.g. future questions. The tenets include: maintain temporal consistency; provide balance between regional and national needs; emphasize the need for field-level checks for temporal consistency; allow for correction of variables with inconsistent temporal measures, recognize the need to re-process delinquent data sets to current standards by allowing as much field checking of previous measurements as possible; and archive all existing raw and computed data sets. These may be useful to others grappling with similar questions.

Many of the questions considered conditions with low occurrence and impact, however, trees with the conditions described can be problematic if they do not have a specific pathway to follow, e.g., contribute to the wrong change component/sub-component or follow a terminal pathway incorrectly. Growing-stock computations account for the majority of the GRM algorithmic pathways of NIMS and are quite complicated due to the number of pathways and junctures.

Efforts to develop nationally consistent and harmonized estimates of volume, wood and carbon weight should continue. Currently, modeled and measured independent variables are used in estimation and vary by region, e.g., height and rotten portion. As harmonization continues, the national system will continue to use models appropriate for the biomes that span FIA regions and become more seamless in application.

During the discussion of tracking trees and attributes over time, a related issue was noted that may need addressed in the future. It was clear that approaches and protocols for re-measuring trees with diameter measurement at the root collar are needed. All four regions of the country utilize root-collar diameter measurements and are challenged by the task of reconciling data and producing meaningful trend estimates.

It was very clear from this experience that the entire NIMS program code lacks the kind of documentation needed to understand the critical estimation components. The little information that is available publicly does not begin to address the need for transparency for details; although there is considerable documentation housed within regional FIA units. This report highlights the major incremental benefit of the annual system: the benefits of consistent rules, transparent methods, and reliable trend estimates for tracking forests in time and space accrue with each time step.

ACKNOWLEDGMENTS

The authors would like to thank John Coulston and Frank Roesch for providing important review comments.

LITERATURE CITED

- Bechtold, William, A.**; Patterson, Paul L., Editors. 2005. The enhanced Forest Inventory and Analysis program – national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department, Forest Service, Southern Research Station. 85 p
- Heath, L.S.**; Hansen, M.H.; Smith, J.E.; Smith, W.B.; Miles, P.D. 2009. Investigation into calculating tree biomass and carbon in the FIADB using a biomass expansion factor approach. In: McWilliams, W.; Moisen, G.; Czaplowski, R., comps. 2008 Forest Inventory and Analysis (FIA) symposium; 2008 October 21-23; Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. CD.
- Scott, C.T.** 1979. Northeastern forest survey board-foot volume equations. Res. Note NE-271. Broomall, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. 2 p.
- Scott, C.T.** 1981. Northeastern forest survey revised cubic-foot volume equations. Res. Note NE-304. Broomall, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. 2 p.
- Westfall, J.A.**, Scott, C.T. (in press) Taper equations for commercial tree species in the Northeastern U.S. *Forest Science*.
- Westfall, J.A.**; Frieswyk, T.; Griffith, D.M. 2009. Implementing the measurement interval midpoint method for change estimation. In: McRoberts, R.E.; Reams, G.A.; Van Deusen, P.C.; McWilliams, W.H., eds. Proceedings of the eighth annual Forest Inventory and Analysis symposium; 2006 October 16-19; Monterey, CA. Gen. Tech. Rep. WO-79. Washington, DC: U.S. Department of Agriculture, Forest Service: 231-236.
- Woudenberg, S.W.**; Conkling, B.L.; O'Connell, B.M.; LaPoint, E.B.; Turner, J.A.; Waddell, K.L. 2011. The Forest Inventory and Analysis database: database description and users manual version 4.0 for Phase 2. Gen. Tech. Rep. RMRS-GTR-245. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. <http://fia.fs.fed.us/library/database-documentation/> Accessed February 2011

Table 1 – After evaluation of simulated results of the pathways that such a growing-stock tree could follow, the following decision rules are recommended

<u>Tree Class at T1</u>	<u>Tree Class at T2</u>	<u>Rule</u>
growing stock	growing stock	no change to the variable
growing stock	rough	no change to the variable
growing stock	rotten	no change to the variable
rough	growing stock	change T1 value of tree class and percent cull to T2 values
rough	rough	no change to the variable
rough	rotten	no change to the variable
rotten	growing stock	change T1 value of tree class and percent cull to T2 values
rotten	rough	change T1 value of tree class and percent cull to T2 values
rotten	rotten	no change to the data

APPLICATION OF AN ASSESSMENT PROTOCOL TO EXTENSIVE SPECIES AND TOTAL BASAL AREA PER ACRE DATASETS FOR THE EASTERN COTERMINOUS UNITED STATES

Rachel Riemann, Ty Wilson, and Andrew Lister

ABSTRACT

We recently developed an assessment protocol that provides information on the magnitude, location, frequency and type of error in geospatial datasets of continuous variables (Riemann et al. 2010). The protocol consists of a suite of assessment metrics which include an examination of data distributions and areas estimates, at several scales, examining each in the form of maps, graphics, and summary statistics. In this study we have applied this protocol to the modeled total and species-level basal area/acre datasets recently completed for the eastern coterminous United States (Wilson et al. in review). We were interested in the answers to two questions: (1) how can assessment results be effectively presented over extensive areas, and (2) what is the accuracy of modeled datasets of much less common forest characteristics such as the presence of an individual species, and what might that tell us about the limitations of the current modeled dataset for other less common variables. Results from this study will help fine-tune the type of assessments applied and how they are presented in the metadata available with all geospatial datasets produced by Forest Inventory and Analysis (FIA).

Keywords: Accuracy assessment, uncertainty, geospatial data, continuous variables, species distribution

INTRODUCTION

Modeled geospatial datasets benefit greatly from detailed accuracy assessment. Every geospatial dataset is a model of real conditions on the ground and thus inevitably contains some error. This error can take the form of truncated distributions, a loss of local variability, and/or an underestimation or overestimation of values that can be random (unsystematic error) or represent a bias (systematic error) across the entire dataset or in some areas. Similarly, the type of error present and its magnitude frequently varies with scale, and by the subpopulation being examined. Such inaccuracies do not usually render a modeled dataset useless, but these errors do affect interpretation and

appropriate use of the dataset, and may suggest different approaches for iterative improvement of the modeled geospatial dataset. In addition, for an assessment to be truly effective it must be consistent, to facilitate the comparison of results between datasets of the same variable, and timely, ideally available as soon as the dataset itself. In a previous study we developed a protocol for assessing geospatial datasets of continuous variables (Riemann et al. 2010). This protocol consists of a suite of assessment metrics that together describe the location of errors, the frequency of errors, the magnitude of errors, and the type/nature of errors (Foody 2002) (Canter 1997), and improves timeliness by taking advantage of USFS Forest Inventory and Analysis' (FIA's) existing extensive plot database as the reference data source.

U.S. Forest Inventory and Analysis (FIA) is in the process of developing a broad set of modeled geospatial datasets of forest characteristics across the entire United States, and needs to provide information on the accuracies of those datasets in the accompanying metadata as soon as datasets are released. One such set of datasets has been produced using an approach developed by Wilson et al. (in review) and will soon be available for the coterminous United States.

In this study we applied the existing assessment protocol to the eastern half of this extensive modeled geospatial dataset, and in particular to datasets of total basal area/acre (ba/acre) as well as six individual tree species ba/acre. We were interested in: 1) how assessment results could be most effectively presented for such extensive areas, and 2) the accuracy of individual species datasets given the wide range in their frequency of occurrence, in their spatial patterns of distribution, and in their level of canopy and/or basal area dominance in the stands in which they occur.

Rachel Riemann, Research Forester/Geographer, Northern Research Station, Troy, NY, 12180
Barry Tyler Wilson, Research Forester, Northern Research Station, St. Paul, MN, 55108
Andrew Lister, Research Forester, Northern Research Station, Newtown Square, PA, 19073

DESCRIPTION OF THE ASSESSMENT PROTOCOL

The protocol recommends a suite of assessments, including:

- assessment of data distributions – at several scales
- assessment of overall agreement of area estimates – at several scales
- examining differences in local variability
- examining spatial and distribution patterns of local differences

First, assessment of data distributions is accomplished by comparing the empirical cumulative distribution functions (ecdf's) of the modeled and reference datasets. A Kolomogorov-Smirnov (KS) statistic can also be used to summarize the largest distance between the two curves (Figure 1a). Second, assessment of overall agreement of estimates is accomplished by comparing a scatterplot of model-derived vs. FIA plot-based estimates against the 1:1 line (Figure 1b). Metrics can be calculated from this scatterplot to quantitatively describe the overall agreement (agreement coefficient, AC), systematic agreement (AC_{sys}), unsystematic agreement (AC_{uns}), and root mean square error (RMSE) (Ji and Gallo 2006). Systematic agreement quantifies the difference between the 1:1 line and the geometric mean functional relationship (GMFR) regression line, which describes the level of bias present. Unsystematic agreement quantifies the level of scatter about the GMFR regression line, which describes the magnitude of remaining random or unexplained error. The GMFR regression line is used instead of the linear regression line because GMFR is a symmetric regression model that assumes both X and Y datasets are subject to error, unlike least squares regression. All three agreement coefficient metrics are symmetric and standardized, facilitating easy comparison between datasets. RMSE values are also symmetric, and are in data units, providing a measure of the magnitude of the error in data units. As many studies have pointed out, dataset accuracy changes with scale (e.g. (Blackard et al. 2008), (Nelson et al. 2009)). Thus, these first two assessments should be calculated at a range of scales to provide information on how dataset accuracy changes with scale. We have recommended choosing that scale at which we have reasonable confidence in FIA estimates (216,500 ha), plus one or two below and above that (Riemann et al. 2010).

The third assessment examines differences in local variability (figure 1c), and the fourth examines the spatial and distribution patterns of local differences between the modeled and reference datasets, (figures 1d,e). These last two assessments can be effective if calculated at a scale at which a sufficient number of FIA plots are available to have reasonable confidence in the FIA plot-based estimates for mean, standard deviation, and a reasonably small confidence interval. When working with hexagons as the spatial unit,

a hexagon 50 kilometers in diameter is 216,500 hectares in size and contains an average of 35 FIA plots (forest and nonforest) per hexagon. A full description of all metrics can be found in Riemann et al. (2010). A complete description of the Agreement Coefficient metrics can be found in Ji and Gallo (2006).

The last two assessments calculate and present assessment results for local areas (i.e. at a fixed scale defining that local area) and can thus be easily expanded to cover large map extents without any loss of descriptive power for local areas. The first two assessments, however, are initially calculated for the dataset as a whole, making them less valuable as the extent of the dataset increases. Thus, when working with datasets as extensive as the Eastern CONTinental United States (ECONUS), calculating these metrics for the dataset as a whole is not sufficient. Over such a large area both the type and magnitude of errors can and will vary by region, and thus the summary metrics should be calculated and available by smaller regions as well as the dataset as a whole, which requires choosing both the regions of assessment and the scale at which it will be assessed.

In this study we calculated assessments using both level 3 ecoregions and 3.5 million ha hexagons to examine any differences resulting from choice of region. We selected the 78,100 ha scale because of its reasonably high comparative accuracies reported in the dataset-wide assessment (Figure 2)— $AC=0.95$, $AC_{sys}=1.0$, $AC_{uns}=0.95$. The 78,100 ha scale also represents a compromise between having a sufficient number of plots within each hexagon (an average of 20) so that the FIA estimate is a reasonably robust estimate of the mean for the area, and having a sufficient number of hexagons within each region so that there are enough points from which to calculate reasonably robust assessment metrics for each region, whether level 3 ecoregions or 3.5 million ha hex-regions are used (Figure 3). With respect to the choice of region, ecoregions have the advantage of including the entire land area, and of dividing the area by one characteristic which could contribute to differences in accuracy, such as different ecosystem types. However, they have the disadvantage of varying widely in area (which means one cannot simply use the histogram to display the amount of area in each error category), of sometimes being very long and narrow and even containing exclusions which inevitably translates into a greater number of summary units (the 78,100 ha hexagons) that include area from neighboring ecoregions. Using a 3.5 million ha hexagon as the region of assessment has the advantage of being equal land area unless we include the edge hexagons that reach beyond the extent of the land area and plot data in the study. They have the disadvantage of not being based on any factor suspected of affecting accuracy other than geographic location, however since there could be many factors, perhaps this is a less important criteria.

DESCRIPTION OF THE DATASETS

In this paper we assess datasets generated for the Eastern CONTinental United States (ECONUS) using an approach developed by Wilson et al. (in review), which is a modification of the gradient nearest neighbor technique developed by Ohmann and Gregory (2002). The approach uses MODIS (MODerate-resolution Imaging Spectroradiometer) composites taken from the entire growing season to take advantage of phenological differences between species, along with climate and topographic variables. The datasets are both modeled and output at a resolution of 250m (6.25 ha, 15.44 acre) grid cells. The technique uses a weighted nearest-neighbor approach, using the 2nd through 7th nearest neighbors, moderated by the proportion of forest pixels from the 2001 National Land Cover Dataset (NLCD2001) within each grid cell. All grid cells have modeled estimates regardless of the proportion of forestland present within them. The approach modeled entire plots, and thus the full suite of variables (volume, individual species basal area, stand size structure, etc.) are essentially modeled together. The plot data used did not record tree data on nonforest plots even if trees occurred (Wilson et al. in review). Knowing basic details about the method used to produce the geospatial dataset being examined provides valuable information about model assumptions, data used, known smoothing applied, characteristics optimized for, etc., that can help interpretation of assessment results and the sources of different types of error found.

RESULTS/DISCUSSION

ASSESSMENT OF MODELED TOTAL BASAL AREA PER ACRE DATASET

Figure 2 presents results from the comparative assessments of data distribution and area estimates across four different spatial scales. From the information provided by the scatterplots and ecdf plots in figure 2, it is apparent that the modeled dataset is closely approximating plot-based estimates for total basal area by the 78,100 hectare scale when the entire dataset is assessed together. Agreement coefficient values are greater than 0.90 by that scale, and KS distance values are very small from the 78,100 to 3.5 million ha scales.

In the choropleth map of local differences between model- and plot-based means at the 216,500 scale (Figure 4), 74 percent of the hexagon means are within the bounds of the 90th CI. Differences larger than that appear to be relatively scattered across the dataset, although there is more tendency for the modeled dataset to overestimate total ba/acre present (23 percent of the hexagons) than underestimate (3 percent of the hexagons), when compared to the plot-based estimate.

The modeled dataset tends to overestimate with respect to plot-based means in areas with low or no total tree ba/acre inventoried by the plots, such as the plains areas in the western side of the study area or southeastern Michigan. This is not surprising, given that basal area is modeled for pixels with tree cover, even if those trees do not fall within FIA's definition of 'forest land.' In the graph of local differences sorted by increasing plot mean (figure 5), there does not seem to be much difference in this pattern across the range of ba/acre values.

With respect to local variability (figure 6), the modeled dataset appears to retain the general pattern of local variability across the study area, but frequently underestimates that variability. This difference will reflect the difference in sample unit size between the two datasets – here between FIA plots measuring the landscape at a 0.06 ha scale, and the modeled dataset describing the landscape at a 6.25 ha scale. However, local variability is considered a sufficiently important characteristic of modeled geospatial datasets to warrant its assessment as a description of the level of local spatial variability present in the modeled dataset, with the plot-based results providing an indication of the smaller-scale variability likely to be present in the real population.

The accuracy presented in figure 2 for the entire area is relatively high, with agreement values at the 78,100 ha scale of 0.95 for AC, 1.0 for AC_{sys} , 0.95 for AC_{uns} , 4.35 (sq. feet per acre) for RMSE, and 0.13 for KS. Figure 7 presents these four assessment metrics for the same 78,100 ha scale by ecoregion and by 3.5 million ha hexagon. It is clear from these results that regional agreement metrics vary widely. For example, while national AC = 0.95, regional AC ranges from less than 0.4 to 1.0. Lowest values predominate in the northern plains region where the lowest total ba/acre is found, however moderately to very low AC values are also found in the northeast and east sections as well. Systematic agreement metric values (AC_{sys}), indicating the level and location of any bias present are much higher overall. However AC_{sys} values still range from 0.76 to 1.0 when calculated by local region, as compared to a national AC_{sys} value of 1.0. Ecoregions with high AC_{uns} values are those with the highest scatter about the GMFR regression line – suggesting those that are currently the most difficult to model given the current set of predictor variables used. Unsystematic agreement (AC_{uns}) values are more similar in range and distribution to AC values, indicating the general dominance of unsystematic error in the overall AC values, with of course a few exceptions. When examined regionally, the magnitude of RMSE values appear to largely track the magnitude of total tree ba/acre present in each local region, with larger errors in areas with higher total ba/acre values. This is entirely understandable given that RMSE values are expressed in data units. The maps of KS distances are

strongly driven by those areas where the plots measured no tree ba/acre and the model estimated some ba/acre greater than zero. Given the fact that FIA plots do not record any tree ba/acre if the plot is defined as ‘nonforest’ even if trees are present, while the data used in the modeling includes tree cover on all lands, it is understandable that this may occur. Thus, it would be helpful if one could calculate the KS distance between the two ecdf’s excluding that difference in the y-intercept, because we may be more interested in differences between the ecdf’s at other places in the plot, rather than the understandable and probably often reasonable differences in the y-intercept due to modeling ba/acre where the plots did not measure any. With the datasets examined here, original data distributions were very closely captured by the modeled dataset, so the only difference was really in the y-intercept. However this is not always the case (see Riemann et al. 2010).

Some differences in results did occur when a different region was used. The most noticeable example was in the systematic agreement (AC_{sys}) maps. Here the northwestern corner of the study area changes from having moderate to relatively high systematic agreement if examined by level 3 ecoregion, to having much lower systematic agreement if examined by 3.5 million ha hexagon. There are several ecoregions that appear very different across many of the maps, such as those along the New York/Pennsylvania border, and a long thin ecoregion down the Appalachian mountains in east central U.S. This is likely due to the small size or long, thin shape of the ecoregions in question, and may be an example of the ecoregions picking up specific areas with different characteristics, while the hexagon includes enough adjacent area to smooth over these differences. Overall, the ecoregion maps indicate that users in the northeast corner may want to improve both the systematic and unsystematic error in many places, whereas the hexagon maps do not draw your attention to that area. Given their equal area and shape, the hexagons may provide a better idea of the spatial patterning of errors, with the cost that errors specifically associated with other region types may not appear as clearly.

ASSESSMENT FOR SPECIES BASAL AREA PER ACRE

Forestland occurs on many FIA plots in United States. Individual tree species, however, represent variables that are much less common. Even sugar maple, a relatively common species, occurs only 7.5 percent of FIA plots in the ECONUS area. Factors affecting how well an individual species is modeled include the number of plots available to model with, whether those plots reflect the full range of variation present over the study area, how dominant that species is where it occurs, and how correlated that species is with respect to the predictor variables used. In this situation, rare species, those with less specific site characteristics,

those in the understory (when working with remotely sensed predictor variables), and those that occur at low densities when they do occur tend to be the most difficult to model accurately when they are modeled independently. One of the characteristics of the nearest neighbor techniques used to generate the modeled datasets being assessed here is that each species is not modeled independently, but rather all species are modeled concurrently. Thus, a relatively rare species which might not have a sufficient number of plots to model well on its own, may achieve a higher accuracy due to its correlation with other species which are more visible or site-specific.

We assessed the modeled ba/acre datasets for six individual tree species, and present four of these species, sugar maple, flowering dogwood, eastern red cedar, and river birch, in more detail in figures 8-11. Results for selected summary metrics for all six species are presented in table 1.

Sugar maple occurs on 7.5 percent of FIA plots in the ECONUS area, has a maximum ba/acre value of 188.7 square feet per acre, and a mean of 14.4 percent ba/acre where it occurs. Figure 8 presents assessment results in terms of the scatterplot across four scales, and the comparison of modeled means to plot-based confidence intervals. The ecdf plot is not shown because it is so dominated by the large number of zero areas over this large an area that it has little story to tell. Assessment results for sugar maple are in general similar to the total ba/acre dataset, with AC values greater than .90 by the 216,500 ha scale, and AC_{sys} values greater than 0.95 by the 8660 ha scale. The percentage of estimates at the 216,500 ha scale falling within, above, and below the 90th plot-based confidence interval are also similar to results for the total ba/acre dataset, although the spatial distribution of those values is of course somewhat different. Modeled estimates for sugar maple at this scale are much more likely to overestimate plot-based estimates in areas where it occurs at lower ba/acre levels, and underestimate plot-based estimates in areas where it occurs at higher ba/acre levels.

Flowering dogwood is an intermediate and understory species that never reaches a very large size. It occurs on only 3.8 percent of ECONUS plots, and of the six species examined it has the lowest maximum ba/acre (42.8 sq. feet per acre) and mean percent basal area/acre (2.6 percent) where it does occur. Yet, despite this, dogwood was relatively well modeled (figure 9, table 1), reaching our target AC and AC_{sys} values set for this study by the 866,025 ha and 78,100 ha scales, respectively. Seventy-five percent of the modeled dataset is within the 90th CI at the 216,500 ha scale, and the model more often overestimates dogwood in the remaining hexagons with respect to the plot-based means. The fact that dogwood does report relatively high accuracies despite its rarity may be due to its correlations

with associated species, although we did not investigate this specifically in time for this study.

Eastern red cedar occurs on 3.2 percent of ECONUS plots, similar to dogwood, although it has higher ba/acre values and represents a larger proportion of the stands where it occurs (figure 10, table 1). Results are similar to dogwood, with the exception that the model much more frequently predicts eastern red cedar in hexagons where the plots record none. Given the habit of cedar to occupy old field locations that may not yet qualify as forestland and thus not be recorded by FIA plots, this may be an example of the model picking up more of the species actually present than the FIA plots are detecting when they record trees on FIA-defined “forestland” only.

River Birch is an example of an extremely rare species, occurring on only 0.4 percent of ECONUS plots (figure 11, table 1). Because of the large number of hexagons without any inventoried or predicted river birch, a high percentage of modeled estimates at the 216,500 ha scale still fall within the 90th CI. However the scatterplots and agreement metrics reveal much higher systematic and random error, reflected in the low AC_{sys} and AC_{uns} (and AC) values, respectively. River birch is an example of a species that has poor overall accuracy in the modeled dataset, probably because of its rarity within the study area, and perhaps in combination with a wide spatial distribution and/or lack of association with other more common species. Regional examination of species assessment results would undoubtedly provide valuable additional information for users and should be added to the standard assessment protocol.

CONCLUSION

Results indicated important regional differences in assessment metrics. For extensive geospatial datasets such as these ECONUS datasets, calculating additional agreement metrics by region better characterizes geographic differences in the magnitude and types of errors present in the modeled geospatial dataset. This may be sufficient basic information for the metadata, particularly when used in combination with dataset-wide scatterplot and ecdf results across several different scales. For application in a specific area, a user may want to additionally examine the scatterplots and ecdf plots at multiple scales for the specific area of interest to gain more insight into accuracy at that location as you move across spatial scales.

There are many factors affecting the accuracy of an individual tree species, one of which is its rarity within the study area. Application of the protocol to individual

species from the modeled pGNN dataset indicates a general tendency toward decreasing accuracy as a species becomes less common, although the threshold seems very low. In this ECONUS-wide assessment, modeled species datasets appeared to be reasonably accurate even when species occurred on only 3–4 percent of the plots, but were substantially less accurate when a species occurred on less than 1 percent of the plots. From our quick examination here of only six species, there did not appear to be a similar relationship between level of accuracy and low basal/area per acre values or low relative dominance. Results of the species assessment provide some indication of the scale(s) at which modeled datasets of rarer variables (e.g. river birch, occurrence of downed wood, etc.) are most consistent with the data from FIA plots. Regional assessment of accuracy will be important with individual species datasets, as assessment results may vary widely from the national values, particularly where a species is locally rare.

This study further develops the minimum information that should be included in the standard metadata available with every FIA geospatial dataset. In addition to indicating the true accuracy of the dataset with respect to the real population on the ground, this assessment protocol provides an explicit description of how summaries generated from a modeled dataset relate to summaries generated directly from the FIA plot data.

ACKNOWLEDGEMENTS

The authors wish to thank Matt Gregory, Mark Nelson, and Kevin Megown for their thoughtful reviews.

LITERATURE CITED

- Blackard, J.A.;** Finco, M.V.; Helmer, E.H. [and others]. 2008. Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sensing of Environment*. 112(4):1658-1677.
- Caners, F.** 1997. Evaluating the uncertainty of area estimates derived from fuzzy land-cover classification. *PE&RS, Photogrammetric Engineering & Remote Sensing*. 63:403-414.
- Foody, G.** 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*. 80(1):185-201.
- Ji, L.;** Gallo, K. 2006. An agreement coefficient for image comparison. *Photogramm. Eng. Remote Sens.* 72(7):823-833.
- Nelson, M.D.;** McRoberts, R.E.; Holden, G.R.; Bauer, M.E. 2009. Effects of satellite image spatial aggregation and resolution on estimates of forest land area. *International Journal of Remote Sensing*. 30(8):1913-1940.

Ohmann, J.L.; Gregory, M.J. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, U.S.A. *Can. J. For. Res.* . 32(4):725-741.

Wilson, B.T.; Lister, A.J.; Riemann, R. in review. A nearest-neighbor imputation approach to large area mapping of tree species distributions using field sampled data. *Remote Sensing of Environment*.

Riemann, R.; Wilson, B.T.; Lister, A.; Parks, S. 2010. An effective assessment protocol for continuous geospatial datasets of forest characteristics using USFS Forest Inventory and Analysis (FIA) data. *Remote Sensing of Environment*. 114(10):2337-2352.

Table 1—Selected assessment results for six individual species, sorted by decreasing agreement (AC and AC_{sys})

species	percent of ECONUS plots on which it occurs	maximum ba/acre value (sq. ft per acre)	mean % ba/acre value where it occurs	scale at which AC > 0.90 (hectares)	scale at which AC _{sys} > 0.95 (hectares)	percent of modeled dataset within 90th CI at 216,500 ha scale	percent of dataset ABOVE 90th plot CI at 216,500 ha scale	percent of dataset BELOW 90th plot CI at 216,500 ha scale
sugar maple	7.5	188.7	14.4	216,500	8660	71	25	3
black cherry	7.9	178.2	5	866,025	78100	68	29	3
flowering dogwood	3.8	42.8	2.6	866,025	78,100	75	24	1
eastern red cedar	3.2	104.6	7.3	866,025	78,100	50	47	2
bitternut hickory	1.6	67.2	5.7	3.5 million	866,025	69	29	1
river birch	0.4	95.6	8.6	AC=0.85 at 3.5 million	3.5 million	67	32	1

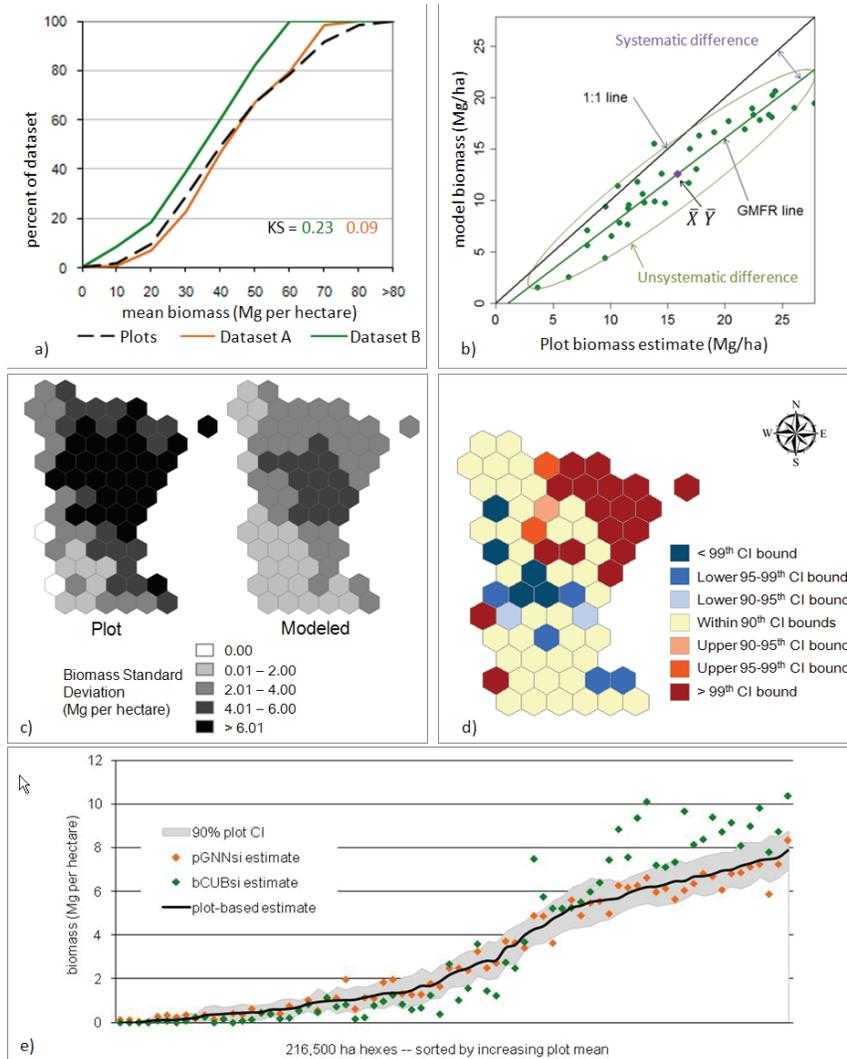


Figure 1—Example assessment protocol used: a) assessment of data distributions with KS distance metrics, b) assessment of agreement between model- and plot-based means—in this example, $AC = 0.80$, $AC_{sys} = 0.84$, and $AC_{uns} = 0.96$, and $RMSE = 3.9$, c) comparing local variability, d) spatial pattern of local differences with respect to plot-based confidence intervals, e) pattern of those differences across the range of biomass values. Example is from assessing modeled datasets of biomass in Minnesota.

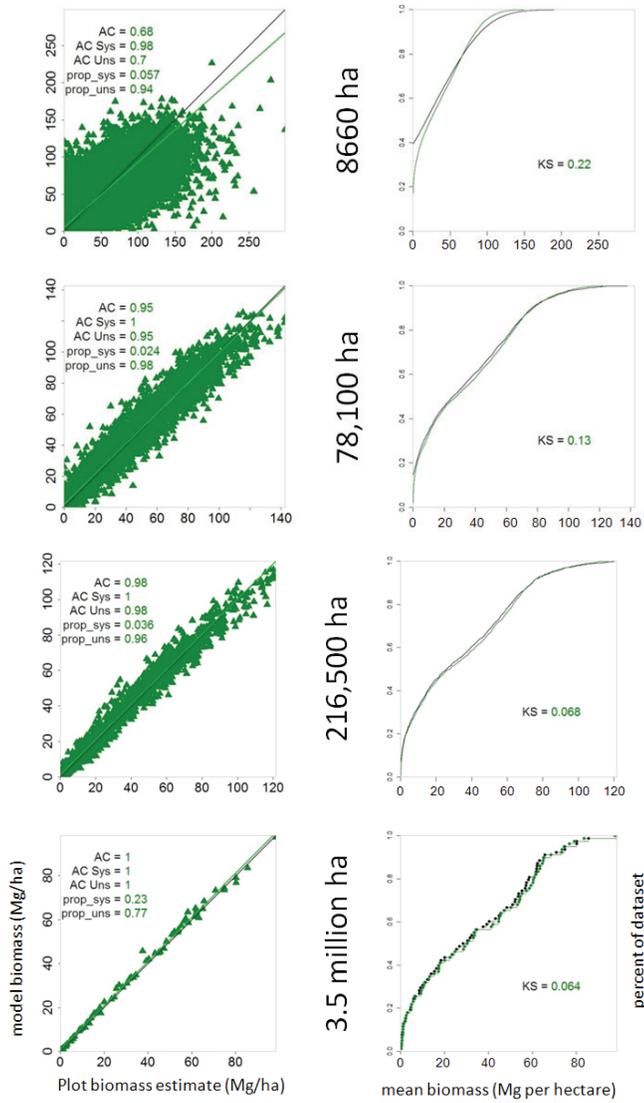


Figure 2—Scatterplots and ecdf plots of ECONUS total ba/acre across four scales.

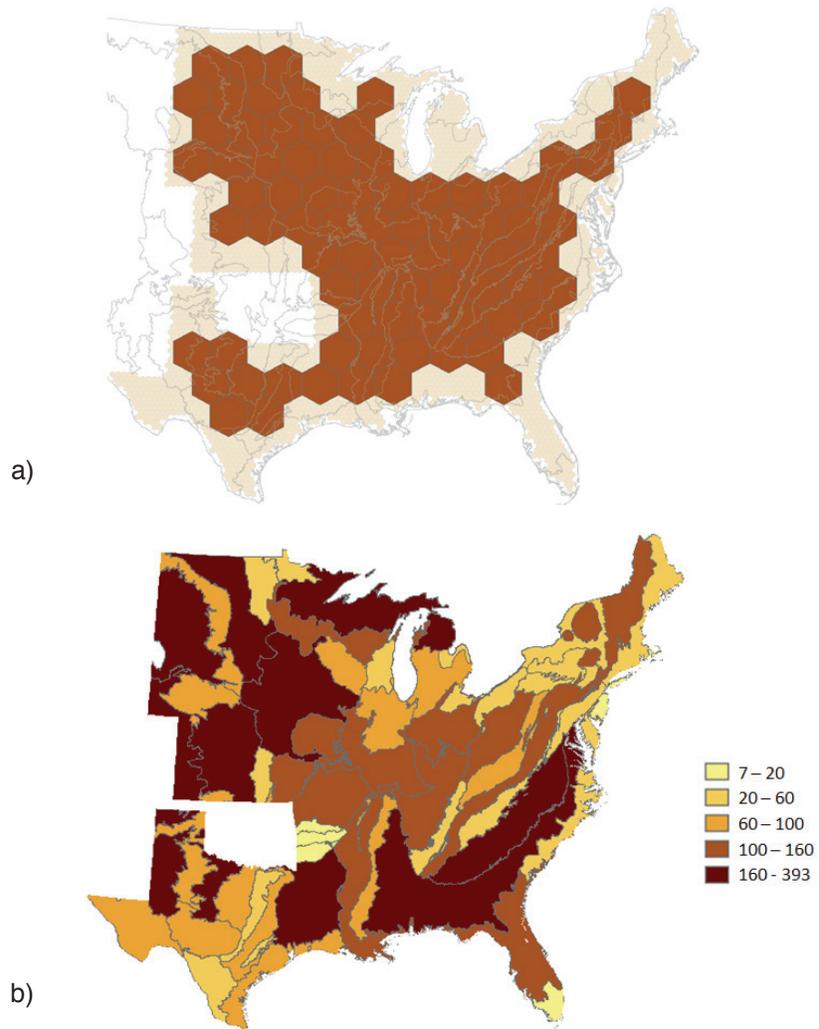


Figure 3—Number of 78,100 ha hexagons within each scale examined: a) 216,500 ha hexagons, where $n = 42-45$, and b) level 3 ecoregions.

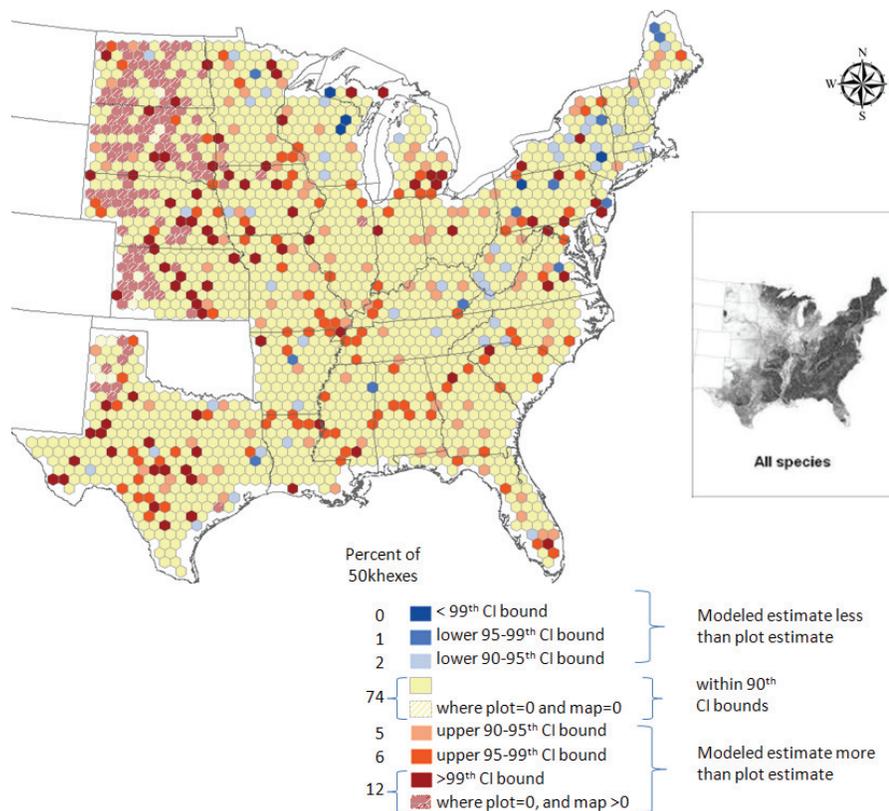


Figure 4—Choropleth map of differences between mapped estimates and plot-based means and confidence intervals for ECONUS at the 216,500 ha scale.

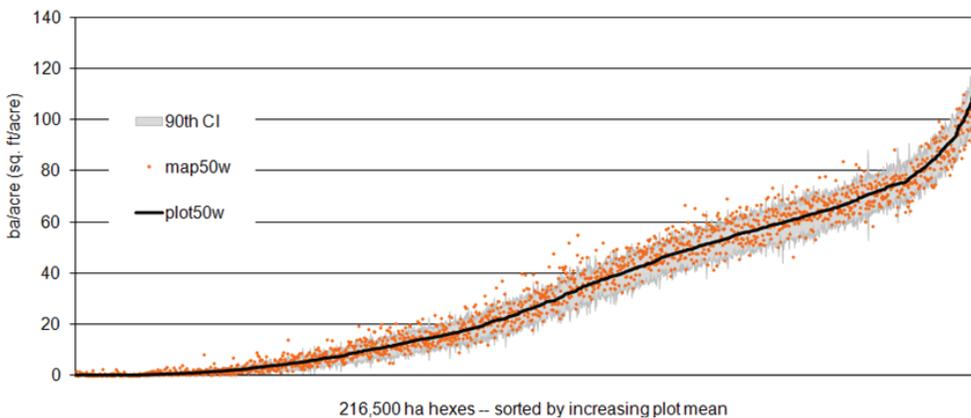


Figure 5—Differences between mapped estimates and plot-based means and confidence intervals for ECONUS at the 216,500 ha scale, as graphed across the distribution of plot mean values.

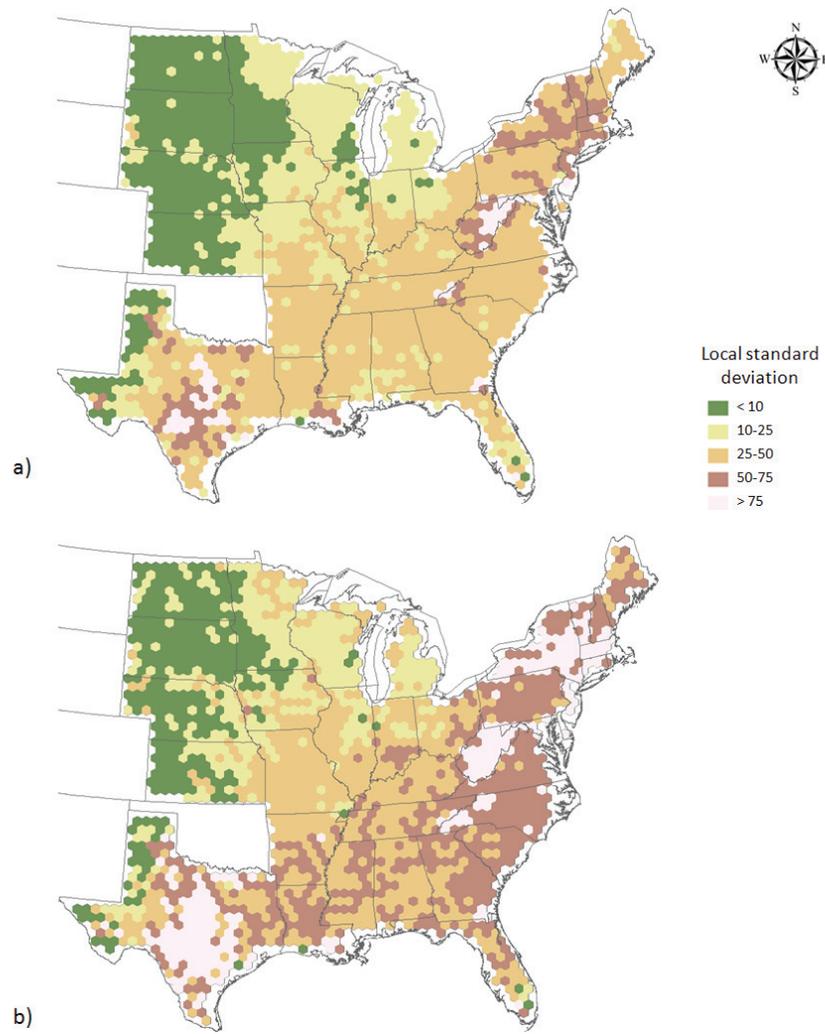


Figure 6—Choropleth map of differences in local spatial variability of ECONUS estimates, as described by the standard deviation of modeled (a) or plot-based (b) values at the 216,500 ha scale.

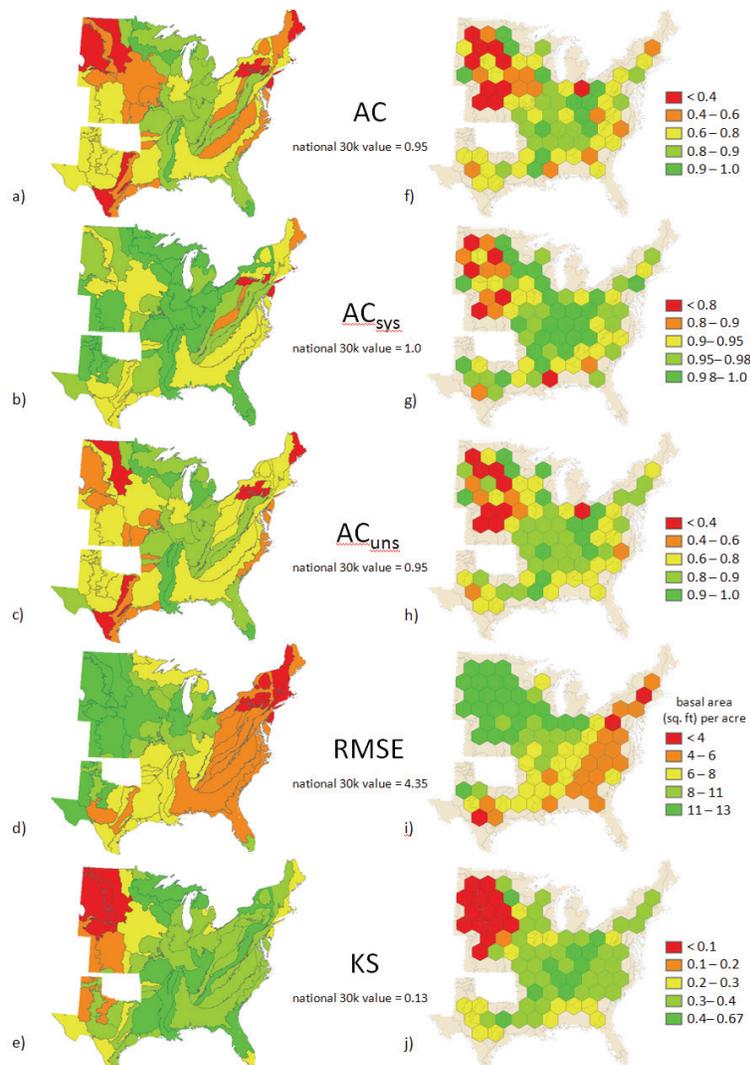


Figure 7—Maps of agreement metrics for the ECONUS dataset, summarized by ecoregion (a-e), and 3.5 million ha hexagon (f-j).

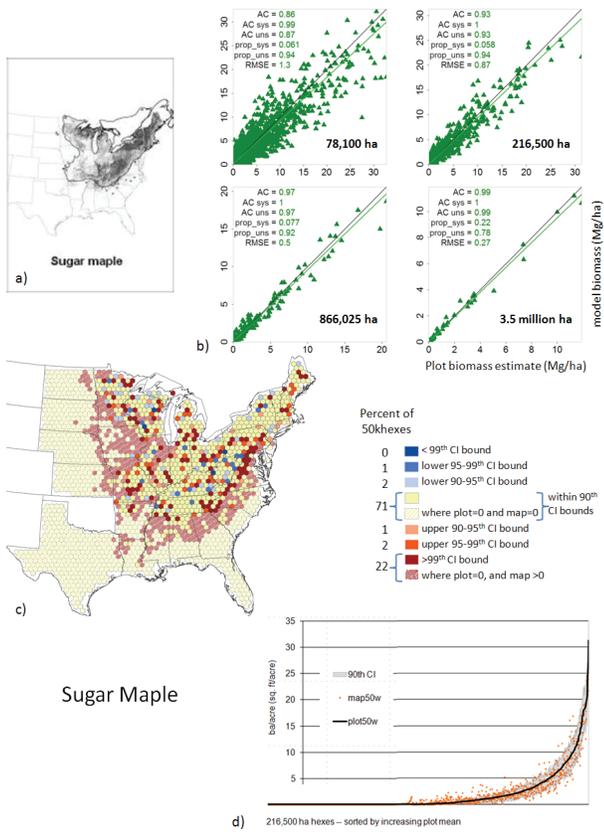


Figure 8—Assessment results for the ECONUS sugar maple dataset: a) map of modeled distribution, b) comparison of model- to plot-based means across four scales, c) magnitude and spatial pattern local differences at the 216,500 ha scale, and d) magnitude and distribution pattern of local differences across the range of sugar maple ba/acre values.

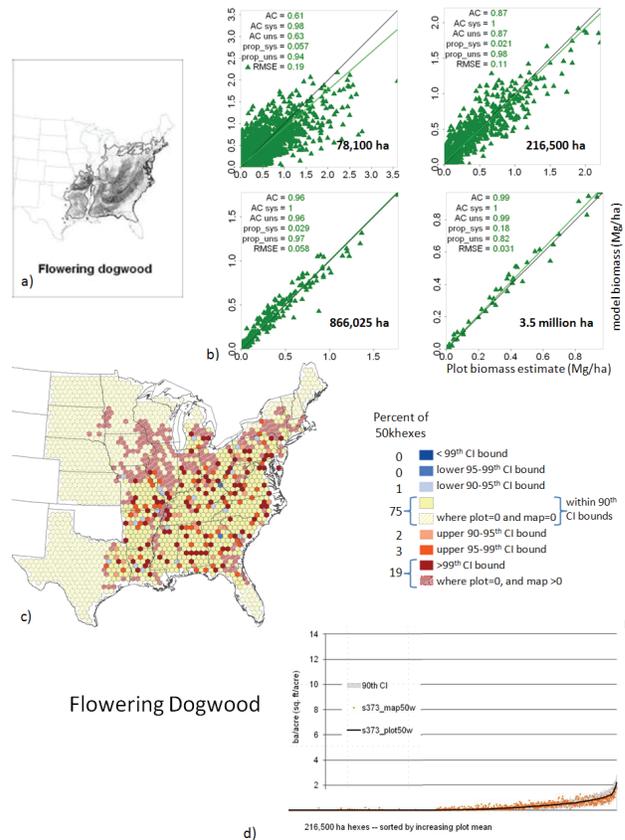


Figure 9—Assessment results for flowering dogwood: a) map of modeled distribution, b) comparison of model- to plot-based means across four scales, c) magnitude and spatial pattern local differences at the 216,500 ha scale, and d) magnitude and distribution pattern of local differences across the range of sugar maple ba/acre values.

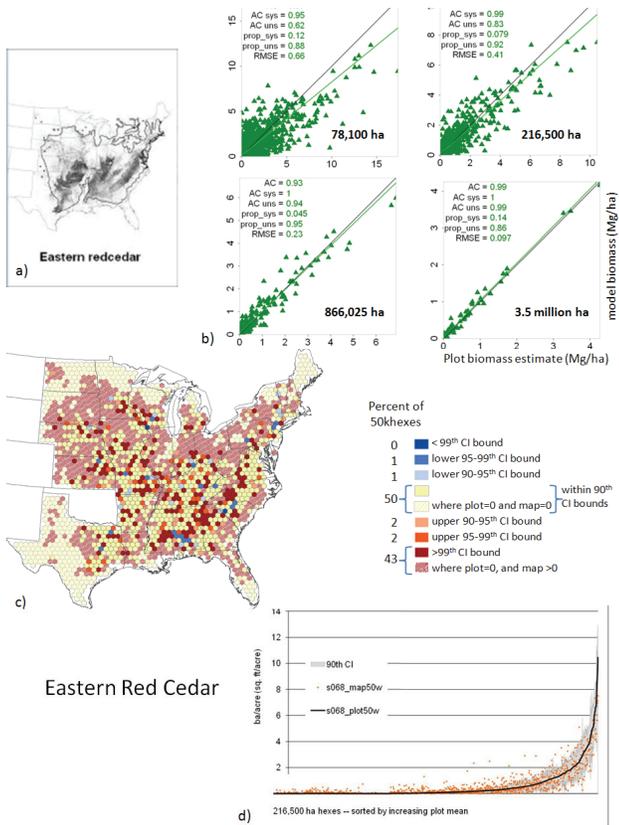


Figure 10—Assessment results for eastern red cedar: a) map of modeled distribution, b) comparison of model- to plot-based means across four scales, c) magnitude and spatial pattern local differences at the 216,500 ha scale, and d) magnitude and distribution pattern of local differences across the range of sugar maple ba/acre values.

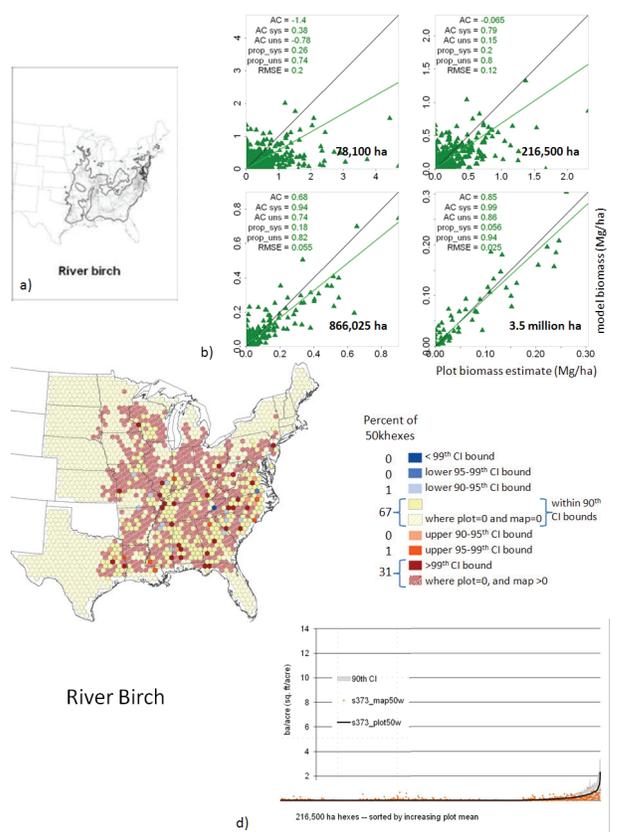


Figure 11—Assessment results for river birch: a) map of modeled distribution, b) comparison of model- to plot-based means across four scales, c) magnitude and spatial pattern local differences at the 216,500 ha scale, and d) magnitude and distribution pattern of local differences across the range of sugar maple ba/acre values.

A DATABASE STRATEGY FOR NEW VARIABLES

B. Tyler Wilson, Ali Conner, Glenn Christensen, John Shaw, Jason Meade, and Larry Royer

ABSTRACT

The introduction of new variables into the annual inventory system of the U.S. Forest Service's Forest Inventory and Analysis (FIA) program can create issues with population estimates since evaluations (or expansion factors) based on a full cycle's worth of data should not be used with new data that have not been collected for a full cycle. This manuscript provides guidance on how to manage evaluations within the National Information Management System Compilation System (NIMS-CS) when new variables/attributes are added to the Forest Inventory and Analysis annual inventory.

INTRODUCTION

In an attempt to be responsive to the changing needs of its users, the U.S. Forest Service's Forest Inventory and Analysis (FIA) program sometimes begins collecting new variables on field plots in mid-cycle. For example, to bring our definition of forest into alignment with international standards, FIA has begun collecting tree canopy cover data on all phase 2 field plots. Because only a fraction of field plots are sampled annually (i.e., 10-20 percent, depending upon the state), there will be a period of time that elapses before a full cycle of data is available for these new variables (i.e., 5-10 years). For the purposes of population estimation (Bechtold et al. 2005), all plots collected within a stratum of an estimation unit for an evaluation period are assigned the same plot expansion factor. Expansion factors are computed by dividing the acreage of each stratum within an estimation unit by the numbers of plots sampled in the stratum for the evaluation period. For new variables that have not yet been collected for an entire cycle, the expansion factors associated with a full cycle of plots will be incorrect. Therefore, FIA needs a strategy to handle these differences in plot expansion factors amongst variables prior to the completion of the first full cycle of sampling.

ORGANIZATIONAL STRUCTURE

The first step in the solution was to identify FIA staff to approach the problem from the perspective of each of the regional programs and the functional areas impacted: information management, techniques research, and analysis. Information management staff process the data collected in the field and maintain the FIA database. Staff of techniques research ensure that the field sample is collected and compiled in such a way to permit meaningful population estimation. Analysts interpret the population estimates and produce annual and 5-year reports on the status of forests in their region of expertise. The authors of this manuscript represent each of these regions and functional areas and were identified as the task team. The second step was for the team to develop and analyze a small set of alternative solutions to the problem. In the final step, the team recommended a strategy for handling new variables to the program managers. In this paper, we present a synopsis of the team's analysis and recommendation.

ALTERNATIVE STRATEGIES

The team proposed and examined three alternative strategies for handling new variables. The first approach is to wait until a full cycle of plots is collected before reporting on a new variable. The benefit of this approach is that it requires no additional effort beyond that required to collect, compile, store, analyze, and report these new variables. However, there are a few drawbacks. Waiting until a complete cycle has been collected would result in unacceptable delays (i.e., between 5 years and 10 years) prior to reporting. This would foster the perception of "gate-keeping" by FIA and would not enhance recent efforts by the organization to promote transparency in its methods.

B. Tyler Wilson, Research Forester, Northern Research Station, Saint Paul, MN 55108
Ali Conner, Supervisory, IT Specialist, Southern Research Station, Knoxville, TN 37919
Glenn Christensen, Forester, Pacific Northwest Research Station, Portland, OR 97205
John Shaw, Forester, Rocky Mountain Research Station, Ogden, UT 84401
Jason Meade, Forester, Southern Research Station, Knoxville, TN 37919
Larry Royer, IT Specialist, University of Nevada, Las Vegas, NV 89154

Table 1—Summary of worked example

RPT YR	Data Example	EVAL_GRP	VALID	EVAL_TYP	Avg. Exp. Factor	# of Plots	
2005	AL8_12345 (2001, 2002, 2003, 2004, 2005) (full cycle data)	012005	010550	EXPALL	6,000	4,000	
			010551	EXPCURR	6,000	4,000	
				EXPVOL	6,000	4,000	
			010553	EXPGROW	6,000	4,000	
				EXPMORT	6,000	4,000	
	010553	EXPREMV	6,000	4,000			
2006	AL8_2345 + AL9_1 (2002, 2003, 2004, 2005, 2006)	012105	010510	EXPALL	30,000	800	
			010511	EXPCURR	30,000	800	
				EXPVOL	30,000	800	
2006	AL8_2345 + AL9_1 (2002, 2003, 2004, 2005, 2006)	012006	010650	EXPALL	6,000	4,000	
			010651	EXPCURR	6,000	4,000	
				EXPVOL	6,000	4,000	
			010653	EXPGROW	6,000	4,000	
				EXPMORT	6,000	4,000	
	010653	EXPREMV	6,000	4,000			
2007	Variable (A) AL8_5 + AL9_1 (2005, 2006)	012206	010620	EXPALL	15,000	1,600	
			010621	EXPCURR	15,000	1,600	
				EXPVOL	15,000	1,600	
2007	AL8_345 + AL9_12 (2003, 2004, 2005, 2006, 2007)	012007	010750	EXPALL	6,000	4,000	
			010751	EXPCURR	6,000	4,000	
				EXPVOL	6,000	4,000	
			010753	EXPGROW	6,000	4,000	
				EXPMORT	6,000	4,000	
	010753	EXPREMV	6,000	4,000			
2008	Variable (A) AL8_5 + AL9_12 (2005, 2006, 2007)	012307	010730	EXPALL	10,000	2,400	
			010731	EXPCURR	10,000	2,400	
				EXPVOL	10,000	2,400	
2008	AL8_45 + AL9_123 (2004, 2005, 2006, 2007, 2008)	012008	010850	EXPALL	6,000	4,000	
			10851	EXPCURR	6,000	4,000	
				EXPVOL	6,000	4,000	
			010853	EXPGROW	6,000	4,000	
				EXPMORT	6,000	4,000	
	010840	EXPREMV	6,000	4,000			
	2008	Variable (A) AL8_5 + AL9_123 (2005, 2006, 2007, 2008)	012408	010840	EXPALL	7,500	3,200
				010841	EXPCURR	7,500	3,200
					EXPVOL	7,500	3,200
010810				EXPALL	30,000	800	
2008	New Variable (B) AL9_3 (2008)	012108	010810	EXPALL	30,000	800	
			010811	EXPCURR	30,000	800	
				EXPVOL	30,000	800	
2009	AL8_5+AL9_1234 5 panels inc Var (A) (2005, 2006, 2007, 2008, 2009)	012009	010950	EXPALL	6,000	4,000	
			010951	EXPCURR	6,000	4,000	
				EXPVOL	6,000	4,000	
			010953 No Var A	EXPGROW	6,000	4,000	
				EXPMORT	6,000	4,000	
	010953	EXPREMV	6,000	4,000			
2009	Variable (B) AL9_34 (2008, 2009)	012209	010920	EXPALL	15,000	1,600	
			010921	EXPCURR	15,000	1,600	
				EXPVOL	15,000	1,600	
2010	AL9_12345 (2006, 2007, 2008, 2009, 2010)	012010	011050	EXPALL	6,000	4,000	
			011051	EXPCURR	6,000	4,000	
				EXPVOL	6,000	4,000	
			011053	EXPGROW	6,000	4,000	
				EXPMORT	6,000	4,000	
	011053	EXPREMV	6,000	4,000			
	2010	Variable (A) AL9_5 (2010)	012110	011013	EXPGROW	30,000	800
				011013	EXPMORT	30,000	800
					EXPREMV	30,000	800
011030				EXPALL	10,000	2,400	
2010	Variable (B) AL9_345 (2008, 2009, 2010)	012310	011030	EXPALL	10,000	2,400	
			011031	EXPCURR	10,000	2,400	
				EXPVOL	10,000	2,400	

The second approach is to modify the code in the reporting tools that calculates estimates to adjust stored expansion factors dynamically to account for incomplete cycles of data for new variables. The benefit is that the underlying data and database structure would remain unchanged. But this approach also has several issues. It would require some complex programming to encode the necessary logic. Furthermore, such business logic is most appropriately stored in the National Information Management System Compilation System (NIMS-CS), along with all of the other compilation procedures. Finally, the sampling errors of estimates produced using less than a full cycle of plots would be larger because of the smaller numbers of plots in the sample, though this problem goes away once a full cycle of plots is collected.

The third approach is to create separate evaluations for new variables within the NIMS-CS. The benefits are that the database structure would not require any modification other than the additional records to be created, which is true of any evaluation. The drawbacks are that this approach requires slightly more compilation time and the need to maintain more records in the population (POP) tables, and thus possible confusion for which evaluation to choose. It also results in larger sampling errors before the collection of a full cycle of plots. However, both of these problems are eliminated once the cycle for the new variable is complete. Because of the simplicity of the approach and the benefit of reporting on new variables prior to the collection of a full cycle of data, in spite of slightly larger sampling errors in the interim, the third approach was recommended for handling new variables. The application of this approach is illustrated in the next section using a specific example.

EXAMPLE USING RECOMMENDED STRATEGY

For the purposes of clarity, let us make a few simplifying assumptions for the example. Assume that FIA will start collecting new variables on phase 2 plots in Alabama, which is on a 5-year cycle, with 800 plots sampled in an inventory year and 4,000 plots sampled in a full cycle. Assume that the average expansion factor per plot is 6,000 at base sampling intensity for a full cycle of 5 inventory years, giving a total of 24 million acres sampled.

We will assume that a full cycle of data is available for Alabama in 2005, which is comprised of 5 panels of plots sampled during inventory years 2001, 2002, 2003, 2004, and 2005. We will designate an evaluation group (12005) as a reference to this set of plots. This evaluation group includes three evaluations: EXPALL (10550), EXPCURR (10551), and EXPGRM (10553). The EXPALL evaluation includes all of the plots, whether sampled or not. Plots may

not be sampled because of hazardous conditions or denied access. The EXPCURR evaluation includes only those plots that were sampled, either via a field or office visit. The EXPGRM evaluation includes only those plots that were sampled at two points in time, thus allowing the calculation of the components of change, broadly categorized as growth, removals, and mortality.

In this example, new variable A is introduced for the 2005 field season. In order to account for the fact that this variable has been collected on only 800 plots and therefore requires a different expansion factor (24 million acres / 800 plots = 30,000 acres/plot) than those variables collected on a full cycle of plots, a second evaluation group is needed (12105). This evaluation group will include two evaluations: EXPALL (10510) and EXPCURR (10511), corresponding to the first two evaluations in the first evaluation group, but using only one panel of plots.

In 2006, all variables are collected on the plots in the panel, including new variable A. For 2006, Alabama again requires two evaluation groups (12006, 12206). Evaluation group 12006 includes plots from 5 inventory years (2002-2006). This evaluation group is comprised of three evaluations: EXPALL (10650), EXPCURR (10651), and EXPGRM (10653). Evaluation group 12206 is created for new variable A, which has now been collected in inventory years 2005 and 2006 on 1,600 plots. This works out to an expansion factor of 15,000 acres/plot. This evaluation group is comprised of two evaluations: EXPALL (10620) and EXPCURR (10621), corresponding to the first two evaluations in the first evaluation group, but using only two panels of plots.

New variable B is introduced in 2008. Because there are different sampling intensities for the original variables, new variable A, and new variable B, Alabama 2008 requires three evaluation groups (12008, 12408, 12108). Evaluation group 12008 includes plots from 5 inventory years (2004-2008). It consists of three evaluations: EXPALL (10850), EXPCURR (10851), and EXPGRM (10853). Evaluation group 12408 is created for new variable A, which has now been collected on approximately 3,200 plots in the inventory years 2005-2008. This works out to an expansion factor of 7,500 acres/plot. This evaluation group is comprised of two evaluations: EXPALL (10840) and EXPCURR (10841), corresponding to the first two evaluations in the first evaluation group, but using only four panels of plots. Evaluation group 12108 is created for new variable B, which has been collected on approximately 800 plots in inventory year 2008, for an average expansion factor of 30,000 acres/plot. It consists of two evaluations: EXPALL (10810) and EXPCURR (10811), corresponding to the first two evaluations in the first evaluation group, but using only one panel of plots.

In 2009, Alabama requires only two evaluation groups (12009, 12209). The reason for this is that new variable A has now been collected for a complete cycle of plots (2005-2009). Evaluation group 12009 includes all of the original variables, as well as new variable A. This evaluation group consists of three evaluations: EXPALL (10950), EXPCURR (10951), and EXPGRM (10953), though variable A is not included in EXPGRM since it has not yet been remeasured on any plots. Variable B has now been collected in inventory years 2008 and 2009 on approximately 1,600 plots, for an average expansion factor of 15,000 acres/plot. Variable B belongs to evaluation group 12209, which consists of two evaluations: EXPALL (10920) and EXPCURR (10921), corresponding to the first two evaluations in the first evaluation group, but using only two panels of plots.

In 2010, the last inventory year of our example, Alabama requires three evaluation groups (12100, 12110, 12310). Evaluation group 12100 includes plots from 5 inventory years (2006-2010). It consists of three evaluations: EXPALL (11050), EXPCURR (11051), and EXPGRM (11053). Evaluation group 12110 includes only one evaluation, EXPGRM (11013), which includes approximately 800 plots that have been measured for variable A at two points in time: 2005 and 2010. This works out to an expansion factor of 30,000 acres/plot. Variable B has now been collected in inventory years 2008-2010 on approximately 2,400 plots, for an average expansion factor of 10,000 acres/plot. Variable B belongs to evaluation group 12310, which consists of two evaluations: EXPALL (11030) and EXPCURR (11031), corresponding to the first two evaluations in the first evaluation group, but using only three panels of plots.

It should be apparent from this example that the maximum number of evaluation groups required to handle new variables is equal to the cycle length, which is 5 years in the eastern states and 10 years in the western states. It should also be apparent that these “extra” evaluations are no longer necessary once a new variable has been collected on a full cycle of plots.

REMAINING ISSUES

As was mentioned earlier, the recommended approach will result in FIA’s reporting tools computing higher sampling errors of estimates for new variables in the interim period prior to the completion of a complete cycle of data collection, with the sampling errors being especially large for the first panel. There would appear to be two ways of dealing with this issue. One is to permit reporting on new variables immediately upon the collection and compilation of the first panel of plots. In this case, it is recommended that the reporting tools somehow highlight the fact that the estimates are not based upon a full cycle of data, as well as compute the estimates and sampling errors. The other method would be to permit reporting only after a threshold percentage of plots (e.g. 40 percent or 60 percent) are collected for a new variable. This would somewhat reduce the initial sampling errors, but the reporting tools should once again highlight that fact.

LITERATURE CITED

Bechtold, W.A.; Patterson, P.L., eds. 2005. The enhanced Forest Inventory and Analysis program—national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p.

Cover Estimation

POTENTIAL APPLICATIONS OF PREFIELD LAND USE AND CANOPY COVER DATA: EXAMPLES FROM NONFOREST AND NONSAMPLED FOREST INVENTORY PLOTS Sara A. Goeking.....	181
REPEATABILITY IN PHOTO-INTERPRETATION OF TREE CANOPY COVER AND ITS EFFECT ON PREDICTIVE MAPPING Thomas A. Jackson, Gretchen G. Moisen, Paul L. Patterson, and John Tipton.....	189
A FRAMEWORK FOR REPORTING TREE COVER ATTRIBUTES IN AGRICULTURAL LANDSCAPES Dacia M. Meneguzzo and Greg C. Liknes.....	193
CHOOSING APPROPRIATE SUBPOPULATIONS FOR MODELING TREE CANOPY COVER NATIONWIDE Gretchen G. Moisen, John W. Coulston, Barry T. Wilson, Warren B. Cohen, and Mark V. Finco.....	195
SAMPLING INTENSITY AND NORMALIZATIONS: EXPLORING COST-DRIVING FACTORS IN NATIONWIDE MAPPING OF TREE CANOPY COVER John Tipton, Gretchen Moisen, Paul Patterson, Thomas A. Jackson, and John Coulston.....	201
ASSESSING ALTERNATIVE MEASURES OF TREE CANOPY COVER: PHOTO-INTERPRETED NAIP AND GROUND-BASED ESTIMATES Chris Toney, Greg Liknes, Andy Lister, and Dacia Meneguzzo.....	209
A TOOL TO DETERMINE CROWN AND PLOT CANOPY TRANSPARENCY FOR FOREST INVENTORY AND ANALYSIS PHASE 3 PLOTS USING DIGITAL PHOTOGRAPHS Matthew F. Winn and Philip A. Araman.....	217
AN ALTERNATIVE METHOD FOR ESTIMATING CROWN CHARACTERISTICS OF URBAN TREES USING DIGITAL PHOTOGRAPHS Matthew F. Winn and Philip A. Araman.....	223
COMPARISON OF LIDAR- AND PHOTOINTERPRETATION-BASED ESTIMATES OF CANOPY COVER Demetrios Gatziolis.....	231
COMPARING ALTERNATIVE TREE CANOPY COVER ESTIMATES DERIVED FROM DIGITAL AERIAL PHOTOGRAPHY AND FIELD-BASED ASSESSMENTS Tracey S. Frescino and Gretchen G. Moisen.....	237

POTENTIAL APPLICATIONS OF PREFIELD LAND USE AND CANOPY COVER DATA: EXAMPLES FROM NONFOREST AND NONSAMPLED FOREST INVENTORY PLOTS

Sara A. Goeking

ABSTRACT

The Forest Inventory and Analysis (FIA) prefield workflow involves interpreting aerial imagery to determine whether each plot in a given inventory year may meet FIA's definition of forest land. The primary purpose of this determination is to minimize inventory costs by avoiding unnecessary ground surveys of plots that are obviously in nonforest areas. Since the initiation of the annual forest inventory, prefield data collection has consisted primarily of a simple visit/non-visit determination, along with a few regionally inconsistent ancillary variables. Therefore very little information was recorded for nonforest areas with trees, such as recreational developments and urban forests. Beginning in the 2012 inventory year, a nonforest land use code and a continuous tree canopy cover value will be implemented for all non-visited and non-sampled plots. The purpose of this paper is to describe the new prefield protocol and illustrate potential applications of the new variables.

INTRODUCTION

Throughout most of the nation, the Forest Inventory and Analysis Program (FIA) employs photo-interpretive techniques to identify plots in the FIA sample grid that are undoubtedly nonforest and thus do not require inclusion in the field inventory (Reams et al. 2005). The process of making these visit/non-visit determinations is referred to as FIA prefield operations, and its purpose is to reduce field costs by avoiding ground surveys of plots that are obviously nonforest. The FIA definition of forest land has both a land use and a tree cover component. Prefield photo interpreters examine each plot location on the most recent high-resolution aerial imagery and evaluate whether trees are present and whether the plot is subject to a land use that precludes it from meeting the FIA definition of forest land. Plots that are clearly and undoubtedly nonforest are removed from the field inventory and recorded as office-generated nonforest plots.

A visual assessment of the land use and tree cover of each FIA plot location is inherent to prefield visit/non-visit determinations. However, since the initiation of the annual forest inventory, prefield data collection has consisted

primarily of a simple visit/non-visit determination, along with a few regionally inconsistent ancillary variables. Nonforest land use has been recorded as a regional variable in all four FIA regions: Interior West (IW), Northern Research Station (NRS), Pacific Northwest (PNW and PNW-AK), and Southern Research Station (SRS). However, this variable is not part of the national protocol and consequently there are gaps in the national database. Very little information has been recorded for nonforest areas with trees, such as recreational developments and urban forests. Therefore the potential for prefield operations to provide additional information, such as estimates of nonforest lands with trees, has not been realized.

FIA TERMINOLOGY

FIA terminology regarding Phase 1 and Phase 2 are often confused, especially with regard to prefield operations. FIA's national sample grid is referred to as the Phase 2 (P2) grid, and prefield photo interpretation is considered part of this P2 inventory (Reams et al. 2005). In contrast, Phase 1 (P1) of the FIA inventory uses photo-interpreted point data to classify satellite imagery, produce forest/nonforest strata, and develop regional estimation statistics (see Bechtold and Patterson 2005 for details on P1 stratification). The photo-interpretive aspects of prefield operations resemble those used for FIA P1 stratification, but not all regions use P2 plot locations to develop their P1 strata. Furthermore, the land use that is recorded for P2 plots may or may not be consistent with the P1 strata.

Within the FIA P2 inventory, a certain percentage of plots is sampled each year, and the number of years required to sample all P2 plots is referred to as a cycle. For FIA regions with a five-year cycle, each inventory year is referred to as a panel and consists of 20 percent of all P2 plots (e.g., in the majority of the NRS and SRS FIA regions). For regions with a 10-year cycle, each inventory year is referred to as a subpanel and consists of 10 percent of P2 plots (IW and PNW FIA regions).

THE PREFIELD INVENTORY: NEW VARIABLES AND METHODS

Beginning with the 2012 panel, the prefield data collection will include nonforest land use and tree canopy cover variables in addition to the basic visit/non-visit determination. These variables will be populated for each P2 plot that meets one of two criteria: 1) the plot was designated as a non-visit plot, or 2) the plot was not sampled due to hazardous conditions, physical inaccessibility, or denial of access by the property owner or manager. Non-visit plots will be populated prior to field data collection for that panel, while non-sampled plots will be populated post-season. Interpreters will also record metadata for the aerial imagery used to populate land use and canopy cover; in most cases interpretations will be based on 1-m leaf-on NAIP imagery, which is collected on a 3-year return cycle (USDA Farm Service Agency 2009). Data quality will be appraised in terms of the repeatability of prefield assessments as well as prefield (i.e., office-based) vs. field assessments. The specific methods and decision criteria used to populate new variables are described below.

NONFOREST LAND USE

Prefield interpreters will record a single nonforest land use class for all non-visit and non-sampled plots, based on the land use that occurs at plot center. Land use is the apparent intent of human activity on the land, as evidenced by land cover, and consistent with the land use classification used by field crews (see list of classes in Table 1). The geometric requirements for the nonforest land use at plot center are the same as those for field-based mapping of conditions (i.e., must be at least 120 feet wide and at least 1 acre in size), and the same exceptions apply (e.g., windbreaks, rights-of-way, etc., are not subject to the minimum size requirements; see USDA 2007). If the nonforest land use at plot center does not meet these requirements, then the land use occupying the majority of the 144' circular plot area will be recorded.

PERCENT CANOPY COVER

Canopy cover is defined as the proportion of the ground surface within a given area that is covered by a vertical projection of all tree crowns minus the area of crown overlap, i.e., overlapping crowns are not double-counted (cf. Jennings et al. 1999). A plant's crown is considered a tree crown if it is an FIA tally tree species; there is no minimum size requirement on individual trees other than their ability to be seen on aerial imagery. Trees whose boles fall outside the plot area are included in canopy cover assessments if their canopies extend into the plot area.

Tree canopy cover will be assessed using dot-count methodology on all non-visited and non-sampled P2 plots

as part of the regular prefield workflow for each panel. Pilot studies have shown dot-counts to be an accurate and efficient method of assessing canopy cover relative to other methods, including image segmentation and ocular canopy assessment (Frescino et al. 2011, Goeking and Liknes 2009). The sample unit will consist of a circle with 144-foot radius around the plot center, and will completely encompass all four subplots of the field plot design. Within each circle there will be 109 systematically distributed dots. The primary axes of the dot grid will be rotated 15 degrees in a clockwise direction from true north to avoid potential alignment with linear features that may be oriented east-west or north-south, such as windbreaks or rights-of-way.

Under the 2012 protocols, prefield interpreters will not assess canopy cover at plots that are designated as visit plots and are sampled by field crews, with the exception of plots that are selected for quality assessment purposes (see below). Therefore, to ensure that the canopy cover variable is populated for all P2 plots, field crews will record canopy cover on all field-sampled plots, including plots that the field crew determines to be nonforest. Additionally, canopy cover on forested plots where trees are tallied will be modeled from stem-map data (see Toney et al. 2009).

QUALITY ASSESSMENT

Quality assessment of prefield data collection has two primary objectives: 1) to assess the relationship between field and prefield canopy cover measurements, and 2) to estimate the repeatability of photo-interpreted land use and canopy cover determinations. Samples will be drawn from each year's P2 panel to meet these objectives. To meet the first objective, a random sample of four percent of all field-visited plots in each panel, by state, will be designated for photo interpretation by a prefield interpreter. Canopy cover data collected by prefield interpreters will be compared to that collected by field crews. To meet the second objective, a second prefield interpreter will examine four percent of all non-visit plots in each year's panel and record nonforest land use, where one exists, as well as tree canopy cover. If a four percent sample equals more than 30 plots for either objective, then only 30 plots will be required for quality assessment.

INTERIOR WEST PILOT DATA: APPLICATIONS OF LAND USE AND TREE CANOPY COVER DATA

PILOT METHODS

The purpose of the Interior West pilot study was to investigate the feasibility of incorporating additional photo-interpreted variables into the prefield data collection process. Interior West FIA prefield specialists recorded land

use and ocular canopy cover data at all P2 plots in the 2009, 2010, and 2011 subpanels (Figure 1). Note that each year's sample is referred to as a subpanel, rather than a panel, because the Interior West FIA samples ten percent of all P2 plots each year and requires ten years to complete a cycle. Photo interpretations were based on the most recent 1-m NAIP imagery for each state. The subpanels for inventory years 2009-2011 consisted of 28,196 P2 plots across eight states, of which 16,503 (59 percent) were designated as non-visit plots. The high percentage of non-visit plots is due to the vast areas of rangeland without trees in the Interior West, as well as smaller percentages of plots that exist in areas with nonforest land uses such as agricultural and cultural development.

PILOT RESULTS

Quality assessment—As mentioned above, one objective of prefield quality assessment, beginning with the 2012 panel, will be to assess field versus prefield canopy cover data. The Interior West prefield pilot included both field and prefield canopy cover data for the 2009 subpanel, enabling quantitative comparison of the two metrics. Field crews recorded ground-based ocular assessments of tree canopy cover for 2,928 plots across five states, as well as transect data for 2,260 of those plots. The transect method consists of four transects on each of the four subplots in the FIA plot design. Each of these 16 transects is 25 feet long, where presence/absence of tree canopy cover is observed every foot. Prefield photo-interpreted canopy cover data exist for all of the plots where field data were collected. Correlations between prefield and field measurements were weak to moderate for both prefield versus ground-based ocular measurements ($r=0.62$) and prefield versus ground-based transects ($r=0.58$).

Ocular assessments of canopy cover using NAIP imagery have been previously shown to be highly variable among users and to overestimate canopy cover relative to less subjective methods such as dot counts and image classification (Frescino and Moisen 2011, Goeking and Liknes 2009). Therefore, the pilot data in this study are likely to be less accurate than canopy cover data collected after dot-count procedures are implemented in the 2012 subpanel. The sample applications of these data are provided below for illustrative purposes only.

Pilot Applications of Land Use Data—The most basic application of the land use data is a simple summary of the number of plots, and the corresponding number of acres represented by those plots, in each nonforest land use class. Table 1 shows the distribution of all non-visit P2 plots in the Interior West. Rangeland (land use code=20) occupies roughly 75 percent of all non-visit plots, and cropland (land use code=11) occupies about 11 percent. Each P2 plot represents slightly more than 5,900 acres

(Woudenberg et al., in press), assuming that an entire cycle has been sampled. Because we have only 30 percent of a cycle included in this pilot study, each plot represents about 19,667 acres. Multiplying the number of plots in each category by this expansion yields a conservative estimate of the total number of acres in that category.

Pilot Applications of Canopy Cover Data—When the new prefield variables are fully implemented in the 2012 inventory year, they will be collected only for non-visit and non-sampled plots. However, for the purposes of this pilot study they were collected for all P2 plots. Therefore, we can compare the frequency distribution of canopy cover at all plots within a state with the distribution of canopy cover at non-sampled plots. Figure 2 shows the distribution of canopy cover at plots that were not sampled due to either denial of access by the land owner or manager, or field crews' decisions that the plots were too hazardous to be sampled safely.

Whether plots are non-sampled due to an absence of a statewide inventory or due to the inability of the field crew to safely and legally access a plot, the assessment of canopy cover at non-sampled plots provides more information than FIA previously collected on those plots. Current FIA estimation procedures assume that non-sampled plots are randomly distributed among the Phase 1 strata, yet this assumption can increase the error of FIA's estimates, particularly in areas with low percentages of forested area (Bechtold and Scott 2005). Future work should focus on the potential incorporation of prefield canopy cover data into statistical estimators of forest area.

Pilot Applications of Combined Land Use and Canopy Cover Data—Land use and canopy cover data can be combined to provide information about the distribution of tree canopy cover among nonforest land uses, to estimate the area of tree cover in nonforest areas, and to estimate the number of plots in inventories of lands that do not meet the FIA definition of forest, such as urban forest or protective forest. Table 1 shows the mean canopy cover of each nonforest land use category. Windbreaks/shelterbelts have the highest mean canopy cover, followed in descending order by Cultural/Urban, Nonforest/chaparral, Recreation, Rights-of-Way, and Rangeland.

Although the Rangeland category averages only two percent canopy cover (Table 1), the large number of such plots and the commensurate acreage they represent indicates that trees on nonforest rangelands may account for a substantial amount of biomass and carbon in tree form. From Table 2, an entire ten-year cycle is estimated to include 2,503 Rangeland plots with canopy cover greater than ten percent. These plots likely do not meet the FIA definition of forest because the trees on the plot do not occupy more than one

acre or do not constitute a stand greater than 120 feet wide. However, these data suggest that substantial areas of the Interior West contain trees yet are too sparsely wooded to qualify as forest using traditional FIA field survey methods. Prefield data collection based on photo interpretation may be able to augment the FIA field inventory to quantify the area covered by nonforest land with trees.

Table 2 also illustrates the estimated number of plots in several other nonforest categories of interest. For example, an urban forest inventory in the eight Interior West states is likely to include about 50 plots in an entire ten-year cycle, or five plots per year. An inventory of all plots with greater than ten percent canopy cover in areas with any developed land use would consist of 110 plots per cycle (11 plots per year), while an inventory of agricultural areas would include 50 plots per cycle (five plots per year).

FUTURE APPLICATIONS OF PREFIELD INVENTORY DATA

As mentioned above, current protocols require prefield assessment of land use and tree canopy cover only at non-visit and non-sampled P2 plots. Prefield interpreters in some FIA regions will also populate the canopy cover variable for all P2 plots in the 2012 panel/subpanel, including forest plots, and this expanded dataset will serve as predictor data for the imminent update of the tree canopy cover layer in the 2011 National Land Cover Database (see Homer et al. 2004). Other potential changes to prefield workflows include the addition of a land cover classification, an updated the land use classification, and implementation of a comprehensive and nationally consistent quality assurance protocol.

CONCLUSIONS

Until recently, prefield photo-interpretation focused primarily on making visit/non-visit determinations by distinguishing potentially forested plots from obvious nonforest plots. Prefield photo interpreters examine every plot in the FIA P2 grid and are thus in a unique position to provide additional information about every plot, regardless of whether it meets the FIA definition of forest land. Tree cover in wooded areas that are not considered “forest” by FIA may contribute to biomass and carbon budgets as well as wildlife habitats in developed or sparsely wooded areas. The land use and canopy cover dataset generated within existing prefield workflows has the potential to expand FIA’s ability to inventory and monitor all lands, and not just those areas that meet the FIA definition of forest land.

ACKNOWLEDGMENTS

Several members of the IW-FIA data collection staff contributed extra effort above and beyond standard IW-FIA prefield protocols for the 2009-2011 subpanels.

LITERATURE CITED

- Bechtold, W. A.;** Patterson, P. L., eds. 2005. The enhanced Forest Inventory and Analysis program – national sampling design and estimation procedures, Gen. Tech. Rep. SRS-80. U.S. Department of Agriculture Forest Service, Asheville, NC.
- Bechtold, W. A.;** Scott, C. T. 2005. The Forest Inventory and Analysis plot design, pp. 27-42. In: Bechtold, W. A. and Patterson, P. L. (eds.). The enhanced Forest Inventory and Analysis program – national sampling design and estimation procedures, Gen. Tech. Rep. SRS-80. U.S. Department of Agriculture Forest Service, Asheville, NC.
- Frescino, T.;** Moisen, G. 2011. Comparing tree canopy cover estimates from digital aerial photography of different resolution with field-based assessments. [This volume.]
- Goeking, S. A.;** Liknes, G. C. 2009. The role of pre-field operations at four forest inventory units: we can see the trees, not just the forest. In: McWilliams, W.; Moisen, G.; Czaplowski, R., comps. 2008. Forest Inventory and Analysis (FIA) Symposium 2008; October 21-23, 2008; Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station. 1 CD.
- Homer, C.;** Huang, C.; Yang, L. [and others]. 2004. Development of a 2001 national land-cover database for the United States, Photogrammetric Engineering & Remote Sensing 70: 829-840.
- Jennings, S. B.;** Brown, N.D.; and Sheil, D. 1999. Assessing forest canopies and understory illumination: canopy closure, canopy cover and other measures. *Forestry* 72(1): 59-73.
- Reams, G. A.;** Smith, W. D.; Hansen, M. H. [and others]. 2005. The forest inventory and analysis sampling frame, pp. 11-26. In: Bechtold, W. A. and Patterson, P. L. (eds.). The enhanced Forest Inventory and Analysis program – national sampling design and estimation procedures, Gen. Tech. Rep. SRS-80. U.S. Department of Agriculture Forest Service, Asheville, NC.
- Toney, C.;** Shaw, J. D.; Nelson, M. D. 2009. A stem-map model for predicting tree canopy cover of forest inventory and analysis (FIA) plots. In: McWilliams, W.; Moisen, G.; Czaplowski, R., comps. 2008. Forest Inventory and Analysis (FIA) Symposium 2008; October 21-23, 2008; Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station. 1 CD.
- USDA Farm Service Agency.** 2009. National Agriculture Imagery Program (NAIP) website. Available online at <http://www.fsa.usda.gov/FSA/apfoapp?area=home&subject=prog&topic=nai> ; last accessed December 6, 2010.
- USDA Forest Service.** 2007. Forest Inventory and Analysis National Core Field Guide, Volume 1: Field data collection procedures for Phase 2 plots, version 4.0. Available online at <http://fia.fs.fed.us/library/field-guides-methods-proc/> Last accessed December 6, 2010.

USDA Forest Service. In press. The Forest Inventory and Analysis database: database description and users manual version 4.0 for Phase 2. DRAFT.

Woudenberg, S. W.; Conkling, B. L.; O'Connell, B. M. [and others]. [In press.] The Forest Inventory and Analysis Database: database description and users manual version 4.0 for Phase 2. Gen. Tech. Rep. RMRS-GTR-245. Fort Collins, CO: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station.

Table 1—Distribution of non-visit Phase 2 plots in each nonforest land use category for 2009, 2010, and 2011 subpanels. Area estimates are based on an expansion of 19,667 acres per plot given a sample of 30 percent of one ten-year cycle (one plot represents about 5,900 acres given a complete ten-year cycle). Mean canopy cover is averaged among plots from all three subpanels

Land Use Class	Code	Number of plots	Percentage of plots	Estimated area in Interior West (acres)	Mean canopy cover (percent)
Agricultural Land	10	581	3.58%	11,426,527	0.17
Cropland	11	1,733	10.67%	34,082,911	0.08
Pasture	12	115	0.71%	2,261,705	0.72
Idle farmland	13	225	1.39%	4,425,075	0.46
Orchard	14	2	0.01%	39,334	0.00
Windbreak/shelterbelt	17	2	0.01%	39,334	19.25
Rangeland	20	12,472	76.80%	245,286,824	2.06
Developed	30	114	0.70%	2,242,038	1.90
Cultural or urban	31	147	0.91%	2,891,049	3.22
Rights-of-way	32	142	0.87%	2,792,714	2.24
Recreation	33	11	0.07%	216,337	2.93
Mining	34	20	0.12%	393,340	1.91
Other undeveloped	40	87	0.54%	1,711,029	1.50
Naturally nonvegetated	41	506	3.12%	9,951,502	1.09
Wetland	42	78	0.48%	1,534,026	0.22
Beach	43	4	0.02%	78,668	0.00
Nonforest/chaparral	45	1	0.01%	19,667	3.00

Table 2—Total number of plots in each canopy cover class and land use category. Estimated number of plots with canopy cover greater than ten percent in one cycle is calculated as the mean annual number of plots from the 2009-2011 subpanels multiplied by ten subpanels

Land Use Class	Code	Canopy cover class					Estimated number of plots with >10% canopy cover, in one cycle
		0%	1%-5%	6%-10%	10%-20%	>20%	
Agricultural Land	10	561	25	2	2	3	17
Cropland	11	1717	33	2	1	3	13
Pasture	12	101	21		1		3
Idle farmland	13	211	16	1	2	1	10
Orchard	14	2					0
Windbreak/shelterbelt	17		2			2	7
Rangeland	20	11500	3467	798	344	407	2503
Developed	30	78	32	5	1	5	20
Cultural or urban	31	85	47	12	6	9	50
Rights-of-way	32	129	29	6	4	5	30
Recreation	33	7	4	2	1		3
Mining	34	19	1	1	2		7
Other undeveloped	40	84	35	8	1	5	20
Naturally nonvegetated	41	456	104	27	2	6	27
Wetland	42	79	5	1		1	3
Beach	43	4					0
Nonforest/chaparral	45		1				0

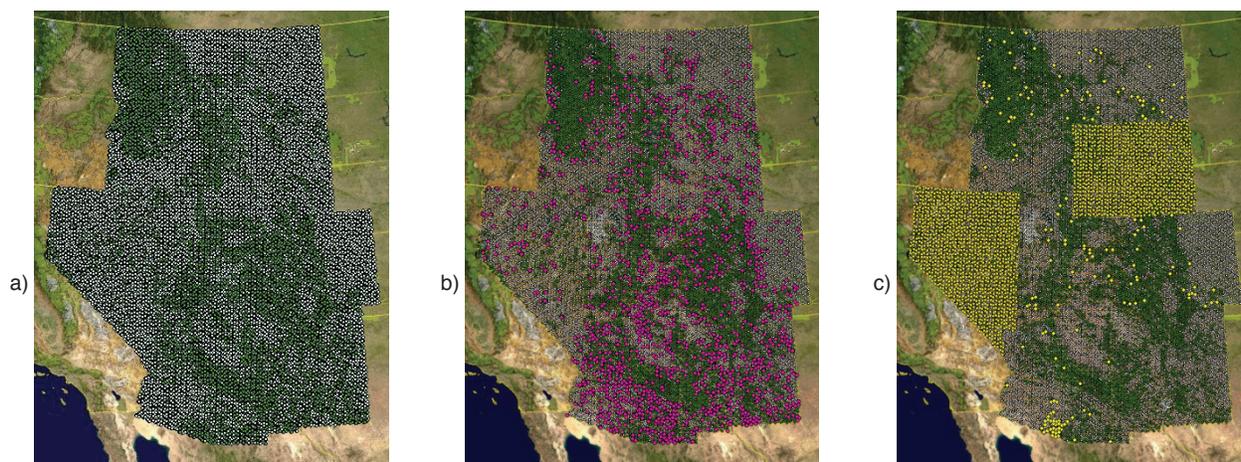


Figure 1—Maps showing plots in the Interior West FIA region in the 2009—2011 inventory years: a) Plots designated for the field inventory, from all three subplots (green dots), b) Plots that were not designated for the field inventory but include trees, from all three subpanels (magenta dots), and c) Non-sampled plots, from the 2009 subpanel only (yellow dots), i.e., plots that were designated for field sampling but no data was collected because the plot was inaccessible/hazardous or access to the property was denied.

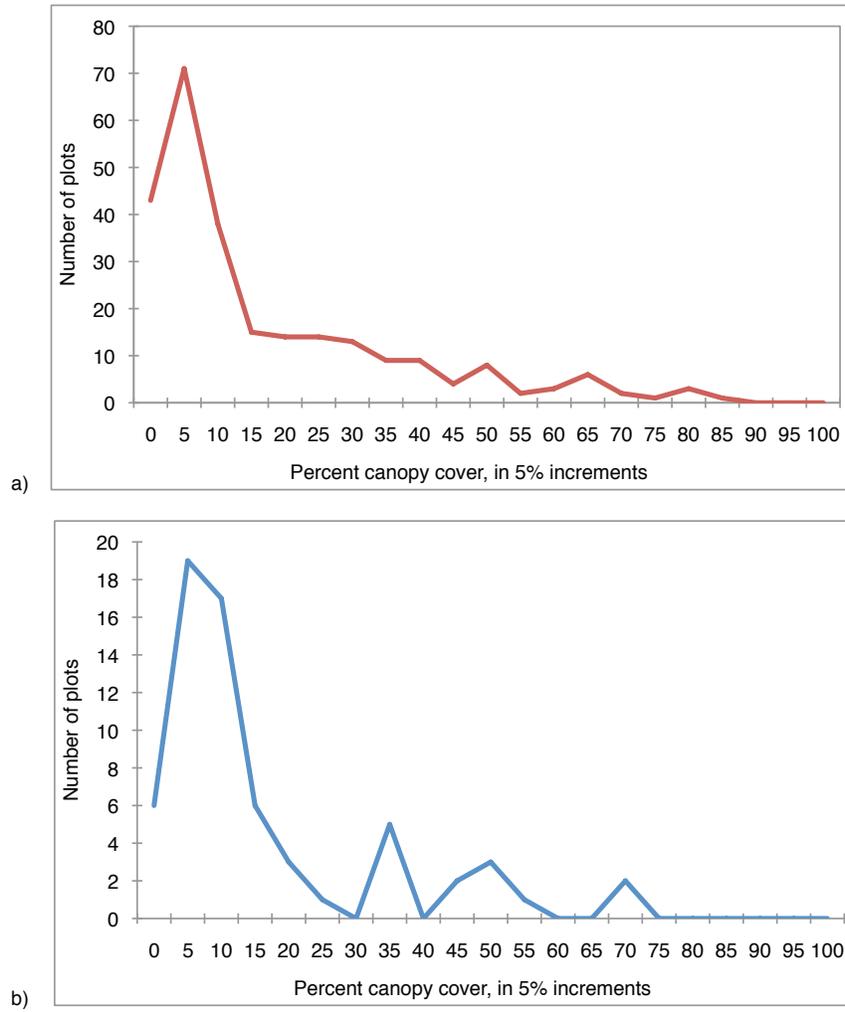


Figure 2—Histogram of tree canopy cover for P2 plots in the 2009 subpanel where a) access was denied by the land owner or manager, and b) the plot was deemed by a field crew to be hazardous and/or inaccessible due to safety concerns.

REPEATABILITY IN PHOTO-INTERPRETATION OF TREE CANOPY COVER AND ITS EFFECT ON PREDICTIVE MAPPING

Thomas A. Jackson, Gretchen G. Moisen, Paul L. Patterson, and John Tipton

ABSTRACT

In this study, we explore repeatability in photo-interpreted imagery from the National Agriculture Imagery Program that was sampled as part of the National Land Cover Database 2011 Tree Canopy Cover pilot project. Data were collected in 5 diverse pilot areas in the US, including one each in Oregon, Utah, Kansas, Michigan and Georgia. Repeatability metrics. The intra-class correlation coefficient as well as repeatability graphics are explored for each pilot area. In addition, we conduct a simulation to illustrate the effect of varying tolerance and compliance rates on the predictive maps of tree canopy cover.

INTRODUCTION

The National Land Cover Database (NLCD) includes a map the tree canopy cover of the entire United States. In support of developing a new 2011 Tree Canopy Cover product, a pilot test was recently launched to investigate numerous questions surrounding this map's development. The pilot used photo-interpreted data from large-scale aerial photography to give individual percent tree canopy estimates. Coupled with spatial predictor variables, these data were then used as the response variable in predictive models of tree canopy cover, addressing questions related to sampling intensity, value of specific predictor variables, modeling technique, and appropriate subpopulations (Tipton and others 2011, and Moisen and others 2011). The goal of this specific manuscript is to quantify the repeatability of the data and calculate the effects of fluctuation in individual interpreters on the mapping of the canopy cover.

METHODS

DATA

An intensive (approximately 1 km) grid of photo plots was established over five pilot areas the size of one Landsat scene, with one each in Georgia, Kansas, Michigan, Oregon, and Utah. On each photo plot, 105 dots were interpreted as landing on a tree or a non-tree using aerial photography acquired through the National Agriculture Imagery Program

(NAIP). To assess repeatability amongst photo interpreters, 208 plots were interpreted by every interpreter. This was a blind study in that the interpreters were unaware of which plots contributed to the repeatability data set. Percent tree canopy cover was obtained on each repeatability plot by averaging over the 105 dots per plot. This data was used to determine the repeatability of the photo interpretations. Also, for the five pilot areas, predictor variables were obtained including Landsat TM reflectance values and various GIS layers such as aspect, slope, digital elevation maps, and other measures important to tree growth. The response variable was the percent canopy cover discussed earlier. These predictors and response variables were used in modeling the effects of fluctuation in individual observers on the mapping of canopy cover.

REPEATABILITY

Each of the five pilot areas contained between two and five observers to interpret the plots, with Georgia having four, Michigan having two, Utah having five, Kansas having five and Oregon having three. Repeatability among photo interpreters was first examined by constructing graphs of the paired canopy cover percentages for all pairs of photo interpreters within each pilot area. Agreement between photo interpreters was also explored by calculating the percent of observations for each pair of interpreters that fell within 10 percent of each other.

There are several different methods to test if data sets are different. A commonly used method for repeatability is the kappa function. Because the kappa function is primarily used for qualitative data and not continuous data as seen in this study, a new metric for the repeatability was investigated. The Interclass Correlation Coefficient (ICC, Equation 1) was first proposed by Pearson in 1901, but was only applicable for comparing two variables.

$$\rho_{ICC} = \frac{2 \sum_{i=1}^{k-1} \sum_{j=i+1}^k \sigma_{ij}}{(k-1) \sum_{i=1}^k \sigma_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^k (\mu_i - \mu_j)^2} \quad (1)$$

Thomas Jackson, Graduate Student, Colorado State University Department of Statistics, Fort Collins, CO, 80526
Gretchen G. Moisen, Research Forester, US Forest Service Rocky Mountain Research Station, Ogden, UT, 84401
Paul L. Patterson, Statistician, US Forest Service Rocky Mountain Research Station, Fort Collins, CO, 80526
John Tipton, Graduate Student, Colorado State University Department of Statistics

Carrasco et al (2003) have expanded on his initial work to incorporate more than two variables, taking into account the mean, variance and covariance of each variable in question. The equation returns a score between 0 and 1, which can be interpreted as an R^2 value; that is, the closer to 1, the better the repeatability between all observers. An ICC value was obtained for each of the five pilot areas.

TOLERANCE AND COMPLIANCE RATES

Most measurement quality objectives (MQO) are stated in terms of a tolerance and a compliance rate (Pollard et al 2006, Westfall 2010). The tolerance limit is the amount of error we are willing to accept from each observer and is obtained based on the accuracy necessary for analytical purposes and the ability of the observer to repeat the findings. The compliance rate is the minimum percentage of observations that should be within tolerance. To calculate the effects of fluctuation in individual observers on the mapping of canopy cover, Random Forests (Breiman 2001) was employed to model tree canopy cover as a function of the predictor variables using data that were perturbed under varying tolerance and compliance rates. The changes in R^2 values were then measured in terms of these perturbations. The predictors and response used in the Random Forests model were the same as those used in Tipton et al (2011). Initial R^2 values were obtained for each of the five pilot areas using an independent test set using the definition of R^2 of sum of squares regression divided by the sum of squares total. Then, 1,000 simulations were run at varying tolerances and varying compliance rates to investigate the effect on the mapping of the canopy cover. For each tolerance and compliance rate, an R^2 value was obtained in the same manner as above. This was accomplished by first setting aside a portion of the original data, perturbing the remaining data based on the tolerance and compliance rates, and obtaining the predictor coefficients for the canopy cover using the Random Forests model on the perturbed data. Once this was accomplished, the portion of data set aside for validation was used to construct the R^2 value.

RESULTS AND DISCUSSION

REPEATABILITY

Figure 1 illustrates the paired canopy cover percentages for all pairs of photo interpreters from the Kansas pilot area and is representative of the results from Michigan, Oregon, and Utah. We can see that the grouping of the points tends to lie on or closely around the one-to-one line (y equals x line), which indicates that there is strong agreement between observers. Another way to look at this, Table 1, illustrates results from the Oregon pilot area and shows that all three interpreters have at least 70 percent of their observations within 10 percent of each other, with better agreement between some observers than others. Figure 2, however, illustrates the lack of a pattern seen in some of the paired

canopy cover percentages for all pairs of photo interpreters from Georgia. Table 2, showing the percent of observations for each pair of interpreters that fell within 10 percent of each in Georgia, further illustrates this. Here we see that three of the observers (observers 1, 3 and 4) are in strong agreement with each other, with observer 2 not being in strong agreement with any of the other three. We can see that observer 2 does not have more than 50 percent of its observations within 10 percent of each other. Observer 2 tends to be the outlier in the Georgia pilot area, with the other three observers being in strong agreement with each other not only in the paired canopy cover plots but within percentages as well. For the other four pilot areas, the proportions are always above 63 percent, with the majority of the calls being above 70 percent, which indicates that the agreement between observers is very high. An ICC value was obtained for each of the five pilot areas and can be seen in Table 3. As discussed above, the agreement of the observers from the Georgia pilot area was shown not to be as strong as seen in the other four pilot areas. This is again seen in the low ICC value from the Georgia pilot area, but since observer 2 is already known to be the outlier from the Georgia pilot area, it can be treated as an outlier and removed from the dataset. With observer 2 removed, the ICC value from Georgia increases from 0.6491 to 0.9830. With the high ICC values returned, this not only echoes the findings from the one-to-one plots and the within percentages that there is strong repeatability across observers, but also gives us a way to quantify the strength of the repeatability.

TOLERANCE AND COMPLIANCE RATES

The effect of the fluctuation was measured using the ratio of the perturbed R^2 value over the original R^2 value. Table 4 shows results for the Georgia pilot area. The tolerances are on the left hand side and indicated as decimals, and the compliance rates are across the bottom and indicated as percentages. The ratios from the other four pilot areas were all within 2 percent of each other. We can see from this table that to develop models that explain at least 80 percent of the variability that the model from unperturbed data provides, we have to have a minimum tolerance limit of 20 percent and a minimum compliance rate of 85 percent.

CONCLUSIONS

With the inclusion of several different metrics for all five pilot areas, we were able to conclude that, with proper training and preprocessing, the repeatability of the study areas is high. Not only are the one-to-one plots and within percentages strong indicators of repeatability, but high ICC scores are able to quantify this fact. The repeated simulations for all pilot areas show that as long as strict guild lines are upheld, large proportions of the true predictive strength can be returned.

LITERATURE CITED

- Breiman, L.** 2001. Random Forests. *Machine learning*. 45:5-32.
- Moisen, G.G.;** Coulston, J.W.; Wilson, B.T. [and others]. 2011. Choosing appropriate subpopulations for modeling tree canopy cover nationwide. In McWilliams, W., Moisen, G.G.; Roesch, F. (eds.). *Forest Inventory and Analysis (FIA) Symposium 2010*; October 5-8, 2010; Knoxville, TN. Proc. SRS-CD. Knoxville, TN: U.S. Department of Agriculture Forest Service, Southern Research Station. 1 CD.
- Carrasco, J.L.;** Jover, J.L. 2003. The concordance correlation coefficient estimated through variance components. IX Conferencia Espanola de Biometria.
- Pollard, J.E.;** Westfall, J.A.; Patterson, P.L. [and others]. 2006. Forest Inventory and Analysis national data quality assessment report for 2000 to 2003. RMRS-GTR-181. U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station.
- Tipton, J.;** Moisen, G.G.; Patterson, P.L. [and others]. 2011. Sampling intensity and normalizations: exploring cost-driving factors in nationwide mapping of tree canopy cover. In McWilliams, W., Moisen, G.G.; Roesch, F. (eds.). *Forest Inventory and Analysis (FIA) Symposium 2010*; October 5-8, 2010; Knoxville, TN. Proc. SRS-CD. Knoxville, TN: U.S. Department of Agriculture Forest Service, Southern Research Station. 1 CD.
- Westfall, J. A.** ed. 2010. FIA national assessment of data quality for forest health indicators. Gen. Tech. Rep. NRS-53. Newtown Square, PA: U.S. Department of Agriculture Forest Service, Northeastern Research Service. p.80

Table 1—Within 10 percent proportions for Oregon Pilot Area

	Obs 1	Obs 2	Obs 3
Obs1	1.00	0.7588	0.7075
Obs 2		1.00	0.8103
Obs 3			1.00

Table 3—Interclass correlation coefficient values for all five plot areas

Pilot Area	ICC Value
Georgia	0.6491
Kansas	0.9126
Michigan	0.9424
Oregon	0.9299
Utah	0.8652

Table 2—Within 10 percent proportions for Georgia Plot Area

	Obs 1	Obs 2	Obs 3	Obs 4
Obs 1	1.00	0.4615	0.9279	0.9087
Obs 2		1.00	0.4327	0.4471
Obs 3			1.00	0.9135
Obs 4				1.00

Table 4—R² ratios for Georgia Pilot Area tolerance and compliance limits

.3	.7771	.7530	.7235	.7004	.6779
.2	.8675	.8468	.8277	.8101	.7893
.1	.9472	.9339	.9195	.9075	.8933
	100%	95%	90%	85%	80%

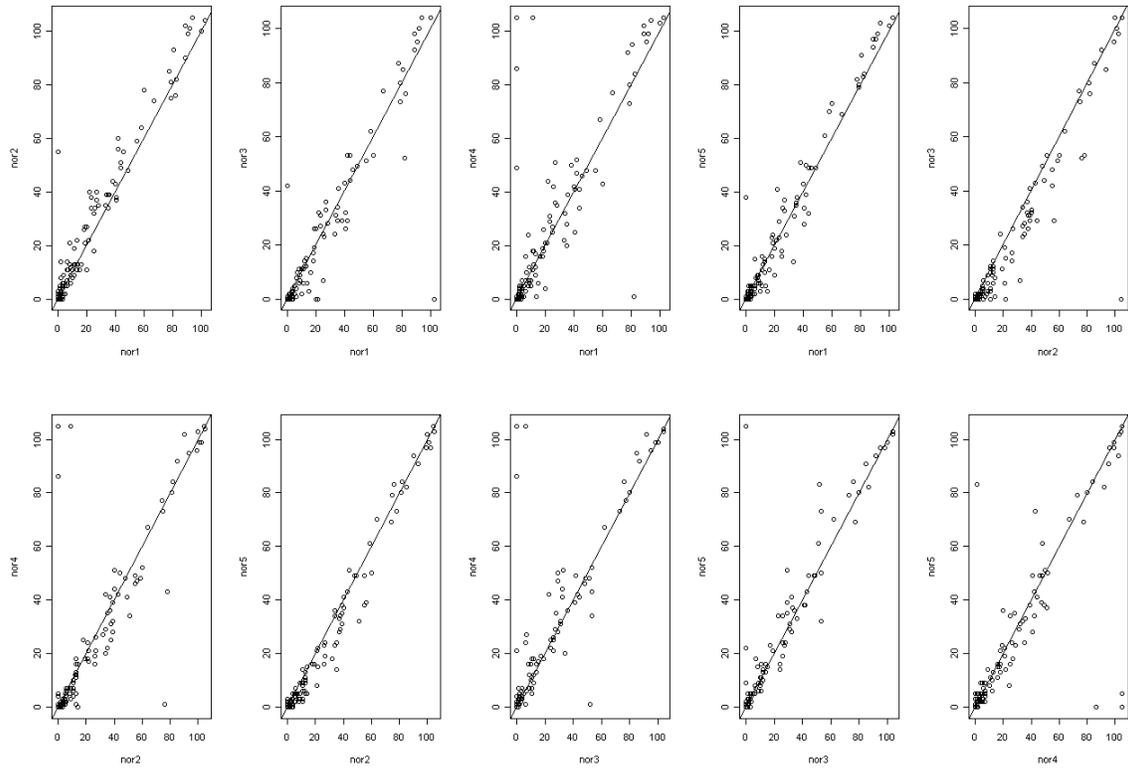


Figure 1—Paired Canopy Cover Plots for all Interpreters in the Kasas Pilot Area Including One-to-One Line

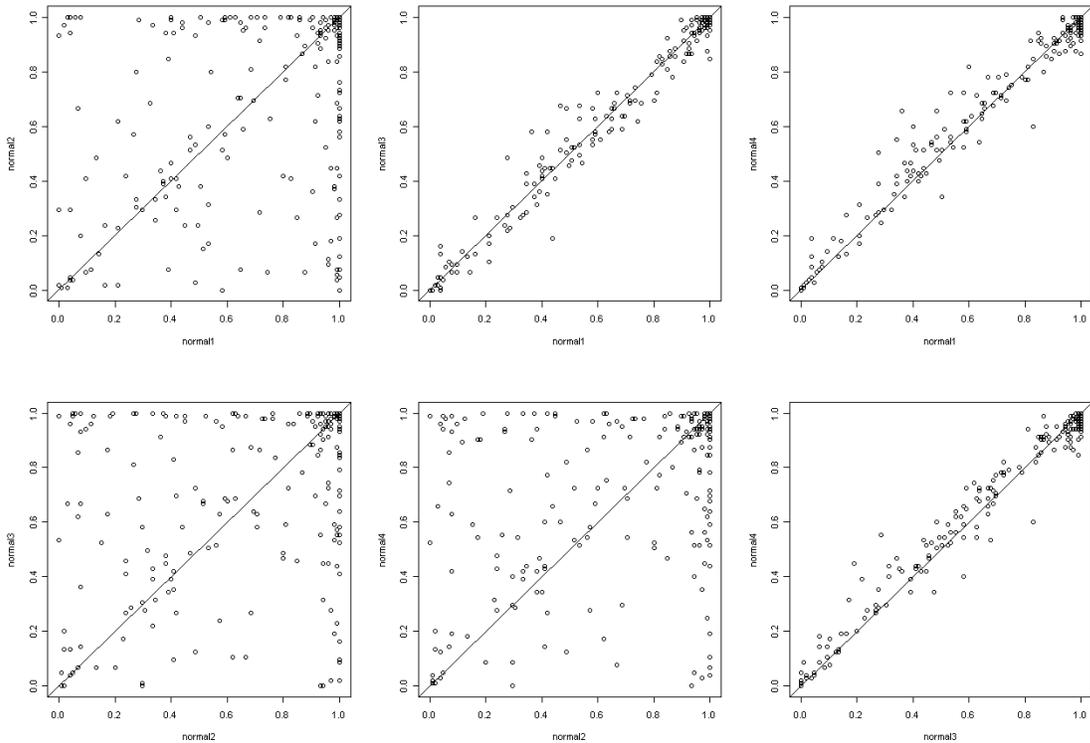


Figure 2—Paired Canopy Cover Plots for all Interpreters in the Georgia Pilot Area Including One-to-One Line

A FRAMEWORK FOR REPORTING TREE COVER ATTRIBUTES IN AGRICULTURAL LANDSCAPES

Dacia M. Meneguzzo and Greg C. Liknes

INTRODUCTION

The definition of forest land used by the USDA Forest Service's Forest Inventory and Analysis program includes area, width, and density requirements. These requirements frequently exclude from the inventory any trees occupying narrow riparian corridors or linear tree plantings (e.g., windbreaks and shelterbelts). With recent attention being paid to such topics as bio-fuel production and carbon sequestration, motivation exists to account for trees outside the forest (TOF). Because much of the tree cover in the Plains States occurs as TOF as opposed to definitional forest land, alternative methods are needed for collecting and reporting information about landscapes dominated by agricultural practices.

METHODOLOGY

Using aerial imagery from the USDA's National Agriculture Imagery Program (NAIP), an object-based image analysis technique was used to identify and quantify tree cover. The technique was applied to 316 3.75' x 3.75' quarter quadrangles covering seven counties in the state of Nebraska. An area of approximately 3,525 square miles was mapped into the following categories: water, tree, agriculture/other vegetation, man-made, and other non-vegetated (e.g., bare soil). In addition to area estimates, a series of descriptive measures was included to describe the spatial arrangement and extent of tree cover. The following metrics were calculated for each of the 7 counties: percent of county area, number of tree-covered patches, average patch size, minimum patch size, maximum patch size, largest patch index (percentage of tree cover area occupied by the largest patch), and patch density.

RESULTS

Sample metrics are presented for Kearney and Nemaha counties in Figure 1 and Table 1. The majority of tree-covered area (76 percent) in Nemaha County exists as large patches, i.e., those exceeding 10 acres in size. In contrast, most of Kearney County's tree cover (82 percent) is comprised of patches that fall into the two smaller size classes. County differences in tree cover are characterized by additional metrics in Table 1. Nemaha County has a higher percentage of tree cover, a higher patch density with larger patches, and a much larger patch index, as 10 percent of the county's tree cover occurs in one continuous patch of forest.

DISCUSSION

A method for mapping the extent and spatial arrangement of tree cover in agricultural landscapes has been developed using high-resolution imagery and object-based image analysis techniques. This method provides a useful alternative for describing TOFs and may complement traditional inventories of definitional forest land. The suite of metrics used here appears to capture important differences between counties with drastically differing composition of tree cover.

Table 1—Metrics describing the extent and spatial pattern of tree cover in Kearney and Nemaha Counties, NE.

TREE COVER METRICS	KEARNEY	NEMAHA
Area of tree cover (acres)	4,387	36,279
Percent of county area	1.3%	13.8%
FIA 2008 estimate of forest land (acres)	441	17,710
Number of patches	11,118	14,629
Average patch size (acres)	0.4	2.5
Minimum patch size (acres)	0.01	0.01
Maximum patch size (acres)	82	3,453
Largest patch index (percentage of tree cover area occupied by the largest patch)	2%	10%
Patch Density (number of patches per square mile of county area)	22	36

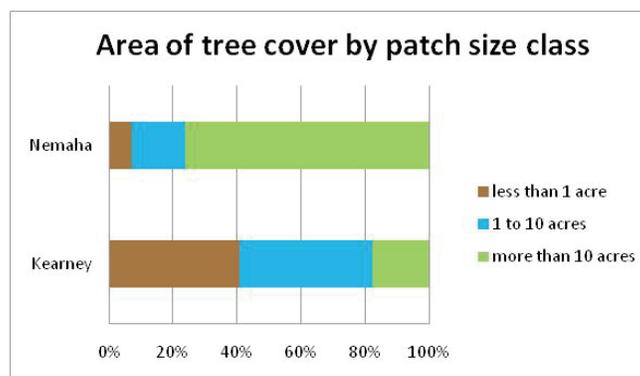


Figure 1—Distribution of tree-covered area by patch size class for Kearney and Nemaha Counties, NE.

CHOOSING APPROPRIATE SUBPOPULATIONS FOR MODELING TREE CANOPY COVER NATIONWIDE

Gretchen G. Moisen, John W. Coulston, Barry T. Wilson, Warren B. Cohen,
and Mark V. Finco

ABSTRACT

In prior national mapping efforts, the country has been divided into numerous ecologically similar mapping zones, and individual models have been constructed for each zone. Additionally, a hierarchical approach has been taken within zones to first mask out areas of nonforest, then target models of tree attributes within forested areas only. This results in many models nationwide, which reduces the number of training points per model, increases the cost of the process, results in numerous seam lines, and complicates validation efforts. Consequently, we use response data based on photo-interpreted aerial photography and spatially continuous predictor data (Landsat imagery, topographic and other ancillary data) in five pilot areas across the country to explore the effect of the choice of modeling subpopulation on models of tree canopy cover. Using Random Forests as our predictive tool, we explore the consequences of modeling pilot areas alone, modeling groups of pilot areas, and modeling hierarchically within each pilot area. Recommendations are made for appropriate modeling subpopulations to be used in a nationwide tree canopy cover map.

INTRODUCTION

The Multi-Resolution Land Characteristics (MRLC, <http://www.mrlc.gov/>) consortium has developed plans for the 2011 National Land Cover Dataset (NLCD) which will include an approximate Anderson Level II classification, percent impervious surface, and percent tree canopy cover. Because it is central to its business needs, the US Forest Service, Forest Inventory and Analysis (FIA) program has assumed responsibility for the latter, and will be developing this Tree Canopy Cover (TCC) layer. Recently a national pilot project was launched to test the use of high resolution photography acquired through the National Agriculture Imagery Program (NAIP) coupled with extensive ancillary data layers through alternative sampling and modeling methodologies in support of this commitment. A number of studies have resulted from initial pilot analyses answering questions about alternative means to observe tree canopy cover (Frescino and others 2011), relationship between photo-based tree canopy cover and canopy modeled from FIA plots (Toney and others 2011), repeatability in

photo-interpretation (Jackson and others 2011), efficient sampling strategies (Jackson and others 2011), and, in this paper, choice of appropriate subpopulations over which to construct predictive models.

Tree canopy cover in the conterminous U.S. is remarkably diverse. Previous nationwide mapping efforts, like that of the US Forest Biomass map (Blackard and others 2008), nationwide forest type and forest type group maps (Ruefenacht and others 2008), as well as Landfire (Rollins and Frame 2006) have tried to accommodate this diversity by using 66 different mapping zones (Homer and Gallant 2001, Figure 1). In these efforts, mapping zones were modeled independently and in some cases forest masks were first developed, then models developed solely for the areas predicted to be forest. With most FIA mapping efforts, the sampling intensity of the training data is fixed at the nominal sampling intensity of the base FIA program (approximately 1 plot per 6000 ac). Therefore developing models for relatively small mapping zones decreases the number of training points available. Additionally, when small mapping zones are used the number of models increases which results in increased cost, seamline issues, and complicated validation approaches. Consequently, we used photo-interpreted data collected in five pilot areas in the conterminous United States to explore the effect of modeling over larger, more geographically diverse areas, as well as the value of empirically masking non-tree areas prior to modeling tree canopy cover.

METHODS

DATA

Photo-interpreted percent tree canopy cover data from NAIP imagery was collected in 5 diverse pilot areas in the United States, including areas in Oregon, Utah, Kansas, Michigan and Georgia. Photo plots were collected on the

Gretchen G. Moisen, Research Forester, US Forest Service Rocky Mountain Research Station, Ogden, UT, 84401

John W. Coulston, Research Forester, US Forest Service Southern Research Station, Knoxville, TN, 37919

Barry T. Wilson, Research Forester, US Forest Service Northern Research Station, St. Paul, MN, 55108

Warren B. Cohen, Research Forester, US Forest Service Pacific Northwest Research Station, Corvallis, OR, 97331

Mark V. Finco, Contract Leader, Red Castle Resources, US Forest Service Remote Sensing Applications Center, Salt Lake City, UT, 84119

FIA grid, intensified 4-fold, and each photo plot consisted of 105 dots distributed in a 90 m square area (Figure 2). Each dot was characterized as being a tree or not-a-tree, then the proportion of tree dots were summarized for each plot. This percent tree cover was used as the response variable in models described below. Predictor variables included Landsat-5 reflectance bands, 30 m elevation, transformed aspect, slope, topographic positional index, land cover from the 2001 NLCD, and Bailey's ecoregions. Because many of the predictor variables originated from 30 m products, assignment to each 90 m plot was accomplished by taking a focal mean over a 3x3 window for continuous variables, and focal majority for the categorical variables. In addition, the standard deviations for all continuous predictor variables within the 3x3 window were included as predictor variables. Following findings presented in Tipton and others (2011), a subset of the total data available equivalent to the intensity of the FIA grid was used for modeling, and an equal size independent test set used for testing in these analyses.

MODEL

Classification and regression trees (Breiman and others 1984) are flexible and robust tools that are well suited to the task of modeling the relationship between a response and a set of explanatory variables for the purposes of making spatial predictions in the form of a map. These are intuitive methods, often described in graphical or biological terms. Typically shown growing upside down, a classification or regression tree begins at its root. An observation passes down the tree through a series of splits, or nodes, at which a decision is made as to which direction to proceed based on values of the explanatory variables. Ultimately, a terminal node or leaf is reached and predicted response is given, the mean of observations in the node for a continuous response, or a vote for a categorical response. (See De'ath and Fabricius 2000 for a thorough explanation, and Moisen 2008 for a simple overview.)

Although classification and regression trees are powerful tools by themselves, much work has been done in the data mining and machine learning fields to improve the predictive ability of these models by combining separate tree models into what is often called a committee of experts, or ensemble. One such tool, Random Forests (Breiman 2001) is receiving increasing attention in the ecological and remote sensing literature. In this technique, a bootstrap sample of the training data is chosen. At the root node, a small random sample of explanatory variables is selected and the best split made using that limited set of variables. At each subsequent node, another small random sample of the explanatory variables is chosen, and the best split made. The tree continues to be grown in this fashion until it reaches the largest possible size, and is left un-pruned. The whole

process, starting with a new bootstrap sample, is repeated 500 or more times. The final prediction is a vote (for categorical responses) or average (for continuous variables) from prediction of all the trees in the collection. All of the following analyses were fit using the "randomForest" library in R (Liaw and Wiener 2002).

SMALL VS. LARGE MAPPING ZONES

Using the training data sets described above, eight models of tree canopy cover were constructed. The first five were individual "pilot area" models for each of Georgia, Michigan, Kansas, Utah, and Oregon which only contained training data from each of their respective areas. The sixth model, called the "East" model used training data from GA, MI and KS, while the seventh "West" model used all the training data from OR, UT, and KS. The eighth model was called a "USA" model and used all the training data in all five pilot areas. These eight models were applied to the test data sets within each of the pilot area and the resulting metrics of the relationship between observed and predicted values in these test sets were compared. The metrics included: correlation, root mean squared error (RMSE), and slope of a regression line. Density plots of observed and predicted, which are like a continuous version of histograms reflecting the relative number of plots by tree canopy cover class, were also compared.

NO-TREE MASK

This analysis involves building two models for each pilot area. First, a binary response of "trees present" versus "no trees present" was modeled as a function of all the predictor variables, again using Random Forests. The probability of tree presence was predicted over the test data and these probabilities were then converted to binary "trees present" or "no trees present" using the prevalence of treed land in each area as the threshold. (See Freeman and Moisen 2008a for a discussion of thresholding options). Using these predictions over the test data, assessments were made of the tree mask using the PresenceAbsence library (Freeman and Moisen 2008b) in R. This first tree presence model was then applied to the training data so that only those training plots predicted to have trees present were included in a continuous model of tree canopy cover, the second model. To validate the effectiveness of combining the first and second models, test data plots predicted to have "no trees present" from the first model were simply given a predicted tree canopy cover of zero, while test data plots predicted to have "trees present" were then assigned tree canopy cover predictions from the second model. This, in effect, empirically masks out areas thought to have no trees present at all. Comparisons of the final predicted versus observed tree canopy values in each pilot area (including treed and non-treed land) were done using metrics as above.

RESULTS AND DISCUSSION

SMALL VS. LARGE MAPPING ZONES

Figure 3 illustrates the effect of increasing mapping unit size on map accuracy metrics, including correlation, root RMSE, and slope of a line fitted between predicted vs. observed values in the independent test set (with intercept term.) Interestingly, little difference between accuracy metrics is noted between the individual pilot area models, and models built for larger areas (East, West and USA models). Note that models built for large areas naturally included many more training plots. The only exception is in cases where model predictions were made over areas whose data were not included in that particular model. For example, the West model predicted over Michigan, or the East model predicted over Utah. In addition, density plots of the true tree canopy cover values in each pilot area were plotted along with densities obtained by applying the four classes of models (pilot, East, West and USA) to that same training data, as illustrated in Figure 4. As with the accuracy metrics, there was little difference in the densities obtained under the four modeling scenarios except in cases where no data from that particular pilot area was used in the model.

NO-TREE MASK

Figure 5 illustrates the results from the tree presence model in UT which were typical of the other pilot areas. The first graph (Figure 5a) is a Receiver Operator Curve (ROC Plot) indicating a strong model fit and high Area Under the Curve (AUC) value of 0.94. Here, sensitivity, or proportion of correctly predicted positive observations, reflects a model's ability to detect a presence, given at least one tree actually occurs at a location. Specificity, or proportion of correctly predicted negative observations, reflects a model's ability to predict an absence where trees do not exist. The second graph in this figure (Figure 5b) illustrates how measures of map accuracy change with different threshold values. In UT, approximately 70 percent of the land area had trees present. This graph illustrates how using prevalence as a threshold to convert probability predictions in to a presence-absence map resulted in maximizing map accuracy.

Plots exploring the effect of first creating a tree presence model prior to modeling tree canopy cover are illustrated in Figures 6 and 7. Figure 6 illustrates the effect of using a no-tree mask on map accuracy metrics, including correlation, root RMSE, and slope of a line fitted between predicted vs. observed values in the independent test set (with intercept term.) Little difference between accuracy metrics is noted between the unmasked, and masked approaches. In Figure 7a, predicted tree canopy cover from a single unmasked model is plotted against the tree canopy cover response from the photo interpretation illustrating the tendency to predict canopy where no trees exist at all (the zero line on the x-axis). Next in Figure 7b, predicted tree canopy cover

from the tree presence model followed by the masked tree canopy model is plotted against the tree canopy cover response from the photo interpretation illustrating a slight reduction in the number observations where canopy was predicted over no-tree areas, but also an increase in errors of false negative (the zero line on the y-axis). Finally in Figure 7c, the predicted probability of having trees present from the tree presence model is plotted against predicted tree canopy cover with no masking, illustrating the strong relationship between masked and unmasked scenarios suggesting most of the necessary information may be contained in a single model. That is, an empirical mask constructed prior to modeling tree canopy cover may not be that effective in improving the final tree canopy cover map. Also shown in 7c is the prevalence-based threshold in blue (~70 percent of the Utah pilot area is treed) above which plots are predicted to have trees. In addition, the pink vertical line illustrates a threshold a user might impose by applying a 10 percent cover threshold to the predicted tree canopy cover. Interestingly, these two thresholding criteria applied to two different models identify very closely to the same sets of plots, again indicating not a lot of additional information is gained by hierarchically modeling tree/no-tree followed by tree canopy cover in an empirical fashion.

CONCLUSION

Random Forests is a flexible and robust tool for mapping tree canopy cover over large geographic areas. Although past nationwide mapping efforts have delineated many small mapping zones across the country, the analyses conducted here suggest that modeling over much larger zones does not compromise model fit. This provides an opportunity to decrease the cost of the mapping process, reduce the numerous seam lines, and simplify validation efforts. Still to be investigated, however, is the effect of modeling over larger units with decreased sampling intensity. This could further reduce sampling costs. In addition, modeling hierarchically by creating an empirical tree presence model prior to modeling tree canopy cover does not completely alleviate the problem of predicting tree canopy cover where no trees exist, and does tend to mask treed areas as no-tree erroneously. However, this does not diminish the importance of applying a variety of regionally-specific masks, such as water and impervious surface masks, to the final product.

Naturally, results from these pilot tests as well as those described in Tipton and others 2011, and Jackson and others 2011 need to be confirmed over larger geographic areas. The NLCD Tree Canopy Cover project is entering the prototype phase. In this prototype, photo interpreted as well as ancillary data are being collected in two diverse areas, one approximately 49 million acres in size in the Interior Western U.S., the other approximately 59 million acres in

size in the Southeastern U.S. Prototype tests will be run to provide yet stronger basis for production mapping which is scheduled to begin in the fall of 2011.

ACKNOWLEDGEMENTS

We are grateful for the tremendous effort given by numerous photo-interpreters in the FIA units. Without them, the NLCD 2011 Tree Canopy Cover pilot project would not have been possible. We also appreciate the staff at the Remote Sensing Application Center for helping us meet very tight timelines. Finally, we extend our thanks the entire pilot team for all the good energy and banter.

LITERATURE CITED

- Blackard, J.;** Finco, M; Helmer, E. [and others]. 2008. Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sensing of Environment*. 112: 1658-1677.
- Breiman, L.;** Friedman, R. A., Olshen, R. A. and Stone, C. G. 1984. *Classification and regression trees*. Pacific Grove, CA, USA: Wadsworth.
- Breiman, L.** 2001. *Random Forests*. *Machine learning*. 45:5-32.
- De'ath, G.;** Fabricius, K. E. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*. 81: 3178-3192.
- Freeman, E.A.;** Moisen, G.G. 2008a. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and Kappa. *Ecological Modelling*. 217:48-58.
- Freeman, E. A.;** Moisen, G.G. 2008b. PresenceAbsence R Library. *Journal of Statistical Software*. 23(11).
- Homer, C. G.;** Gallant, A. 2001. Partitioning the conterminous United States into mapping zones for Landsat TM land cover mapping. USGS White Paper available at <http://landcover.usgs.gov>.
- Jackson, T.;** Moisen, G.G.; Patterson, P.L.; Tipton, J. 2011. Repeatability in photo-interpretation of tree canopy cover and its effect on predictive mapping. In McWilliams, W., Moisen, G.G.; Roesch, F. (eds.). *Forest Inventory and Analysis (FIA) Symposium 2010*; October 5-7, 2010; Knoxville, TN. Proc. SRS-CD. Knoxville, TN: U.S. Department of Agriculture Forest Service, Southern Research Station. 1 CD.
- Liaw, A.;** Wiener, M. 2002. Classification and Regression by randomForest. *R News*. 2(3): 18-22.
- Moisen, G. G.** Classification and regression trees. 2008. In: Sven Erik Jørgensen and Brian D. Fath (Eds.), *Ecological Informatics, Encyclopedia of Ecology*. Volume 1: 582-588.
- Rollins, M.G.;** Frame, C.K., tech. eds. 2006. *The LANDFIRE Prototype Project: nationally consistent and locally relevant geospatial data for wildland fire management*. Gen. Tech. Rep. RMRS-GTR-175. Fort Collins: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station. 416 p.
- Ruefenacht, B.;** Finco, M.V.; Nelson, M.D. [and others]. 2008. Conterminous U.S. and Alaska forest type mapping using forest inventory and analysis data. *Photogrammetric Engineering and Remote Sensing*. 74(11):1379-1388.
- Tipton, J.;** Moisen, G.G.; Patterson, P.L. [and others]. 2011. Sampling intensity and normalizations: exploring cost-driving factors in nationwide mapping of tree canopy cover. In McWilliams, W., Moisen, G.G.; Roesch, F. (eds.). *Forest Inventory and Analysis (FIA) Symposium 2010*; October 5-7, 2010; Knoxville, TN. Proc. SRS-CD. Knoxville, TN: U.S. Department of Agriculture Forest Service, Southern Research Station. 1 CD.



Figure 1—Mapping/modeling zones (Homer & Gallant, 2001) used in previous NLCD mapping efforts.

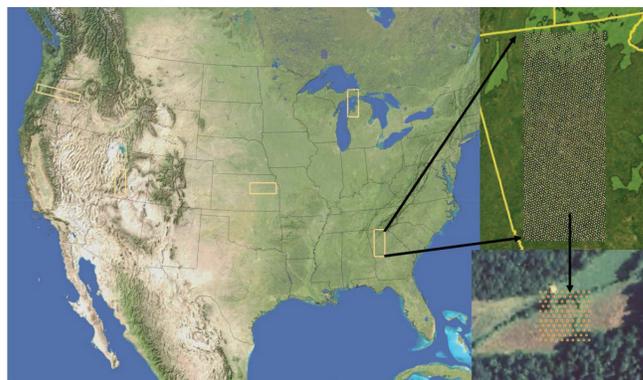


Figure 2—Five pilot areas including one each in Georgia, Michigan, Kansas, Utah, and Oregon. Photo-based sample plots were interpreted at 4 times the FIA grid intensity within each plot area. Each photo plot consisted of 105 photo points used to estimate percent tree canopy cover on the plots.

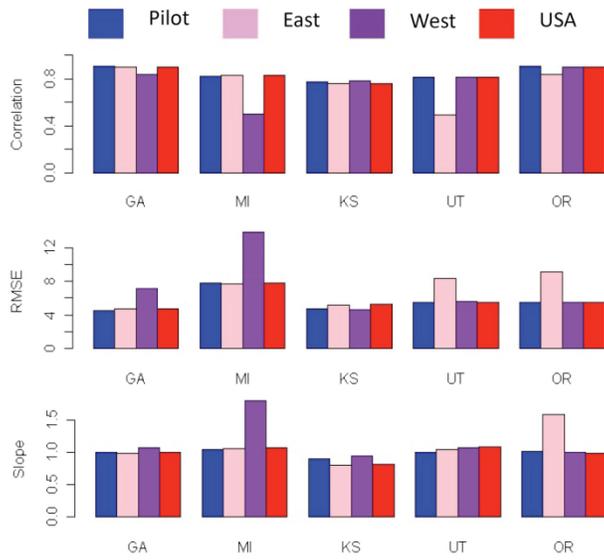


Figure 3—Correlation, RMSE, and slopes obtained in each of the five pilot areas when applying five different tree canopy models to independent test data. “Pilot” models (blue) included only training data from each individual pilot area. The “East” model (pink) included data from Georgia, Michigan, and Kansas. The “West” model (purple) included data from Oregon, Utah, and Kansas. And the “USA” model (red) included data from all the pilot areas.

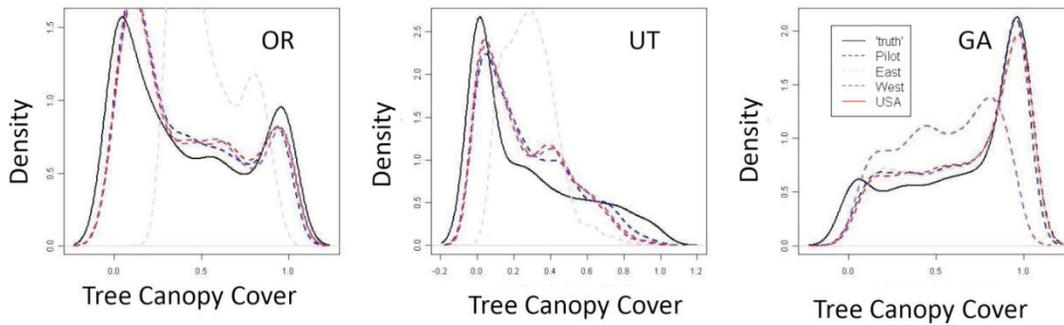


Figure 4—Density plots of tree canopy cover in independent test sets in three pilot areas, a. Oregon, b. Utah, and c. Georgia. Solid black lines reflect the “truth” from photo-interpreted data. Dotted blue lines reflect prediction from the individual pilot area models, then dotted pink, purple, and red from East, West, and USA models respectively.

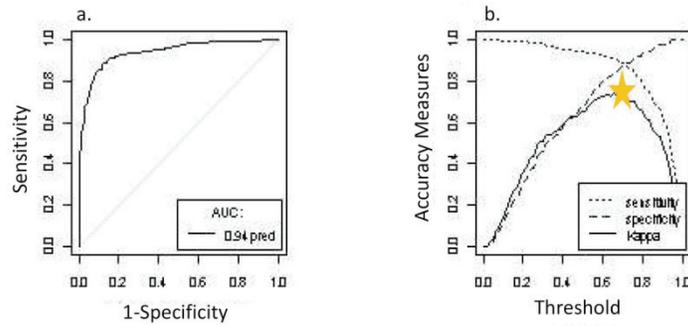


Figure 5—Results from the tree/not-tree model in UT. Plot a) is a Receiver Operator Curve (ROC Plot). Plot b) illustrates how using prevalence (indicated by the yellow star) as a threshold to convert probability predictions into a presence-absence map resulted in maximizing map accuracy.

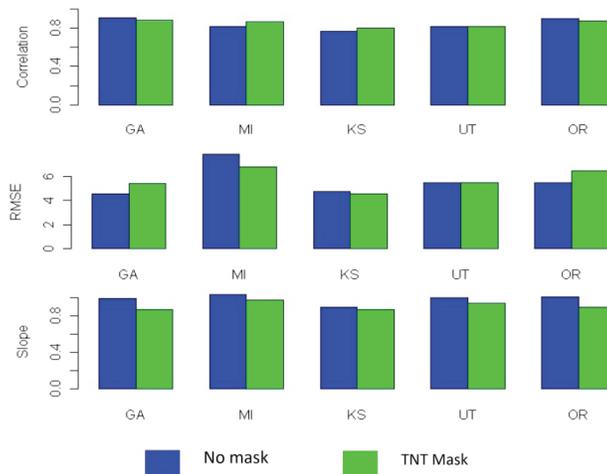


Figure 6—Correlation, RMSE, and slopes obtained in each of the five pilot areas when applying a tree canopy model without a mask (blue) versus a tree canopy model with an empirical mask (green) to an independent test set.

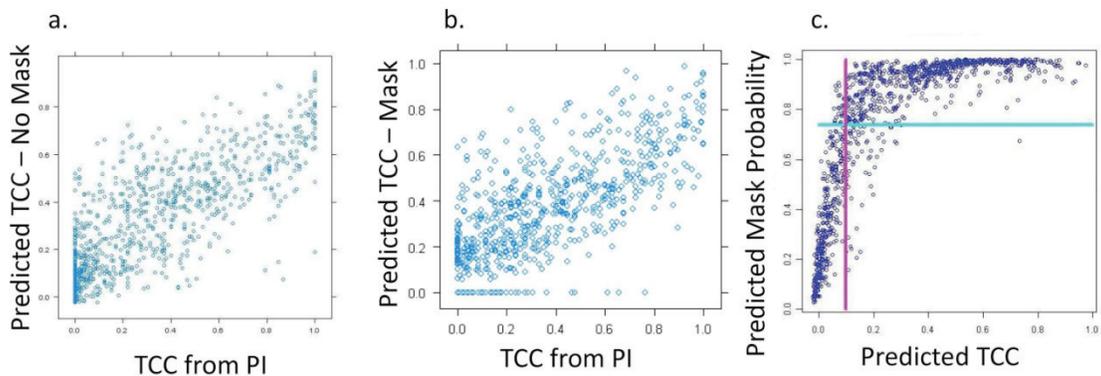


Figure 7—Scatter plots exploring effect of first creating a tree/no-tree mask prior to modeling tree canopy cover. In 7a, predicted tree canopy cover from a single unmasked model is plotted against the tree canopy cover response from the photo interpretation. In 7b, predicted tree canopy cover from the tree/no-tree model followed by the masked tree canopy model is plotted against the tree canopy cover response from the photo interpretation. In 7c, the predicted probability of having trees present from the tree/no-tree model is plotted against predicted tree canopy cover with no masking, with the prevalence-based threshold shown horizontally in blue and threshold a user might impose by applying a 10 percent cover threshold to the predicted tree canopy cover shown vertically in purple.

SAMPLING INTENSITY AND NORMALIZATIONS: EXPLORING COST-DRIVING FACTORS IN NATIONWIDE MAPPING OF TREE CANOPY COVER

John Tipton, Gretchen Moisen, Paul Patterson, Thomas A. Jackson,
and John Coulston

ABSTRACT

There are many factors that will determine the final cost of modeling and mapping tree canopy cover nationwide. For example, applying a normalization process to Landsat data used in the models is important in standardizing reflectance values among scenes and eliminating visual seams in the final map product. However, normalization at the national scale is expensive and logistically challenging, and its importance to model fit is unknown. Cost also increases with each location sampled, yet appropriate photo sampling intensity relative to the FIA grid has yet to be explored. In addition, cost is also affected by how intensively the photo plots themselves are sampled with a dot count, and the effect of reducing the number of dots on predictive models is also unknown. Using intensively sampled photo plot data in 5 pilot areas across the United States, we address these three cost factors by exploring the effect of a normalization process of Landsat TM data on model fits of tree canopy cover using Random Forests regression, the relationship between the sampling intensity of photo interpreted plots and model fit, and the relationship between the number of dots for each photo interpreted location and model fit.

INTRODUCTION

The National Land Cover Database (NLCD, <http://www.mrlc.gov/>) for 2011 will contain a map of tree canopy cover that will be a spatially explicit map-based data on percent tree canopy cover is used for forest management, estimates of timber production, determining the potential for and extent of fire danger and other management issues across the United States. The 2001 NLCD provides map-based estimates of percent tree canopy cover along with land cover and percent impervious cover (Homer and others 2004). The NLCD is a periodic product with an update cycle of approximately five years. However, because of funding constraints the percent tree canopy estimates were not updated for the 2006 NLCD. For the 2011 NLCD the U.S. Forest Service Forest Inventory and Analysis program (FIA) will take the lead on developing the percent tree canopy cover layer.

FIA is uniquely positioned to lead the development of the 2011 NLCD percent tree canopy cover layer. First, FIA uses a probabilistic sample design that covers all lands (forest and non-forest) and can be easily intensified for geospatial modeling purposes. Second, the FIA program is beginning to make percent tree canopy cover estimates for all sample locations. This provides an opportunity to leverage data collected as part of the FIA program to develop predictive models used to produce percent tree canopy map products. To this end, a pilot study was carried out in 2010. The pilot study was designed to answer specific research questions and estimate costs for developing the 2011 NLCD percent tree canopy cover map.

Creating a tree canopy cover product that encompasses the entire country presents many questions that must be answered before prototype or production mapping can begin. Consequently, a pilot project was launched that included five study areas, one each in Georgia, Michigan, Kansas, Oregon, and Utah. Within each study area, over two thousand photo plots were photo-interpreted by an interpreter looking at a grid overlaid on an aerial photo of each plot. At each of the 105 points on the grid, the interpreter determined if the point was a tree or not, and this response was used to calculate percent tree cover.

Using data from the pilot project, several issues are addressed in this paper to support production of mapping of tree canopy cover nationwide. First, the number of samples plays an important role in the quality of the model. It is important to find a balance between the quality of model fit and concerns of cost. Second, normalization of Landsat TM images is important because adjacent Landsat scenes on a map are not taken on the same day. Because of this, when a mosaic of multiple images is constructed, there will be

John Tipton, Colorado State University, Fort Collins, CO 80525

Gretchen G. Moisen, Research Forester, US Forest Service Rocky Mountain Research Station, Ogden, UT, 84401

Paul Patterson, Statistician, US Forest Service Rocky Mountain Research Station, Ogden, UT, 84401

Thomas A. Jackson, Colorado State University, Fort Collins, CO 80525

John W. Coulston, Research Forester, US Forest Service Southern Research Station, Knoxville, TN, 37919

seams in the image where the raw reflectance values for one image are not equal to the reflectance values of the adjoining image. Normalization of one image to another using the overlap between two images will remove the visual seam in a map, but the effect of normalization on how well a model predicts percent tree canopy cover has not been explored. Third, at each sample locations an estimate of percent tree canopy cover was made using a simple dot grid approach. The pilot study design used 105 dots however, if the same information can be obtained with fewer points, we can trim costs and maintain the quality of the model. Consequently in this paper, we explore the effects of sample size, normalization and number of dots on predictive models of tree canopy cover.

METHODS

Percent tree canopy cover data was collected for five study areas in the coterminous United States (Figure 2). The standard FIA sampling grid (1 plot per 2400 ha) was intensified fourfold to 1 plot per 600 ha using the techniques described by White et al. (1992). At each sample location a 105 point dot grid covering a 90m by 90m area was developed. At each of the 105 points, a photo interpreter determined if the point was a tree or not, by examining high resolution digital aerial photography collected in 2009 (USDA 2009). The percent tree canopy cover for each sample location was defined as the number of points intersecting tree crowns divided by 105 and was used as the dependent variable for random forest model development.

The independent variables came from a variety of sources but they were primarily Landsat 5 data and vegetation indices derived from Landsat data (e.g. normalized difference vegetation index, tasseled cap). Additionally, digital elevation models and derivatives (e.g. slope, aspect) were also used as potential independent variables for random forest model development. The Landsat data were available as normalized mosaics and non-normalized mosaics. Because each study area covered multiple Landsat scenes differences in spectral values among scenes arise because of differing collection data and atmospheric effects. The non-normalized data had no correction for these effects. The normalized data accounted for these effects by standardizing reflectance values from a target scene to a reference scene based on the overlap among scenes.

The specific modeling tool used was Random Forests, implemented in R using the library RandomForests (Liaw and Wiener 2002). Random Forests is a machine learning process that uses decision trees for classification and regression. The algorithm computes many trees, with each tree getting a “vote,” with the final model being

a majority decision (categorical variables) or average (continuous variables). For each node in those trees, a subset of explanatory variables is randomly selected and a dichotomous split in the data is made based on the largest decrease in the MSE of the data. To get the final model, the process is run for 500 trees, and the results are averaged. Each tree is constructed using a randomly selected set of the data where approximately one-third is held “out of bag” and can, therefore, be used as a validation data set and as a measure of model fit. Our measure of model fit is called pseudo R^2 and it represents the relative amount of variation in the data that is explained by the model. Pseudo R^2 is calculated as $1 - \text{MSE}/\text{Var}(y)$ where the pseudo R^2 is calculated individually for every tree in the forest, then averaged over all trees to compute the final value.

To investigate the question related to sample size, we performed an iterative sampling process where, for each iteration, plots were randomly sampled from our study site, a model is fit using the RandomForest command, and the measure of model fit (pseudo R^2) is recorded. Then, for the next iteration, the number in the sample was increased by 20 plot locations and so on until the number in the iterative sample equaled the total sample size for the study site. When plotting the pseudo R-squared values against the number of study site samples, we applied a lowess smoothing curve for each of the study site locations to get a visual indication of the asymptotic behavior. From this method, we were able to get estimates of the variance of the fit of the model as well as to determine the asymptotic behavior of model fit relative to sample size.

The simulations described above were performed for both the data set that was normalized (corrected for differences in Landsat scenes) and for the data set that was not normalized. This allowed us to also explore the asymptotic behavior of the model fit relative to normalization.

The final question had to do with the number of dots used for the photo interpretation grid. For each study site we sampled 500, 1000, and 1500 study locations and calculated the percent tree cover based on randomly sampling a number of photo dots. We started with sampling one dot, and then fit a Random Forest model and recorded the pseudo R^2 . The process was then iterated, increasing the number of dots by one each time. In the plot of model fit versus the number of dots, we applied a lowess smoothing curve to see patterns in the simulations and to get a visual indication of the asymptotic behavior relative to number of dots. Also, estimates of the number of man-hours needed to complete a prototype of the same size with different number of sample plots and numbers of dots were produced. This assumed 3 minutes for loading each sample plot picture and another 3 minutes to count all 105 dots.

RESULTS

From these simulations we were able to get an understanding of what intensity sampling intensity provides the most information for the least cost. In Figure 2 we call attention to the smoothed curve of pseudo R^2 versus the number of sample plots for the non-normalized data in Oregon. Looking at the spread of the simulated model fits we see that between 1000 and 2000 sample plots the variation in simulated pseudo R^2 drops off quickly. This is of interest because the default FIA sampling intensity grid for this study site is approximately 1500. A similar pattern is seen in Figure 3, in which the variation in simulated pseudo R^2 drops off quickly between 1000 and 2000 sample plots for the other four study sites.

Figures 1 and 2 also show the effect of normalization on model fit. When looking at the plot of sample size versus pseudo R^2 for Oregon in Figure 2 we see that there is little difference in the fit of our model with regards to whether the data was normalized or not normalized. When looking at the four plots in Figure 3 we see the same pattern in Georgia, Utah, and Michigan, but we have different results in Kansas. In the Kansas plot we see that the normalized data model outperforms the non-normalized data model, but the difference is small (at 4000 sample points the difference in pseudo R^2 between the normalized and non-normalized models is about .03). These results indicated that normalization plays a very minor role in the quality of model fit, and we made the decision to consider only the non-normalized data set for the rest of the analyses.

In Figures 3 and 4 we are looking at the plots of pseudo R^2 versus number of dots on the photo grid. By looking at the plots of number of dots versus pseudo R^2 in Figure 4 we see that in Oregon we are not getting more information by including more than 40 dots. This is evidenced by the inflection in the lowest smoothing curve on the plot. The same pattern is repeated in Figure 5 for the other study sites. By combining the recommendations of using non-normalized data and roughly 1000 sample plots per study site we are able to make estimates of the amount of man-hours needed to complete a study site of similar size. Figure 6 shows the amount of person-hours needed versus the number of photo interpretation grid dots for 500, 1000, and 1500 sample plots. Using our assumptions that each image takes three minutes to load and three minutes to calculate tree cover using all 105 dots, we plotted the number of photo grid dots versus time for 500, 1000, and 1500 sample plots. From this we can see that if we used 1000 sample plots with 40 dots we would expect one person to finish all five study areas in about 12 weeks.

DISCUSSION

By looking at the smoothed curves for the non-normalized data in Figures 1 and 2 we see the relationship between the number of sample plots and the precision of the model fit as measured by pseudo R^2 . We see that between 1000 and 2000 sample plots the variation in pseudo R^2 decreases rapidly versus the number of sample plots when compared to larger sample sizes. This suggests diminishing returns in model fit when increasing the number of sample plots beyond values in the 1000 to 2000 range. This suggests that we can get good model relative to cost in the 1000 to 2000 sample plot range, which also happens to be approximately the FIA standard sampling intensity grid for each study site.

Choosing to use only non-normalized data to fit a Random Forests model has major implications for the budget of the project. Normalization is an expensive and time consuming process, especially on a scale the size of the entire United States. Our results indicate that the Random Forests model performs equally well using either normalized or non-normalized data. From this result, we are able to make recommendations to get a higher quality product for less cost. However, the visual effects of not normalizing are still under investigation.

Because a human observer will be used to measure percent tree cover in the final product, using fewer dots will decrease the time the observer will spend on each photo, which will decrease the overall cost of the project. Since it appears that we gain little in terms of model fit when considering more than 40 dots, this suggests that we can reduce the person-hours needed for the prototype.

CONCLUSION

Because there are limited resources available it is important to get an understanding of the behavior of the sampling protocols and model fits relative to the costs of the process. The recommendations in this paper give guidelines for the next prototype phase of the NLCD Canopy Cover project.

ACKNOWLEDGEMENTS

We would like to thank everyone who has worked with this pilot project for collecting the data, taking the time to photo interpret the images, and provide counsel on this project. Also we would like to thank Dr. Jean Opsomer for his time and assistance.

LITERATURE CITED

Liaw, A.; Wiener, M. 2002. Classification and Regression by randomForest. R News 2(3), 18—22.

Breiman, L. 2001. “Random Forests.” Machine Learning. 5-32.

Breiman, L.; Friedman, R. A.; Olshen, R. A.; Stone, C. G. 1984. Classification and regression trees. Pacific Grove, CA, USA: Wadsworth.

Homer, C.; Haug, C.; Yang, L. [and others]. 2004. Development of a 2001 national land-cover database for the United States. Photogrammetric Engineering and Remote Sensing. 70, 829-840.

U.S. Department of Agriculture. 2009. National Agriculture Imagery Program, Salt Lake City; U.S. Department of Agriculture, Farm Service Agency, Aerial Photography Field Office. Information available: <http://www.apfo.usda.gov/FSA/apfoapp?area=home&subject=prog&topic=nai>

White, D.; Kimerling, A.J.; Overton, W.S. 1992. Cartographic and geometric components of a global sampling design for environmental monitoring. Cartography and Geographic Information Systems 19: 5-22.



Figure 1—Location and extent of the five pilot study areas.

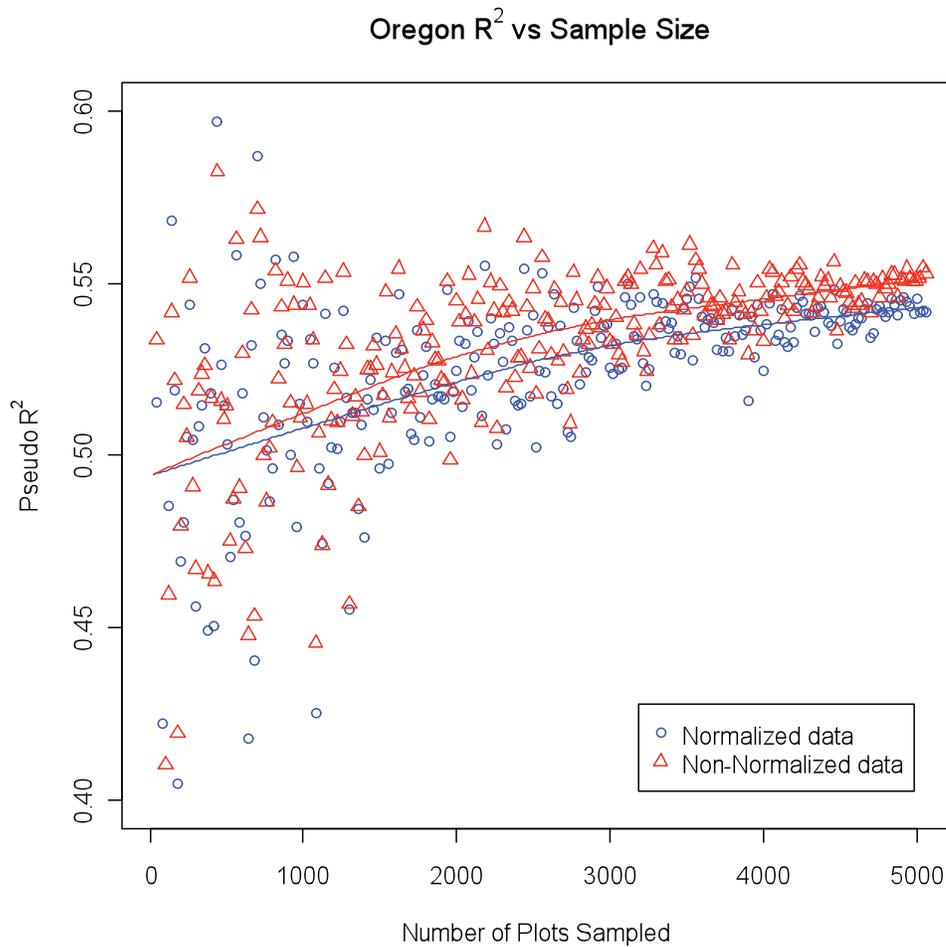


Figure 2—Shows the pseudo- R^2 values plotted against the number of plots sampled for Oregon for both the normalized and non-normalized data sets with the solid lines representing a lowess smoothing curve.

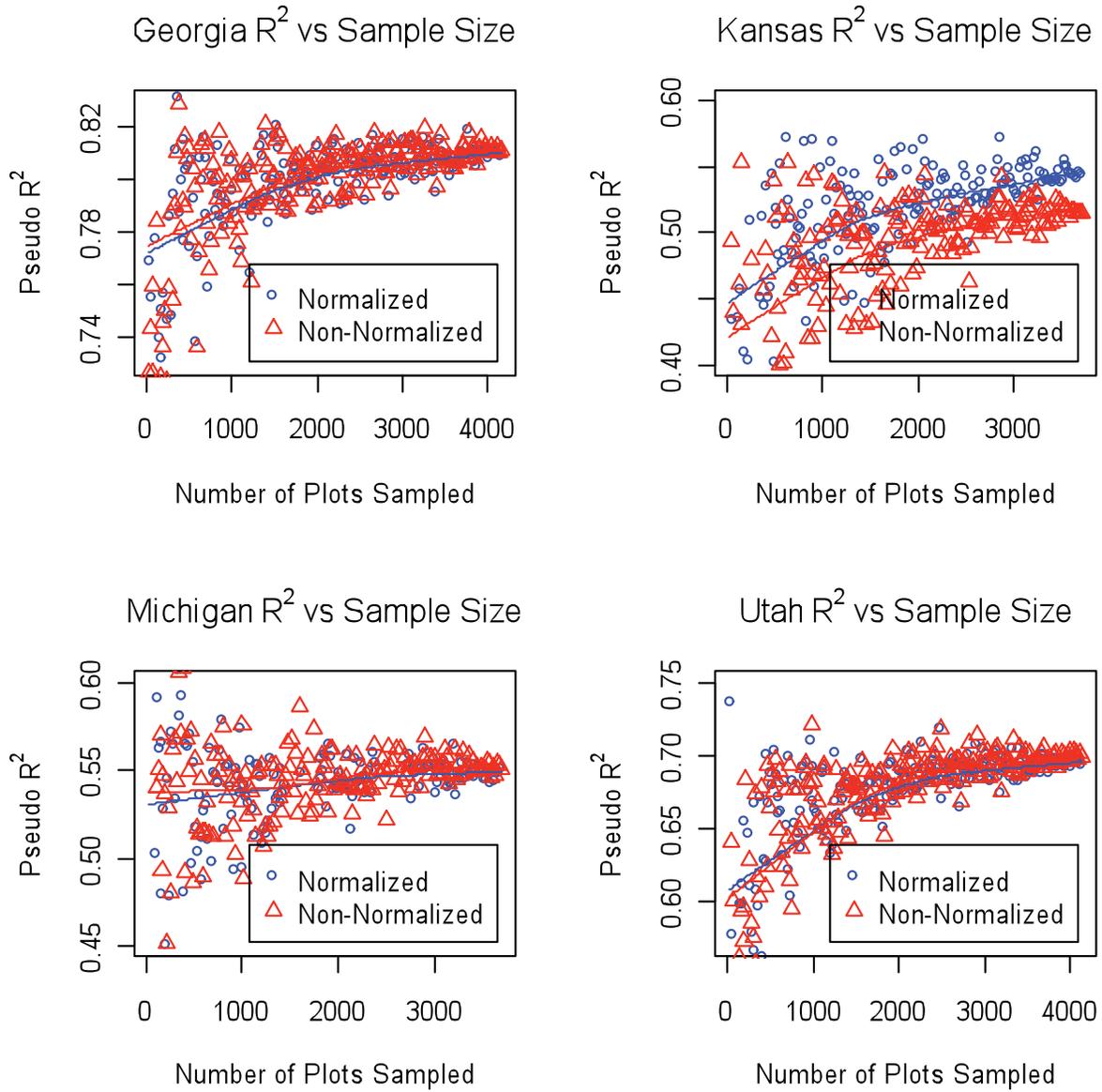


Figure 3—Shows the pseudo-R² values plotted against the number of plots sampled for Georgia, Kansas, Michigan, and Utah for both the normalized and non-normalized data sets with the solid lines representing a lowess smoothing curve.

Oregon R^2 vs Sample Size

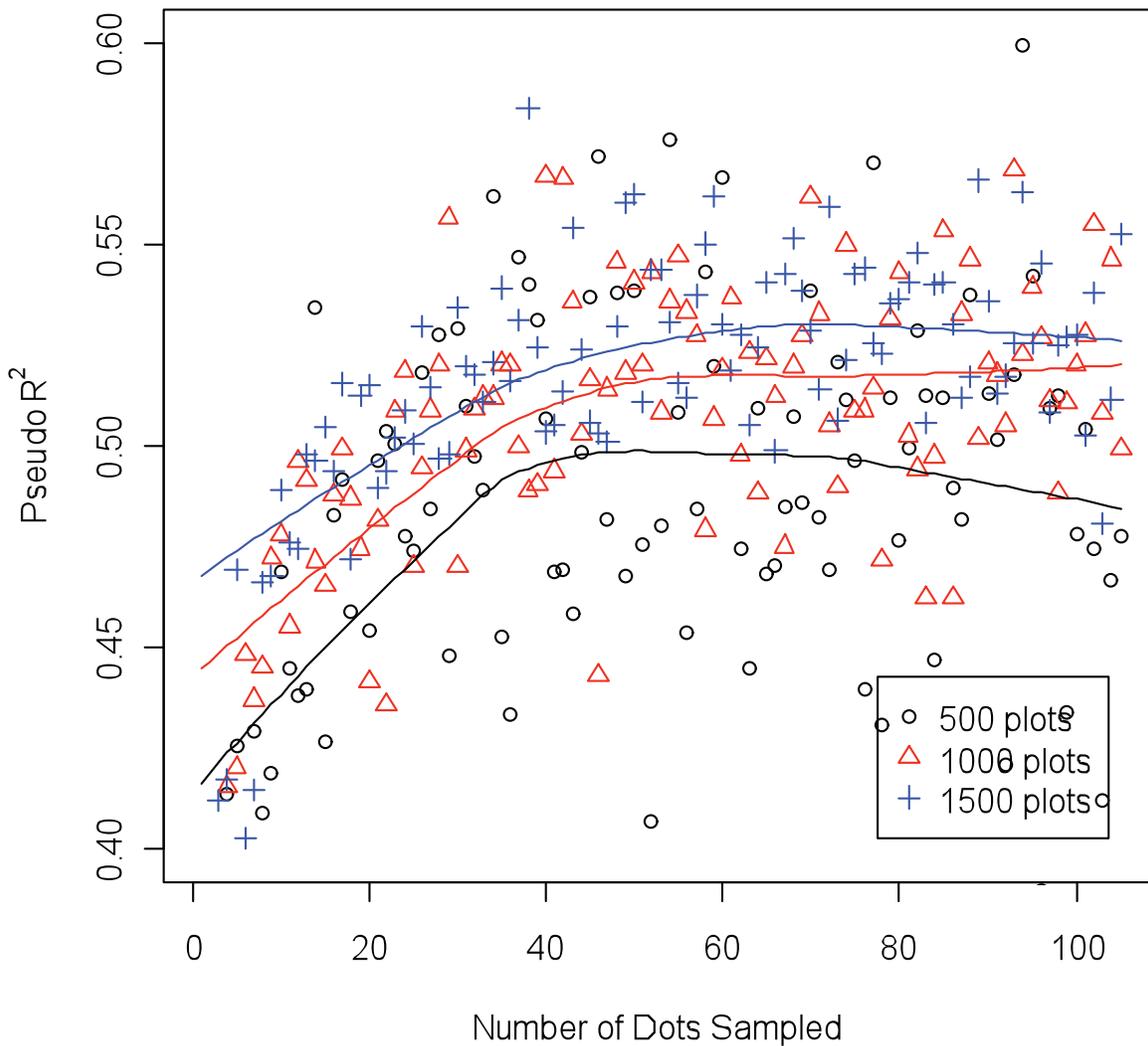


Figure 4—Shows the pseudo- R^2 values plotted against the number of dots sampled for Oregon, for both the 500, 1000, and 1500 sample plots with the solid lines representing a lowest smoothing curve.

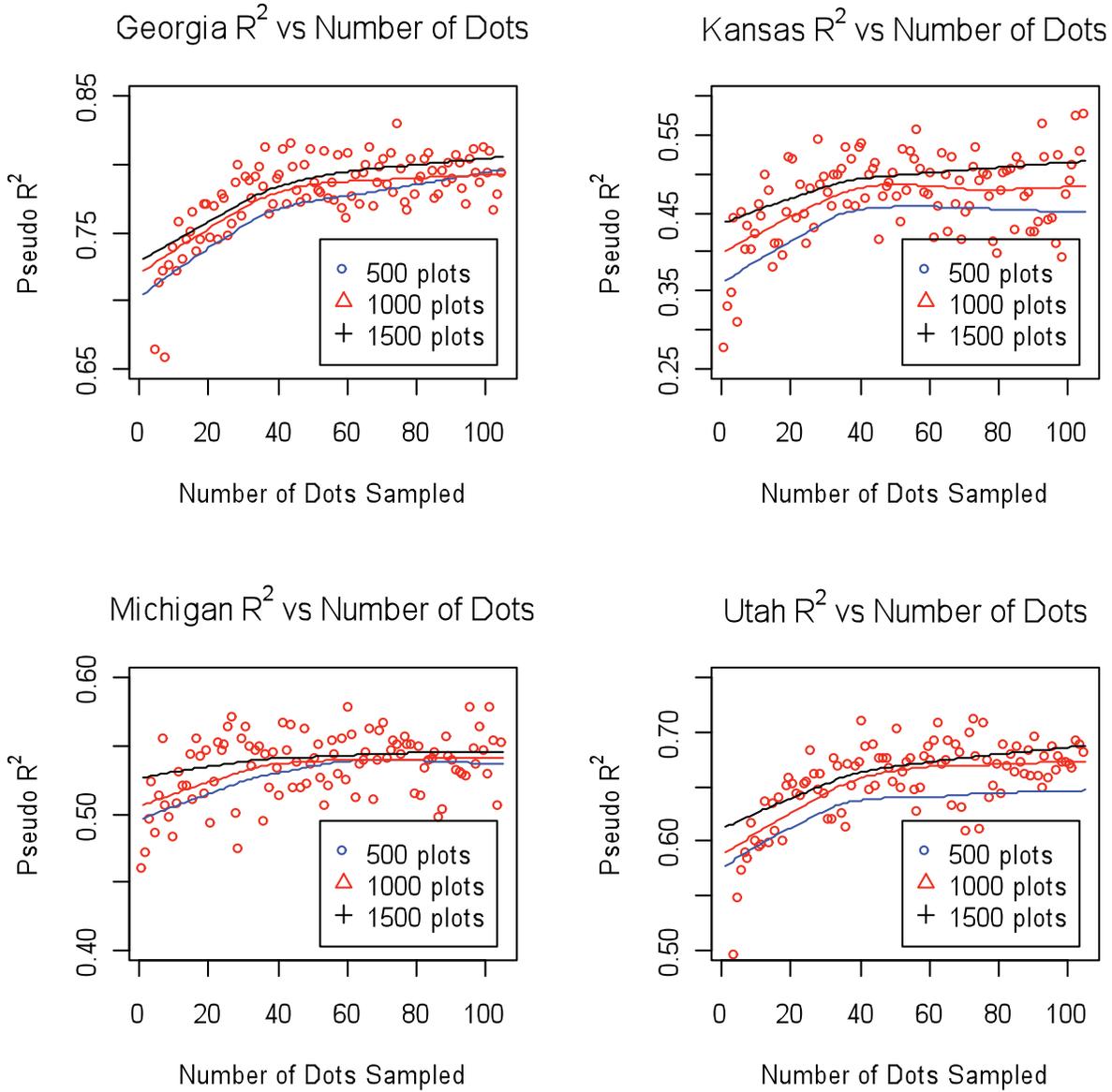


Figure 5—Shows the pseudo-R2 values plotted against the number of dots sampled for Georgia, Kansas, Michigan, and Utah, for both the 500, 1000, and 1500 sample plots with the solid lines representing a lowest smoothing curve.

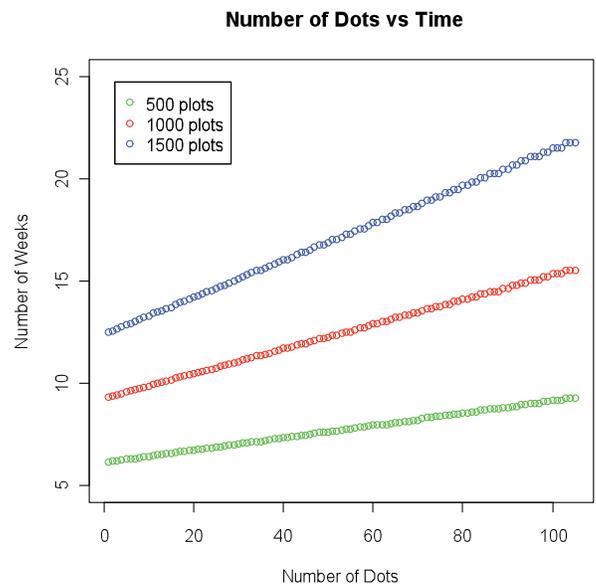


Figure 6—Shows the amount of time to complete a prototype of similar size toon the five study sites versus the number of dots used in photo interpretation for 500, 1000, and 1500 sample plots.

ASSESSING ALTERNATIVE MEASURES OF TREE CANOPY COVER: PHOTO-INTERPRETED NAIP AND GROUND-BASED ESTIMATES

Chris Toney, Greg Liknes, Andy Lister, and Dacia Meneguzzo

ABSTRACT

In preparation for the development of the National Land Cover Database (NLCD) 2011 tree canopy cover layer, a pilot project for research and method development was completed in 2010 by the USDA Forest Service Forest Inventory and Analysis (FIA) program and Remote Sensing Applications Center (RSAC). This paper explores one of several topics investigated during the NLCD pilot. We compared estimates of tree canopy cover derived by photo-interpretation (PI) of 1-m resolution NAIP imagery to modeled estimates based on field-measured tree data collected on FIA plots in five study areas in Georgia, Michigan, Kansas, Utah, and Oregon, and to direct measurements of canopy cover by line intercept on FIA plots in Utah only. Photo-interpreted NAIP overestimated tree canopy cover (+10 to +20 percent canopy cover) at forested FIA plot locations compared with ground-based estimates derived from stem-mapped tree data or line intercept field measurements. Oblique viewing angles at sample locations away from the image nadir, and excessive shadowing in some NAIP images, could be the primary reasons for overestimation of canopy cover by PI. We also examined canopy cover estimates derived from NAIP imagery using an automated algorithm implemented in image processing software, as an alternative to manual PI by humans. This initial test showed that automated PI of NAIP images by image analysis could be a feasible approach for generating canopy cover data at reduced time and cost, but the current rule set exacerbated the problem of overestimation.

INTRODUCTION

The National Land Cover Database (NLCD) comprises a suite of 30-m resolution map layers depicting land cover characteristics for the United States, developed by the Multi-Resolution Land Characteristics Consortium (www.mrlc.gov). The NLCD 2001 product suite included a percent tree canopy cover layer based on circa 2001 LANDSAT imagery. A decadal update to the NLCD 2001 products is scheduled to begin production in fall 2011. In preparation for the development of an NLCD 2011 tree canopy cover layer, a pilot project focused on research and method development was completed in 2010 by the USDA Forest Service Forest Inventory and Analysis (FIA) program and the USDA Forest Service Remote Sensing Applications Center (RSAC).

Mapping tree canopy cover at continental scales involves developing empirical models to relate percent canopy cover from a set of reference locations to predictor variables derived primarily from satellite imagery. Reference data on tree canopy cover can be obtained by different methods including direct field measurements on FIA plots where available (e.g., Interior West FIA line intercept, USDA Forest Service 2007), estimates derived from models based on tree-level data (e.g., stem-mapping, Toney and others 2009), and estimates derived by sampling high-resolution imagery using either human- or computer-based interpretation methods. In addition to differences in cost and processing time, each method of canopy cover observation entails different sources of error and bias relative to other methods. As a result, frequency distributions of tree canopy cover observations across the landscape may differ appreciably among different methods. The distribution of tree canopy cover in the reference data should largely determine the distribution of predicted canopy cover depicted in final map products, assuming reasonably good models.

This paper explores one of several topics investigated during the NLCD pilot. Our objective was to describe characteristics of and relationships between alternative measurements of tree canopy cover for reference locations, specifically, to compare estimates of canopy cover derived by photo-interpretation (PI) of 1-m resolution aerial imagery with estimates derived from field-measured tree data collected on FIA plots. We also examined canopy cover estimates derived from aerial imagery using an automated algorithm implemented in image processing software, as an alternative to manual PI by humans.

Chris Toney, US Forest Service, Rocky Mountain Research Station, Forest Inventory and Analysis, Missoula, MT 59808
Greg Liknes, US Forest Service, Northern Research Station, Forest Inventory and Analysis, St. Paul, MN 55108
Andy Lister, US Forest Service, Northern Research Station, Forest Inventory and Analysis, Newtown Square, PA 19073
Dacia Meneguzzo, US Forest Service, Northern Research Station, Forest Inventory and Analysis, St. Paul, MN 55108

METHODS

COMPARISON OF PHOTO-INTERPRETATION WITH GROUND-BASED CANOPY COVER ESTIMATES

For the PI canopy cover estimates, a grid of PI locations was established across five study areas located in Georgia, Michigan, Kansas, Utah, and Oregon as part of the pilot effort (figure 1). The PI grid was designed as an intensified FIA grid so that a subset of the PI locations was coincident with the FIA sample locations within each pilot area. The aerial imagery used was natural color, 1-m resolution and obtained from the National Agricultural Imagery Program (NAIP) (USDA Farm Service Agency 2009). The PI method used 90 m x 90 m dot grids on 2009 NAIP imagery at each sample location, with 105 regularly spaced points in each grid classified by human interpreters as either tree canopy or not tree canopy. Shadowed areas were left to the interpreter's judgment, and all 105 points in each grid were used for the canopy cover estimate. Percent canopy cover for the dot grid was calculated as the number of points classified as tree canopy divided by 105. In each pilot area, we included only the subset of PI locations that were coincident with entirely forested FIA plots (all conditions on the plot classified as forest), since FIA plots (or portions of plots) classified as nonforest lack any ground measurements from which a canopy cover estimate could be derived. We also included only the FIA plots that had been sampled within two years of the NAIP images (plots measured in 2007, 2008, and 2009 if available) so that the plot measurements would be reasonably concurrent with the NAIP imagery.

Ground-based estimates of tree canopy cover for the forested FIA plots were obtained by stem-mapping individual trees within each plot, and predicting the dimensions of each tree crown from stem diameter using published equations. Details of the stem-map model were described by Toney and others (2009). Canopy cover from the stem-map model is defined as an estimate of the vertically projected canopy cover of live FIA tally trees on the plot that are 1-inch diameter and larger. The stem-mapping approach uses the spatial data on individual trees that are available for FIA plots, unlike the canopy cover estimate in the Forest Vegetation Simulator (Crookston and Stage 1999) which assumes that trees are distributed randomly within the stand for the purpose of overlap adjustment. The stem-map model was developed and validated with line-intercept field measurements of canopy cover (USDA Forest Service 2007) on approximately 12,000 plots from the Interior West FIA unit (Idaho, Montana, Wyoming, Nevada, Utah, Colorado, Arizona, and New Mexico). Model predictions were compared to field measurements of canopy cover on 1,454 plots that were not used in model development. The mean absolute difference between field-measured and model-predicted values was

± 7.9 percent canopy cover, with mean bias of -0.7 percent canopy cover. The relationship between field-measured and predicted values was linear with approximately constant variance and a correlation coefficient $r = 0.875$.

Since field-measured canopy cover was available from Interior West FIA but not from the other regional FIA units, we also compared PI canopy cover with line intercept canopy cover in the Utah pilot area. Line intercept canopy cover in Utah was measured with four 25-foot transects in each of four subplots per FIA plot, arranged in cardinal directions beginning 1 foot from the subplot centers. The length of crown interception of live tally trees was recorded along each transect. Canopy cover was calculated by FIA condition class within the plots, by dividing the total live crown interception length by the total length of transect within each condition (400 feet total transect length in plots where all conditions are forested). For plots in which more than one forest condition was delineated by field crews, plot-level canopy cover was calculated as a weighted average of condition-level canopy cover, weighted by the proportion of the total plot area that each condition occupied. Measurement precision of the line-intercept canopy cover was assessed by FIA using blind check plots during 2000-2003 (Pollard and others 2006). A target tolerance of ± 10 percent canopy cover was specified for the measurement. Blind check data showed that measurements were within tolerance 88 percent of the time, and were within 2x tolerance 99.1 percent of the time ($n = 101$ plots).

PI canopy cover was compared with ground-based canopy cover (i.e., stem-map modeled canopy cover, along with line intercept field measurements in Utah only) by qualitative analysis of scatterplots. The mean absolute difference was calculated as

$$\text{mean_absolute_difference} = \left(\sum |PI_i - G_i| \right) \div n$$

where n is the number of plots, PI_i is the PI canopy cover for plot i , and G_i is the ground-based canopy cover for plot i . Bias was assessed with the mean difference:

$$\text{mean_difference} = \sum (PI_i - G_i) \div n$$

AUTOMATION BY OBJECT-BASED IMAGE ANALYSIS

An object-based image analysis (OBIA) approach was used to estimate tree canopy cover at 488 PI locations in the Georgia pilot area, as an initial test of the feasibility of automating the PI work flow. The OBIA method involves two steps: image segmentation and classification. Image segmentation divides an image into "spatially continuous, disjoint, and homogeneous regions" (de Jong and van deMeer, 2004), which are referred to as image objects. In

OBIA approaches, classification occurs on the image objects as opposed to individual pixels. The classification process involved separating image objects into “tree canopy” and “not tree canopy” categories. As a final step, percent canopy cover was derived by calculating the area occupied by image objects belonging to the “tree canopy” category and dividing by the total area of each image chip (i.e., the area sampled by the 90-m x 90-m dot grids).

Images were processed with eCognition Developer (v. 8.0.1), and both image segmentation and classification steps were included in a single rule set. A multi-resolution segmentation algorithm was used to segment each image (scale parameter: 30; shape criterion: 0.1), and utilized red, green, blue, near-infrared, and NDVI input bands. In the classification phase, NDVI was used to discriminate between vegetated and non-vegetated parts of each image chip. In order to separate tree cover from other vegetation, a series of expressions that used spectral and texture information were combined to identify tree-cover image objects.

RESULTS AND DISCUSSION

COMPARISON OF PHOTO-INTERPRETATION WITH GROUND-BASED CANOPY COVER ESTIMATES

Differences in size and shape of the sampled areas were responsible for some of the variability in the relationship between PI canopy cover derived from the 90-m dot grids, and canopy cover estimates derived from tree measurements on the four 24-foot radius subplots comprising each FIA plot location (figure 2). However, there was a consistent pattern of higher canopy cover estimates by PI compared with ground-based estimates. At forested FIA plot locations in the Georgia pilot area, PI canopy cover was higher than stem-map canopy cover by an average of 19, with a mean absolute difference of ± 21 percent canopy cover (figure 3a). Eighty-five percent of the PI estimates at forested plots in Georgia were in the highest cover class of 90 to 100 percent, with a median value of 98 percent. Stem-mapped canopy cover in Georgia was more evenly distributed between 50 and 100 percent with a median value of 80 percent. The pattern was similar in the Michigan pilot area where 83 percent of the PI estimates at forested plot locations were in the highest cover class with a median value of 100 percent, while the stem-mapped cover was more evenly distributed between 50 and 100 percent with a median value of 75. The mean difference was 24 percent canopy cover (figure 3b). At forested plot locations in the Kansas pilot area, PI canopy cover was higher than stem-map canopy cover by an average of 20, with a mean absolute difference of ± 25 percent canopy cover (figure 3c). Variability and mean differences were lower in the western pilot areas. At forested plot locations in the Utah pilot area, PI canopy

cover was higher than stem-map canopy cover by an average of 13, with a mean absolute difference of ± 18 percent canopy cover (figure 3d). As expected, a similar relationship was seen between PI canopy cover and field-measured line intercept canopy cover on the Utah plots. PI canopy cover was higher than line intercept canopy cover by an average of 12, with a mean absolute difference of ± 17 percent canopy cover (figure 3e). At forested plot locations in the Oregon pilot area, PI canopy cover was higher than stem-map canopy cover by an average of 13, with a mean absolute difference of ± 16 percent canopy cover (figure 3f).

A tendency for overestimation of tree canopy cover derived from NAIP imagery was expected due to off-nadir view angles and excessive shadowing in some images (Guess 2010), but the magnitude of differences relative to ground-based estimates has not been previously quantified. The viewing angle between camera and ground is variable within a NAIP image depending on distance from the flight line. Portions of an image near the flight line provide vertical viewing (near nadir), but with increasing distance from the flight line view angles become oblique (off-nadir). Trees viewed at oblique angles appear to occupy a greater area than their canopies actually cover with a vertical projection, and openings can be obscured (Guess 2010). Tree canopy measurements are expected to be most accurate near the image nadir (Korpela 2004). Shadowing can also be heavy in NAIP images due to sun angle and terrain. Excessive shadowing may lead to overestimation of canopy cover since shadows can make the canopy appear denser than it actually is, and shadows may cause additional difficulty in discerning non-tree background vegetation from tree canopy. View angles vary from location to location within images, and shadowing varies from image to image, so these sources of bias could be highly variable in magnitude. The current NAIP products do not include flight line data, so analysis of, and possible adjustment for these sources of bias do not appear feasible at present.

It is possible that the stem-map model underestimates canopy cover in some eastern forest types. The stem-map model has two components. A geometric component involves overlaying the crowns of stem-mapped trees ≥ 5 inches diameter on the subplot boundaries to calculate vertically projected canopy cover accounting for overlap. Crown dimensions are predicted from stem diameters using equations from Bechtold (2003, 2004), Bragg (2001), Gill and others (2000), and others. Trees ≥ 1 inch diameter but < 5 inches diameter are denoted as saplings in FIA protocols and are only measured in one 6.8-foot radius microplot within each FIA subplot. Since saplings cannot be stem-mapped across the entire plot, the contribution of saplings to total canopy cover is estimated with a regression equation that includes predictor variables characterizing stand structure and the spatial pattern of overstory trees. This

empirical component of the stem-map model was developed using field data from the Interior West FIA unit, and has not been validated in eastern forest types. To the extent that eastern forest types tend to differ structurally from interior west types, especially if they tend to be more multi-layered, the stem-map model could underestimate the total canopy cover of trees ≥ 1 inch diameter on the plot. If this is true, then the magnitude of overestimation from photo-interpreted NAIP is probably overstated for the Georgia, Michigan, and Kansas pilot areas (+19 to +24 percent canopy cover on average). It is possible that the actual overestimation for these pilot areas is more similar to that reported for the western pilot areas (+12 to +13 percent canopy cover on average) where the stem-map model has known reliability.

OBJECT-BASED IMAGE ANALYSIS

The OBIA method consistently overestimated tree canopy cover relative to human photo interpretation using a dot grid (figure 4). OBIA canopy cover was higher than PI canopy cover by an average of 11 (mean difference ranging from 0 to 94). During the segmentation process, continuous image objects that included shadowed areas between or adjacent to individual tree crowns were created. As a result, it was expected the OBIA would tend to estimate a higher percentage of canopy cover than a human interpreter in areas with less dense canopies, which could be exacerbated in images with excessive shadowing due to low sun angles. Other issues such as variations in color and brightness among images and the limited contextual information due to the small image extents led to some difficulty in discriminating trees from other vegetation in some images (figure 5). Some of these problems could be overcome with additional development work on the OBIA approach. Once a rule set is developed, the approach requires little manual intervention and processing time is fast. In contrast to the stem mapping approach, the OBIA method does not rely on the availability of in situ tree data, and unlike human photo interpretation it is fully repeatable.

CONCLUSIONS

The use of photo-interpreted NAIP as reference data for continental mapping of tree canopy cover has important operational advantages. The generation of PI data does not require in situ tree measurements. In contrast, most FIA sample locations classified as nonforest by definition are not field-visited, even though many of these nonforest locations can have significant tree cover (e.g., urban and residential areas). Likewise, PI data can be collected relatively efficiently from an intensified grid, so that (at least low levels of) intensification are probably feasible within time and cost constraints, resulting in larger sample sizes and increased map accuracy. PI data from NAIP

generally can be produced so that it is concurrent with the satellite imagery being classified. In contrast, the FIA grid is sampled on a 5-year (eastern U.S.) or 10-year (western U.S.) inventory cycle. Plot measurements that are disjoint in time from image acquisition dates may no longer reflect ground conditions accurately for the time period of interest. The spatial registration of reference locations with pixels in the satellite imagery also should be as accurate as possible to support the development of predictive models. The horizontal accuracy of NAIP is currently specified as “inspected locations match photo-identifiable ground control points with an accuracy of within 6 meters at a 95 percent confidence level” (USDA Farm Service Agency 2009). The 90-m x 90-m dot grids used for PI in the NLCD pilot were positioned coincident with 3x3 pixel blocks on the 30-m LANDSAT imagery to be classified. Error rates for FIA plot coordinates have not been described systematically and could exceed NAIP specifications, and FIA plot footprints cannot be “snapped” to specific pixel configurations.

Photo-interpreted NAIP appears to overestimate tree canopy cover at forested FIA plot locations compared with ground-based estimates derived from stem-mapped tree data or line intercept field measurements. The magnitude of overestimation is likely in the range of 10 to 20 percent canopy cover on average, but with moderate to high variability that may be related to characteristics of the NAIP imagery. Oblique viewing angles at sample locations away from the image nadir, and excessive shadowing in some NAIP images, could be the primary reasons for overestimation of canopy cover. Map products developed from PI reference data derived from NAIP may depict canopy cover in forested areas that is too high on average, and the variability of mapped canopy cover in some forested areas could be artificially low if nearly all PI samples are in the highest cover class (90 to 100 percent canopy cover). Future research should consider the possibility of obtaining flight line information for NAIP images to test the feasibility of adjustments to PI canopy cover for view angle and sun elevation.

An initial test showed that automated PI of NAIP images by object-based image analysis could be a feasible approach for generating canopy cover data at reduced time and cost. An algorithm-based method also has the advantage of being fully repeatable. However, the rule set used in the present study exacerbated the problem of overestimation, resulting in percent canopy cover values higher by 11 on average compared with human PI of NAIP. Additional development work on the OBIA approach, ideally in conjunction with research on the effects of view angle and sun angle mentioned above, is warranted considering the potential for large gains in efficiency.

ACKNOWLEDGMENTS

Portions of this work were performed in collaboration with the LANDFIRE program in support of tree canopy cover mapping in LANDFIRE Refresh. Demetrios Gatzolis and Chengquan Huang provided helpful comments on the manuscript.

LITERATURE CITED

Bechtold, W.A. 2003. Crown-diameter prediction models for 87 species of stand-grown trees in the eastern United States. *South. J. Appl. For.* 27(4):269-278.

Bechtold, W.A. 2004. Largest-crown-width prediction models for 53 species in the western United States. *West. J. Appl. For.* 19(4):245-251.

Bragg, D.C. 2001. A local basal area adjustment for crown width prediction. *North. J. Appl. For.* 18(1):22-28.

De Jong, S.M.; van der Meer, F.D. (eds.). 2004. *Remote Sensing Image Analysis: Including the Spatial Domain*. Dordrecht, The Netherlands: Kluwer Academic Publishers. 359 p.

Gill, S.J.; Biging, G.S.; Murphy, E.C. 2000. Modeling conifer tree crown radius and estimating canopy cover. *For. Ecol. Manage.* 126:405-416.

Guess, A. 2010. Using NAIP Imagery to Estimate Tree Canopy Cover - A Preliminary Evaluation. Forest Service Remote Sensing Tips, RSAC-5001-TIP1: USDA Forest Service Remote Sensing Applications Center, Salt Lake City, UT. 4 p.

Korpela, I. 2004. Individual tree measurements by means of digital aerial photogrammetry. *Silva Fennica Monographs* 3. 93 p.

Toney, C.; Shaw, J.D.; Nelson, M.D. 2008. A stem-map model for predicting tree canopy cover of Forest Inventory and Analysis (FIA) plots. In: McWilliams, W.; Moisen, G.; Czaplewski, R. (comps.) *Proceedings, 2008 Forest Inventory and Analysis (FIA) symposium; October 21-23, 2008; Park City, UT. Proc. RMRS-P-56CD*. Fort Collins, CO: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station. 1 CD.

USDA Farm Service Agency. 2009. National Agricultural Imagery Program (NAIP) fact sheet. http://www.fsa.usda.gov/Internet/FSA_File/naip_2009_info_final.pdf. Date accessed: December 3, 2010.

USDA Forest Service. 2007. Interior West Forest Inventory and Analysis field procedures, version 3.01. <http://www.fs.fed.us/rm/ogden/data-collection/field-manuals.shtml>. Date accessed: November 30, 2010.

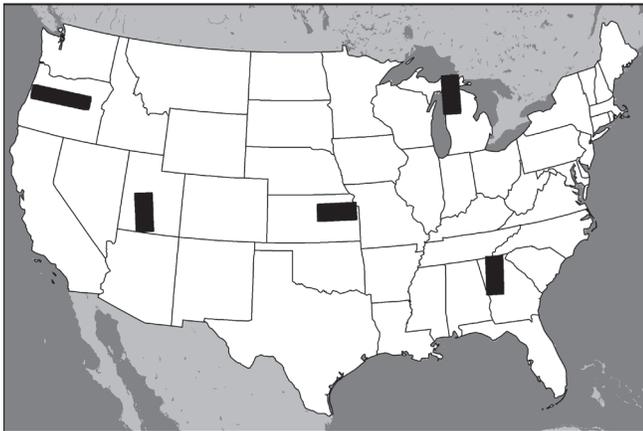


Figure 1—Locations of five pilot study areas (black shaded) in Georgia, Michigan, Kansas, Utah, and Oregon.

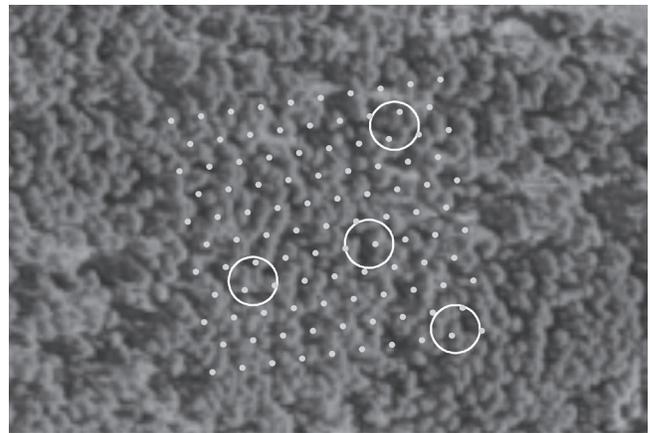
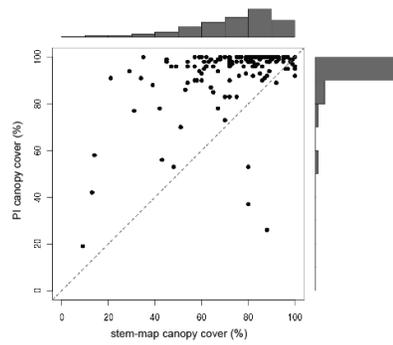
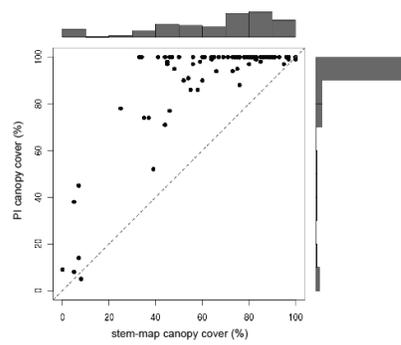


Figure 2—Example of a 90 m x 90 m grid containing 105 photo-interpretation points on NAIP imagery used to estimate tree canopy cover, along with the 24-foot radius subplots of a coincident FIA sample location.

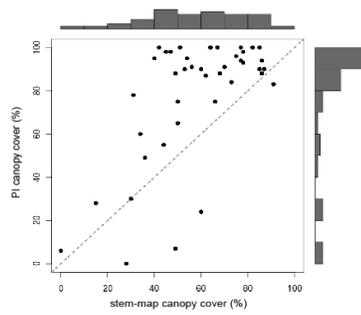
a. Georgia



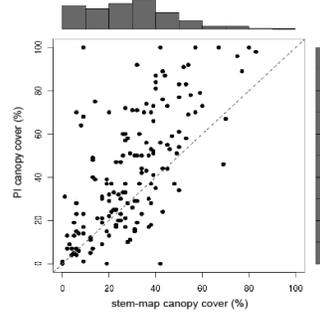
b. Michigan



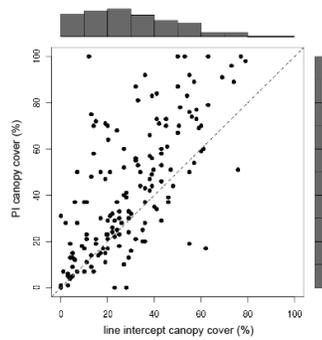
c. Kansas



d. Utah



e. Utah



f. Oregon

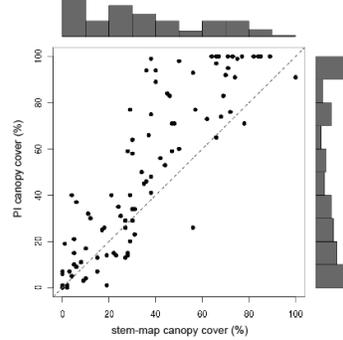


Figure 3—Comparison of tree canopy cover estimated by photo-interpretation (PI) of NAIP images with canopy cover estimated from stem-mapped tree data at forested FIA plot locations in five pilot study areas in the US, and with canopy cover measured in the field by line intercept on forested FIA plots in the Utah pilot area only. The histograms for each variable are in the margins and the dashed line is the 1:1 line. a) 144 plot locations in Georgia, b) 89 plot locations in Michigan, c) 42 plot locations in Kansas, d) 159 plot locations in Utah (stem-mapped canopy cover estimates), e) 158 plot locations in Utah (line intercept canopy cover measurements), and f) 98 plot locations in Oregon.

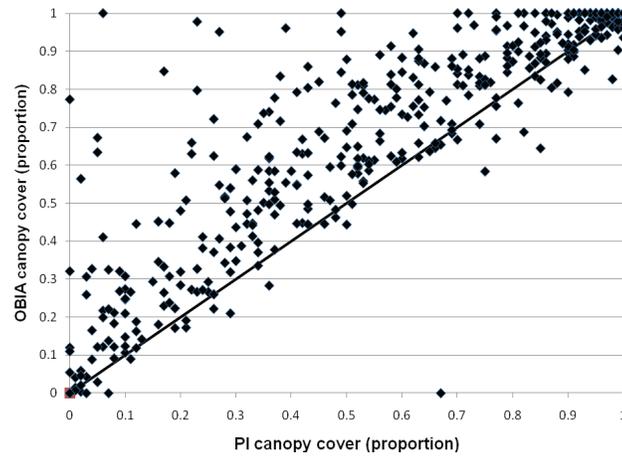


Figure 4—Comparison of tree canopy cover as estimated by human photo-interpreters (PI) interpreters using a dot grid with canopy cover estimated by an object-based image analysis (OBIA) automated mapping approach for 488 locations in Georgia, USA.

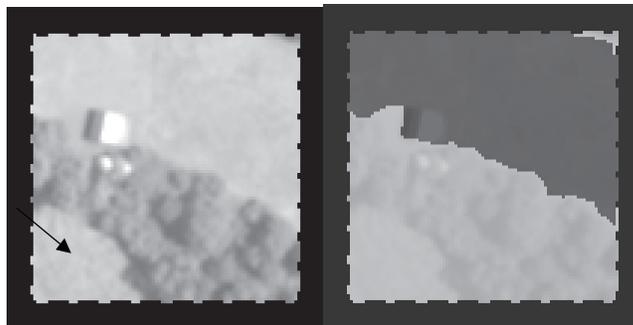


Figure 5—NAIP image chip on the left from Georgia, USA. The image on the right shows the classified image created using an object-based image analysis approach (dark gray = not tree canopy, light gray = tree canopy). The arrow indicates an area where grassy lawn was mistakenly classified as 'tree.'

A TOOL TO DETERMINE CROWN AND PLOT CANOPY TRANSPARENCY FOR FOREST INVENTORY AND ANALYSIS PHASE 3 PLOTS USING DIGITAL PHOTOGRAPHS

Matthew F. Winn and Philip A. Araman

ABSTRACT

The USDA Forest Service Forest Inventory and Analysis (FIA) program collects crown foliage transparency estimates for individual trees on Phase 3 (P3) inventory plots. The FIA crown foliage estimate is obtained from a pair of perpendicular side views of the tree. Researchers with the USDA Forest Service Southern Research Station have developed a computer program that uses a different approach to estimate transparency utilizing digital photographs. The program can compute individual crown transparency as well as canopy transparency (multiple crowns) from vertical photographs taken below the canopy. The pictures and results can be stored for multiple year evaluations of each plot.

INTRODUCTION

The Forest Inventory and Analysis (FIA) program of the USDA Forest Service is charged with the task of conducting large-scale vegetation surveys on forestland throughout the U.S. For every 6,000 acres of land, a permanent sample plot has been established where FIA field crews periodically collect data on forest type, site attributes, tree species, tree size, and overall tree condition. These sample plots are referred to as Phase 2 (P2) plots. On a subset of the P2 plots, forest health attributes are also collected. These plots are referred to as Phase 3 (P3) plots and there is approximately one P3 plot for every 16 P2 plots. Forest health attributes measured on P3 plots include tree crown conditions, lichen communities, understory vegetation, down woody debris, and soil attributes. As part of the crown condition assessment, the following crown measurements are collected: uncompact live crown ratio, crown diameter, light exposure, foliage absent, density, foliage transparency and dieback. This paper will introduce a software tool developed by the USDA Forest Service Southern Research Station that can assist FIA with the transparency estimation portion of P3 inventories. The software, called ForestCrowns, measures individual crown transparency and canopy transparency from digital photographs taken vertically from the ground.

FIA FOLIAGE TRANSPARENCY ESTIMATION

FIA defines foliage transparency as the amount of skylight visible through a side view of the live, normally foliated portion of the crown (USDA Forest Service 2007). The “normally foliated” portion of the crown is where there is visible foliage, normal or damaged, or remnants of its recent presence. FIA crew members determine foliage transparency by first projecting a two-dimensional outline around the tree crown that extends from the live crown base to the top and outward to the branch tips. This imaginary outline will be the transparency rating region and can be thought of as shrink wrapping a side view of the live crown. Excluded from the rating region are dead branches in the lower live crown, snag branches, crown dieback (recent branch mortality), and areas where foliage is expected to be missing. Once the rating region has been established, a transparency reference card (figure 1) is used to estimate the amount of skylight that is or would be penetrating the foliated crown (expressed as a percentage of the total foliated crown area). Typically, an estimate is obtained by two crew members standing at perpendicular viewpoints from the tree and averaged.

FORESTCROWNS SOFTWARE

The ForestCrowns software tool provides an alternative method of measuring transparency using standard or fisheye digital photographs. Instead of estimating transparency from a side view, ForestCrowns uses an upward view of the crown and can estimate individual crown transparency as well as canopy transparency. Transparency, as defined by ForestCrowns, is the amount of skylight visible through all physically-present crown structures, including leaves, branches and fruit. The program can assess an entire image or select areas of an image and can also analyze multiple images together. Batch processing is also available, which

provides quick full-image analysis of multiple images. Assessment input and results can be written to a database for storage and imported back into the program later.

ForestCrowns consists of two windows: the Image window and the Data window (figure 2). The Image window contains the crown/canopy image to be analyzed, as well as, options for adding/removing images from an assessment, deleting previous assessments, and selecting the analysis region. The Data window consists of two tabs: *Properties* and *Assessment*. The *Properties* tab is where the user enters the input data associated with the analysis, including: tree ID and species (for individual crown analysis), location, photo date, analysis date, and comments. The *Assessment* tab is where the transparency estimates are displayed following analysis. Other features found in the Data window include the database menu and the batch processing option.

PHOTOGRAPHING THE CROWN/CANOPY

The foundation for the ForestCrowns analysis is the digital image of the crown/canopy. Prior to obtaining the photograph, the best photo location should be determined based on understory vegetation and lighting conditions. Dense understory vegetation can block the view of the canopy and produce inaccurate transparency results when the image is processed with ForestCrowns. Poor lighting conditions and shooting directly at the sun can also produce erroneous transparency estimates. The optimum photo location is away from dense understory vegetation and when the sun is not directly overhead.

One advantage of using photographic records is that differences in canopy and crown transparency can be detected in subsequent inventories. In order for the results to be comparable, however, the location and orientation of the photograph should be consistent from one inventory to another. Once the photo location has been established, it is documented using a combination of GPS coordinates and distances to adjacent trees. In addition, a permanent metal pin can be placed in the ground at the photo location. The camera is then mounted to a tripod, centered over the pin, and leveled to insure that the camera angle is truly vertical. The radial orientation of the camera is documented and subsequent photographs are taken at the same orientation. If a fisheye lens is used, care should be taken to insure that the photographer is below the camera and not included in the picture. During subsequent inventories, if new lower vegetation exists that impedes the photographic view of the canopy, either have another crew member hold the vegetation out of the frame of view or, as a last resort, relocate the photo location close to the original.

INDIVIDUAL CROWN ANALYSIS

The first step in analyzing an individual tree crown using the ForestCrowns software is to upload the photo and enter the

input data (figure 2). Next, areas within the crown that have sky in the background are delineated on the image using the rectangular selection tool, the elliptical selection tool, or a combination of both (figure 3). The greater the proportion of tree crown that is delineated, the more accurate the overall transparency estimate will be. After each area is drawn, the transparency for that region will be displayed in the *Assessment* tab of the Data window. In addition to the individual transparency values, the weighted average transparency value for all regions is displayed in the lower right corner of the window. For this example, the individual transparency values range from 6.87 percent to 32.07 percent and the overall transparency value is 16.06 percent.

CANOPY ANALYSIS

ForestCrowns can determine canopy transparency from a single image or multiple images. To estimate transparency from a single image, the photo is first uploaded and input data entered as was done with the individual crown analysis. Next, the option to analyze the entire image is chosen under the selection menu. The transparency value of the entire image is then displayed under the *Assessment* tab (figure 4). For this example, the overall transparency value is 17.62 percent. Additional images can be added to the assessment by clicking on the “Add Images” button. Once the additional images have been uploaded, each image is assessed as in the above example. The Assessment tab will show the transparency values for each individual assessment as well as the combined average transparency value for all photos. To quickly analyze a large quantity of photos for full canopy transparency, the user can run the batch processing function. Finally, to assess photographs taken with a fisheye lens, the elliptical selection tool is used to select the entire circular photo region prior to processing (figure 5). The transparency for the fisheye image example is 13.41 percent.

DISCUSSION

Foliage transparency, one of the key crown variables collected by FIA on P3 inventory plots, serves as an indicator of overall tree health. High transparency values, relative to what is normal for a species, indicate that a tree has less leaf area to capture sunlight for photosynthesis. Some of the factors that can cause an increase in transparency are disease, insect damage, or drought. Foliage transparency is important for classifying tree health, but it is one of the most difficult variables to measure.

Current FIA estimates of transparency can be very subjective (Ghosh and others 1995; Innes 1988). Some of the factors that can influence a field crew’s assessment are: background vegetation, foreground vegetation, tree height (distance to crown), weather and lighting conditions, observer training/experience, and observer perception. Background vegetation, which occurs in just about every

forest setting, blocks light and makes it difficult to see light coming through the crown of the tree being assessed. In addition, dense forest environments create a situation where it is difficult to determine which foliage belongs to which crown. Foliage that does not belong in the assessment may be included and foliage that should be in the assessment may be omitted. Another problem associated with dense forests is foreground vegetation, which can prevent the observer from viewing the full crown. Tall trees can also pose a problem for transparency estimation due to increased distance and viewing angle. Weather conditions can affect ratings by altering the amount of light available while making an assessment. Crown ratings done on sunny days can differ from crown ratings done for the same tree on overcast or rainy days. Finally, observer bias can significantly add to discrepancies in transparency ratings. Training and experience are just two of the variables that can add to observer bias, but the bottom line is that not everyone sees the same things in the same way.

Because of the subjectivity involved in current FIA crown rating procedures, we propose a more objective method of estimating transparency for individual tree crowns using digital photographs and the ForestCrowns software. Though the protocols for viewing and rating the crown are significantly different, ForestCrowns provides an unbiased estimate of transparency. The main procedural difference is that FIA uses a side view of the tree and rates the foliated area only, whereas ForestCrowns uses an upward view of the crown and includes all crown structures in the rating. Though some of the same variables that adversely affect FIA ratings can still come into play, the one important variable that is removed from the equation is observer bias. Another advantage of using a photograph is that it serves a permanent visual record of the tree crown at that time. Results of crown assessments from multiple inventory years can also be compared to detect changes in individual crown transparency.

One subplot-level variable not currently collected by FIA is canopy cover, or inversely, transparency. Canopy transparency can be a good predictor for understory plant survival, growth, and succession, as well as many other sub-story ecosystem functions. Adding a measure of canopy transparency to FIA inventories would not significantly

increase data collection time and would provide additional useful data to land managers. What we propose is that four digital photographs are taken at each FIA plot [one at the center of each subplot (figure 6)], and then the images analyzed using the ForestCrowns software to determine transparency. Photographs can be taken with a standard or fisheye camera lens. In addition to providing an accurate estimate of transparency, the images can also be used to detect gaps in the canopy, which can indicate blow-downs, removals, or other tree mortality.

CONCLUSION

As part of FIA P3 inventories, crews measure a variety of health related tree variables, including foliage transparency. However, due to observer bias and other limiting factors, transparency estimates using current FIA guidelines can be very subjective. This paper proposes an alternative method of estimating crown transparency using digital photographs and crown analysis software developed by the USDA Forest Service Southern Research Station. The software, ForestCrowns, analyzes standard or fisheye photographs taken vertically from the ground and provides an accurate estimate of crown transparency. In addition to individual crown analysis, the software can also perform canopy transparency analysis. Our recommendation is to collect and analyze photos at each FIA subplot center. The photos would serve as permanent records of the canopy condition at the time, and the canopy transparency estimates derived from ForestCrowns would be a valuable addition to FIA inventory data. A future study is planned that will examine comparisons between crown cover estimates obtained from aerial photographs and ground-based estimates using the ForestCrowns software.

LITERATURE CITED

- Ghosh, S.;** Innes, J.L.; Hoffmann, C. 1995. Observer variation as a source of error in assessments of crown condition through time. *Forest Science*. 41(2): 235-254.
- Innes, J.L.** 1988. Forest health surveys: problems in assessing observer objectivity. *Canadian Journal of Forest Research*. 18(5): 560-565.
- USDA Forest Service.** (2007). Phase 3 Field Guide - Crowns: Measurements and Sampling, Version 4.0. <http://www.fia.fs.fed.us/library/field-guides-methods-proc/>. [Date accessed: November 24, 2010].

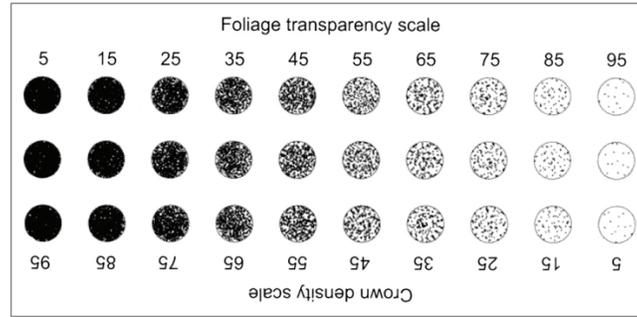


Figure 1—Reference card used by FIA field crews to determine foliage transparency of tree crowns.

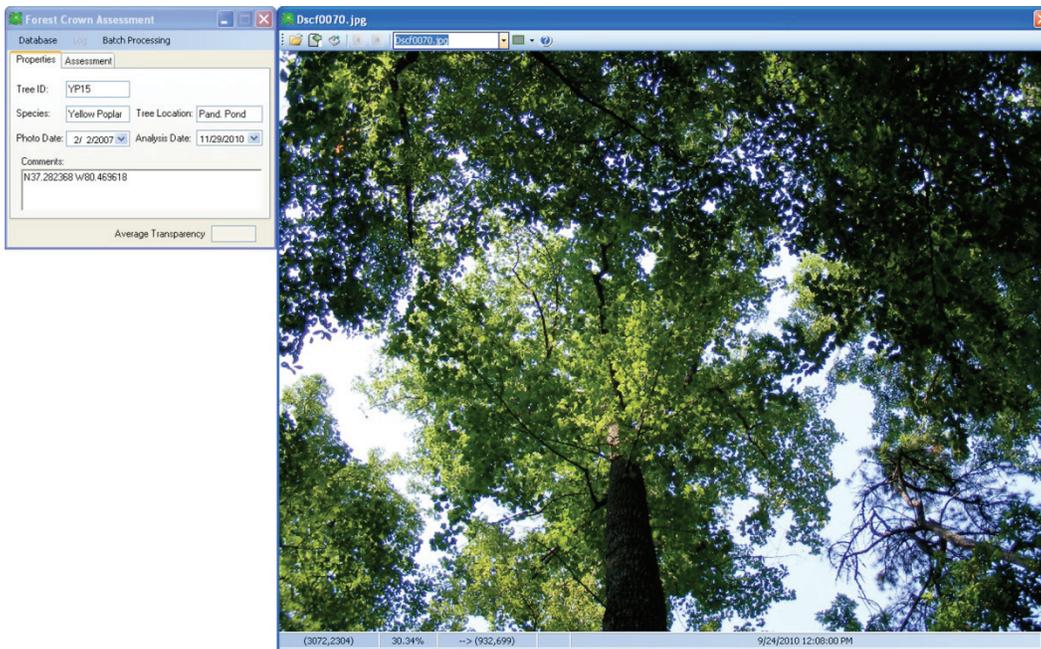


Figure 2—Screen shot of the ForestCrowns computer program showing the Data window with input parameters on the left and the Image window on the right.

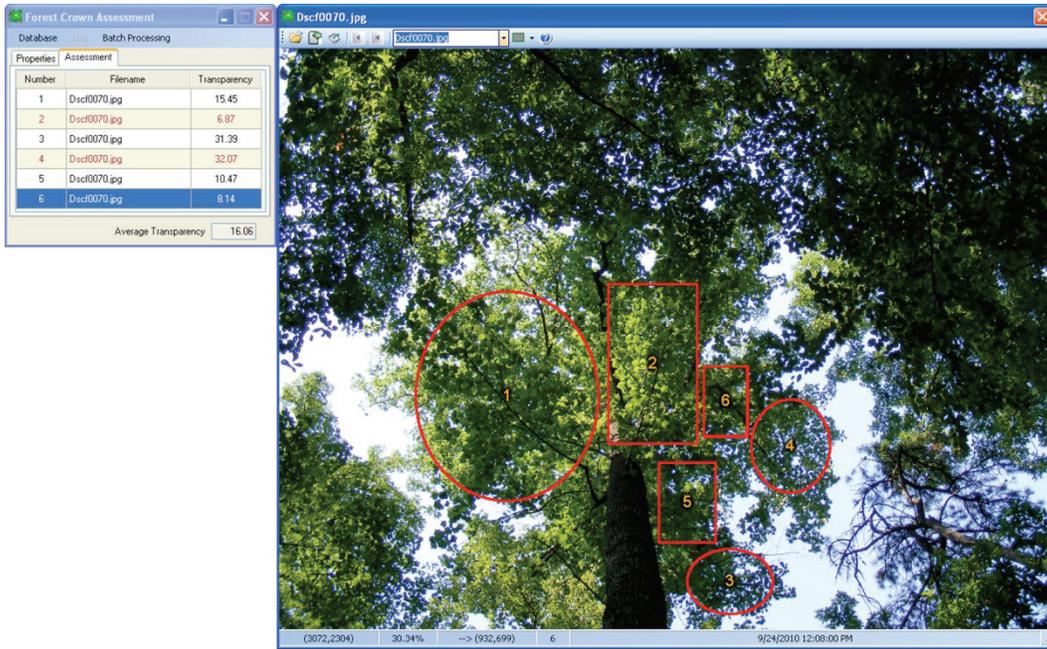


Figure 3—Sample crown delineation and transparency results for individual crown analysis in ForestCrowns.

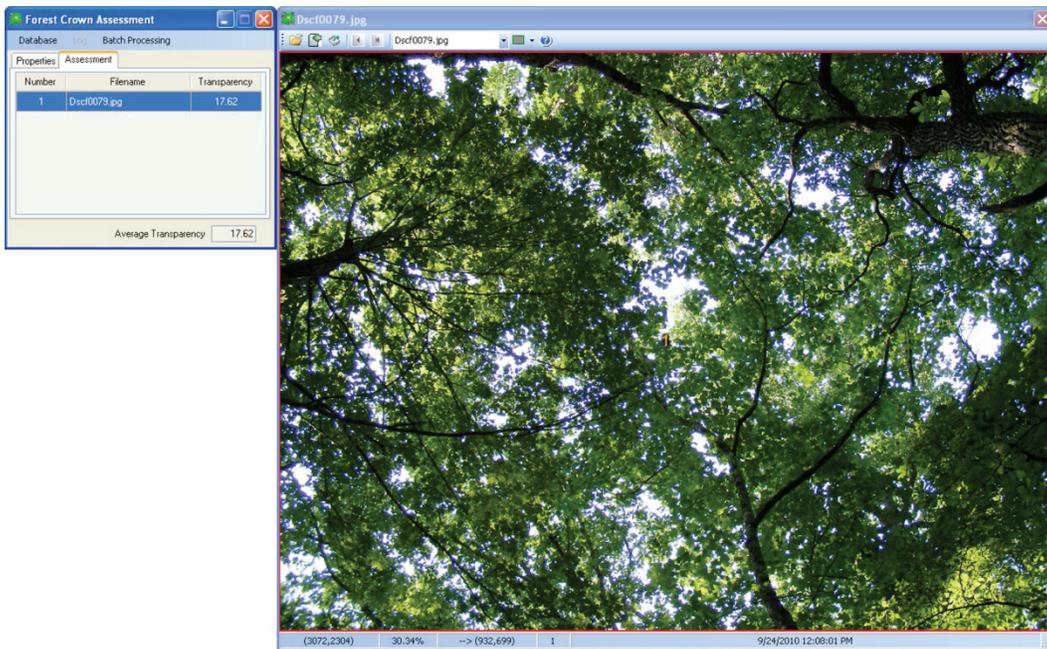


Figure 4—Delineation area and transparency results for canopy analysis using ForestCrowns.

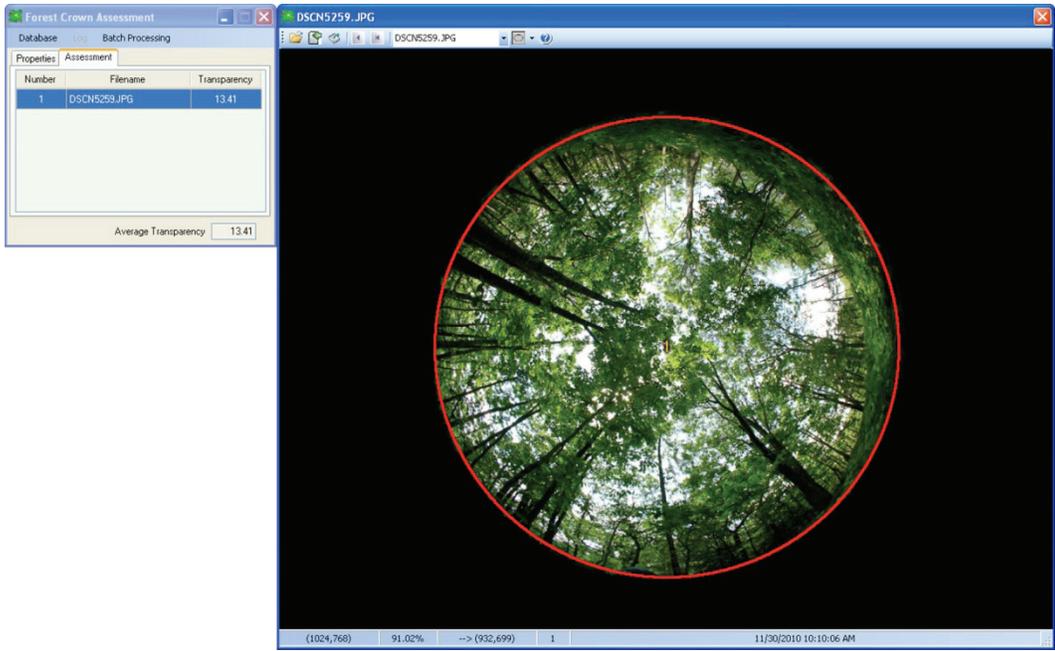


Figure 5—Delineation area and transparency results for canopy analysis of fisheye photograph using ForestCrowns.

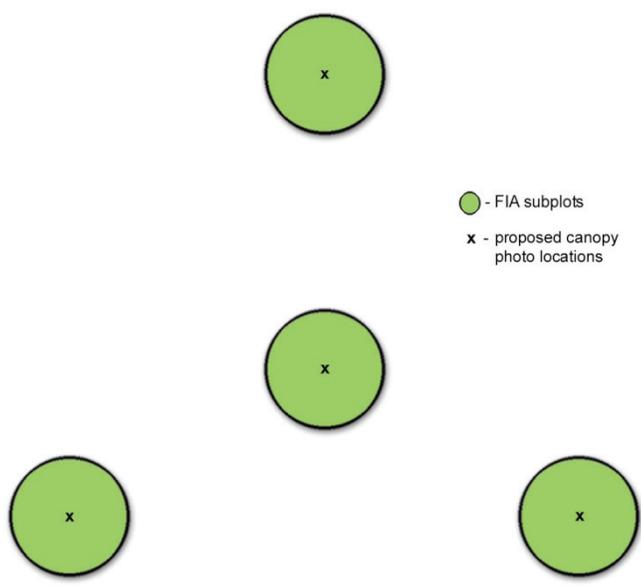


Figure 6—FIA subplot arrangement and proposed canopy photo locations.

AN ALTERNATIVE METHOD FOR ESTIMATING CROWN CHARACTERISTICS OF URBAN TREES USING DIGITAL PHOTOGRAPHS

Matthew F. Winn and Philip A. Araman

ABSTRACT

The USDA Forest Service Forest Inventory and Analysis (FIA) program has concluded that statewide urban forest inventories are feasible based on a series of pilot studies initiated in 2001. However, much of the tree crown data collected during inventories are based on visual inspection and therefore highly subjective. In order to objectively determine the crown characteristics of urban trees and assure reliability of the data, researchers with the U.S. Forest Service Southern Research Station have developed a computer software tool called UrbanCrowns that can potentially be used to assist with crown data collection on urban FIA plots. In addition to its operational use, the software can also be used as a training tool. UrbanCrowns analyzes a single, side-view digital photograph of an urban tree and computes crown measurements similar to those collected during FIA inventories. UrbanCrowns output includes estimates of crown height, live crown ratio, crown diameter, crown density, transparency, and crown volume.

INTRODUCTION

Urban forests provide a wide range of aesthetic, health, economic, and environmental benefits to urban communities (Nowak and others 2007). For this reason, many municipalities conduct periodic inventories to quantify and characterize urban tree resources. These inventories are very useful when developing management plans at the local level but not suited for broader management planning at the state, regional, or national scale. Even though there is much urban inventory data available, the methods used to obtain the data and the type of data collected are not standardized among communities. In order to meet the needs of state and federal resource managers and to expand the range of data collection, the Forest Inventory and Analysis (FIA) program of the USDA Forest Service initiated a series of urban forest inventory and health monitoring pilot studies in 2001. The purpose of the pilot studies was to determine the feasibility and logistics of conducting statewide urban forest inventories.

FIA has long been responsible for conducting national forest inventories, but urban areas have always been excluded.

This is because data collected are exclusively on lands that meet the specific definition of “forest” and urban forests do not meet this definition. The purpose of the urban inventory is to account for the trees in urban areas not measured during traditional FIA inventories. Since results from the pilot studies have shown that urban forest inventory and health monitoring data collection and analysis is feasible (Cumming and others 2008), annual FIA urban inventories are likely to continue sometime in the future if funded.

The data collected on urban sampling plots include the standard FIA data as well as additional urban health monitoring variables, which are grouped into four general data types: plot, trees, tree damages, and tree crowns. The tree crown variables include: uncompact live crown ratio, crown light exposure, crown position, crown density, crown dieback, foliage transparency, foliage absent, and crown diameter. Of all the data types, the tree crown measurements are the most difficult to obtain as well as the most subjective. For this reason, an alternative method of measuring urban tree crown characteristics has been developed and is presented here.

The USDA Forest Service Southern Research Station has developed a computer software tool to address many of the problems associated with urban tree crown measurements, both for FIA and municipal tree inventories. The software, called UrbanCrowns, analyzes digital photographs of urban trees and produces various crown metrics important for urban tree inventories (Winn and others 2010). Output produced by UrbanCrowns includes: crown height, crown diameter, live crown ratio, crown volume, crown density, and transparency. The program supports both broadleaf and coniferous tree analysis. The next sections describe the methodology currently used for FIA urban data collection, followed by a description of the UrbanCrowns software.

CURRENT FIA URBAN DATA COLLECTION

FIA field crews periodically visit urban sampling plots and collect standard FIA data as well as additional urban forest health monitoring variables on trees and tree crowns. Crown variables collected during the inventory include: uncompact live crown ratio, crown light exposure, crown position, crown density, crown dieback, foliage transparency, foliage absent, and crown diameter. FIA definitions and field measurement procedures for select crown variable are shown below (Schomaker and others 2007).

UNCOMPACTED LIVE CROWN RATIO

Uncompact live crown ratio is the length of a tree that supports live foliage relative to the actual tree length. First, the live crown base and the live crown top are established. The live crown base is the point of the lowest live foliage and the live crown top is the highest point of the crown containing live foliage. Uncompact live crown ratio is then determined by dividing the distance between the live crown base and the live crown top by the total length of the tree. A reference card is typically used by FIA crews to estimate ratios (figure 1). The card is held parallel to the tree and moved closer and farther from the crew member's eye until the zero mark is at the live top of the tree and the 99 is at the base of the tree. The point at which the live crown base intersects with the reference card scale shows the uncompact live crown ratio.

CROWN DENSITY

Crown density, expressed as a percentage, is the amount of crown stem, branches, twigs, shoots, buds, foliage, and reproductive structures that block light penetration through the crown. This includes dead branches and dead tops. First, a vertically symmetrical outline is visualized around the tree crown that extends from the live crown base to the top and outward to the branch tips. If the top is broken or missing, it is visually reconstructed before determining the density rating outline. The area within the imaginary outline is then compared to the crown density-foliage transparency card to determine the density (figure 1). Density is typically estimated by two crew members at perpendicular views of the tree. If ratings differ by more than 10 percent, they discuss the reasons for their ratings and the final rating is derived by averaging the two crew members' final ratings.

CROWN DIEBACK

Crown dieback is recent mortality of branches with fine twigs that begins at the terminal portion of a branch and proceeds toward the trunk. Only branch mortality occurring in the upper and outer portions of the crown is considered dieback. To estimate dieback, crew members first visualize an outline around the crown, extending from the live crown base to the top and outward to the branch tips. Next, the area

classified as dieback is determined and compared to the total live crown area. Dieback is expressed as the percentage of the total live crown area that is affected. Crew members then compare their estimates and reach an agreement as to what's recorded.

FOLIAGE TRANSPARENCY

Foliage transparency is the amount of skylight visible through microholes in the normally foliated portion of the live crown. Large holes, dieback, and dead branches are excluded from the estimate. Foliage transparency differs from density in that it ignores stems, branches, and fruits in the crown. Each crew member first draws an imaginary two-dimensional outline around the live tree crown, similar to the region used to estimate dieback but in this case, dieback regions are excluded. Crew members then use the crown density-foliage transparency card (figure 1) to estimate the amount of skylight penetrating the foliated crown (expressed as a percentage of the total foliated crown area). Crew member estimates are compared, adjusted if necessary, and averaged to determine the final rating.

CROWN DIAMETER

Crown diameter is the average width of the crown, extending from the drip line on one side of the tree to the drip line on the opposite side of the tree. The drip line is determined by projecting a vertical line upward from the ground to the outermost branch tip. Crew members measure the diameter at the widest part of the crown using a tape and then again at 90 degrees from the widest point.

SRS URBANCROWNS SOFTWARE

UrbanCrowns is a software tool developed by the Southern Research Station that could potentially be used by FIA to obtain crown metrics on urban sampling plots. As opposed to FIA data collection methodologies commonly used, UrbanCrowns offers an objective approach to evaluating urban tree crowns. The software quickly calculates crown height, crown diameter, live crown ratio, crown volume, crown density, and transparency from a single digital photograph and several field measurements. The steps involved in the analysis process, including photographing the tree, gathering field data, and analyzing the tree image, are described below.

ACQUIRING THE TREE PHOTO

The first step in the analysis procedure is to acquire a single ground-based digital photograph of the subject tree. The UrbanCrowns computer program is designed to be used with basic digital imagery, so the use of specialized camera equipment is not necessary. The only requirements for the program is that the entire tree is visible and centered in the

photograph, there are no obstructions between the camera and the tree, and that a portion of the tree crown is free from background vegetation or buildings (meaning only sky in the background). This area will be used to estimate crown transparency which uses color contrasts to identify the crown. This means that the program cannot easily distinguish between foreground and background vegetation as well as filter out buildings and other man-made obstructions, depending on the contrast in color. Weather conditions can also affect the photograph, as they influence the amount of available light. Optimum lighting occurs on clear sunny days when the sun is high in the sky, and the camera should never be pointed directly at the sun. Since the photo is the basis for all analyses within the program, careful consideration should be taken when choosing a photo location. Realistically, it may not be possible to find a vantage point that satisfies all of the photographic requirements. If this is the case, the photographer should find the best photo location for the conditions but understand that it may affect the UrbanCrowns output.

UrbanCrowns allows the use of only one photo by making the assumption that the tree crown is relatively symmetrical and that transparency and density are constant regardless of the vantage point. However, a single photograph of a significantly asymmetrical crown (such as a tree pruned around a power line) may not provide an accurate representation of true crown volume. If this is the case, a second photograph of the tree should be taken perpendicular to the first. Results can then be averaged to obtain the best possible estimates for that tree.

FIELD MEASUREMENTS

In addition to photographing the tree, several tree measurements must also be collected. These measurements are necessary in order to scale the photograph within the UrbanCrowns program. First, the angles (in degrees) to the top and base of the tree must be measured using a clinometer or other vertical angle measuring device. The measurements should be taken from the same location and height at which the photograph was taken. The program also requires the horizontal distance (in feet) from the photo location to the trunk of the tree. This can be determined using a laser or sonic rangefinder or a tape measure. Several instruments are currently available that measure both horizontal distance and vertical angles from a single location, and though they cost a bit more than traditional measuring tools, they can significantly reduce the data collection time at each tree. Though not required by the program, the azimuth from the photo location to the tree should also be recorded. By combining the azimuth with the horizontal distance measurement, it will be possible to return to the original photo location during subsequent inventories. The program has the capability to store the azimuth and any other tree or site information (species,

location, weather conditions, etc...) within that tree's data file.

PROGRAM OVERVIEW

The UrbanCrowns software is comprised of two main windows: the Tree Image window and the Data Control window (figure 2). The Tree Image window contains the uploaded photograph of the tree to be analyzed. All image controls are located in this window, including: opening, rotating, scaling, printing, and saving the image. The Data Control window is used for inputting data, initiating assessment, viewing output, and managing the database. Within the Data Control window, there are two tabs: *Information* and *Assessment*. The *Information* tab is where the user enters the input parameters for the tree image. The *Assessment* tab is where the results of the analysis are displayed once the image is processed.

INPUT DATA

The first step in analyzing a tree crown is to upload the desired photo into the UrbanCrowns program. Once the photo has been uploaded, the field data and other input parameters are entered under the *Information* tab of the Data Control window (figure 2). The input consists of: tree ID, tree species, photo location, photo date, azimuth to tree, horizontal distance, angle to the top of the tree, angle to the base of the tree, and user comments.

REFERENCE LINES

The next step is to draw a series of reference lines on the photo (figure 3). The first reference line (shown in yellow) extends from the base of the tree stem to the top of the tree crown, following the lean of the tree. This line, combined with the angle and horizontal distance measurements entered earlier, is used to scale the photograph (determine the actual area represented by each pixel). The second reference line (shown in pink) is a polygon drawn around the portion of the tree crown that is free from background vegetation or other obstructions. This is the area that will be used by the program to determine transparency and density. The final reference line (shown in blue) is another polygon drawn around the entire tree crown and is used to estimate crown volume. Note that neither of the polygons needs to be drawn close to the crown in areas that have a clear background. When the image is processed, the program shrink-wraps the selection regions so that they conform to the unobstructed outline of the tree crown.

CROWN ASSESSMENT

Once the field measurements have been entered and the reference lines have been drawn, the image is processed by clicking on the *Assess* button in the Data Control window. The results of the analysis are then displayed under the *Assessment* tab (figure 4). For this example, the results are as follows: tree height and length are 42.6 feet, crown

height is 40.1 feet, crown diameter is 47.4 feet, crown ratio is 94 percent, crown density is 94.8 percent, transparency is 5.2 percent, and crown volume is 43,041 feet³. Though the true crown volume can't be determined without destructive sampling, all other output for this example are within 5 percent of actual measurements obtained on site.

After an image is processed, a pop-up window will appear that contains a contour image of the tree (figure 5). The contour image shows the shrink-wrapped crown assessment areas that were used in the analysis. A summary page can then be printed that shows the input parameters, the output data, and the tree image including reference lines. The same information can also be stored in a database and later imported back into the program. Other available program options include saving the image and reference lines as a JPG file, saving the contour image as a JPG file, and saving the input and output data to a text file.

DISCUSSION

Many of the crown variables collected by FIA field crews during urban tree inventories are provided as output in the UrbanCrowns computer program. Overlapping variables include: crown diameter, uncompact live crown ratio, transparency, and crown density. UrbanCrowns does not provide an estimate of light exposure, foliage absent, or dieback but it does provide an additional measurement of crown volume. Each of the overlapping variables, as well as the volume estimate produced by UrbanCrowns, is discussed in more detail below.

CROWN DIAMETER

Of the four variables, crown diameter is the only one measured directly by field crews. Since both methods of determining crown diameter rely on objective measurements, no improvement in accuracy is expected by using UrbanCrowns. In fact, if only one photograph is taken, the FIA method of averaging two perpendicular diameter measurements is probably a better representation of the overall crown diameter. One advantage of using the software, however, is that the diameter can be measured at a later date in the office, freeing up valuable data collection time in the field.

UNCOMPACTED LIVE CROWN RATIO

Whereas crown diameter is measured directly at the tree by field crews, uncompact live crown ratio is measured indirectly from a distance using the crown density-foliage transparency reference card. However, the fact that it's a ratio and not a measured distance means that diminished accuracy is not likely. Both FIA field estimates and UrbanCrowns output should be comparable. Similar to crown diameter, the advantage of using UrbanCrowns

to determine crown ratio is that the measurement can be performed at a later time, requiring less time for measurements in the field.

TRANSPARENCY

The methods used by FIA crews and by UrbanCrowns to determine transparency are somewhat similar but also have some differences. FIA estimation starts with creating an outline of the live crown that extends from the base to the top of the live crown and out to the branch tips. Dieback, dead branches, and large holes are excluded from the rating area. Crew members then estimate the amount of light penetrating the normally foliated portion of the rating region, ignoring stems, branches, and fruits.

Instead of rating the entire live crown, UrbanCrowns determines transparency from partial crown analysis (using only the portion of the crown that has no obstructions in the background) and assumes that the transparency of the partial crown is representative of the full crown transparency. Since UrbanCrowns cannot distinguish between foreground vegetation and background vegetation, the program is only able to analyze full crowns if the entire crown is free from background interference. However, realistically, it is difficult to find a photo vantage point that would allow for full crown analysis. Once the rating region has been established, UrbanCrowns measures the amount of light penetrating the crown, but unlike the FIA methods, does not exclude any crown structures such as branches or large holes. For this reason, UrbanCrowns does not use the FIA terminology "foliage transparency" but instead, "crown transparency."

In summary, FIA measures the amount of skylight visible through the full foliated crown whereas UrbanCrowns measures the amount of skylight visible through all crown structures for a portion of the crown. Though these methods seem very different, there is one major advantage of using UrbanCrowns to compute transparency; it is objective. FIA transparency estimation, on the other hand, is highly subjective. Some of the factors that can contribute to discrepancies in visual estimation include: observer bias, observer training, weather and lighting conditions, background vegetation, and obstructions. With the increased likelihood of obtaining inaccurate transparency estimates using FIA methods, UrbanCrowns could be a viable alternative.

CROWN DENSITY

FIA defines crown density as the amount of crown structures blocking light penetration through the crown. The area used to rate crown density differs from the transparency rating region in that it assumes a symmetrical crown. Areas that would normally have crown structures if it were a perfectly health tree are included in the outline. UrbanCrowns, however, uses the same partial crown rating region that

the program uses to determine crown transparency. In fact, UrbanCrowns defines crown density as the complement of crown transparency (calculated by subtracting crown transparency percent from 100 percent). In other terms, crown density is a measure of the amount of crown structures blocking light within the partial crown rating region. Again, the methods used by FIA and UrbanCrowns to determine crown density are quite different. However, like foliage transparency, FIA density estimates are highly subjective while UrbanCrowns provides an objective estimate of crown density.

CROWN VOLUME

Accurate estimates of crown volume can be beneficial for predicting ecosystem functions such as carbon sequestration, rainfall interception, pollution removal, and surface temperature reductions. Crown volume of urban trees is typically calculated using allometric equations with crown length and crown diameter as the dependent variables (Schomaker and others 2007). One problem with using this method to determine crown volume is that it does not take the shape of the crown into account. Even by incorporating a shape variable into the equation, the assumption is still made that the tree is vertically symmetrical when viewed from the side. For an accurate estimate of crown volume, the true shape of the crown needs to be considered.

The UrbanCrowns software is unique in that it can provide an accurate estimate of crown volume using only one photograph of the tree, and it does so without assuming vertical crown symmetry. Instead, it assumes that if you take a cross section anywhere in the crown, that cross section is circular. In a recent study comparing UrbanCrowns volume output to crown volume estimates obtained through traditional methods, a high correlation was found between the two (Patterson and others, in press). Though destructive sampling is necessary to determine the true accuracy of the program, the results suggest that the volume estimates produced by UrbanCrowns are reasonably accurate. To determine crown volume, the program first calculates the actual width and height represented by each row of pixels within the crown selection region. An imaginary cylinder is generated for each row of pixels that has a height equal to the calculated height of one pixel and a diameter equal to the calculated width of the row (figure 6). The volume estimates for each row of pixels are summed to obtain the volume estimate for the entire crown selection region. This estimate includes tree structure and void areas, so the volume is then multiplied by the crown density to get a volume estimate that includes tree structures only. This method of determining volume works well for crowns that are relatively symmetrical around the stem. For crowns that are significantly asymmetrical, such as trees pruned around power lines, it may be necessary to analyze a second

photograph taken perpendicular to the first and calculate the average volume.

TRAINING

In addition to UrbanCrowns' potential use in FIA urban inventories, the software could also serve as a valuable training tool for FIA field crews and trainers. Regional trainers currently train, test, and certify field crews for their ability to measure trees within the tolerances specified by the FIA program. For objective crown measurements, such as crown diameter and live crown ratio, the training is straightforward and success is probably not trainer dependent. However, for more subjective measurements such as foliage transparency and crown density, using multiple regional trainers to teach field crews can introduce individual trainer bias into the rating procedures. This can result in crown ratings that are not consistent at a broader level. Though there is no substitute for field training, the software could supplement current training programs and provide a standardized platform for learning crown measurement procedures, particularly with the more subjective crown measurements. In addition, the program can also be used to pre-analyze test trees used to certify trainers and field crews. This would provide objective estimates of crown features that student ratings can then be compared to.

CONCLUSION

As urban areas expand, statewide urban forest inventories (such as those conducted by the Forest Inventory and Analysis program) will become more and more important. However, current methodologies used by FIA to conduct crown assessments of urban trees are highly subjective. The UrbanCrowns software developed by the USDA Forest Service Southern Research Station could be a useful tool for FIA urban inventories, as it provides an objective and efficient means of determining various crown metrics. In addition to the program's potential for operational use, the software can also be a valuable training tool for field crews. Crown measurements generated by UrbanCrowns include: crown diameter, uncompact live crown ratio, transparency, crown density, and crown volume.

LITERATURE CITED

- Cumming, A.B.;** Twardus, D.B.; Nowak, D.J. 2008. Urban forest health monitoring: large-scale assessments in the United States. *Arboriculture & Urban Forestry*. 34(6): 341-346.
- Nowak, D.J.;** Wang, J.; Endreny, T. 2007. Environmental and economic benefits of preserving forests within urban areas: air and water quality. In: de Brun, C.T.F., eds. *The economic benefits of land conservation*. San Francisco, CA: The Trust for Public Land: 28-47.

Patterson, M.F.; Wiseman, P.E.; Winn, M.F. [and others]. [In press].
 Effects of photographic distance on tree crown attributes calculated using UrbanCrowns image analysis software. *Arboriculture & Urban Forestry*.

Winn, M.F.; Lee, S.M.; Araman, P.A. 2010. UrbanCrowns: crown analysis software to assist in quantifying urban tree benefits. In: Laband, D.N., ed. *Proceedings, Emerging Issues Along Urban/Rural Interfaces III: Linking Science and Society*. Auburn, AL: Auburn University, Center for Forest Sustainability: 86-91.

Schomaker, M.E.; Zarnoch, S.J.; Bechtold, W.A. [and others]. 2007.
 Crown-condition classification: a guide to data collection and analysis. Gen. Tech. Rep. SRS-102. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 78 p.

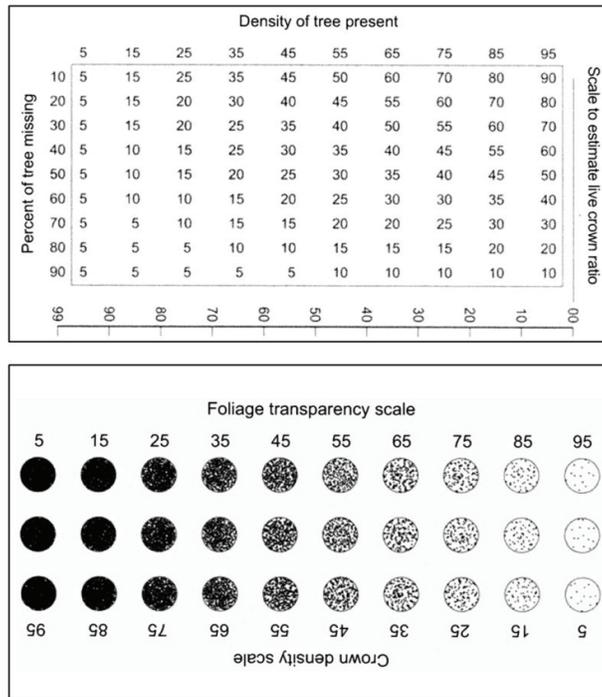


Figure 1 — Front and back of reference card used by FIA field crews to estimate crown density, foliage transparency, and uncompacted live crown ratio.

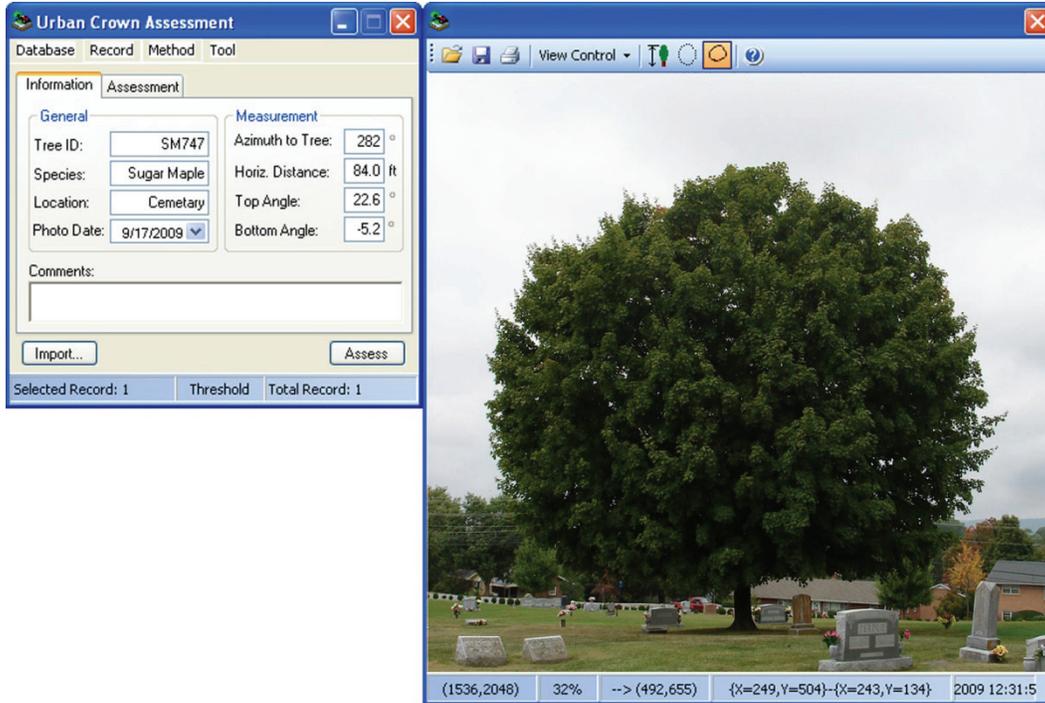


Figure 2—Screen capture of the UrbanCrowns program showing the Data Control window with input parameters on the left and the Tree Image window with uploaded image on the right.

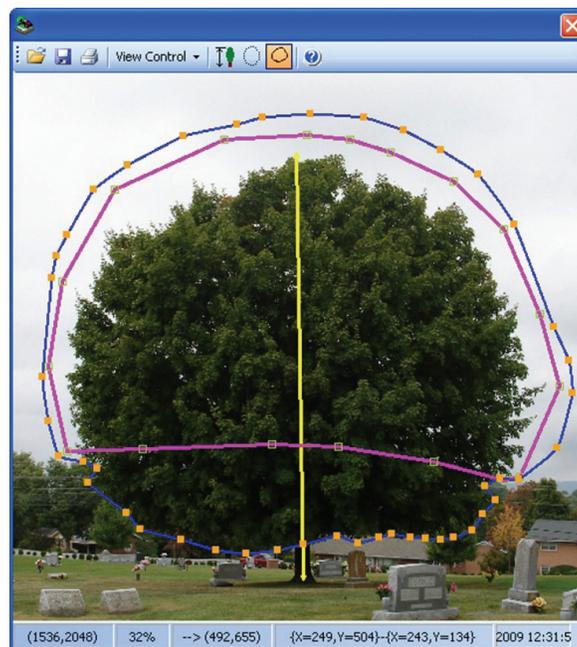


Figure 3—Screen capture of the UrbanCrowns program showing reference lines drawn on the photo to calculate tree height (yellow), transparency (pink), and crown volume (blue).

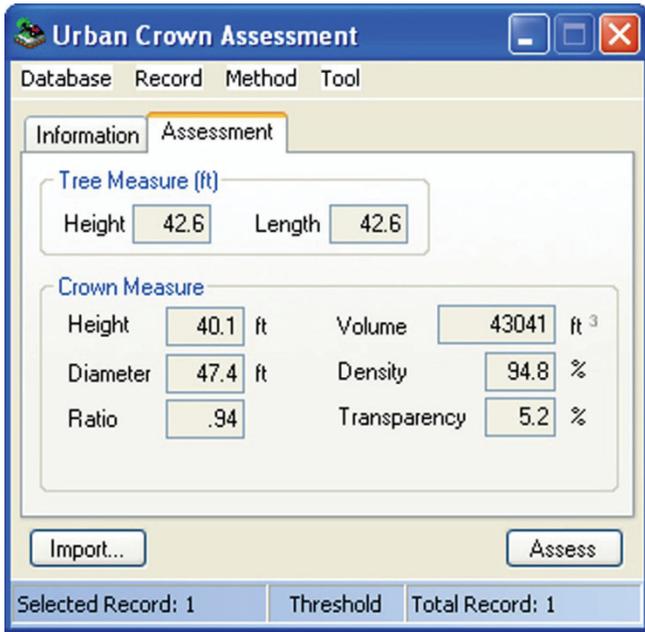


Figure 4—Screen capture of the UrbanCrowns *Assessment* tab showing the post-processing results of the crown analysis.

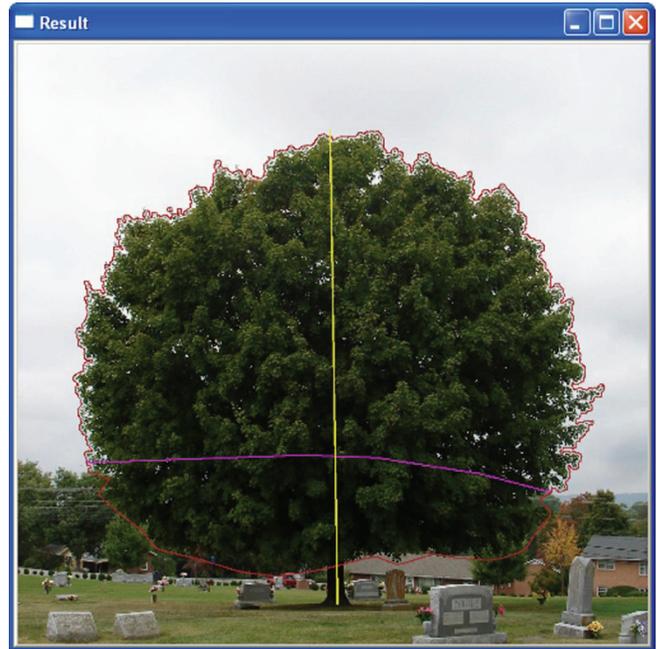


Figure 5—Contour image generated by UrbanCrowns that shows the transparency and full crown regions used in the analysis.

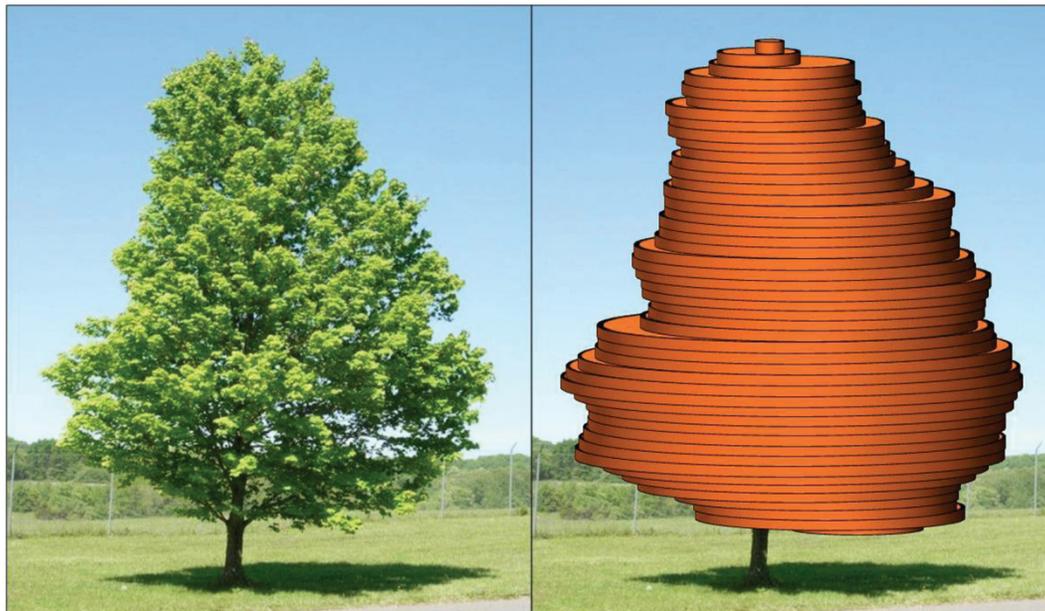


Figure 6—Illustration showing the method used by UrbanCrowns to determine crown volume.

COMPARISON OF LIDAR- AND PHOTOINTERPRETATION-BASED ESTIMATES OF CANOPY COVER

Demetrios Gatzliolis

ABSTRACT

An evaluation of the agreement between photointerpretation- and LiDAR-based estimates of canopy cover was performed using 397 90 x 90 m reference areas in Oregon. It was determined that at low canopy cover levels LiDAR estimates tend to exceed those from photointerpretation and that this tendency reverses at high canopy cover levels. Characteristics of the airborne imagery used, and, to a lesser extent, the density of the sampling point pattern employed and the occasional photointerpretation error inflated estimate discrepancies. Where available, LiDAR data can potentially be used to quantify the magnitude of error embedded in estimates of canopy cover obtained via photointerpretation.

INTRODUCTION

Forest canopy cover is an important ecological indicator that is known to affect, among many other phenomena, near-ground solar radiation (Zou and others, 2007), tree regeneration (Stancioiu and O'Hara, 2006), and wildlife habitat (Ganey and others, 2008). It also plays a key role in estimating forest stand attributes from remotely sensed data (Jennings, 1999). The importance of canopy cover for national forest inventory operations has increased since the Food and Agriculture Organization (FAO, 2000) established the 10 percent canopy cover threshold as the universal criterion defining forest land. Prompted by this development, the Forest Inventory and Analysis (FIA) Program of the U.S. Forest Service has recently decided to adopt canopy cover as forest land determinant and it is now participating in an effort designed to model canopy cover across the conterminous U.S. Model predictions are based on Thematic Mapper imagery and ancillary data and will be organized in raster layers. A 5-year updating schedule is envisioned. Canopy cover estimates serving as training data for model development are obtained by manual photointerpretation (PI) of high-resolution airborne imagery.

The term 'canopy cover' adopted by FIA follows the definition suggested by Avery and Burkart (1994) according to whom it is the percent forest area occupied by the vertical projection of tree crowns. In this definition, tree crowns are considered opaque or solid objects and it is implied that canopy cover estimates obtained in the field should

only involve observations performed in the exact vertical direction. Dot count (Rautiainen and others, 2005), line intercept (Gregoire and Valentine, 2007) and moosehorn (Fiala and others, 2006) sampling techniques meet this requirement; hemispherical photography (Korhonen and Heikkinen, 2009), a popular alternative, does not, but, reportedly, the effects of the oblique angle view can be minimized by photograph post-processing. All these approaches for field estimation of canopy cover are logistically infeasible for a project with national scope. Estimates based on remotely sensed data are perhaps the only plausible alternative.

Spectral imagery acquired by airborne or satellite platforms conducive to unbiased estimation of canopy cover should have sufficiently fine spatial resolution that allows the identification of individual tree crowns or crown clusters and the delineation of between-crown openings (gaps), and narrow field of view centered at nadir (Korpela, 2004). Where the latter requirement is not met, trees depicted in high-resolution imagery exhibit substantial 'layover' or radial displacement of their crown tops relative to their bases that is intensified as the distance from the image's nadir point increases. This displacement leads to partial obstruction of portions of a tree's crown or of nearby canopy gaps, either by the tree in question or by its neighbors. Consequently, the minimum size of canopy gaps that can be reliably identified in such imagery increases with the distance from the nadir point, ultimately leading to bias in the estimation of canopy cover. Solar illumination and terrain conditions can inflate the bias.

High-density Light Detection and Ranging (LiDAR) data are far less susceptible to bias in part because they are independent of solar illumination and terrain conditions but primarily due to the fact that laser pulses are capable of penetrating tree crowns. LiDAR instruments emit short pulses of light propagated as a narrow beam towards illuminated objects and record the amount of energy that is backscattered to the sensor and the length of time that has elapsed. By processing this information the laser instrument identifies points, also known as returns or

echoes, precisely georeferenced in space, that correspond to the locus of the backscattering. Pulses illuminating hard objects (bare ground, building roofs, etc.) generate a single return. Objects that are not solid, for example tall vegetation, typically generate more returns along the pulse's propagation trajectory. LiDAR data over forested landscapes comprise large sets of returns known as the return 'clouds' that represent sampling of terrain and vegetation materials. Therefore, unlike spectral imagery, LiDAR pulses can sample the portion of a tree crown, even its lower components, positioned away from the flight line of the airborne platform.

This study compares canopy cover estimates obtained via photointerpretation to those assessed from corresponding high-density LiDAR data across a variety of topographic, physiographic, and forest management conditions in Oregon.

METHODS

The 31500 km², 75 km wide study area extends from the coastal mountains of Oregon, across the Willamette Valley and the Cascades, eastward to the nearly the Idaho border (Figure 1a), and it is sometimes known as the Oregon transect. It is one of the five pilot study areas selected for the national canopy cover project undertaken by FIA. Forests on the coastal mountains and the western half of the Cascades typically present with high canopy cover which is progressively reduced in the eastern part of the Cascades until the open forests of eastern Oregon are reached. Within the study area, 397 reference areas, each covering 90 x 90 m and centered on FIA plot locations were identified as contained in high-density LiDAR acquisitions in the 2008 – 2010 period. These reference areas will be henceforth mentioned as 'plots.' In each plot, a regularly-spaced 105 point grid was superimposed on 1-m airborne National Agriculture Imagery Program (NAIP) data acquired in 2009 (Figure 1b). Using the NAIP imagery as reference, experienced photointerpreters labeled each of the 105 points in each plot either as belonging either on a tree crown or background objects. Estimates of plot canopy cover were obtained as the ratio of tree points to the total.

To obtain the LiDAR-based estimates of plot canopy cover, the elevation value of each return was first converted to above-ground height by using a digital elevation model (DEM) also generated from the LiDAR data. All returns with height equal to or larger than a threshold were labeled as trees and the remaining ones as background returns. Three height thresholds (1, 2, and 3 m) were considered. Subsequently, raster representations of tree and background return frequencies were computed. Raster cells containing at least one return labeled tree were assigned a value of 1 while

cells with only background returns were assigned a value of 0. Cells with no returns were assigned a 'nodata' value and were excluded from further consideration. To ensure that the frequency of nodata cells, and therefore their effect on the canopy cover estimates, is minimized, the resolution of the raster frequency representation was set to the mean laser (footprint) spacing between spatially adjacent pulses. The plot estimates of canopy cover were calculated as the ratio of the value 1 cells to the sum of value 1 and 0 cells. This method for computing canopy cover estimates from laser data was evaluated using precise delineations of tree crowns detailed in Gatzolis and others (2010) and was found to not deviate by more than 3 percent from the field estimates, at least where the density of the LiDAR data exceeded 8 returns per square meter.

To account for registration discrepancies between the LiDAR and NAIP data, all returns on and in the vicinity of a plot were jittered 200 times in two dimensions by using random azimuths and distances drawn from a -5 to 5 m uniform distribution. The magnitude of the jittering was determined by measuring the mean adjustment required to achieve spatial registration by ocular means. The mean LiDAR-based plot canopy cover was finally calculated from the 200 plot-jittering instances.

RESULTS AND DISCUSSION

The scatterplot of PI- vs. LiDAR-based canopy cover indicates that at low cover levels, PI tends to produce lower estimates (Figure 2) than LiDAR. At high canopy cover levels this tendency reverses. A second-order polynomial regression of PI on LiDAR estimates exhibits coefficient of determination $R^2 = 0.787$ with the regression fitted line crossing the 1:1 one at canopy cover of approximately 35 percent. This is in part because at very low canopy cover levels, trees in the landscape can be considered rare events that are not sampled adequately by the point pattern used. At high canopy cover, it is the openings or gaps within the crowns that are rare and undersampled. Given the 1 m resolution of the NAIP imagery, the horizontal footprint of either a small tree or canopy opening would have to exceed 4 m², twice the square of the resolution, before it can be identified clearly. To both comply with the minimum identifiable object size requirement and avoid bias due to undersampling of rare events, the density of the point pattern would have to increase by at least an order of magnitude above the present level, an option which is logistically infeasible.

Figure 2 features 3 plots with LiDAR estimates higher than 60 percent and corresponding PI estimates lower than 30 percent and another 3 plots with LiDAR estimates lower than 35 percent and PI estimates higher than 75 percent.

These plots and many others have been examined carefully using various ancillary data in an effort to identify the source of such large discrepancies between the estimates. It was determined that for all six plots photointerpretation error was responsible for the discrepancies. Among the most challenging plots ranked those with uniform hardwood tree crowns mistaken for grass or brush and those covered with snow at the time of the NAIP acquisition. Two other plots with large estimate discrepancy had sustained insect infestation, a condition not anticipated by the LiDAR-based canopy cover estimation procedure which lead to overestimation. Smaller discrepancies were attributed to poor imagery quality, such as hazy conditions and lack of sufficient contrast.

In addition to the undersampling of openings, the PI overestimation of canopy cover compared to LiDAR was attributed to oblique NAIP imagery. In plots or stands with canopy cover higher than 50 percent, or even lower but with trees growing in clusters, the effects of imagery obliqueness are more pronounced. While only a small percentage of pulses had viewing angle greater than 10 degrees, for about 1/3rd of the study area the effective view angle of the NAIP imagery exceeded that angular threshold. In the presence of tall vegetation, steep terrain and fairly low sun elevation angle, conditions that are actually the norm rather than the exception in much of the Pacific Northwest, crown openings are partially or completely obstructed from view. Unless the airborne imagery is acquired with long focal length lens, its information content may not be compatible with unbiased estimation of canopy cover regardless of the diligence and skill of the photointerpreter or the sampling intensity.

Overlays of the sampling point pattern with the NAIP imagery questioned the choice of regularity in the former for several plots examined. The arrangement of points in the pattern yields a 9-m distance between a point and its immediate neighbors. This point spacing is a multiple of the planting distance for many commercial forests in the Pacific Northwest. Although certainly not an issue in 'natural' forest stands, systematic sampling can have unintended implications where sample points happen to consistently lay on crowns or canopy openings. Alternatively, random sampling point pattern could perhaps be employed in regions with substantial component of commercial forests.

Modifying the object height threshold that separates trees from background objects was found to have a small overall effect on the agreement between PI and LiDAR estimates of canopy cover. For height threshold equal to 1 m, 2 m, and 3 m the root mean square discrepancy between the two types of estimates was 14.98, 15.20, and 15.92 percent respectively. For 21 plots, 5.3 percent of the total, increasing

the threshold from 1 to 3 m resulted in a more than 4 percent change in estimate discrepancy. It should be noted that within the study area there was hardly any portions with non-tree vegetation of mean height larger than each of the thresholds specified and for the majority of the open forests in eastern Oregon there is little or no understory. Such conditions facilitate precise estimation of canopy cover, at least for the LiDAR-based approach. In different biomes and dominant cover types, the distinction between tree and non-tree vegetation might be less clear. While LiDAR data do describe the vertical structure of vegetation, we are yet to see in literature methodologies and applications capable of accurately and consistently discerning bushes and brush from tree overstory. In such conditions, how well the height threshold selected represents the vegetation profile will likely determine the accuracy of the estimates obtained.

Assuming that the LiDAR-based estimates of canopy cover are either unbiased or, if not, only marginally biased, the results of this study suggest that the PI-based estimates contain substantial bias at least for plots with low or high true canopy cover. Considering that the primary objective for the PI effort is to support the national canopy cover project, it should be concerning that the bias in the PI estimates will propagate through the modeling function and likely bias the outputs. Given the model structure types considered for the national project, it is unlikely that one can assess *a priori* the effect of the bias in the input to any bias in the output. Perhaps the only viable option is to repeat the modeling effort once with PI estimates as input and once with their LiDAR equivalent and compare the outputs, at least in regions where high-density LiDAR data is available. Such a comparison could lead to useful insights towards methodological improvement in the PI process and in the structure of models employed for future implementations of the national canopy cover project.

CONCLUSION

An evaluation of the agreement between PI- and LiDAR-based estimates of canopy cover was performed using a large number of plots across a variety of vegetation and topographic conditions. The evaluation indicates that the agreement between estimates relates to the value of canopy cover. There is sufficient evidence to suggest that the PI approach tends to underestimate low and to overestimate high canopy cover. In addition to bias, PI estimates appear to be imprecise as well, in part because of the characteristics of the airborne imagery used. The magnitude of the bias can be quantified where high-density LiDAR data is available. Additional investigations are needed to determine if bias removal or reduction can be achieved.

LITERATURE CITED

- Avery, T.E.;** Burkart, H.E. 1994. *Forest Measurements*, 4th edition, McGraw-Hill, New York, USA, 331 pp.
- Fiala A.C.S.;** Garman, S.L.; Gray, A.N. 2006. Comparison of five canopy cover estimation techniques in the western Oregon Cascades. *Forest Ecology and Management*, 232:188–197.
- Food and Agriculture Organization.** 2000. On definitions of forest and forest change. Forest Resources Assessment Program. Working paper 33. Rome, Italy. 15p. URL: <ftp://ftp.fao.org/docrep/fao/006/ad665e/ad665e00.pdf> (last accessed December 15, 2010).
- Gatziolis, D.;** Fried, J.S.; Monleon, V. 2010. Challenges to estimating tree-height via LiDAR in closed-canopy forests: a parable from western Oregon. *Forest Science*, 56(2):139-155.
- Gregoire, T.;** Valentine, H. 2007. Sampling strategies for natural resources and the environment. *Applied Environmental Statistics*, Volume 3. Chapman & Hall/CRC, London, UK. 474p.
- Ganey, J.L.;** Cassidy, R.H.; Block, W.M. 2008. Estimating canopy cover in forest stands used by Mexican spotted owls: Do stand-exam routines provide estimates comparable to field-based techniques? Res. Paper RMRS-RP-72WWW, US Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado. 8p.
- Jennings, S.B.;** Brown, N.D.; Sheil, D. 1999. Assessing forest attributes and under-storey illumination: Canopy closure, canopy cover, and other measures. *Forestry*, 72(1):59-74.
- Korhonen, L.;** Heikkinen, J. 2009. Automated analysis of in situ canopy images for the estimation of forest canopy cover. *Forest Science*, 55(4):323-334.
- Korpela, I.** 2004. Individual tree measurements by means of digital aerial photogrammetry. *Silva Fennica Monographs*, 3. 93 p.
- Rautiainen, M.;** Stenberg, P.; Nilson, T. 2005. Estimating canopy cover in Scots pine stands. *Silva Fennica*, 39(1):137-142.
- Stancioiu, P.T.;** O'Hara, K.L. 2006. Regeneration growth in different light environments of mixed species, multiaged, mountainous forests of Romania. *European Journal of Forest Research*, 125: 151–162, DOI 10.1007/s10342-005-0069-3.
- Zou, C.B.;** Barron-Gafford, G.A.; Breshears, D.D. 2007. Effects of topography and woody plant canopy cover on near-ground solar radiation: Relevant energy inputs for ecohydrology and hydrogeology. *Geophysical Research Letters*, L24S21, doi:10.1029/2007/2007GL031484.

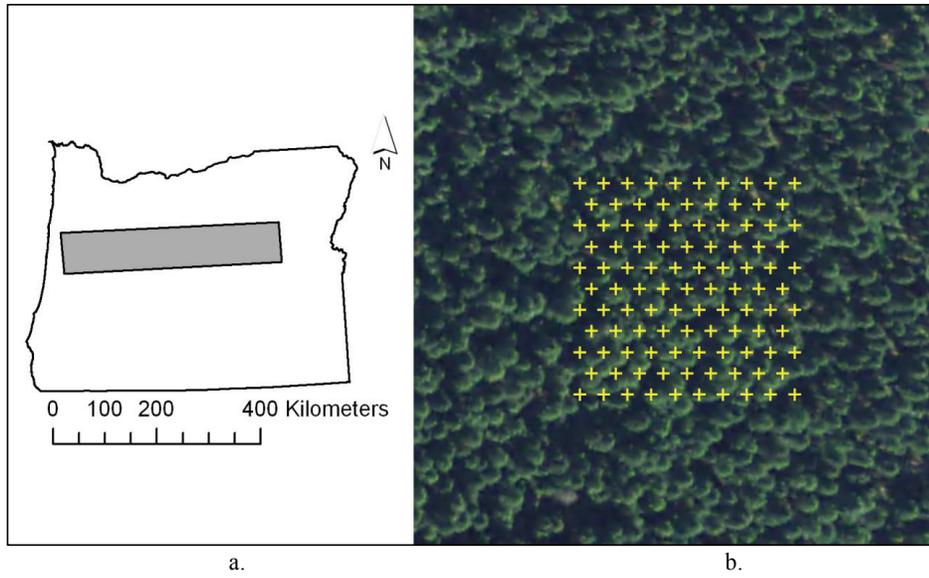


Figure 1— a. Study area (shaded rectangle) and State of Oregon boundary, b. 90 x 90 m sampling point pattern on NAIP panchromatic imagery for a randomly selected location.

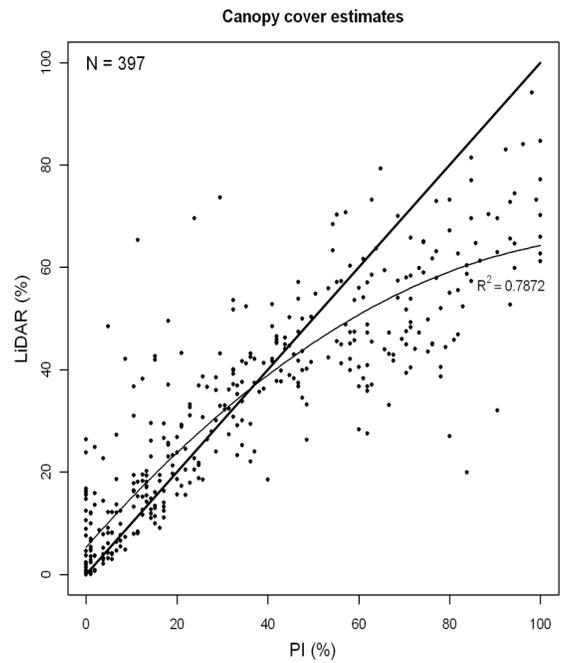


Figure 2— Scatterplot of LiDAR-vs. photointerpretation-based canopy cover estimates with 1:1 (thick) line and second-order regression fit (thin line).

COMPARING ALTERNATIVE TREE CANOPY COVER ESTIMATES DERIVED FROM DIGITAL AERIAL PHOTOGRAPHY AND FIELD-BASED ASSESSMENTS

Tracey S. Frescino and Gretchen G. Moisen

ABSTRACT

A spatially-explicit representation of live tree canopy cover, such as the National Land Cover Dataset (NLCD) percent tree canopy cover layer, is a valuable tool for many applications, such as defining forest land, delineating wildlife habitat, estimating carbon, and modeling fire risk and behavior. These layers are generated by predictive models wherein their accuracy is dependent on the quality of the data used to train the models. This analysis compares several different methods for estimating live tree canopy cover, including ocular, image segmentation, and dot count assessments from digital aerial photography, as well as field-based measurements.

INTRODUCTION

Tree canopy cover is defined as the proportion of the ground covered by a vertical projection of all the live tree canopies. A spatially-explicit representation of live tree canopy cover, such as the 2001 National Land Cover Dataset (NLCD) percent tree canopy cover layer, is valuable for many natural resource applications including: wildlife habitat models (Allen 1982; Kroll and Haufler 2006; Zarnetske and others 2007; Koy and others 2005), atmospheric carbon estimates (Nowak and Crane 2002), and fire applications such as FARSITE (Finney 1998). For strategic level forest inventories, such as the Forest Inventory and Analysis (FIA) program, percent canopy cover is also a very important measurement used for defining forest land.

The NLCD originated in 1992 from the Multi-Resolution Land Characteristics (MRLC) consortium. This multi-agency program was formed specifically to acquire Landsat data across the conterminous U.S. and to generate a 30-m pixel land cover map. In 2001, a second-generation land cover map was produced from more current Landsat imagery purchased by the MRLC. In addition, 30-m pixel maps of imperviousness and percent tree canopy cover were developed. A third generation map of land cover and second generation maps of imperviousness and percent tree canopy are currently in progress and slated for release in 2011.

The NLCD 2001 map of tree canopy cover was produced by modeling a response of tree canopy cover as a function of an extensive database of predictor layers, including 30-m resolution Landsat satellite imagery, digital elevation models, and other ancillary data that were meaningful to the model. Here, the tree canopy cover response (training) data were acquired using an automated classification of 1-m digital orthophoto quadrangles (DOQs), along with extensive post-processing hand editing. Models were developed by mapping zones, dividing the landscape into relatively homogenous regions with respect to landform, soil, vegetation, spectral reflectance, and characteristics of the imagery (Homer and others 2004).

A number of accuracy assessments of the NLCD 2001 tree canopy layer have been conducted, but their findings are inconsistent. For example, Homer and others (2004) reported mean absolute error averaging 10.8 percent based on cross-validation of per-pixel estimates from three different mapping zones. An evaluation of zonal estimates from the tree canopy cover layer compared to photo-interpreted estimates from Google Earth imagery indicated an underestimation of tree canopy cover by an average of 9.7 percent consistently across the conterminous United States (Nowak and Greenfield 2010). The Landscape Fire and Resource Management Planning Tools (LANDFIRE) project found the canopy cover values to be too high for use in existing fire models (Scott 2008). Questions of the accuracy of the NLCD tree canopy layer have led to a reassessment of the model's response data and potential alternative methods for the third generation. As the leading agency for national-level tree data, the United States Department of Agriculture, Forest Service, Forest Inventory and Analysis (FIA) program was identified as a logical candidate for leadership in the third generation product.

This paper compares several techniques for measuring live tree canopy cover for use as training data in predictive mapping efforts, such as the NLCD. Using data collected

throughout the state of Nevada, the underlying objectives are to: (1) assess variability in ocular estimates of tree canopy cover from multiple observers; (2) compare photo-based methods and field-based methods for measuring canopy cover, an extension from Goeking and Liknes' 2009 analysis; and (3) compare photo-based measurements from 1-m resolution NAIP imagery to higher-resolution imagery acquired in Nevada.

METHODS

Analyses were conducted throughout the state of Nevada using a subset of field-sampled FIA plots from the 2004 and 2005 inventory years. A total of 150 plots, or approximately 45 percent, were randomly selected from the population of 328 FIA plots collected during these years that sampled at least one forested condition. Field sampling determined that 128 of the 150 plots had a woodland forest type, 8 plots had a timber forest type, and 12 plots were nonstocked (see http://socrates.lv-hrc.nevada.edu/fia/ab/issues/pending/glossary/Glossary_5_30_06.pdf for definitions). Two plots were removed from the analysis because of geographical inconsistencies. On each of the remaining 148 plots, three photo-based methods and two field-based methods were applied to construct alternative measures of live percent tree canopy cover. This process followed that of Goeking and Liknes (2009), where similar methods of remotely estimating crown cover were compared to field transect data across five states in the interior west: Arizona, Colorado, Idaho, Montana, and Utah, in an attempt to expand the utility of the pre-field operations for the national FIA program.

The photo-based methods involved interpreting one-acre circular plots, coinciding with the 148 FIA plot locations, using three methods applied to large scale aerial photographs: an ocular estimate; a dot count method; and an image segmentation method using Feature Analyst software. Ocular estimates were collected by three photo interpreters for each type of photography. Three interpreters were used to examine variability in the subjective measurements among interpreters based on findings from Goeking and Liknes (2009), where ocular estimates varied widely among three photo interpreters. A crown cover callibration key was used to assist in determining the crown coverage within the acre circle. Although the individual observer values were analyzed for variability, the average value of the three interpreters was used as the single ocular estimate for each plot to compare to other methods. The dot count method involved photo-interpreting 50 randomly distributed points within the acre circle, with a restriction of having a minimum distance of 2 meters between points. Percent cover was calculated by counting the number of points that fell on live tree crowns and dividing by the total number

of points. The image segmentation procedure entailed digitizing a few live tree crowns, or polygons, within the acre circle and using these polygons as training data in the Feature Analyst extension to ArcMap. Feature Analyst® is an automated feature extraction software that uses inductive learning algorithms and techniques to model object recognition (Opitz and Blundell 2008). Training information is assigned by the user and the software automatically generates a model, correlating the known data to target objects, and applies the model to the entire area of interest.

These three photo-based methods were applied to two different types of large scale aerial photography to help understand the effect of resolution on estimating canopy cover. The types of photography included: 1.0-meter (39-inch) resolution, National Agriculture Imagery Program (NAIP) natural color, orthorectified photography for year 2006 that is freely available from the USDA Forest Service Image Server extension of ArcMap (http://fswb.rsac.fs.fed.us/imageserver/image_server_home.html); and 0.15-meter (6-inch) resolution, natural color, georeferenced, direct-to-digital or scanned-digital photography that was acquired by contract for a photo-based inventory pilot study throughout the state of Nevada in years 2004 and 2005 (NPIP; Frescino and others 2009).

The field-based methods included a line-transect method and a modeled method based on field measurements obtained from FIA's extensive database of field measurements. For the field-transect estimates, live tree crown cover was measured using sixteen 25-ft transects, totaling 400 feet, with intercepts of all live trees 1.0 inch and greater recorded at one-foot intervals (O'Brien 1989). These data were aggregated to the plot level. For the field-model estimates, predictive models of tree canopy cover were previously generated from field measurements of tree species and diameter using over 12,000 FIA plots across the Interior West (Toney and others 2009). These models were applied to the plots in this study using the FIA field plot measurements as parameters in the models.

The three photo-based methods (ocular ("oc"), dot count ("dot"), and Feature Analyst ("fa")) applied to two scales of aerial photography ("NAIP" and "NPIP") plus the two field based methods (field transects ("Field") and models developed by Chris Toney and others 2009 ("CTmodel")) gave a total of 8 different estimates of tree canopy cover over the 148 plots, summarized in Table 1. Analyses comparing these different estimates included exploratory displays of data distributions using boxplots, histograms, and scatterplots. Simple linear regression analyses were conducted using combinations of the eight canopy estimates to describe the relationship between the different methods. The regression line slopes that were visually closer to the 1:1 line indicated the canopy measures were more closely

related. The differences in the slope and intercept values were examined for understanding potential bias across the range of percent canopy values. The Pearson correlation coefficient was also shown to measure the association between variables.

RESULTS

VARIABILITY IN OCULAR ESTIMATES

Figures 1 and 2 show the results of the ocular estimates from each observer using both the NAIP photography and the NPIP photography. Similar to findings from Goeking and Liknes (2009), the ocular estimates are highly variable among interpreters with an overall mean difference of 22 percent for estimates using NAIP photography and 16 percent for estimates using NPIP photography (Figure 1). Figure 2 illustrates a number of results. First, scatterplots and regression lines for each pair of ocular estimates are shown in the lower diagonal half of the figure, where the method associated with each axis is the label in the diagonal box corresponding to the row and column of interest. For example, the scatterplot seen in row 4, column 2 of Figure 2 is a graph of observations obtained from observer #2 using NAIP photography in the x-axis, versus observer #1 using NPIP photography in the y-axis. Similarly, the upper diagonal half of the graphic displays the correlations between the estimation pair labeled in the diagonal box corresponding to the row and column of interest. Scatterplots of observers using NAIP photography (seen in the (row, column) pairs of (2,1), (3,1), and (3,2)) reveal regression slopes are closer to 1 than those from observers using NPIP photography ((5,4), (6,4), and (6,5)). In general, estimates using the NAIP photography are higher and more variable than the estimates using NPIP photography (Figures 1 and 2) and estimates from observer #2 are higher than the other observers using both photography sources. We used the mean estimate by plot from the three observers for comparison with the other methods for the rest of the analyses.

ESTIMATION METHOD COMPARISON

Figures 3 and 4 show results comparing all methods using both the NAIP and NPIP photography, including the mean estimates from the ocular method. The modeled estimates (CTmodel) are highly correlated (0.84) with the field transect estimates (Field) but, in general, the modeled estimates are slightly lower than the field transect estimates, with an overall mean difference of 4 percent (Figure 3). These differences are more emphasized in the higher canopy ranges (Figure 4).

All other estimation methods using both NAIP and NPIP photography tend to be lower than the field transect method, except the dot count method using NPIP, having an overall

mean difference of 3 percent; and the mean ocular estimates using NAIP photography, having an overall mean difference of 9 percent. Similar results are seen when comparing the modeled versus field estimates (Figures 3 and 4). The estimation method most highly correlated to the field transect method is the mean ocular estimate using NPIP photography followed by the dot count method using NPIP photography. Again, similar results are found by comparing the modeled and field estimates (Figure 4).

PHOTOGRAPHY COMPARISON

When looking at the differences in estimation methods using only NAIP photography, the mean ocular estimates tend to be higher overall than all the other methods, with an overall mean of 39 percent (Figure 3). Regardless, the regression slope is the closest to 1 compared to the other methods when related to the field transect method. The Feature Analyst estimates are generally higher than the dot count method, although the dot count method has no estimates greater than 50. The correlation is highest between the mean ocular method and the dot count method (Figure 4).

The estimates using NPIP photography show much different results than the estimates using NAIP photography. In general, the correlations are higher, the regression slopes are closer to 1, and the regression intercepts are smaller. The highest correlation (0.84) is between the ocular method and the dot count method, with the ocular method having slightly lower estimates. The correlation is also high between the ocular method and the Feature Analyst method (0.81) with a regression slope closer to 1.0, followed by the dot count and the Feature Analyst method having correlation of 0.80, although the Feature Analyst estimates are slightly lower overall (Figure 4).

When comparing the different types of photography by estimation method, the mean ocular estimate shows the highest correlation of 0.86, followed by the Feature Analyst method at 0.72, and the dot count method at 0.68. Conversely, the dot count method has the regression slope closest to the 1:1 line. For both the ocular method and the Feature Analyst method, the NAIP estimates are quite a bit higher than the estimates using NPIP photography, especially at the higher end, where as the dot count method shows the NAIP estimates slightly lower than the NPIP estimates (Figure 4). The average overall mean for the NAIP estimates is 29 percent compared to the average overall mean of the NPIP at 28 percent (Figure 3).

DISCUSSION

OCULAR COMPARISON

The first objective of this paper was to assess consistency between photo-interpreters' estimates of tree canopy cover using the ocular method. Ocular estimates, although the

fastest of the methods analyzed, may be quite biased and/or inconsistent depending on the experience level of the observers. In this study, the ocular estimates were shown to be highly variable between observers, with the estimates using NAIP photography almost always higher when compared to estimates from NPIP photography, especially at higher canopy cover values. Possible reasons for this are discussed under the third objective below. The correlations were slightly higher and the regression slopes closer to 1.0 for the NAIP estimates compared to the NPIP estimates, suggesting a tighter relationship between NAIP observers, but the higher intercepts suggest more bias between these same observers (Figure 2).

Similar results of bias can be seen when comparing the mean ocular estimates with estimates from the other methods. The ocular estimates using NAIP photography tended to be consistently higher than the other methods. The NPIP estimates, on the other hand, had lower estimates and higher correlations between estimates from the other methods, with regression slopes, in general, closer to 1.0 (Figure 4).

METHOD COMPARISON

The second objective of the paper was to compare photo-based methods and field-based methods for measuring canopy cover. For other comparison studies of canopy cover, the field-based method is often used as a control or a source of truth (Goeking and Liknes 2009; Paletto and Tosi 2009; O'Brien 1989). We found all estimation methods using both NAIP and NPIP photography to be lower than the field-transect method as well as the field-modeled method, except the dot count method using NPIP and the ocular method using NAIP, which tended to have slightly higher estimates (Figure 4). Perhaps this reflects the fact that small trees (down to 1.0 inches in diameter) are included in the field-based methods, but small trees are difficult to detect in the photos, especially in NAIP. Alternatively, there is a potential bias in the field-based methods.

For the NAIP estimates, the ocular method tended to be higher than the other methods and the dot count estimates were on the low side with no estimates greater than 50. The NPIP estimates showed higher correlations compared to all methods, indicating greater consistency between methods (Figure 4). The resolution and quality of the photography play a big factor in these results and are discussed in the following section.

PHOTOGRAPHY COMPARISON

The third objective of the paper was to compare photo-based measurements from 1-m resolution NAIP imagery to a higher resolution imagery acquired for NPIP. Using different resolution photography added interesting value to the analysis with the higher resolution photography from

NPIP representing a potential greater source of truth than the lower resolution NAIP photography. Here, we found the NAIP to estimate higher cover relative to the estimates from the NPIP photography.

In general, the lower resolution, NAIP photography has more shadows, especially when the total canopy cover is high or the terrain is steep. It is harder to distinguish tree versus shrub lifeform characteristics of a vegetative object, as well as to discriminate seedlings and saplings from shrub lifeforms. It is also harder to see regeneration of tree species in areas supporting larger trees or areas that have no recognizable trees present. These characteristics lead to overestimations and errors when estimating canopy cover.

Figure 5 shows an example of a plot where the lower resolution photography led to a large discrepancy in the cover estimate. For this example plot, one interpreter said there was zero percent cover using the dot count method and three percent cover using the Feature Analyst method, where another observer estimated 96 percent using the dot count method and 10 percent using the Feature Analyst method. The plot consists of dense aspen (*Populus tremuloides*) cover, and this species may be confused with shrubs at lower resolutions. In general, the resolution and quality of the photo are more influential to the photo interpreter than the method itself. Here, the Feature Analyst method lends itself better to an automated process that is less subjective than the interpreter's eye. Figure 6 presents another example of differences in photo resolution which affect percent canopy calls. For this plot, one observer said there was 0 percent cover using the dot count method and another observer said 12 percent for the same plot. The plot has a large percentage of dead trees that were not noticeable on the lower resolution image.

CONCLUSION

These analyses illustrate that consistent and accurate measurement of tree canopy cover is challenging. Photo-interpreter experience level, method of canopy estimation, and scale of photography all play interrelated roles in determining the quality and consistency of tree canopy estimates over training or sample plots. Certainly, there are many advantages to using NAIP photography for a national project like the NLCD 2011 tree canopy cover map: it is free, it is relatively high resolution, it has extensive continuous coverage, it is updated frequently, and photo quality continues to improve through time. It is especially effective when used through the ArcGIS Image Server extension (<http://www.esri.com/software/arcgis/serverimage/index.html>), where it is easy to move from plot to plot. However, more research is needed to understand the limitations in the use of this photography, to develop

methods to improve training methods to ensure consistency between interpreters, and to explore the potential of automated classification algorithms to improve objectivity in interpretations.

ACKNOWLEDGMENTS

A special thanks goes out to the photo interpreters for this study: Jeremy Hamblin, Val Nelson, Buzz Hettick, Cyndi Crooker, Jim Barney, Vachel Carter, and Laura Anzalene who spent their winter looking at photos. We are also grateful to Andy Lister, Sara Goeking, and Liz Freeman for their willingness to provide speedy reviews during the holiday crunch.

LITERATURE CITED

- Allen, A. W.** 1982. Habitat suitability index models: fox squirrel. U.S. Department of Interior, Fish Wildl. Serv. FWS/OBS-82/10.18. 11 pp.
- Finney M.A.** 1998. Farsite: Fire Area Simulator-model development and evaluation. USDA Forest Service, Rocky Mountain Research Station, Ogden, UT. Research Paper RMRS RP 4.
- Frescino, T. S.;** Moisen, G. G.; Megown, K. A.; Nelson, V. J.; Fressman, E. A.; Patterson, P. L.; Finco, M.; Brewer, K.; Menlove, J. 2009. Nevada Photo-Based Inventory Pilot (NPIP) Photo Sampling Procedures. Gen. Tech. Rep. RMRS-GTR-222. Fort Collins, CO: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station. 30 p.
- Goeking, S.A.;** Liknes, G.C. 2009. The role of pre-field operations at four forest inventory units: we can see the trees, not just the forest. In: McWilliams, W.; Moisen, G.; Czaplowski, R., comps. Forest Inventory and Analysis (FIA) Symposium 2008. October 21-23, 2008. Park City, UT. Proceedings RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station. 12p.
- Homer, C.;** Huang, C.; Yang, L.; Wylie, B.; Coan, M. 2004. Development of a 2001 national land-cover database for the United States. Photogrammetric Engineering and Remote Sensing 70: 829-840.
- Koy, K.;** McShea, W.J.; Leimgruber, P.; Haack, B.N.; Aung, M. 2005. Percentage canopy cover – using Landsat imagery to delineate habitat for Myanmar’s endangered Eld’s deer (*Cervus eldi*). Animal Conservation 8:289-296.
- Kroll, A.J.;** Haufler, J.B. 2006. Development and evaluation of habitat models at multiple spatial scales: a case study with the dusky flycatcher. Forest Ecology and Management 229:161-169.
- Nowak, D.J.;** Crane, D.E. 2002. Carbon storage and sequestration by urban trees in the USA. Environmental Pollution 116:381-389.
- Nowak, D.J.;** Greenfield, E.J. 2010. Evaluating the national land cover database tree canopy cover and impervious cover estimates across the conterminous United States: a comparison with photo-interpreted estimates. Environmental Management 46: 378-390.
- O’Brien, R.A.** 1989. Comparison of overstory canopy cover estimates on forest survey plots. Ogden, UT : U.S. Department of Agriculture Forest Service, Intermountain Research Station.
- Opitz, D.;** Blundell, S. 2008. Object recognition and image segmentation: the Feature Analyst® approach. In: Object-based image analysis spatial concepts for knowledge-driven remote sensing applications. Springer Berlin Heidelberg.
- Paletto, A.;** Tosi, V. 2009. Forest canopy cover and canopy closure: comparison of assessment techniques. European Journal of Forest Research 128: 265-272.
- Scott, J.** 2008. Review and assessment of LANDFIRE canopy fuel mapping procedures. LF Bulletin.
- Toney, C.;** Shaw, J.D.; Nelson, M.D. 2009. A stem-map model for predicting tree canopy cover of Forest Inventory and Analysis (FIA) plots. In: McWilliams, Will; Moisen, Gretchen; Czaplowski, Ray, comps. Forest Inventory and Analysis (FIA) Symposium 2008; October 21-23, 2008; Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station. 19p.
- Zarnetske, P.L.;** Edwards Jr., T.C.; Moisen, G.G. 2007. Habitat classification modeling with incomplete data: pushing the habitat envelope. Ecological Applications. 17:1714–1726.

Table 1—Variable short names by estimation method type

Method Type	Short Name	Description
Field-based	Field	IW-FIA field transects
	CTmodel	Toney and others (2009) field-based stem-map models
Ocular	NAIP-oc	Mean ocular estimate using NAIP photography
	NPIP-oc	Mean ocular estimate using NPIP photography
Dot count	NAIP-dot	Dot count estimate using NAIP photography
	NPIP-dot	Dot count estimate using NPIP photography
Feature Analyst	NAIP-fa	Feature Analyst estimate using NAIP photography
	NPIP-fa	Feature Analyst estimate using NPIP photography

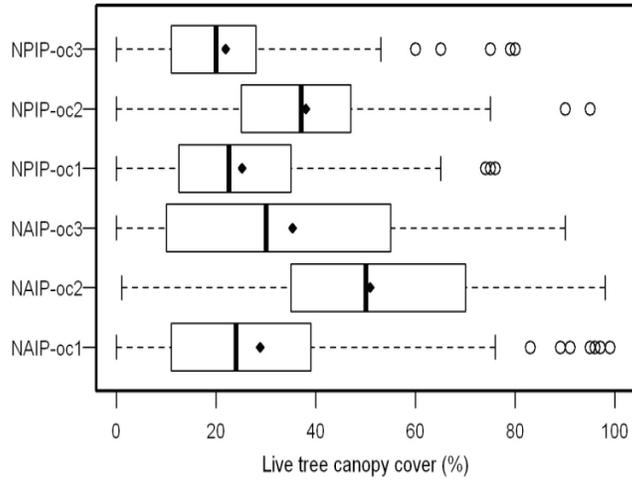


Figure 1—Boxplots of ocular estimates of percent live tree canopy cover using both NAIP and NPIP photography. The points represent the overall mean value. See Table 1 for short name descriptions. Numbers in the name correspond to the three different observers.

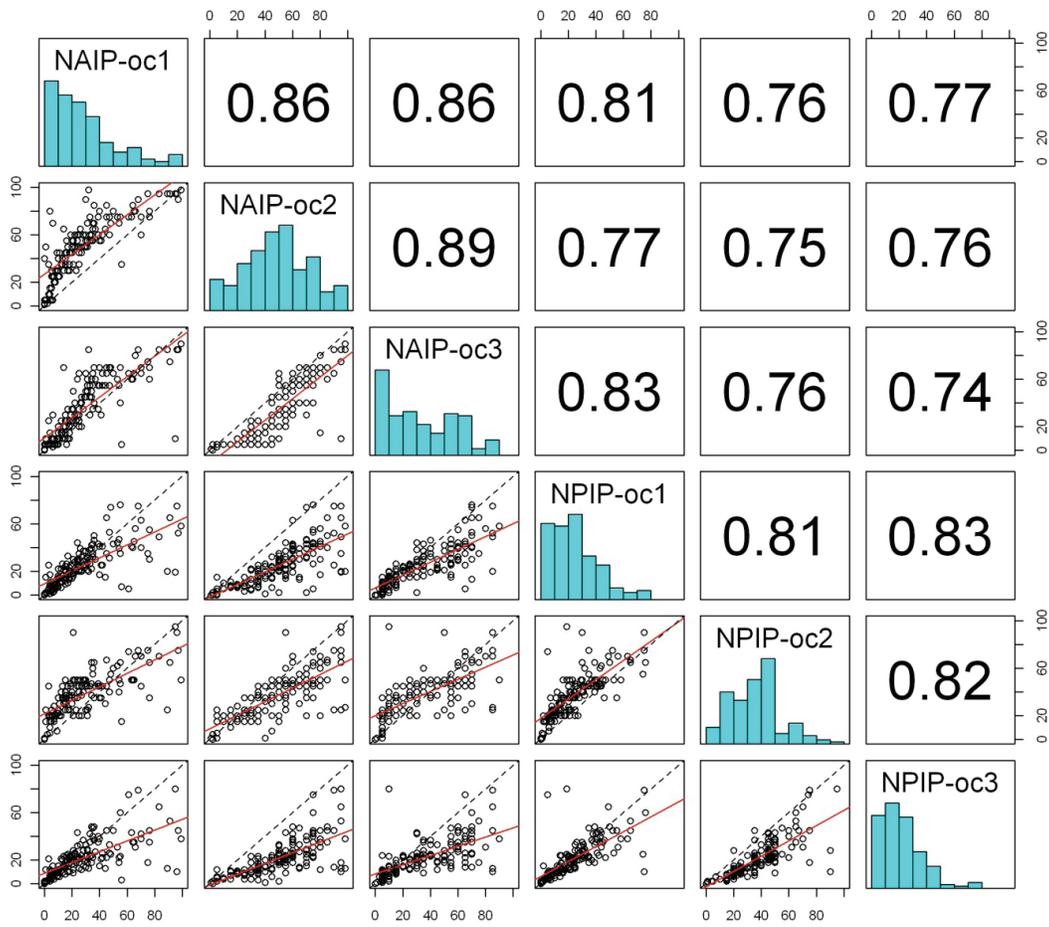


Figure 2—Pairwise comparisons of ocular estimates of percent live tree canopy cover using both NAIP and NPIP photography. See Table 1 for short name descriptions. The diagonal boxes display histogram distributions of each of the eight estimates. The left side shows scatterplot distributions for pairs of estimates with percent canopy cover on each axis. For each scatterplot, the dotted black line represents the 1:1 line and the red line is a linear regression line. The right side displays the Pearson correlation coefficient for each pair.

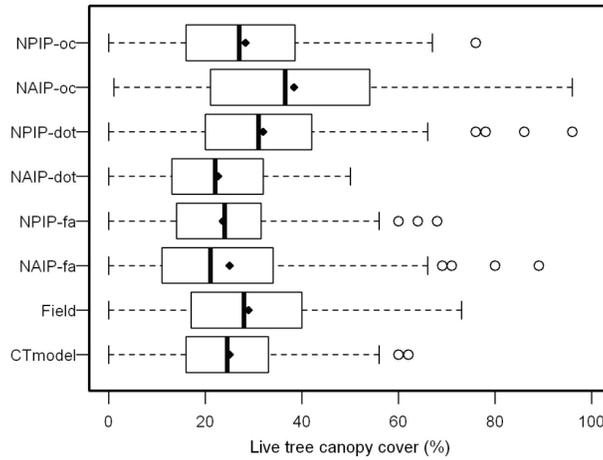


Figure 3—Boxplots of all estimation methods of percent live tree canopy cover using both NAIP and NPIP photography. The points represent the overall mean value. See Table 1 for short name descriptions.

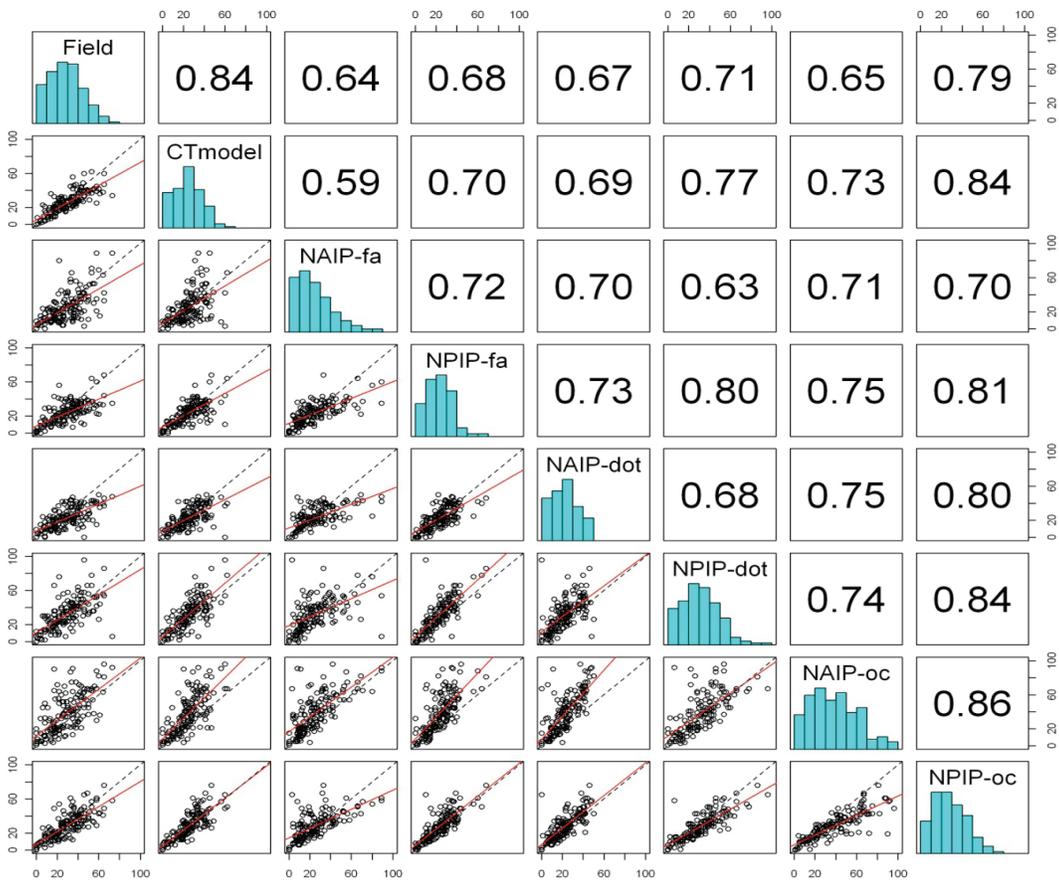


Figure 4—Pairwise comparisons of estimates from all methods including the mean ocular estimates of percent live tree canopy cover using both NAIP and NPIP photography. See Table 1 for shortname descriptions. The diagonal boxes display histogram distributions of each estimate. The left side shows scatterplot distributions between all of the eight estimates with percent canopy cover on each axis. For each scatterplot, the dotted black line represents the 1:1 line and the red line is a linear regression line. The right side displays the Pearson correlation coefficient for each pair.

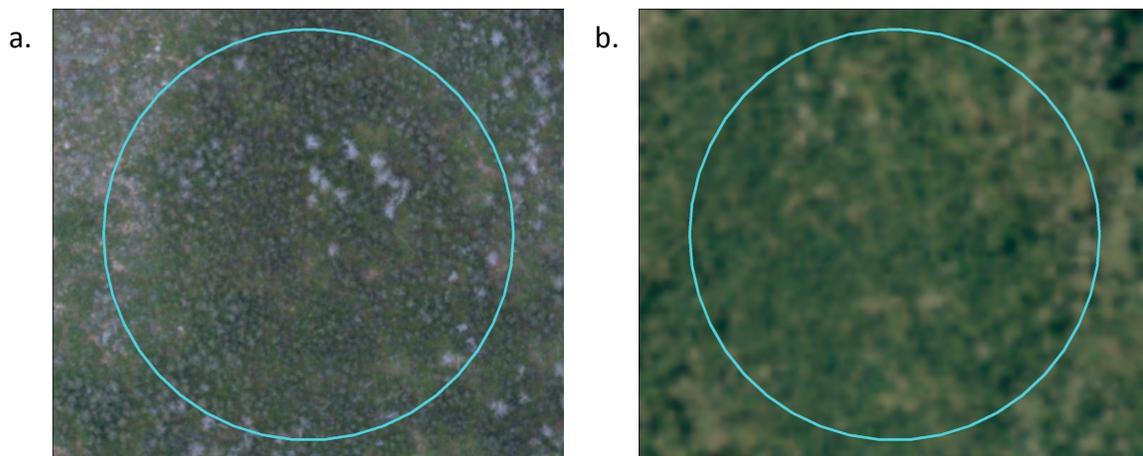


Figure 5—Plot 07-1352. a. NPIP photography with 1-ac plot boundary overlay. b. NAIP photography with 1-ac plot boundary overlay. The plot is covered with aspen (*Populus tremuloides*) seedlings.

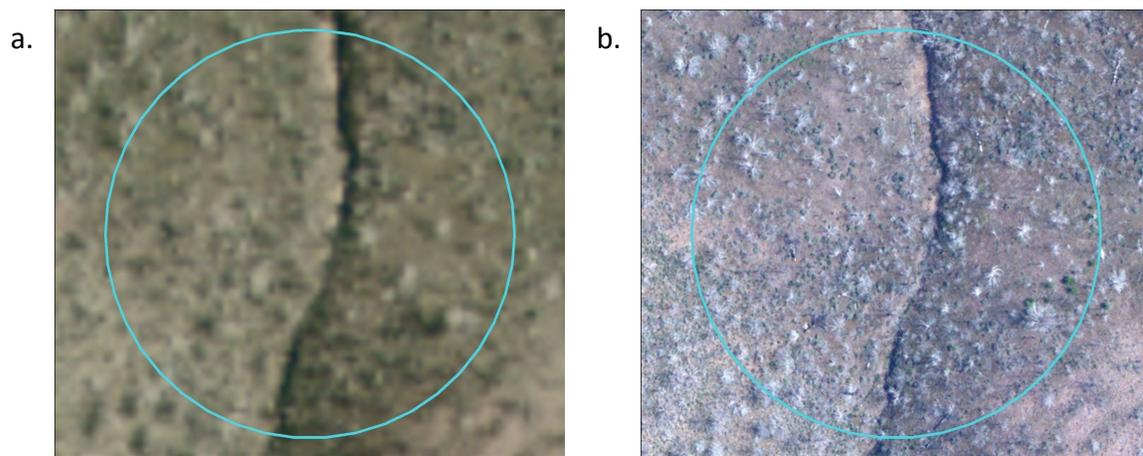


Figure 6—Plot 17-153. a. NPIP photography with 1-ac plot boundary overlay. b. NAIP photography with 1-ac plot boundary overlay. The plot is covered with dead juniper (*Juniperus* spp.).

Carbon and Biomass

ASSESSING THE ACCURACY OF CROWN BIOMASS EQUATIONS FOR THE MAJOR COMMERCIAL SPECIES OF THE INTERIOR NORTHWEST: STUDY PLAN AND PRELIMINARY RESULTS David L.R. Affleck and Brian R. Turnquist.....	247
FOREST BIOMASS SUPPLY FOR BIOENERGY IN THE SOUTHEAST: EVALUATING ASSESSMENT SCALE Christopher S. Galik and Robert C. Abt.....	255
LONG-TERM SIMULATIONS OF FOREST MANAGEMENT IMPACTS ON CARBON STORAGE FROM LOBLOLLY PINE PLANTATIONS IN THE SOUTHERN U.S. Huei-Jin Wang, Philip J. Radtke, and Stephen P. Prisley.....	265
STAND DENSITY INDEX AS A TOOL TO ASSESS THE MAXIMIZATION OF FOREST CARBON AND BIOMASS Christopher W. Woodall, Anthony W. D'Amato, John B. Bradford, and Andrew O. Finley.....	293
THE ZERO INFLATION OF STANDING DEAD TREE CARBON STOCKS Christopher W. Woodall and David W. MacFarlane.....	297

ASSESSING THE ACCURACY OF CROWN BIOMASS EQUATIONS FOR THE MAJOR COMMERCIAL SPECIES OF THE INTERIOR NORTHWEST: STUDY PLAN AND PRELIMINARY RESULTS

David L.R. Affleck and Brian R. Turnquist

ABSTRACT

Fueled by the insistencies of wildfire mitigation, bioenergy development, and carbon sequestration, there is growing demand for reliable characterizations of crown and stem biomass stocks in conifer forests of the Interior Northwest, United States (western Montana, northern Idaho, and eastern Washington). Predictive equations for crown biomass have been developed for this region but they have limited empirical support and supply markedly different predictions. This paper provides a methodological overview and preliminary results from an on-going study aimed in part at describing the accuracy of existing tree biomass equations for the Interior Northwest. Crown biomass estimates obtained from destructive sampling of 81 trees exhibited considerable variation around predictions from commonly used crown biomass equations based on DBH (diameter at breast height, 1.37 m). Some of this variation is attributable to within-tree sampling error, but initial results suggest that an appreciable proportion is due to variation in crown dimensions within DBH classes. Continuing data collection efforts will permit statistical descriptions of the accuracy of existing equations, as well as a basis for developing more integrative and precise tree biomass equations.

INTRODUCTION

The management of western North American conifer forests is increasingly attentive to the quantity and distribution of non-merchantable biomass in tree crowns and small-diameter trees. The aggregate mass and distribution of foliage have long been recognized as important determinants of tree and stand growth (see Long and Smith 1990). Likewise, in intensively managed systems, considerable research has focused on stand tending practices to control conifer crown architecture and thus wood quality (e.g., Waring and O'Hara 2005). However, it is the potential of conifer foliage and non-merchantable branch wood in processes other than stem development that have become central to the management of public and private forests across the inter-mountain western USA. Specifically, these forests are increasingly being managed to mitigate wildfire risk, to provide bio-energy stocks, or to sequester atmospheric carbon. Foliage and branch wood distributions

strongly affect wildfire behavior and, by the same token, form the primary constituents of bioenergy feedstocks. Thus, multiple emerging management goals have generated converging demands for accurate characterizations of conifer crown biomass, its distribution by component and branch size, and even its vertical distribution on the bole (see e.g., Dymond and others 2010, Keyser and Smith 2010, Reinhardt and others 2006).

BACKGROUND

Numerous studies undertaken across western North America have reported conifer biomass relationships and developed allometric equations (see reviews by Jenkins and others 2004, Ter-Mikaelian and Korzukhin 1997). Yet many of these studies have been confined to individual stands or have drawn data only from a particular subset of forest conditions, rendering the results unsuitable for widespread application. In practice, the biomass equations used in decision support for forest and fuels management in the Interior Northwest (i.e., from eastern Washington to western Montana) come primarily from a pair of studies carried out by Brown (1978; see also Brown and Johnston 1976) and by Jenkins and others (2003).

In 1978, Brown published a set of species-specific crown biomass equations for Rocky Mountain conifers. The equations were developed largely from dominant and codominant tree data collected in Idaho and Montana, but additional data from separate studies undertaken in Nevada and California were also incorporated. Brown developed predictive equations for multiple crown biomass components (foliage, dead branches, live branches of various size classes) but not for stem wood or stem bark. Separate equations were developed for 11 conifer species, including interior Douglas-fir (*Pseudotsuga menziesii*), western larch (*Larix occidentalis*), ponderosa pine (*Pinus ponderosa*), and lodgepole pine (*P. contorta*). Brown developed log-linear predictive equations based solely on

tree DBH (i.e., diameter at breast height, 1.37 m) as well as equations based on DBH, height, and dominance. His crown biomass equations have been integrated into the Forest Vegetation Simulator's Fire and Fuels Extension (Crookston and Dixon 2005, Reinhardt and Crookston 2003) and thus are now widely used in stand development simulations and fire behavior modeling.

The biomass equations of Jenkins and others (2003) were developed to provide a consistent basis for estimating tree biomass at large scales (e.g., at the regional or national level). Their DBH-based biomass equations were derived through meta-analysis of published biomass allometries (including the equations of Brown 1978) rather than from direct measurement of tree biomass. Based on similarities in equation form, Jenkins and others developed broad-based total aboveground biomass equations for species groups (e.g., all *Pinus* species; *Cupressaceae* plus *Larix* species) or, in the case of Douglas-fir, for both coastal and interior variants. Furthermore, since the study's primary emphasis was on total aboveground tree biomass, Jenkins and others (2003) developed a single set of component ratio equations to fractionate the total for any species into foliage, branch wood, and other tree biomass components. These component ratio equations are now used for tree biomass reporting in the Forest Inventory and Analysis (FIA) program (U.S.D.A. Forest Service 2010) and therefore find widespread application across the West.

The behavior of predictions from Brown's (1978) DBH-based equations and of those from Jenkins and others' (2003) crown biomass ratio equations are illustrated in Fig. 1. Both sets of equations were fit in log-linear form and while Brown's published equations incorporate a correction factor for logarithmic transformation, the equations from Jenkins and others do not. Within each of the 4 species shown, the predictions from these equations follow a similar exponential form but differ in magnitude. This is not surprising given the differences in the equations' derivations, intended spatial scales of application, and biological supports. As noted, the equations of Jenkins and others (2003) were intended for application across the continent and provide identical predictions for ponderosa and lodgepole pine; Brown (1978) focused exclusively on interior tree populations and estimated distinct allometric relationships for the two pine species in Fig. 1.

Figure 2 illustrates the magnitude of the differences between the predictive equations across a range of tree DBHs. The differences are appreciable, particularly for larger trees and for ponderosa pine, where crown biomass predictions from Brown's equations are consistently about 40 percent larger than those from the equations of Jenkins and others.

OBJECTIVES

The research presented here is part of a more extensive, ongoing study of conifer biomass distributions in the Interior Northwest. The discrepancies evident in Fig. 2 between biomass equations applied in this region are large and consequential for applications involving fuels management, bioenergy feedstock estimation, and carbon sequestration. There is a clear need for assessments of the validity and scope of these equations. To date, there has been little work to validate Brown's (1978) equations (but see Gray and Reinhardt 2003, Keyser and Smith 2010) and no evaluation of the bias or accuracy of the equations developed by Jenkins and others (2003) when applied to the major commercial conifer species of the Interior Northwest. The objectives of this study are therefore to:

1. formulate and implement efficient tree biomass data collection strategies for the major commercial conifer species in the Interior Northwest;
2. describe the bias and accuracy of existing tree biomass equations by species, across stem and crown components, and as a function of whole-tree dimensions; and,
3. develop and evaluate new equations for tree biomass as well as its distribution across components and over the vertical profile of the stem.

This paper provides an overview of the data collection strategies that were developed and presents preliminary results regarding the accuracy of the crown biomass equations described above.

SAMPLING METHODS

Biomass equation validation and development efforts require sizable samples for individual species, preferably distributed across the region of interest and its forest habitat types. This is complicated by the high cost and destructive nature of tree biomass assessment. Stem biomass determination necessitates bole weight or wood density measurements. Biomass assessment of crown components demands defoliation of individual branches and the separation of branch wood into various size classes. Green tree materials also need to be oven-dried to obtain dry weights. To mitigate the high cost of tree-level biomass assessment and collect a large sample of trees, this study implemented a three-phase biomass sampling strategy to select stands, trees, and finally individual branches or stem discs along the boles of selected trees.

STAND AND TREE SELECTION

Second-growth stands across the Interior Northwest were selected to ensure broad geographic support (Fig. 3). Spatial coverage and dispersion across forest habitat types (Pfister and Arno 1980) were the primary factors in stand selection, but no formal systematic or random mechanism was applied. Only stands with no treatment history over the previous decade were candidates for sampling. Stand selection was also conditioned by the availability of permits for destructive sampling. Stands selected in 2009 and 2010 were located on federal, State, tribal, and private forest lands.

Within selected stands, sample points were located systematically at 100 m intervals on the Universal Transverse Mercator (UTM) grid. At each sample point a narrow angle gauge (2.3–4.6 m²/ha basal area factor) was used to identify candidate sample trees. Candidate trees were then barred if they were not among the species of interest or had damaged or missing crowns. Up to two of the remaining candidate trees at a sample point were then selected uniformly at random for destructive sampling.

TREE BIOMASS ASSESSMENT

Individual trees were sub-sampled to estimate stem, branch wood, and foliage biomass. Trees were felled and then randomized branch sampling (RBS; Gregoire and Valentine 2008) was employed to select 5 live branches with probability proportional to branch cross-sectional area. For selection purposes, the branches making up the live crown were artificially clustered into 1-m intervals. That is, beginning at the lowest live branch, all branches found within successive 1-m segments on the bole were treated as distinct whorls so that in addition to branch basal diameters only a single stem diameter (at the top of a 1-m segment) was needed. RBS focuses sampling efforts on the larger diameter branches that account for the majority of the crown biomass. The corresponding estimators capitalize on the strong allometric relationships between branch mass and branch basal area (Fig. 4) to provide precise and unbiased estimates of whole-crown biomass.

The selected branches were separated into size-class components so that separate biomass estimates could be obtained for foliage and for branch wood within the 0–0.64 cm, 0.64–2.5 cm, and 2.5+ cm diameter classes (corresponding to 1-hour, 10-hour, and 100-hour time lag fuel classes). Dead and epicormic branches encountered along the live-branch selection paths were also cut and weighed. All live branch material as well as selected bolts of dead branch wood were oven-dried at 105°C. Drying times varied by component and were determined by evaluating the time needed to achieve a constant weight.

Though not discussed below, data were also collected to estimate the stem biomass of selected trees. Discs were

cut from the downed tree at a systematically selected set of heights or, in some stands, at heights determined by merchantability criteria (e.g., at the tops and bottoms of the first two logs). Cross-sectional area and wood density were measured on the discs and calibration estimators (see Gregoire and Valentine 2008) based on regional tree taper equations were then used to obtain whole-stem biomass estimates. A more thorough description of the stem and crown sampling procedures can be obtained from the authors.

RESULTS AND DISCUSSION

In 2009, biomass data were collected from 81 trees in 11 stands in western Montana and eastern Washington (Fig. 3). Data from Engelmann spruce (*Picea engelmannii*) and grand fir (*Abies grandis*) were collected but the bulk of the data were from ponderosa pine, Douglas-fir, western larch, and lodgepole pine. The size-class distribution of the 2009 sample trees of these four species is shown in Fig. 5. Within each of the species, the sample trees spanned a wide range of DBH. The ponderosa pine sample was also well distributed across crown ratio classes but in other species high crown ratios were rarely observed at larger DBHs. This is broadly consistent with the growing conditions of these species. However, data spanning the DBH, height, and crown ratio domains are needed to characterize variation in crown biomass across these dimensions and to assess the utility of the DBH-based equations from Brown (1978) and Jenkins and others (2003).

Figure 6 shows the relationship between tree DBH and estimated total crown mass for the 4 most commonly selected species in the 2009 sample. Total crown mass includes the mass of foliage, live and dead branch wood, and the stem above a 5 cm top. The crown mass estimates are based on subsamples (drawn by RBS) from the crowns of individual sample trees and are thus subject to sampling error. In Fig. 6, this tree-level sampling error is conflated with among-tree differences in crown biomass potentially attributable to variations in tree height, tree crown length (or crown ratio), stand stocking, stand species composition, and site productivity, in addition to intrinsic heterogeneity. Only the conditioning effect of tree DBH is shown in Fig. 6 with the result that considerable variation in crown mass is evident. This is particularly true for larger trees and for ponderosa pine, where crown biomass estimates for trees above 40 cm DBH range from 123–482 kg.

Predictions from the DBH-based equations of Brown (1978) and Jenkins and others (2003) are superimposed on the data in Fig. 6, as are smoothed loess regression curves. Though little data are presently available for lodgepole pine, the pine and Douglas-fir predictions from Jenkins and others'

equations appear to track the empirical trends more closely than those from Brown's equations. The opposite is true for the limited western larch dataset. In all cases, however, there exists substantial variation around the crown mass predictions for large-DBH (i.e., above 30 cm) trees.

As more data are made available for these and other species, more exacting assessments of the overall bias and conditional bias (see e.g., Reynolds and Chung 1986) of Brown's (1978) and Jenkins and others' (2003) prediction equations will be undertaken. At this preliminary stage, our interest is primarily in describing the sources and magnitudes of variation in crown biomass estimates around predictions. Figure 7 focuses on the performance of the predictive equations of Jenkins and others (2003) for Douglas-fir and ponderosa pine. The trend lines in Fig. 7 are smoothed loess regressions. In the case of Douglas-fir, the trend line identifies a DBH-class (approximately 15-35 cm DBH) for which the predictions exceed the observed crown mass estimates. On ponderosa pine, the empirical trend is more consistent but runs strictly above 0 percent, reflecting the tendency for this equation to consistently understate crown biomass relative to the levels observed. For both species the sample trees' crown biomass estimates diverge on the order of -100 to +50 percent from predictions.

Individual trees are drawn as solid or open circles in Fig. 7 according to whether their destructive sampling estimates respectively exceed or fall short of the DBH-based predictions from Jenkins and others' equations. This symbology is carried through to Fig. 8 where the sample trees' crown ratios are plotted against DBH. By this means, Fig. 8 shows that crown biomass estimates falling short of predictions are predominantly observed on trees with lower crown ratios within their respective DBH classes, and vice versa. This result accords with both dimensional and ecological considerations. After tree DBH, dimensions related to crown length should have the greatest impact on total crown mass. Likewise, in untreated stands, crown length and ratio reflect the past growing conditions of the tree and thus integrate the influences of stand density and species composition.

Future analyses will focus on the importance of crown ratio, tree height, and stand density in modifying foliage, branch wood, and total crown mass. In doing so, these analyses will provide information on the bias and accuracy of biomass predictions based only on tree DBH as well as on the potential need for predictive biomass equations integrating other tree and stand characteristics.

SUMMARY AND FUTURE RESEARCH

Management of conifer forests in the Interior Northwest for wildfire fuels reduction, bioenergy extraction, or carbon sequestration requires reliable estimates of tree and crown biomass. The bias and accuracy of the predictive equations currently applied in the region have not been evaluated and in many cases these equations supply markedly different predictions (Fig. 2). Based on a preliminary dataset of 81 trees selected from across the region in 2009, existing DBH-based biomass equations broadly follow the empirical trends in crown biomass but fail to account for considerable variation in individual-tree estimates. Some of this variation is attributable to within-tree sampling error. However, it is anticipated that a substantial portion of this variation is due to among-tree differences in crown length, tree height, and stand conditions. In particular, exploratory analyses of the 2009 data point to crown ratio as an important modifier of crown biomass in Douglas-fir and ponderosa pine (Fig. 8).

The present study is on-going and as more data become available for these and other species it will be feasible to statistically assess the presence of trends in the bias and accuracy of existing crown biomass equations as a function of tree DBH, tree height, crown ratio, and stand density. To do so, sampling procedures should ensure that trees selected for destructive biomass sampling span a broad range of tree sizes, crown lengths, and stand conditions. Future analyses will also examine variations in stem biomass for the commercial species of the region and the accuracy of existing stem biomass prediction algorithms, including those used in FIA reporting.

ACKNOWLEDGMENTS

Funding for this project was provided by the Joint Fire Science Program (10-1-02-13), the Spokane Tribe of Indians, the Inland Northwest Growth & Yield Cooperative (INGY), and the USDA Forest Service Region 1 Renewable Resource Management office. The project would not have been possible without the cooperation of many people with the USDA Region 1 National Forest System, the Spokane Tribe of Indians, Plum Creek Timber Co., the Montana Dept. of Natural Resources and Conservation, the University of Montana, the Confederated Salish and Kootenay Tribes, and the Coeur d'Alene Tribe. The authors thank the INGY Technical Advisory Committee for its assistance in developing and implementing the project, as well as the contributions of dedicated field personnel. Dennis Ferguson and John Goodburn provided helpful comments in reviewing earlier drafts of this manuscript.

LITERATURE CITED

- Brown, J.K.** 1978. Weight and density of crowns of Rocky Mountain conifers. Res. Pap. INT-197. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station.
- Brown, J.K.;** Johnston, C.M. 1976 Debris Prediction System. Fuel Science RWU 2104. Missoula, MT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Northern Forest Fire Laboratory.
- Crookston, N.L.;** Dixon, G.E. 2005. The Forest Vegetation Simulator: A review of its structure, content, and applications. Computers and Electronics in Agriculture 49: 60-80.
- Dymond, C.C.;** Titus, B.D.; Stinson G.; Kurz, W.A. 2010. Future quantities and spatial distribution of harvesting residue and dead wood from natural disturbances in Canada. Forest Ecology and Management 260: 181-192.
- Gray, K.L.;** Reinhardt, E.D. 2003. Analysis of algorithms for predicting canopy fuel. In: Proceedings of the 2nd International Wildland Fire Ecology & Fire Management Congress and 5th Symposium on Fire and Forest Meteorology, 16-20 Nov 2003, Orlando FL. American Meteorological Society, Boston MA.
- Gregoire, T.G.;** Valentine, H.T. 2008. Sampling Methods for Natural Resources and the Environment. Chapman & Hall/CRC, Boca Raton, FL.
- Jenkins, J.C.;** Chojnacky, D.C.; Heath, L.S.; Birdsey, R.A. 2003. National-scale biomass estimators for United States tree species. Forest Science 49: 12-35.
- Jenkins, J.C.;** Chojnacky, D.C.; Heath, L.S.; Birdsey, R.A. 2004. Comprehensive Database of Diameter-Based Biomass Regressions for North American Tree Species. Gen. Tech. Rep. NE-319. Newton Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station.
- Keyser, T.;** Smith, F.W. 2010. Influence of crown biomass estimators and distribution on canopy fuel characteristics in ponderosa pine stands of the Black Hills. Forest Science 56: 156-165.
- Long, J.N.;** Smith, F.W. 1990. Determinants of stemwood production in *Pinus contorta* var. *Latifolia* forests: the influence of site quality and stand structure. Journal of Applied Ecology 27: 847-856.
- Pfister, R.D.;** Arno, S.F. 1980. Classifying forest habitat types based on potential climax vegetation. Forest Science 26: 52-70.
- Reinhardt, E.D.;** Crookston, N.L. 2003. The Fire and Fuels Extension to the Forest Vegetation Simulator. Gen. Tech. Rep. RMRS-116. Ogden, UT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Reinhardt, E.D.;** Scott, J.; Gray, K.; Keane, R. 2006. Estimating canopy fuel characteristics in five conifer stands in the western United States using tree and stand measurements. Canadian Journal of Forest Research 36: 2803-2814.
- Reynolds, M.R., Jr.;** Chung, J. 1986. Regression methodology for estimating model prediction error. Canadian Journal of Forest Research 16: 931-938.
- Ter-Mikaelian, M.T.;** Korzukhin, M.D. 1997. Biomass equations for sixty-five North American tree species. Forest Ecology and Management 97: 1-24.
- U.S.D.A. Forest Service.** 2010. The Forest Inventory and Analysis Database: Database Description and Users Manual Version 4.0 for Phase 2 (Draft Revision 3). Arlington, VA: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis Program.
- Waring, K.M.;** O'Hara, K.L. 2005. Ten-year growth and epicormic sprouting response of western larch to pruning in western Montana. Western Journal of Applied Forestry 20: 228-232.

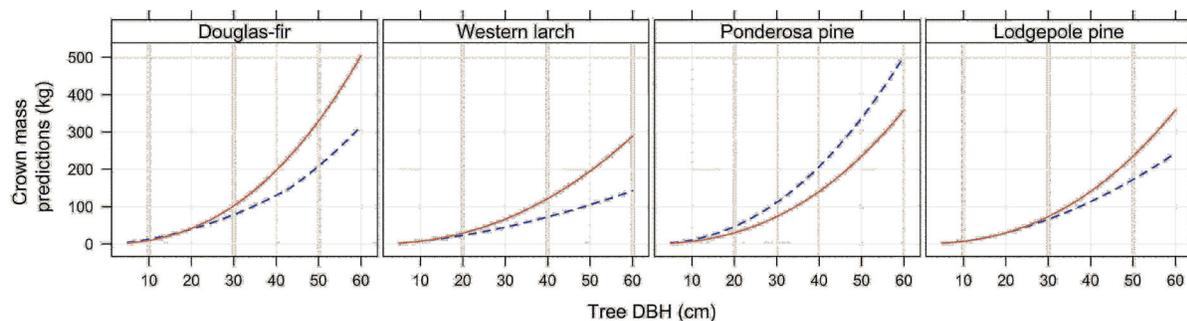


Figure 1—Diameter-based crown biomass equations from Jenkins and others (2003; solid lines) and Brown (1978; dashed line); predictions are of oven-dry mass.

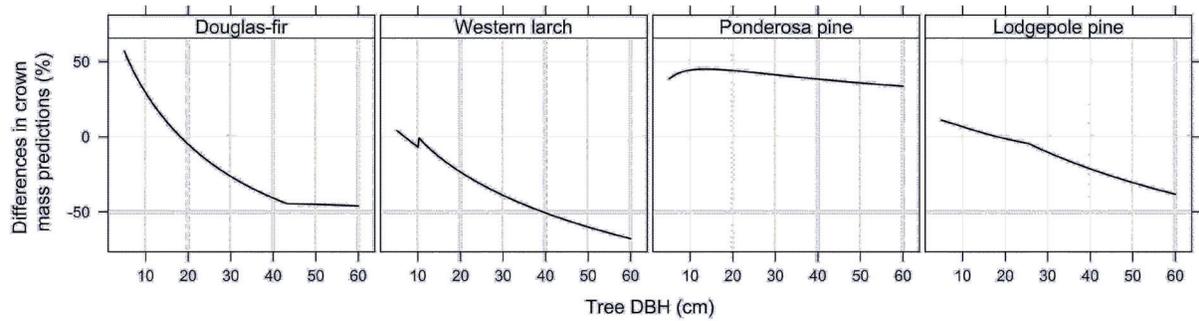


Figure 2—Percent difference in the diameter-based crown biomass equations from Jenkins and others (2003) and Brown (1978) as a function of tree diameter.

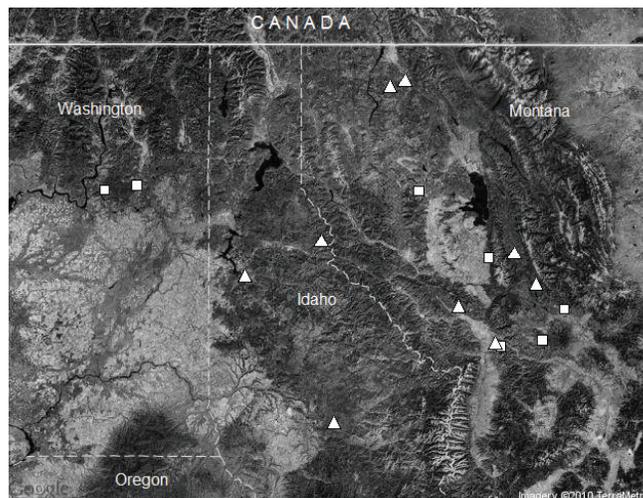


Figure 3—Geographic distribution of 2009 (squares) and 2010 (triangles) sample stands; satellite imagery from Google Maps.

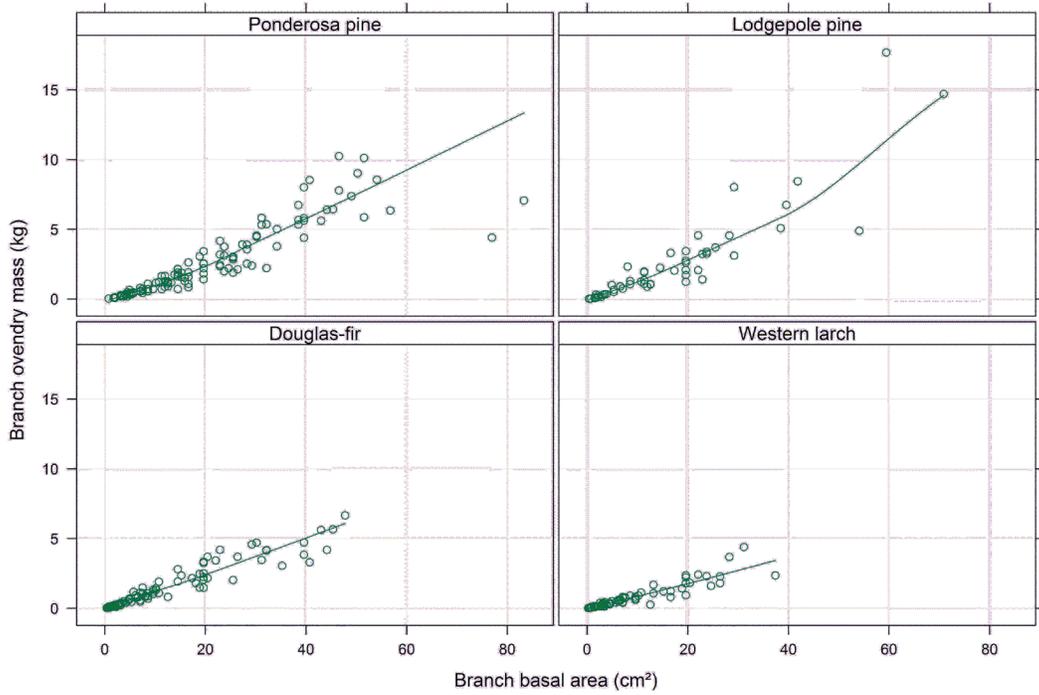


Figure 4—Allometric relationships for 352 live branches selected by randomized branch sampling (one ponderosa pine branch with basal area 170 cm² not shown); Pearson correlations were at or above 0.90 for all four species.

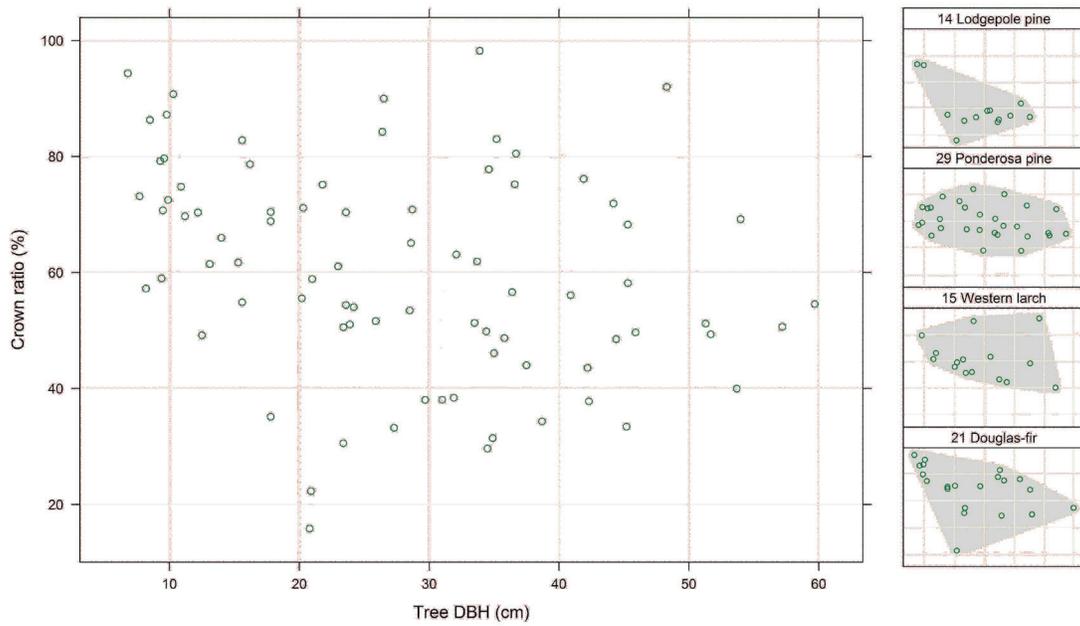


Figure 5—Overall and species-specific size distributions of sample trees selected in 2009.

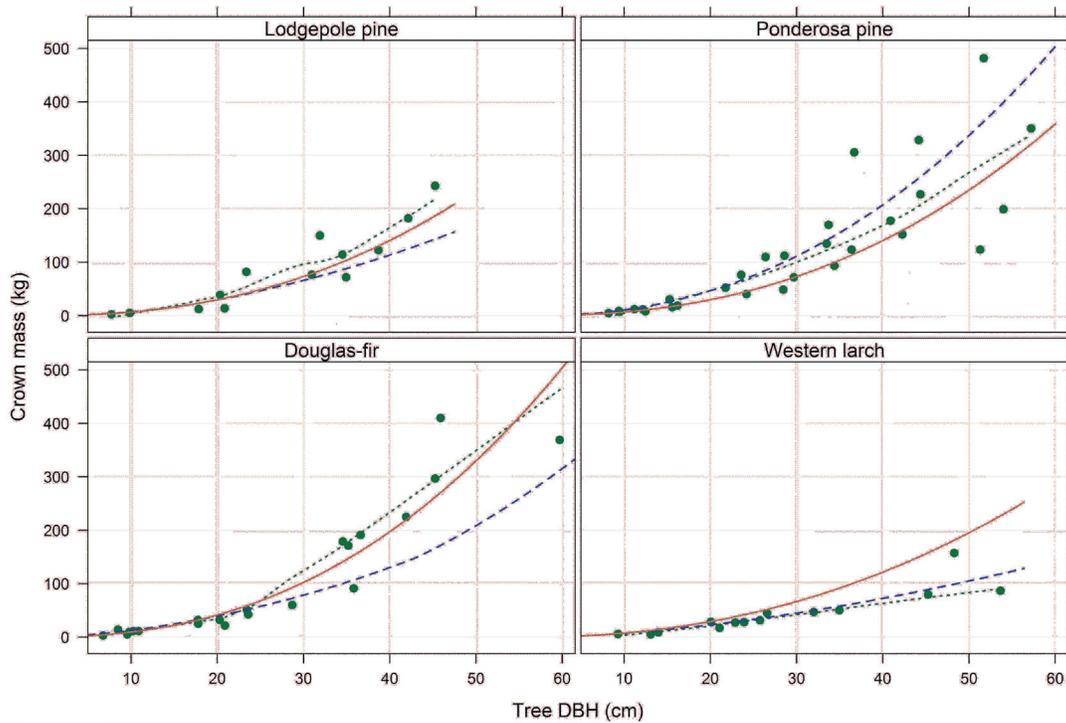


Figure 6—Estimated oven-dry crown biomass of sample trees with loess smoothed trend (dotted line) and with diameter-based crown biomass equations from Jenkins and others (2003; solid lines) and Brown (1978; dashed line).

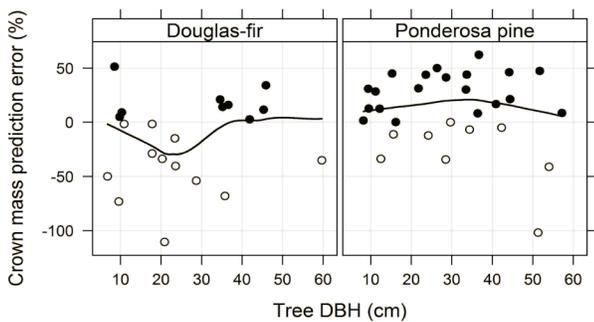


Figure 7—Crown biomass prediction errors as a percentage of estimated mass; solid circles denote trees with crown mass estimates higher than predicted from the equations of Jenkins and others (2003) while open circles denote trees with estimates below predictions.

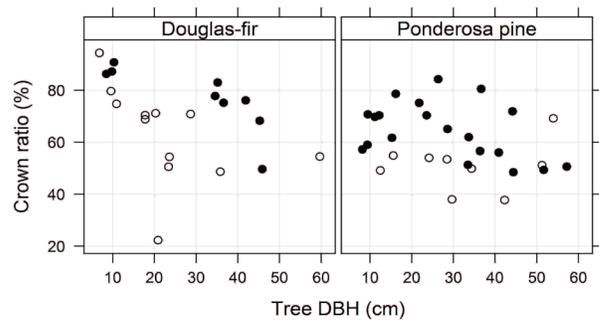


Figure 8—Tree size distribution and crown mass prediction errors associated with the equations of Jenkins and others (2003); solid circles denote trees with higher than predicted crown mass estimates while open circles denote trees with estimates below predictions.

FOREST BIOMASS SUPPLY FOR BIOENERGY IN THE SOUTHEAST: EVALUATING ASSESSMENT SCALE

Christopher S. Galik and Robert C. Abt

ABSTRACT

This study evaluates the potential impacts of expanded forest biomass use in the Southeast from present year through 2036, focusing on the forest supply, industrial, and GHG emissions implications of maximizing biomass co-firing with coal. We model demand scenarios at the state, subregional, and regional levels, and assess the influence of study scale on the observed results. We find that pricing effects are greatest under a state-level assessment scenario, followed by subregional and regional assessments. This has important impacts on observed shifts in forest land use and forest stand carbon, with the state-level assessment resulting in the greatest amounts of forested acreage and carbon relative to the other assessment scales. State-level assessments also experience the lowest relative displacement of pulpwood capacity of the three scales considered, with spatial and temporal dynamics of resource allocation playing a strong role in our findings. If forested acres, forest carbon, and aggregate displacement are the only issues of concern, then these results would suggest that a program encouraging the use of forest biomass for renewable energy production may be best implemented at the state level, rather than at some larger scale. Given the wide variety of other environmental, economic, and social objectives that must be satisfied, however, continued careful evaluation of the multiple impacts of increased forest biomass use is necessary.

INTRODUCTION

Interest in renewable sources of energy is increasing for a variety of reasons, including the mitigation of climate change and furtherance of energy independence. This interest is reflected in an increasing number of proposed and enacted regulations, programs, and initiatives at both state and federal levels. In the Southeastern United States, attention is often focused on the role that forest biomass can play in meeting these and other policy objectives (English and others, 2004). Although abundant, the supply of forest resources in the region is subject to a number of ecological, institutional, and economic constraints.

Despite these multiple factors and potential limitations, the impact of renewable energy policy has not always been carefully evaluated prior to program inception (Sedjo and Sohngen, 2009). Even though multiple studies have estimated potential biomass supply (e.g., Pennock and Doron, 2009; Perlack and others, 2005; see also

Gronowska and others, 2009), estimates of aggregate supply are complicated by the potential interplay between increasing resource demand for biofuel and bioenergy and the competition with current users of the resource (e.g., Lundmark, 2006; Galik and others, 2009; Abt and others, 2010a; Abt and others, 2010b). Also complicating matters are the various scales at which policy to encourage the use of biomass are being discussed (e.g., State, federal) and the differing market responses that could occur as a result. For example, an increase in demand for forest biomass could induce greater harvest activity, as well as displacement of existing pulpwood capacity. A land use response would likewise be expected (Abt and others, 2010b). Should the scale at which biomass demand is evaluated itself influence the magnitude of the resulting price response, we would likewise expect harvest, forest land use, and displacement responses to vary by assessment scale.

In the analysis that follows, we consider the role that assessment scale factors into the potential near-term impacts of expanded forest biomass use in the Southeast. In particular, we focus on the forest supply, industrial, and greenhouse gas (GHG) emissions implications of maximizing biomass co-firing with coal. To better understand the spatial issues at play, we model demand scenarios at the state, subregional, and regional levels from present year through 2036, and evaluate the influence of assessment scale on the observed results. For each assessment level, we examine the effect of increased biomass demand on weighted pine pulpwood prices, potential displacement of existing pine pulpwood capacity, forested acreage, and forest carbon. We identify the important trends and tradeoffs that emerge, and conclude with areas of future research needed to improve and expand this and related work.

METHODS

This analysis builds off of the methodology and findings of Abt and others (2010b), which assumes that the increased

Christopher S. Galik. Research Coordinator, Nicholas Institute for Environmental Policy Solutions, Duke University. Durham, NC 27708
Robert C. Abt. Professor of Forest Resource Economics, Department of Forestry and Environmental Resources, North Carolina State University. Raleigh, NC 27695

demand for bioenergy is driven by a maximization of co-firing potential (on a direct injection basis) at existing coal-fired boilers in a ten-State Southeastern region. Below we discuss the establishment of three separate policy assessment scales, estimation of baseline forest product demand, and the calculation of biomass co-fire demand. We then discuss the modeling of forest biomass supply response and the estimation of forest stand carbon.

ASSESSMENT SCALE

The analysis is conducted at three assessment scales: individual State, subregional, and regional levels. The region evaluated includes Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, and Virginia. In the State-level analyses, we assume that co-firing demand is met purely with in-state biomass resources. For regional assessments, we assume that demand for any plant may be met with biomass sourced from anywhere within the entire ten-State region. At the subregional level, we identify seven semi-distinct clusters within which woody biomass may be sourced to meet co-firing demand (Figure 1).

The implementation of the regional definitions has important market implications. Both demand (which is determined by co-firing boiler capacity) and supply (which is determined by harvest level and associated logging residuals in the short run and by forest type and age class structure in the long run) are defined by a particular regional specification. For example, State level analysis reflects the probability that State-level renewable energy policies will differ. Co-firing demand in the seven subregions is met by the overlapping supply basins within each cluster that are by definition distinct from other clusters. The State and subregional analyses treat the markets as independent; supply and demand pressures in one State or subregion do not affect neighboring ones. These are useful for gauging regional comparative advantage where demand will be affected by State policy and clustering of boiler capacity and resource prices and industry displacement will be affected by the distribution of resources and associated industry capacity. Alternatively, results from the regional analysis are useful for gauging the region-wide market effects assuming that price differences between markets will be minimized by market pressures. This may represent a longer term view of market behavior that may be dominated by regional differences in the short-run. In this analysis, we have chosen to look at the aggregate impact of market and resource metrics. This approach allows us to simply identify the impact of regional scale, but it does so at the expense of exploring regional variation.

BIOMASS DEMAND

We assume here that the forest products industry experienced a 30 percent decline in demand for four

separate product classes (pine pulpwood, pine sawtimber, hardwood pulpwood, hardwood sawtimber) from 2006 to 2010. We also assume a recovery to near pre-recession demand levels by 2013, beyond which demand remains constant at pre-recession levels through the end of our projection. We then add to this base level of biomass demand the estimated amount of biomass used under a maximization of coal co-fire capacity in the region. Year-2007 coal co-firing capacity for individual facilities is determined at the boiler level using boiler configurations contained in the eGRID database (U.S. Environmental Protection Agency, 2008) and maximum coal co-firing capacity on a direct injection basis for each boiler type as indicated in various technical sources (e.g., Grabowski, 2004; Federal Energy Management Program, 2004). Gross energy demand is converted to woody biomass demand using a wood-to-energy conversion rate of 9,000 BTU per pound of dry biomass is the estimated energy content of wood. The value used here falls in the upper-middle portion of estimates of biomass energy content (For example, see http://bioenergy.ornl.gov/papers/misc/energy_conv.html [Retrieved March 26, 2010]); we acknowledge that different types of biomass have different energy content and that additional energy will be necessary to dry green biomass to achieve this energy content on a per-pound basis.

It is also important to note that we make no assumptions about the type of policy that drives this emerging market, only that the incentive to use biomass exists and that co-firing represents among the quickest and easiest paths to do so. Indeed, co-firing is often thought to represent a cost-effective path to biomass utilization in the near term (Robinson and others, 2003; De and Assadi, 2009; Lintunen and Kangas, 2010). While we do not specifically consider the added effects of biomass use in dedicated, low-GHG generation facilities or for the production of pellets or other wood fuels, the near-term demand for co-firing provides a point estimate from which we may begin to assess the various sectoral and temporal tradeoffs likely to accompany an expanding forest biomass energy market.

BIOMASS SUPPLY

Supply effects of maximization of co-firing in the Southeast are assessed using the Sub-Regional Timber Supply (SRTS) model. SRTS models product demand as a function of product stumpage price and demand shifts through time; greater description of the SRTS model and its application may be found in Abt and others (2009) and Prestemon and Abt (2002). To simulate the impact of added biomass demand, we conduct a baseline run of traditional wood-using industries to derive estimates of logging residuals. For all runs, we consider two separate scenarios of residual utilization, one in which utilization of residuals increases over time and peaks at a 50 percent utilization rate in 2020, and one in which utilization peaks at 25 percent. Utilized

residuals are assumed to reduce biomass demand for roundwood, with the net remaining roundwood demand shifting demand for both pine and hardwood pulpwood proportionately.

GHG IMPLICATIONS

Finally, we consider the GHG implications of our modeled scenarios. As in previous analyses, we convert SRTS inventory output into estimates of forest carbon using forest-type-specific relationships and equations defined by the U.S. Forest Service (see Foley, 2009; Abt and others, 2010b). Although Abt and others (2010b) include an estimate of emissions displaced from a reduction in coal usage as compared to observed shifts in on-the-ground forest biomass carbon sequestration, we limit our discussion here to the GHG implications of shifts in planting and harvest behavior.

RESULTS AND DISCUSSION

The results of the analysis are described below, beginning with an overview of total biomass co-fire demand. We then describe the effects of increased biomass demand on pulpwood prices, displacement, forested acreage, and forest carbon for each of the three assessment scales evaluated here. We conclude with a brief discussion of the implications of these findings for bioenergy policy.

When maximizing co-firing capacity at existing coal-fired boilers, we estimate an aggregate average co-firing rate of approximately 10.1 percent. This translates to an annual biomass consumption of approximately 532 million MMBTU or just over 59 million green tons of wood. Disaggregating the larger region into seven distinct supply subregions and into the ten individual States, we find that biomass co-fire demand varies significantly across the Southeast (Table 1).

We next examine the relative shift in the price of pine pulpwood, the biomass component most likely to be affected by increases in near-term demand. Because our price output for each State or subregion is reported as a price shift relative to 2006 prices (and not as an absolute value), we use the weighted average of these price shifts over time to make comparisons across the different assessment scales. The weighted average is based on the relative price shift reported for each State or subregion, multiplied by the volume of removals in that State or subregion. When summed across all States or subregions and divided by total removals, this yields a single metric for each assessment scale, allowing direct comparison. There is of course some detail lost in this aggregation process; this is further discussed below in the context of pulpwood capacity displacement.

Figure 2 compares the relative pine pulpwood price shifts over time. The regional assessment (South-wide) has the

smallest shift, followed by the subregional assessment. The greatest relative shift is generally found in the State-level assessment. The pattern holds for both residue utilization scenarios, though is slightly shifted lower in the 50 percent scenario. This is expected, as the greater amount of available residues both satisfies a greater portion of the biomass demand and tempers price increases in the process. There is also a temporal component, as prices tend to diverge somewhat as magnitude of biomass co-fire demand increases over time and is sustained through the later years of the assessment.

The shift in resource pricing is expected to have several affects. One area of concern is feedstock allocation, or the source of biomass from which increasing levels of demand are met. Although previous work indicates that significant differences in feedstock source exist when comparing different portions of the study to each other (Abt and others, 2010b), we are more concerned here with aggregate feedstock contributions across three assessment scales. Specifically considered in the analysis were shifts in hardwood and pine pulpwood harvests, displacement of existing hardwood pulpwood capacity, displacement of existing pine pulpwood capacity, harvest residues from hardwood or pine pulpwood, and residuals stemming from harvest of roundwood specifically for biomass.

Focusing on the displacement component, we find that relative displacement likewise varies by assessment scale. Further exploration of the findings, however, expose the potential hazards of weighting or aggregating findings at different assessment scales for comparison at the regional level. This is because harvest patterns in the SRTS model shift over time based on relative price pressure, which is itself driven on the supply side by changes in inventory. Inventory change across regions is driven by harvest and growth distributions and age class structure of the inventory. On the demand side, regional variation is driven by the regional co-firing capacity tempered by the availability of residuals. Co-firing capacity tends to be associated with population while residual availability is associated with the distribution of forest product industry capacity. The interrelationship between harvest and co-firing demand drive the displacement impacts.

The overall displacement pattern increases as co-firing demand increases from 2015 to 2030 (Figure 3). After 2030 wood supplies increase as the trees planted in response to initial price increases enter the product supply inventory. For both State and subregional assessment scales, the available supply is constrained relative to the regional one. Figure 3 also shows that both the State and subregion assessment scales follow a similar pattern. The higher displacement for the subregion assessment stems a very high correlation ($r = +.98$) between initial harvest levels (industry location) and

co-firing demand. Across States, the correlation is much smaller ($r = +.46$).

For the regional assessment, the same overall pattern of displacement over time occurs, but another dynamic exists since supply can react to differences in price pressure across regions. Because the regions with the highest initial pine inventory increases also have the highest concentration of current industry as reflected by harvest ($r = +.61$), biomass harvest in the regional assessment shifts to areas with existing industry (e.g., Alabama, Florida, Georgia). This results in a higher displacement trend during the 2015 to 2030 period. Over time the increased harvest from co-firing reduces the comparative advantage of these regions and correlation of additional harvest with initial harvest declines (Figure 4). Since this harvest is also located in areas with the largest planting response to prices, displacement drops faster as new plantations come online.

We would also expect to see a land use response consistent with the ranking of resource pricing shifts. Stated another way, if relative price increases are greater the smaller the assessment area, we would expect forest land use response to be greater with a State-level policy requirement than in the presence of a subregional and regional one. This is reflected in the findings, with the State-level assessment possessing the greatest relative shift in forested acres relative to the baseline, non-co-fire scenario, followed by the subregional assessment, and finally the regional assessment (Figure 5). As before, the influence of residue harvest efficiency is evident, with a lower amount of forest acres being added relative to the baseline in scenarios in which greater residue availability is assumed.

Finally, we examine the effects of assessment scale on aggregate forest stand carbon. As with forested acreage, the greatest amount of forest carbon storage is found under a State-level assessment (Figure 6), followed by subregional and regional levels. We note, however, that some reductions in forest carbon relative to the baseline scenario are found for all three policy scenarios, but that the storage under a State-level policy scenario is greatest. Even though forested acreage increases relative to a non-co-fire scenario, total forest carbon falls due to management intensification and a shift towards younger (and smaller), faster growing trees. Net carbon sequestered under a given assessment scale and for a given year is a function of both this shift in intensity, as well as the total number of forested acres across the landscape.

While forest stand carbon provides one measure of the GHG performance of a policy scenario, also important to consider are corresponding reductions in GHG emissions from fossil fuels. We do not specifically evaluate this component here, but hypothesize that a State-level policy

scenario would result in greater net emission reductions relative to both the regional and subregional assessment scales. The amount of fossil GHG emissions foregone through the use of forest biomass in place of coal is the same across all policy assessment scales, as is the volume of biomass being removed from the forest. This means that the carbon dynamics of the standing forest would likely determine the net GHG performance of the scenario. Using a similar level of biomass demand and a similar method to calculate the effects of this demand, Abt and others (2010b) find that GHG emissions relative to a non-co-firing scenario are possible at both region-wide and subregional assessment scales. In the current study, estimated State-level forest carbon exceeds the carbon stored for either of these, implying that GHG reductions would likely exceed those reported in Abt and others (2010b) for regional and subregional levels.

CONCLUSION

The results presented here indicate that the scale at which bioenergy policy is enacted can influence the resource management trends, even if the amount of energy produced as a result is unchanged. This suggests that if acreage, displacement, and carbon are the only metrics of concern, then a program encouraging the use of forest biomass for renewable energy production may be most efficiently implemented at the State level, rather than at some larger scale. Given the wide variety of other environmental, economic, and social objectives that must be satisfied, however, continued careful evaluation of the multiple impacts of increased forest biomass use is necessary. There are also site-specific opportunities and tradeoffs to consider, such as the local and community impacts of increased harvest activity, construction and operation of new energy facilities, or shifts in activity in existing mills, none of which are directly addressed here.

Furthermore, we make no assumption about the type of policy that is put in place to achieve this increase in biomass demand; we assume only that the demand exists. Because different types of policies can affect markets in various ways, we do not suggest that this finding is universal. In situations where compliance with bioenergy targets are mandated and energy producers are less sensitive to pricing increases than existing industrial users of biomass resources, however, the price effects induced by the increase in demand for woody biomass are greater the smaller the assessment scale.

An important concluding point is the impacts of increased bioenergy demand for woody biomass demand vary over time and space, and that the scale of the policy applied also influences modeled effects. As assessment scale decreases from the region to the subregion to the state level, price

increases lead to increased forest acreage retention and moderated carbon loss. Our displacement estimates at different scales are complicated by spatial and temporal dynamics of biomass allocation, and break from this trend somewhat, with subregional and regional assessments experiencing greater relative displacement than at the state level.

The empirical findings with regard to shifts in pricing, acreage, and carbon confirm the otherwise intuitive relationship between assessment scale and increases in biomass demand. Further investigation of the spatial and temporal dynamics necessitated by our initially counterintuitive findings with regard to displacement likewise provide interesting perspective on resource allocation over time. As such, we believe that the analysis adds much-needed information and perspective to the debate over the role that biomass utilization is to play in state, regional, and national climate and energy policy. More research is needed to fully grasp the economic, social, and environmental implications of increased forest biomass utilization. In particular, the potential market for harvest residues remains uncertain, with multiple questions remaining with regard to the cost-effectiveness of their harvest, transportation, and use. The potential for dedicated energy crops, including switchgrass and short rotation woody biomass, to displace forest biomass as a fuel of first resort is likewise uncertain at the present time. Further work to evaluate the economics of these key questions is necessary to improve our understanding of biomass market response to policy drivers.

ACKNOWLEDGEMENTS

We would like to thank Karen Abt and Justin Baker for their immensely helpful, insightful, and constructive comments on an early version of this manuscript. Any remaining errors are the sole responsibility of the authors.

LITERATURE CITED

- Abt, R.C.;** Cabbage, F.W.; Abt, K.L. 2009. Projecting southern timber supply for multiple products by subregion. *Forest Products Journal* 59:7-16.
- Abt, R.C.;** Abt, K.L.; Cabbage, F.W.; Henderson, J.D. 2010a. Effect of policy-based bioenergy demand on Southern timber markets: a case study of North Carolina. *Biomass and Bioenergy* 34: 1679-1686.
- Abt, R.C.;** Galik, C.S.; Henderson, J.D. 2010b. The Near-Term Market and Greenhouse Gas Implications of Forest Biomass Utilization in the Southeastern United States. Climate Change Policy Partnership, Duke University, Durham, NC. 49p.
- De, S.;** Assadi, M. 2009. Impact of cofiring biomass with coal in power plants - A techno-economic assessment. *Biomass and Bioenergy* 33:283-293.
- English, B.C.;** Jensen, K.; Menard, J. [and others]. 2004. Economic Impacts Resulting from Co-firing Biomass Feedstocks in Southeastern United States Coal-Fired Plants Select Paper prepared for presentation at the American Agricultural Economics Association annual meeting, Denver, CO, July 1-4, 2004. 22p.
- Federal Energy Management Program.** 2004. Biomass cofiring in coal-fired boilers. U.S. Department of Energy, National Renewable Energy Laboratory, Golden, CO. 40p.
- Galik, C.S.;** Abt, R.C.; Wu, Y. 2009. Forest biomass supply in the Southeastern United States—implications for industrial roundwood and bioenergy production. *Journal of Forestry* 107:69-77.
- Grabowski, P.** 2004. Biomass Cofiring. Retrieved October 9, 2006, from <http://www.biomass.govtools.us/pdfs/PGCofiring.pdf>.
- Gronowska, M.;** Joshi, S.; MacLean, H.L. 2009. A review of U.S. and Canadian biomass supply studies. *BioResources* 4:341-369.
- Lintunen, J.;** Kangas, H.-L. 2010. The case of co-firing: The market level effects of subsidizing biomass co-consumption. *Energy Economics* 32:694-701
- Pennock, C.;** Doron, S. 2009. Southern Bioenergy Roadmap. Southeast Agriculture and Forestry Energy Resources Alliance, Research Triangle Park, NC. 136p.
- Perlack, R.D.;** Wright, L.L.; Turhollow, A.F. [and others]. 2005. Biomass as feedstock for a bioenergy and bioproducts industry: The technical feasibility of a billion-ton annual supply. U.S. Department of Energy, Oak Ridge National Laboratory, Oak Ridge, TN. 78p.
- Prestemon, J.P.;** Abt, R.C. 2002. Timber products supply and demand, p299-325, in: D. N. Wear and J. G. Greis (Eds.), Southern forest resource assessment, U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC.
- Robinson, A.L.; Rhodes, J.S.; Keith, D.W. 2003. Assessment of potential carbon dioxide reductions due to biomass-coal cofiring in the United States. *Environmental Science and Technology* 37:5081-5089.
- Sedjo, R.A.;** Sohngen, B. 2009. The Implications of Increased Use of Wood for Biofuel Production. Issue Brief # 09-04. Resources for the Future, Washington, D.C. 13p.
- U.S. Environmental Protection Agency.** 2008. eGRID2007 Version 1.1. Retrieved May 21, 2009, from <http://www.epa.gov/cleanenergy/energy-resources/egrid/index.html>.

Table 1 – Total co-fire demand (green U.S. tons) in each State and supply subregion.

Subregion	Co-fire Demand (gr. U.S. tons)	State	Co-fire Demand (gr. U.S. tons)
Florida	5,989,003	Alabama	9,203,234
Gulf Coast	20,277,996	Arkansas	2,826,722
Highland Rim	4,534,124	Florida	6,943,748
Mid-Atlantic	21,323,771	Georgia	9,708,566
Middle Valley	639,694	Louisiana	2,977,997
North Valley	3,050,383	Mississippi	2,298,985
South-Central Valley	3,373,239	North Carolina	8,622,239
		South Carolina	4,582,469
		Tennessee	7,110,574
		Virginia	4,913,675
Southeast Regional Total	59,188,213	Southeast Regional Total	59,188,213

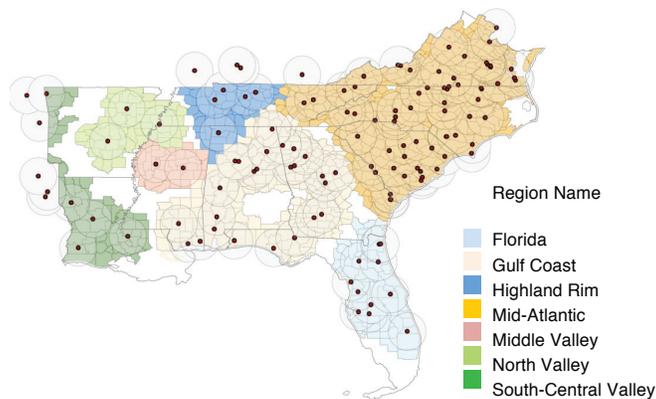


Figure 1 – Supply regions used for the assessment of spatially explicit resource impacts. Coal-burning facilities potentially affecting counties in the Southeast region are indicated; circles represent 50-mile supply radii used to highlight those counties potentially providing biomass to a given facility.

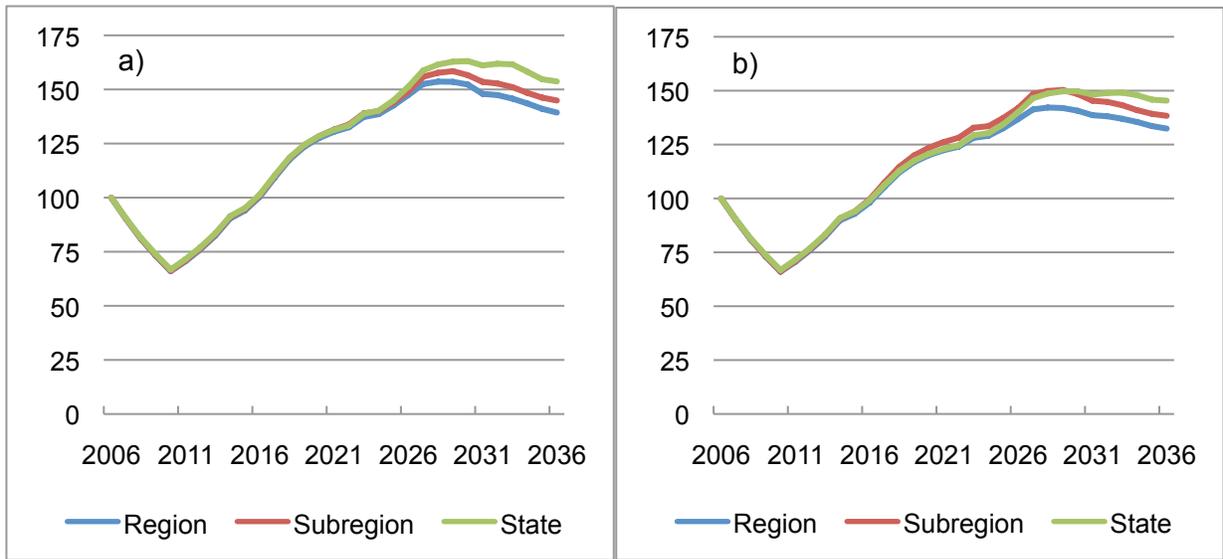


Figure 2—Weighted shift in pine pulpwood prices for a) 25 percent and b) 50 percent residue utilization scenarios. Values indicate magnitude of shift relative to 2006 pine pulpwood prices (100 = 2006 pine pulpwood prices).

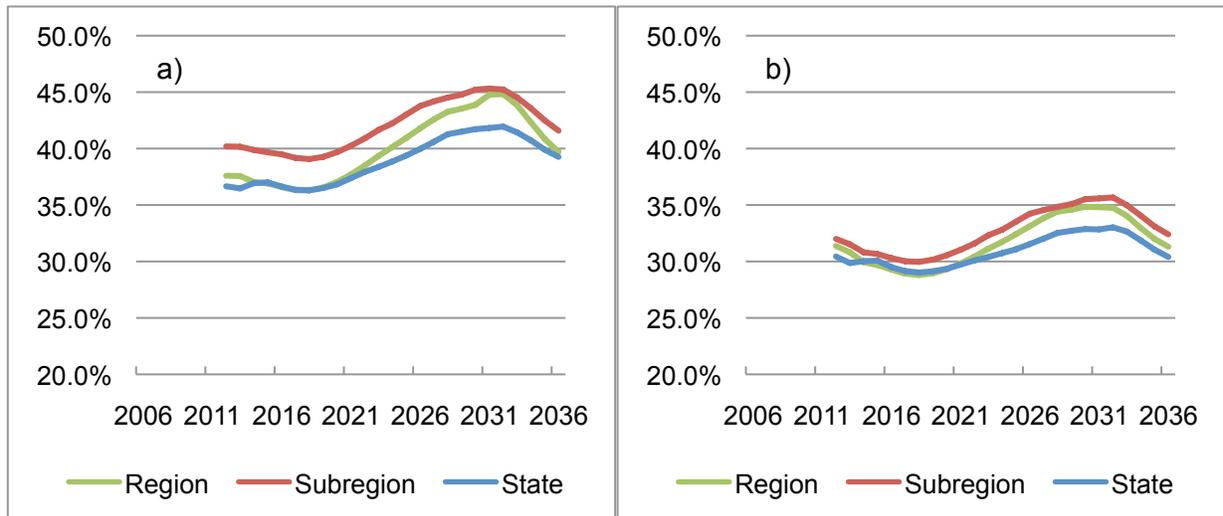


Figure 3—Percent of demand met from displacement of existing soft and hardwood pulpwood capacity, for a) 25 percent and b) 50 percent residue utilization.

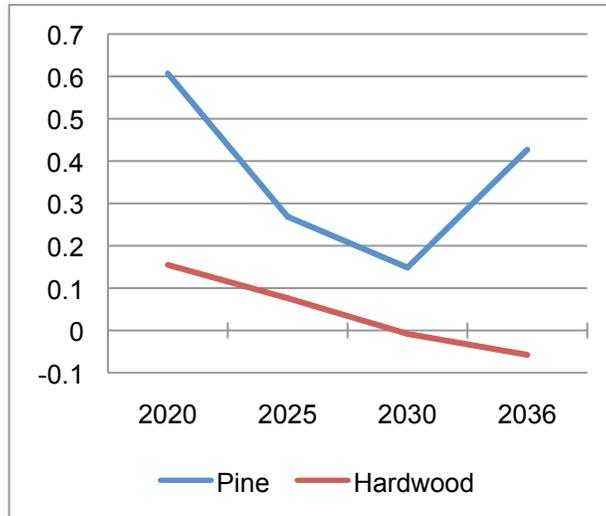


Figure 4—Spatial correlation between starting pulpwood harvest and increased pulpwood harvest due to biomass demand. The example presented here is for a 50 percent residue utilization scenario.

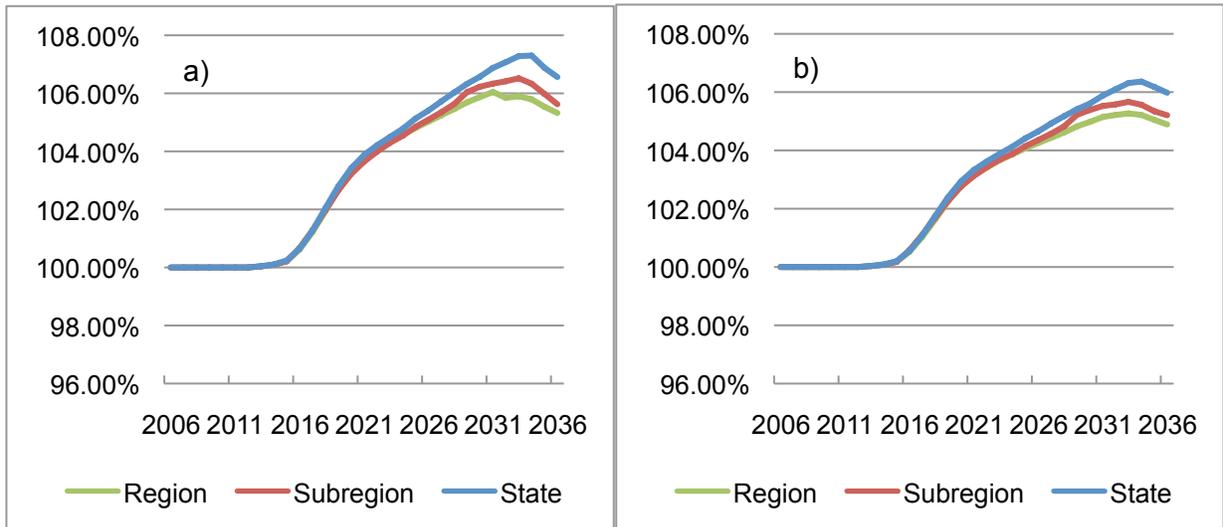


Figure 5—Percent acreage differential, co-fire scenario versus baseline, for a) 25 percent and b) 50 percent residue utilization scenarios. Values above 100 percent indicate an increase in forested acreage relative to baseline conditions.

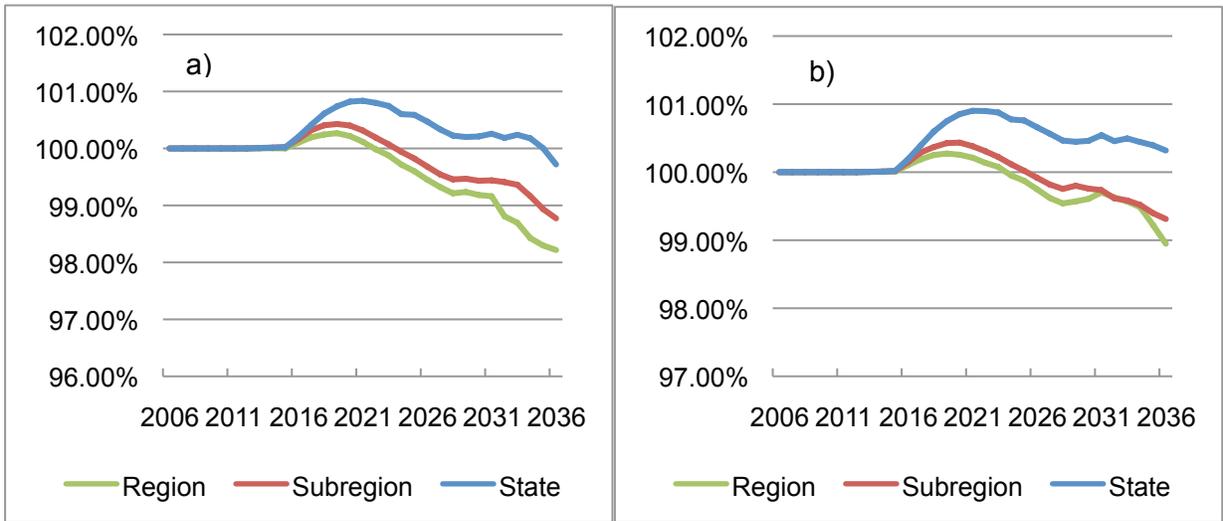


Figure 6—Percent carbon differential, co-fire scenario versus baseline, for a) 25 percent and b) 50 percent residue utilization scenarios. Values above 100 percent indicate an increase in forest carbon storage relative to baseline conditions.

LONG-TERM SIMULATIONS OF FOREST MANAGEMENT IMPACTS ON CARBON STORAGE FROM LOBLOLLY PINE PLANTATIONS IN THE SOUTHERN U.S.

Huei-Jin Wang, Philip J. Radtke, Stephen P. Prisley

ABSTRACT

Accounting for forest components in carbon accounting systems may be insufficient when substantial amounts of sequestered carbon are harvested and converted to wood products in use and in landfill. The potential of forest offset – in-woods aboveground carbon storage, carbon stored in harvested wood, and energy offset by burning harvested wood – from loblolly pine plantations was evaluated for greenhouse gas (GHG) mitigation over a half-century period. The in-woods carbon in well-managed loblolly pine plantations across the South totaled 341 million metric tons. This is equivalent to 20 percent of total energy-consumed GHG emission in the United States in 2006. Present-day carbon storage in southern pine plantations averaged 30.54 Mg·ha⁻¹ (± 2.54 percent) for in-woods carbon. Annual wood production was 62.1 and 45.9 million green metric tons from pulpwood and sawtimber yield, respectively, with roughly one-fourth of the green weight being carbon. The carbon storage in wood products increased steadily over the half-century projection and showed no sign of leveling off, while the storage in plantations was found to remain constant or increase slightly over time. An additional 11 million metric tons of harvested carbon was used for energy per year on average, equivalent to 25 percent of annual forest-products-industry renewable energy use in U.S.A. Intensified application of fertilizers and herbicide and genetic improvement showed the potential to increase total storage in in-wood and harvested carbon pools as much as 30 percent, and energy offset up to 40 percent. Reducing management intensity greatly increased in-woods carbon storage potential, but eliminated the wood-products carbon sink.

INTRODUCTION

Forest ecosystems in the United States sequester 140-300 million metric tons (Mg) of carbon per year, or between 18 percent and 39 percent of the equivalent CO₂ emissions from the Nation's coal-fired power plants (Pacala and others, 2001; Heath and Smith, 2004; U.S. Environmental Protection Agency, 2007b). Despite scientists' knowledge that U.S. forests are an important terrestrial carbon sink, challenges remain in estimating the magnitudes of carbon storage attributed to forests in different geographic regions and in quantifying the magnitudes of fluxes for various forest carbon pools (Houghton and others, 1999; Schimel

and others, 2000; Pacala and others, 2001; Janssens and others, 2003). One challenge involves incorporating uncertainty into estimates, so that decision-makers can plan in accordance with the quality of information in-hand (Gong, 1998; McKenney and others, 2004). Another challenge is to account for carbon sequestered in wood removed from forests as wood and paper products that may persist for long periods of time (Skog and Nicholson, 1998; Perez-Garcia and others, 2005). Such information is generally not a standard component in forest carbon estimates (Heath and others, 2003); however, both concerns are essential in making decision or plans for managed forest ecosystems, including the loblolly pine (*Pinus taeda* L.) plantations extensive throughout the southern United States.

Carbon stored above ground in loblolly pine plantations includes both merchantable and non-merchantable trees and vegetation, along with dead wood and plant detritus (Smith and others, 2004a). Regarding "long-lived" aboveground carbon pools, i.e. those in which carbon remains sequestered for decades or more, separate accounting is often made for live trees and coarse woody debris (CWD) based on the differing biological and ecological processes acting on each. Live trees sequester carbon on temporal scales of several decades, corresponding to rotation lengths. Carbon in CWD may persist in forests for years to decades depending on the relative rates of accumulation and decomposition (Duvall and Grigal, 1999; Vanderwel and others, 2008; Radtke and others, 2009). While aboveground carbon stored in live trees can be reliably assessed and projected over time and space, accumulations of CWD are considerably variable across landscapes and depend significantly on disturbance and management (Duvall and Grigal, 1999; Fridman and Walheim, 2000; Campbell and others, 2008).

In evaluating the potential of managed forest ecosystems such as loblolly pine plantations in mitigating atmospheric

Huei-Jin Wang, Graduate Research Assistant, Virginia Tech, Department of Forest Resources and Environmental Conservation, Blacksburg, VA 24061, U.S.A.
Philip J. Radtke, Associate Professor
Stephen P. Prisley, Associate Professor

GHG accumulations from the burning of fossil fuels, accounting for carbon stored in live trees and CWD is insufficient because substantial amounts of sequestered carbon are harvested and converted to end-use wood products, e.g. building materials, furniture, and paper products, or used as a fuel source to displace GHG emission from fossil fuels (Birdsey and Heath, 1995; Smith and others, 2006). Although harvested wood is not a part of in-woods carbon pools, the linkages between management activities, forest carbon sequestration, and the timing and amount of wood harvested are inextricable. Wood products may persist longer than plantation rotation lengths, and the amount of carbon remaining in wood products – products in use and landfills – contributes significantly to carbon sequestration over time (Skog and Nicholson, 1998). Moreover, the magnitudes and rates of carbon remaining in wood products depend on the timing, intensity, and extent of harvesting activities, which affects what products the harvested wood is allocated to and life spans of wood in these products. On the other hand, wood processing at mills, e.g. drying, peeling, slicing, and sawing, uses energy from burning wood residues and pulping liquors that reduces some need for using fossil fuels. Such energy sources currently supply 1.5 percent of the total energy consumption in the U.S.A. (Perlack and others, 2005). Compared to the combustion of fossil fuels, bioenergy from harvested wood is relatively carbon-neutral and can be renewable (Schiermeier and others, 2008). Reliable accounts of long-term carbon mitigation potential from these managed ecosystems should not fail to take harvested carbon into account (Smith and others, 2006). As demand for wood products grows, so too will plantation management intensity. Both factors will likely impact the amount of atmospheric carbon sequestered in southern U.S. forests and the wood products derived from them. Effective policy-making, planning, and management will require good information to ensure that these factors are accurately accounted for in optimizing carbon sequestration that can be supported by southern U.S. forests (Wear and Greis, 2002).

Plantation management in the U.S. South is expected to increase in intensity in order to provide more raw materials to meet rising societal demands for wood resources (Prestemon and Abt, 2002). Loblolly pine plantations comprise 9.7 million hectares of southern U.S. timberland, roughly 65 percent of the southern plantation area, and their area is projected to increase by 67 percent in the next thirty years (Prestemon and Abt, 2002; Wear and Greis, 2002; Smith and others, 2004c). Through woody and herbaceous vegetation control and fertilization, site characteristics are being actively managed to enhance productivity (Allen, 2001). Intensive site preparation, including bedding, disking, subsoiling, ripping, or combinations of these treatments, can efficiently reduce competition from non-commercial hardwood species (Morris and Lowery, 1988). In addition,

herbicide application can improve seedling establishment and early growth (Nilsson and Allen, 2003). Fertilization has become an important silvicultural tool in treating nutrient-deficient midrotation stands for increasing volume growth (Fox and others, 2007). Planting genetically-improved growing stock has become a standard management tool to increase growth efficiency, with gains in volume growth averaging 10 to 30 percent over unimproved planting stock at harvest (Li and others, 1999; McKeand and others, 2003). Tree breeding and other efforts to improve genetic properties of plantation growing stock are increasingly producing commercially available families and genotypes for increased volume production in loblolly pine (McKeand and others, 2003; Allen and others, 2005; McKeand and others, 2006). Intensive management operations appear to have potential for sequestering greater carbon, and projections of management scenarios will provide an insight on dynamics of in-woods and products-based carbon pools.

Recently, national-scale inventory-based carbon assessments have been augmented to account for carbon stored in aboveground forest pools, as well as the carbon stored in wood products (Skog and Nicholson, 1998; Heath and others, 2003; Jenkins and others, 2003; Smith and others, 2003). To date, such assessments have not directly considered the resolution, intensity, nor timing of management activities prescribed at forest stand scales. Because management is typically carried out on the scale of forest stands, carbon accounting at the same scale will allow for tracking of the full range of management and harvesting activities (Harmon, 2001). In addition, stand-level accounting can be scaled up with increasing certainty, while downscaling of national-scale estimates generally leads to greater uncertainty (Freese, 1967; Smith and others, 2004a). Here, predictions will be made at the resolution of individual forest stands for greatest flexibility in prescribing management conditions. Results will be aggregated to state and regional scales to make broader geographic assessments, presumably with a relatively high degree of precision (Smith and others, 2004a). The resulting analyses should serve the information needs of individuals ranging from those who develop policies for climate change mitigation, to those who set long-term regional goals for carbon sequestration, to those who aim to increase the total carbon stored in the wood grown on and products harvested from their forest lands.

The goal of this research was to assess impacts of forest management on carbon storage in loblolly pine plantations across the southern United States over the next half-century. Of specific interest here are the in-wood carbon pools of aboveground live tree and CWD, and pools of carbon in wood products produced from southern forests. To preserve information related to stand-level management activities, extensive field-plot inventory data were coupled with stand-

level prediction models to reduce uncertainty in estimates and facilitate aggregation across different spatial and temporal scales. Four specific objectives were pursued as a part of the overall goal:

Objective 1—Estimate the amount of carbon stored aboveground in live trees and CWD at scales ranging from individual stands to the entire southern United States.

Objective 2—Predict the annual production of harvested wood under operational management over a 50-year span, distinguishing between wood harvested for use in solid wood and paper products, and accounting for trends related to management intensity;

Objective 3—Project in-woods carbon pools and carbon disposition in harvested wood over a 50-year time span, linking inventory-based data and management activities to existing models of growth and yield and accounting for the lifespan of wood products;

Objective 4—Evaluate long-term effects of intensive management of loblolly pine in the U.S. South, including competing vegetation control, fertilization, and planting of genetically improved growing stock, on carbon sequestration and storage.

MATERIALS AND METHODS

DATA

The primary data source used in addressing the study objectives is the database of forest inventory records available online from the USDA Forest Service Forest Inventory and Analysis (FIA) program (Forest Inventory and Analysis, 2009a). The FIA data used here are composed of two-phase sample data collected using double-sampling for stratification (Smith, 2002; Reams and others, 2005). Phase I data begin with the interpretation and classification of remote-sensing imagery. Strata weights are estimated for each remote-sensing class, and areas of interest, such as the areal extent of loblolly plantations, can be estimated by aggregation based on strata weights. Phase II field plots are established on subsets of Phase I strata to provide field observations of forest conditions and conventional timber-based measurements on trees larger than 2.54 cm diameter at breast height (DBH). The spatial sampling intensity of FIA field plots is one plot per 2,430 hectares, and each field plot comprises a cluster of four 7-m fixed-radius subplots, occupying a 0.067-ha area (Bechtold and Scott, 2005). Within each subplot is nested a 2-m radius microplot where detailed measurements of small trees (< 2.54 cm DBH) are made.

Phase II inventory data obtained, from 2005 – 2007 survey data for loblolly pine plantations of 11 southern states (Figure 1, Table 1), were used as the source of information for stand information, including plot datasets, plot-condition datasets, tree datasets, seedling datasets, and site-tree datasets (Forest Inventory and Analysis, 2009a). Plot datasets bridged Phase I data and plot-condition datasets to estimate forestland areas represented by each plot given its growing condition. Plot datasets provided plot geographic coordinates, remeasurement period (yr), a unique plot identification code and previous plot conditions if any remeasurement occurred. Field observations from plot condition datasets included plot conditional classes, condition status codes, condition proportions, subplot proportions, stand origin codes (natural stands or plantations), stand origin species, stand ages, treatment codes, year of treatment, and year of inventory. Conventional timber-based variables from tree datasets measured in subplots included tree status codes (live or removed), species, DBH, height, and live/removed cubic-foot volumes. Site-tree data included site index relevant measurements, i.e. height and age of dominant or codominant sample trees. Seedling data measured in microplots provided information on planting density.

The FIA data were screened to identify conditions consistent with “well-managed” loblolly pine plantations such as those used in the development of the FASTLOB growth-and-yield model developed by the Forest Modeling Research Cooperative at Virginia Tech (Amateis and Burkhart, 2009). Only those plantations having ≤ 20 percent of the stand basal area comprised of hardwood species and those having ages between 0 and 50 years were defined as “well-managed” and subsequently included in the analyses. These conditions were consistent with the data used to develop FASTLOB and its computer implementation (Ralph Amateis, personal communication, March 1, 2010). Among 12.4 million hectares of planted loblolly pine forest, a set of 5,480 FIA inventory plots matched the screening conditions and the total area was 11.2 million hectares, including 3,139 plots on which the screened condition was observed on the entire plot, and 2,341 on which the screened condition was observed on a portion of the plot.

STAND-LEVEL GROWTH-AND-YIELD MODEL

The FASTLOB model was developed to reflect management activities common to loblolly pine plantations established from the late 1950s to early 1990s (Amateis and Burkhart, 2009). FASTLOB uses site index (base age 25 years), age, stem density, amount of competing vegetation, thinning operations, fertilization, and other stand characteristics to project merchantable yields (pulpwood and sawtimber) and in-woods biomass by component, including stem and bark,

branches and bark, foliage, and CWD, at different ages. Not only projections but also predictive values for initial growing stock can be obtained while inputs are established. Stand-level equations that comprise the nucleus of FASTLOB project dominant height, survival and basal area, and serve as a baseline thinned and unthinned model for stands. In addition, model inputs including information of latitude and longitude provide more precise locale-specific predictions if data are available. FASTLOB is presently used in ongoing forest management across the private sector of the South.

QUANTIFY CURRENT FOREST CARBON POOLS

Coupled with FIA stand attributes, FASTLOB was used to initialize current stand-level forest carbon pools, but an indication of how close the estimate from FIA is to the population parameter was not readily available through applying FIA area expansion factors to scaling up stand-level estimates to state and southwide levels (Scott and others, 2005). “Forest carbon pools” in this study refer to the carbon content (one-half the mass of oven-dry biomass) in aboveground live trees and CWD, including standing snags and downed-woody material. Variances of in-woods carbon estimates were used to characterize the uncertainty of current forest carbon pools.

Bootstrap variance estimation and its corresponding Monte Carlo approximation were used to compute the estimate of in-woods carbon mass (live trees and CWD) at various regional scales (Booth and Sarkar, 1998). Because the probability density function of the population distribution was unknown, a nonparametric approach was applied to assess various regional-level carbon quantities. In the application of bootstrap sampling, predictive values of current in-woods carbon mass from FASTLOB initialization, weighted with representative areas for each FIA plot, were treated as a substitute for the population of in-woods carbon. Then, from these 5,480 observations (the number of FIA plots in the dataset), bootstrap samples of size 5,480 were selected with replacement from the FIA dataset. An estimate of in-woods carbon was obtained from each bootstrap sample at state and southwide levels. Two thousand bootstrap samples from the data were generated in total (Booth and Sarkar, 1998). Standard errors and the 2.5th and 97.5th percentiles of the confidence interval for the in-woods carbon were then approximated from the bootstrap sample distributions.

ASSUMPTIONS OF BASELINE MANAGEMENT

Management conditions considered here included the area and density of planting, timing and intensity of thinnings, ages to harvest (rotation ages), and silvicultural activities associated with high-intensity management. Final (clearcut) harvests are simply referred to as “harvest” in this study, in contrast to wood harvested by thinning, which is

referred to simply as “thinning.” Maximum-likelihood was used in analyzing FIA data to estimate parameters for management-related inputs including planting densities, levels of residual stems per unit area, and ages for thinning. Log-normal distributions were fitted to planting density and residual stem density. A gamma distribution was fitted to approximate the distribution of thinning ages for subsequent simulations. Empirical cumulative distribution functions (ECDFs) and Quantile-Quantile (Q-Q) plots were used to evaluate quality of fit for empirical frequencies with those fitted to density functions.

Distribution functions were fitted to 2005 – 2007 measured plot attributes from FIA to simulate inputs for simulations to be consistent with real-world conditions of planting density, timing, and intensity of thinning (Figure 2, Figure 3, Figure 4). Mean and median planting densities of 1,473 and 1,349 trees•ha⁻¹, respectively, coincided with planting spacings typical of southern U.S. pine plantations and a lognormal distribution function fitted to FIA data (Figure 2). No relationship existed between age of thinning and site index. Therefore, age of thinning from FIA records was fitted to a gamma distribution function (Figure 3). Post-thinning residual densities were simulated by a lognormal distribution (Figure 4). All three of these distribution functions represented the general shape and scale of the FIA data for planting density, thinning age and residual density, although some lack-of-fit was noted, especially in the upper tails of these right-skewed distributions.

Rotation length, the plantation age at final harvest, was needed to schedule operations on individual stands; however, rotation length was only directly observed on a small number ($n = 22$) of the FIA phase II field plots – namely those that had been visited at two different times and were harvested between visits. In these data an inverse relationship between site index and rotation length was noted (Figure 5A). Their mean rotation length was 27.5 years ($s = 6.1$), over plantations that averaged 18.50 m in site index ($s = 2.18$). Although the relationship between rotation length and site index was relatively weak, a trend describing it (Figure 5A) was used to predict rotation length for the full set of FIA phase II plots where rotation lengths had not been observed. Predicted rotation lengths by plantation area averaged 27.5 years using this approach, with 80 percent of plantation area having rotation lengths between 23 and 32 years (Figure 5B). Dividing the total area of plantations by the mean predicted rotation length indicated an annual harvest area over time of 406,000 ha, which was roughly consistent with published report of 524,000 ha planted in loblolly and shortleaf pines in the southern U.S. in 1998 – including those subjected to all levels of management intensity (Moulton and Hernandez, 2000; Smith and others, 2004c).

SIMULATION OF SILVICULTURAL OPERATIONS

Loblolly pine plantations were assumed to be managed primarily for timber benefits over the 50-year simulation. With regard to management objectives, plantations were categorized into two populations throughout the commercial range of species, those that would be thinned at some point during a rotation, and those that would remain unthinned up to the point of their final harvest. An area of 288,623 ha was set as the target for the area of thinnings to be simulated each year, based on the estimated annual area of thinning in FIA plantation area. The same area was targeted for final-harvest operations in previously-thinned stands each year so that the area of thinned plantations would remain constant over time. The area to be harvested annually from never-thinned stands was set at 117,377 ha as an initial target value, so that the area harvested from thinned and never-thinned plantations would target a total of 406,000 ha per year, as was determined in the previous section.

In simulation of area harvested annually from either thinned or unthinned stands, it was necessary to assign the annual area harvested to various plantation ages. Much as growth and yield share an inherent relationship the plantation area harvested in various age classes over time has a cumulative effect on the age distribution of plantation growing stock (Clutter and others, 1983). To reflect this relationship, the mathematical derivative of plantation area with respect to age across the South was used in assigning an age distribution to the area annually harvested. To implement this method, plantation area was first expressed as a function of stand age to match the empirical conditions characterized from the FIA database.

Graphs of plantation area by age for thinned and unthinned stands showed distinct trends of declining area beginning around age 22 for thinned stands, and age 16 for those that were never thinned (Figure 6). These values were used to establish the minimum ages for final harvesting, i.e. the minimum rotation lengths, in thinned and unthinned plantations, respectively. Then the first derivatives of area with respect to age were calculated to represent suitable functions of harvest area (i.e. change in plantation area) by plantation ages. These first derivatives of area harvested from thinned and unthinned stands were defined by functions, Eq. [1] and Eq. [2], respectively:

$$y = c_1 \times (1307x^2 - 133100x^3 + 3606000x^4 - 28876000x^5) \times \exp(15.83 - 1307x^{-1} + 66550x^{-2} - 1202000x^{-3} + 7219000x^{-4}) \quad [1]$$

$$y = c_2 \times 0.13 \times \exp(15.16 - 0.13x) \quad [2]$$

where

x = stand age (yrs)

y_x = total area harvested at age x

c_1 and c_2 are refined factors through simulations

To focus on changes in plantation area that were due to removals by harvesting, only the declining portions of the age class by area distributions were considered (Figure 6). Thus, in accord with the FIA data it was assumed that final harvesting for thinned stands took place no sooner than 22 years after planting in loblolly pine plantations, and no sooner than 16 years for unthinned stands.

All thinnings were simulated based on a thinning intensity of 20 percent removals by row thinning and an additional ≥ 5 percent reduction in stem density removed by thinning from below. Following thinning, a minimum of 6 years was required in any particular stand before final harvest was allowed in order to capture the volume growth response to the thinning treatment. Timings and total area of plantation thinnings were specified by the gamma-model-specified distribution of stand ages at thinning, along with the target for total area to be thinned each year across the South. End-of-rotation harvest timing and area also targeted an age-distribution and total area. A time period for harvesting, site preparation and subsequent planting was assumed to be one year; therefore, artificial regeneration was simulated to follow an end-of-rotation harvest with a one-year fallow period.

SIMULATION ANNUAL OPERATIONS

Forest management regimes span decades for a rotation, and individual stands experience all stages of the forest management cycle including final harvest, site preparation, regeneration, and thinning. Concerning stable production of timber harvests from year to year, total southern loblolly pine plantations were treated as a single entity and management activities were manipulated through coordinating all stands. Final harvests were assumed to be operated on 406,000 ha annually, i.e. 288,623 ha from previously-thinned stands and 117,377 ha from never-thinned stands. Regarding changes in plantation area on rotation ages, Eq. [1], Eq. [2], and rotation ages modified from FIA data (input rotation ages) were programmed into simulations of area harvested annually. Intermediate simulation results were used to refine the two constants c_1 and c_2 in Eqs. [1] and [2], respectively, along with the specified target area for annual harvesting in unthinned plantations. The sequence of steps performed in the simulation algorithm follows (Figure 7): (1) if stand age is equal to its predicted rotation age or greater, then this stand becomes one candidate to be harvested; (2) with regard to the size of candidates' representative area, candidates with large areas have top priority to be harvested; (3) select candidates from the pool of candidate stands to meet requests from each age-class area of Eq. [1] or Eq. [2]; (4) if total area from Step 3 meets the target harvest area, then stop; (5) otherwise, more candidate stands harvested are needed. In this step, number of overdue years of predicted rotation age is used instead as the criterion for choice of

candidate stands to be harvested. Select candidates from more overdue years to meet target harvest area; (6) means of simulated rotation ages and areas harvested by year are evaluated whether simulation underperforms or not; (7) if underperformance occurs, refine c_1 in Eq. [1] or c_2 in Eq. [2]; (8) re-run steps 1-7 for the next 50-year-simulation iteration until simulation output is in good shape. After final harvest and a one-year fallow period for site preparation, all stands were established and their planting densities followed the lognormal-model-specified distribution.

Assignment of stands to be treated by thinning, or remain unthinned during the lengths of their rotations was made using a Bernoulli distribution with values 1 for thinning, and 0 for no thinning. The probability that a stand would be thinned (p) was defined by the total area harvested from thinned stands divided by total harvest area across the South from the previous year in the simulation. As previously noted, the area to be treated by thinning annually was set at 288,623 ha and its age structure was defined by Eq.[1]. This simulation required four component inputs including target area, Eq. [1], input rotation ages, and the gamma-model-specified distribution of stand ages at thinning. The steps of the algorithm procedure follow (Figure 8): (1) if stand age is equal to its gamma-specified age or greater, then this stand becomes one candidate to be thinned; (2) candidates have top priority to be thinned if their representative area are large; (3) select candidate stands to meet demands of future harvest areas from Eq. [1] coupled with predicted rotation ages; (4) if total area from Step 3 meets target thinned area, then stop; (5) otherwise, select more candidate stands to meet the target area. Number of overdue years of thinned age serves as the criterion for choice of candidate stands to be thinned. From large overdue years, select candidate stands to meet target thinned area.

HARVESTED WOOD PRODUCTION OVER TIME

Projections of future production of timber products (i.e. pulpwood and sawtimber) were made under the baseline management scenario described above, which was determined from FIA data. Simulated variables including areas harvested either from thinned or unthinned stands, thinned areas, rotation ages, and ages for thinning were linked to FASTLOB to generate timber products estimates. Pulpwood was defined as 15.24 cm (6 in) DBH and larger and minimum diameter top was 10.16 cm (4 in) outside bark; and sawtimber was defined as 22.86 cm (9 in) DBH and larger to a minimum 17.78 cm (7 in) top diameter outside bark using the International 1/4-inch log rule. Green weights outside bark for both types of timber products were predicted for comparison to regional analyses that express production on the basis of weight (Bullock and Burkhart, 2003). For validation purposes, primary-mill survey results from 2006-2008 were obtained from FIA timber product output (TPO) reports of pulpwood and sawtimber

production from roundwood (e.g. Cooper and Becker, 2009; Johnson and others, 2010).

To assess the potential role of wood products in mitigating GHG emission from fossil fuel, i.e. carbon pools and energy offset, the method for calculating harvested carbon by Smith and others, (2006) was used. The amount of carbon in wood products each year was estimated, including products in use and products in landfill, through 2056, beginning with wood harvested in 2006. Carbon remained in harvested wood products was expressed as metric tons per hectare ($\text{Mg}\cdot\text{ha}^{-1}$) even though the disposition of carbon over time for such wood products are not directly linked to forest area. With regard to renewable energy consumption from wood residues and pulping liquors generated by the forest products industry, the amount of emitted carbon by year was estimated. Year-to-year changes in stocks of carbon sequestered in the wood-products pool was estimated to evaluate whether this pool is a carbon sink, balance, or source.

The carbon content in harvested wood was estimated using green weight of pulpwood and sawtimber production from FASTLOB output and moisture content (MC) of sapwood 110 percent (Glass and Zelinka, 2010). Disposition of carbon in harvested wood products for products in use, products in landfill, and energy offset was estimated as follows: (1) *Ovendry weight* = *Green weight*/(*MC*+1); 50 percent of this is carbon mass; (2) allocate sawtimber and pulpwood to primary wood products (e.g. lumber, plywood, panels, and paper) according to region and category in Table D6 of Smith and others, (2006); (3) compute carbon amount of primary products remaining in use or in landfill each year based on Tables 8 and 9 of Smith and others, (2006), respectively; (4) estimate amount of carbon associated with energy recapture using Table D7 of Smith and others, (2006) (See Smith and others, (2006) for details).

IN-WOODS CARBON OVER TIME

To evaluate long-term effects of baseline management on sequestering carbon and maintaining in-woods carbon, FASTLOB was used to project biomass of aboveground live trees and mass of CWD in a 50-year timeframe since 2006. Rate of change of sequestering carbon was computed to assess whether the managed forest was a carbon-balanced system or not. FASTLOB has embedded prediction equations that estimate biomass for various components (Baldwin and others, 1997; Landsberg and Waring, 1997; Radtke and others, 2009). Carbon mass was assumed to be 50 percent of biomass (Smith and others, 2003).

INTENSIVE MANAGEMENT SCENARIOS

With regard to an increase in management intensity in the southern plantations, two management intensity scenarios were developed to estimate potential loblolly pine growth

and yield and the corresponding effects of management on carbon storage. The two management scenarios included (1) scenario 1: intensive site preparation, herbicide application, and mid-rotation fertilization; and (2) scenario 2: the management regime from scenario 1 plus planting of genetically improved growing stock. The term “genetically improved” here assumes that growing stock came from third- or fourth-generation seed orchards which have not previously been deployed in the South (McKeand and others, 2003).

The intensive management regimes 1 and 2 were used according to the embedded functionality of the FASTLOB modeling system. For completeness, an overview of the FASTLOB implementation for intensive management is given here. Growth responses to intensive silviculture in FASTLOB are added to baseline-management predictions. According to research that showed growth responses to intensive site preparation and herbicide application varying from site to site, the effect of competing vegetation control on growth and yield was modeled in FASTLOB by increasing site index by 0 to 1.5 m (Siry and others, 2001; Nilsson and Allen, 2003). A uniform distribution was used to simulate random site index increases within this range for each stand. In accord with common mid-rotation fertilizer applications of 28 P kg•ha⁻¹ and either 224 or 196 N kg•ha⁻¹, the amount of N fertilizer applied in a given stand was set to follow a Bernoulli distribution with p (224 N kg•ha⁻¹) = 0.58, and $1 - p$ (196 N kg•ha⁻¹) = 0.42 (Albaugh and others, 2007). For unthinned stands, the timing of fertilization was assumed to take place between ages 13 to 20 and no harvesting within six years of fertilizing; for thinned stands fertilization was performed after thinning. Timing assumptions for fertilization were primarily based on published studies varying management intensities that Siry and others, (2001) assumed fertilization at age 15 years for medium intensity and 5-to-10 years for high intensity; Allen and others, (2005) assumed age 17 years for medium intensity and 5-to-21 years for high intensity; Liechty and Fristoe (2010) used ages 17-to-22 years for timing of mid-rotation fertilization. Genetically improved stock was assumed to increase volume by 10 to 20 percent at harvest ages and this increase corresponded to a 1.5- to 3-m site index gain (McKeand and others, 2006). Site index gains due to planting of genetically improved seedlings were simulated by generating a uniform random variate on the interval [1.5, 3.0] for each stand.

RESULTS

ESTIMATES FOR CURRENT CARBON POOLS

In well-managed loblolly pine forestland across the South, the estimate of in-woods carbon mass total exceeded 340

million Mg (1 Mg = 1 metric ton or approximately 1.1 U.S. tons) (Table 2). The mean of area-weighted averaged carbon was 30.54 Mg•ha⁻¹. State-by-State in-woods carbon totals varied from 3.3 to 53.7 million Mg, and 21.30 to 35.51 Mg•ha⁻¹ for carbon means per hectare by accounting for forestland area (Table 2). Carbon total stocks in Tennessee and Florida were significantly less than those in the other nine States, largely due to their comparatively small plantation areas. Aside from the effects due to its small plantation areas, Tennessee had relatively low carbon stocks of 21.30 Mg•ha⁻¹, in part because of its comparatively low average basal area (Table 1). In general, States with the lowest average plantation ages had the lowest yields per hectare, while those with the highest plantation ages had higher yields (Table 1, Table 2). The percentages of aboveground live trees and CWD, contributing to the in-woods aboveground carbon pool, were about 93 percent and 7 percent, respectively (Table 3, Table 4).

BOOTSTRAP RESULTS

Sampling distributions for in-wood carbon quantities (i.e. carbon total and carbon per hectare) in loblolly pine plantations across the South appeared to be consistent with a normal distribution, with the bootstrap-simulated means being approximately equal to estimates from FASTLOB (Figure 9, Figure 10). The simulated results for standard errors and the 2.5th and 97.5th percentiles of distributions were given in Table 2, Table 3, and Table 4. Bootstrap confidence intervals for southwide carbon spanned ±2.80 percent for total carbon mass and ±2.54 percent for carbon per hectare (Mg•ha⁻¹) in the in-woods pool, respectively. Variances in live-tree carbon were ±2.77 percent and ±2.44 percent for carbon total and per hectare, respectively, and those of CWD carbon quantities were ±7.38 percent and ±7.30 percent.

State-level uncertainties for estimates of carbon quantities were assessed using the same set of bootstrap samples (Table 2, Table 3, Table 4). Compared to southwide estimates, State-by-State estimates were relatively imprecise. Uncertainty in the in-woods estimates was primarily contributed by variability from live-tree pools. Despite the larger dispersion of CWD pools across States, because of their smaller size, CWD pools contributed less to overall in-woods variability. Tennessee and Florida had greater variance of carbon estimates, mainly because of the relatively small numbers of FIA field plots in loblolly pine plantations in those States. Therefore, their standard errors of estimated totals and means for in-woods carbon were relatively large compared to other States' estimates. Excepting Tennessee and Florida, 95 percent bootstrap confidence intervals for States' carbon means in well-managed plantation forestland did not exceed ±15 percent of the estimated values.

AGE-CLASS SIMULATIONS OF AREA

The target area for annual harvesting from unthinned stands was set to 112,583 ha following test simulations used to determine whether this value was consistent with the constant c_2 in Eq. [2]. Hypothetical distributions of harvest area by age classes (Figure 11), were multiplied by constants $c_1 = -2.48$ in Eq. [1] and $c_2 = -0.65$ in Eq. [2], which ensured consistency between target harvest areas, Eqs. [1] and [2], and the predicted distribution of rotation ages. The derivative functions or harvest area by age classes reflect the assumed restriction of final harvesting in thinned plantations to those ≥ 22 years of age and unthinned plantations ≥ 16 years. In addition, these hypothetical distributions, especially the harvest curve for thinned stands [1], agreed with the predicted distribution of rotation ages (Figure 5B, Figure 11).

In plotting the area of simulated thinning and final harvest operations in each year of the simulation (Figure 12A), two periods, each spanning about 10-years, reflected relatively low projected areas of thinning (2015 – 2025) and final harvest (2025 – 2035) activity. These periods corresponded to a decade of relatively low establishment of loblolly pine plantations across the South in the 1990s, which is reflected in the relatively low area of 5 to 15 year old plantations in the initial age-class distribution (Figure 13A). At the end of the 50-year simulation, the same pattern was not evident in the age-class structure of loblolly pine plantation area across the South (Figure 13B).

Simulated results of year-by-year areas operated by thinning and harvesting, and their corresponding mean ages for operations were plotted in Figure 12. The annual area of final harvest averaged 400,000 ha over the 50-year simulation, including 290,000 ha ($\pm 9,600$) harvested from thinned stands and 110,000 ha ($\pm 3,300$) from unthinned stands. Rotation lengths in the simulations ranged between 26 and 33 years. Accounting for the occurrence of projected thinnings, simulated rotation ages of thinned stands averaged about one year more than those of unthinned stands, at 28.2 and 27.4 years, respectively. The annual area of thinning operations averaged 280,147 ha with a standard deviation of 27,000 ha over 50 years, with an average age of thinning = 18.0 years ($s = 0.7$ yrs).

PROJECTED TIMBER PRODUCTION

An example of the effect the simulated thinning regime had on stand-level volume and biomass accretion over the 50-year projection period can be compared with that of a stand not subjected to thinning (Figure 14). In both thinned and unthinned simulated stands, all aboveground volume and CWD was set to zero prior to the artificial regeneration of the stands. As is typical of most models that project growth and yield after thinning in plantations, volume

was immediately reduced at the time of thinning, and then allowed to re-accumulate over time until final harvest. In the years immediately following thinning, standing volume growth rates exceeded the rates realized before thinning for a time; however, volume production at final harvest was lower in thinned stands than their unthinned counterparts. The period of no apparent volume or biomass that occurs between rotations is a minor artifact of the way volume accretion is estimated in FASTLOB. In particular, the youngest age at which any volume outputs are generated is five years after planting.

Timber production southwide for each year was computed as an aggregate of all stand-level projections. Results showed that through carrying out thinning operations, stands supplied one-fourth timber production annually including pulpwood and sawtimber, and final harvest three fourths, drawn from Figure 15. Further, thinnings primarily produced pulpwood; and final harvests produced pulpwood and sawtimber. Annual total pulpwood yield was 62.1 million green metric tons, ranging from about 49 to 76 million green metric tons, 38 percent from thinning and 62 percent from final harvest. However, total sawtimber production of 36 – 60 million green metric tons was almost 100 percent made up by final harvests. Mean projected annual pulpwood production was nearly equivalent to 2006 – 2008 TPO reported pulpwood production. For sawtimber the projected mean was about 35 percent lower than the TPO value (Figure 15).

CARBON POOLS AND FLUXES

Figure 16. showed the effects of annual thinning and final harvesting activities on reductions of carbon from the in-woods pool. Intra-annual increases in the trend represented net growth through the growing season, while intra-annual decreases represented removals. Timing of removals was arbitrarily set to follow the annual growth each year, without detailed consideration of the timing of growth and removals within any given year. Considering both additions and losses of carbon in the wood-products pool, which includes products in use and in landfills, harvested wood products created a sink of 6 to 9 million metric tons of carbon per year (Figure 17). Compared to the landfill pool, fluxes of sequestered carbon in the products-in-use fluctuated more from year to year, especially in pulpwood products because of their relatively short lifetimes. For a long run, landfills added more carbon in the accounting system with reference to annual positive carbon fluxes.

EFFECTS OF VARIOUS MANAGEMENT INTENSITIES

Regarding increasing demands of wood products, intensive management might provide opportunities for GHG mitigation. With the intensive approaches, the amount

of carbon stored in all individual pools was substantially increased. Overall, applying fertilizers and herbicide, and deploying genetically improved growing stock increased 15 percent of carbon stocks, respectively (Figure 18). However, the increased magnitudes varied among pools. The more-intensive scenario (scenario 1) produced carbon gains 20 percent in sawtimber-in-use, and 10 percent in pulpwood-in-use, landfill, and in-woods pools, respectively. For the most-intensive scenario (scenario 2), sawtimber-in-use had a 40 percent increase; pulpwood-in-use and landfill had a 25 percent increase, respectively; and in-woods had a 35 percent increase in carbon stocks by comparing to the baseline scenario. For both intensive-management scenarios, carbon stocks in sawtimber-in-use grew much faster than the other pools, primarily due to gains in volume growth that increased long-lifetime sawtimber production (Figure 19).

Beginning with applying more intensive silvicultural approaches in 2006 and following each year, southwide-level timber yield responded to such applications with a time lag at least four years (Figure 19). Use of fertilizers and herbicide enabled substantial increase in pulpwood yields from 2013 and sawtimber yields from 2010. Genetic improvements increased pulpwood yields from 2021 and sawtimber yields from 2027. As expected, with increasing yield, annual energy recapture from wood products increased 20 percent and 40 percent for the more- and most-intensive scenarios, respectively, compared to the burning wood products of 11 million metric tons of carbon per year from the base scenario (Figure 20).

DISCUSSION AND CONCLUSIONS

Regional forest carbon storage in loblolly pine plantations was modeled as an aggregate of stand-level estimates based on FIA data and FASTLOB, which served as a baseline for assessing the potential of managed extensive forests to increase carbon storage. As of 2006, aboveground carbon pools held an estimated 341 million metric tons of carbon, an amount equivalent to 20 percent of GHG emissions from energy consumed in the United States in 2006 (U.S. Environmental Protection Agency, 2008). This estimate corresponded to an average of 31 Mg of carbon accumulated on each hectare of planted loblolly pine across the South. Sources other than planted loblolly pines are excluded from these estimates. Live trees comprised 93 percent of the projected aboveground carbon, with the remaining 7 percent stored in CWD. Smith and others, (2004b) reported that carbon content in aboveground woody pools ranged between 43 and 60 Mg•ha⁻¹ in southern loblolly-shortleaf pine forests. Their comparatively high estimates included some 45 percent natural forests, by area, compared to only plantations considered here (Forest Inventory and Analysis, 2009a). Presumably, the relatively low management

intensity in natural forests allows for greater accumulations of in-woods carbon than what is accomplished in well-managed plantations. Smith and others, (2004b) also reported that CWD comprised 12 percent of in-woods carbon, an amount higher than was found here. This difference can also be attributed to differences in management intensity between their study data set and the one used here. Compared to the 11.17 million hectares of “well-managed” loblolly pine forests studied here, the FIA loblolly-shortleaf forest type comprised 25.2 million hectares of forestland (Forest Inventory and Analysis, 2009a).

Uncertainties for baseline carbon assessments were approximated by a bootstrap procedure that showed error rates of 1.40 percent for total carbon across the South and 1.27 percent for carbon mass per hectare. The relatively small sampling error rates confirm that in-woods carbon estimates from FIA survey data can be highly precise (Figure 9, Figure 10). Smith and Heath (2001) reported an error rate of 6.5 percent for carbon mass stored in aboveground softwoods of maple-beech-birch forests for area of 10⁵-10⁷ ha, based on growing stock used by FIA (Smith and others, 2003; Smith and others, 2004a). Bootstrap error rates for loblolly pine plantation area estimates from the same FIA data used here (details not shown) verified that the FIA-mandated maximum sampling error rate of 1.91 percent for one million hectares of forestland was not exceeded (Forest Inventory and Analysis, 2009a). These results support the widely-held understanding of bootstrap sampling as a state of the art method for quantifies uncertainty in complex statistical analyses such as the regional carbon estimates generated here.

Rotation lengths varied between 26 and 33 years for stands projected over the course of the baseline simulation, based on the targets established by the weak relationship between site index and rotation length noted in FIA data, and also accounting for target harvest levels, thinning, and the modeled age-distributions of thinning and harvesting operations over the region. Rotation lengths were generally consistent with optimal ages to harvest based on financial returns or experts' insight that final harvests occur between ages 25 and 35 years (Siry, 2002; Huang and Kronrad, 2006; Carino, 2009). Year-to-year simulated averaged ages of thinning between 17 and 20 years agreed with pulpwood harvest ages in southern pine plantations from 2000 through 2010 (Fox and others, 2004). In addition, the dip in projected annual areas for thinnings and final harvests reflected past conditions. According to Conner and Hartsell (2002), industry ownership decreased throughout the South between 1989 and 1999, to the point where the removals of growing stock in 1999 exceeded the year's annual growth. Since then the area of southern pines planted has increased, in part because of conversion of some nonforested land

area to pine plantations (Conner and Hartsell, 2002). The projected trends here reflect both the decrease in growing stock prior to 1999 and the subsequent increase reported by Conner and Hartsell (2002).

FIA initial area conditions most strongly influenced projection results during the first 30-years of the 50-year simulation period. Beyond 30 years, the assumptions embedded into the simulation, notably those assumptions related to areas managed over time, exerted stronger influence on projection results. This can be seen in the dip observed in the FIA age-class distribution (Figure 12, Figure 15) that affects areas projected to be available for thinning and final harvest, along with timber production, particularly sawtimber yields, through 2035. After 2035 projected timber production and areas harvested or thinned became relatively stable over time, presumably the result of the repeated application of modeling assumptions that fail to replicate variations that would occur under real-world conditions. In addition, the simulation assumed the area of plantation forestry will remain constant across the South for 50 years and that age distributions of growing stock and harvested wood will remain stable over time. Trends in demographics, land uses, timber supply-and-demand relationships, and timber price are all known to affect timber resources, but were deemed to be outside the scope of this study (Adams and others, 2003). The simulation methods developed here could be improved upon by accounting for future dynamics of number of planted hectares, financial returns, and individual ownerships and their associated management objectives.

Projected sawtimber yields here were lower than reported TPO values by about 35 percent. In contrast pulpwood projections matched TPO reported values almost exactly. Sawtimber output in TPO reports are derived from the loblolly-shortleaf pine forest type, which includes natural and planted pine forests with all levels of management intensity. Management goals for such forests may be considerably different than what are defined here as “well-managed” loblolly pine plantations. For example, goals may include management for aesthetics, wildlife habitat, and recreational uses for a portion of the stand’s lifetime, with sawtimber harvesting taking place once economic returns become a motivating factor (Guldin, 2004). On the other hand the fact that pulpwood production results here strongly agree with TPO pulpwood production implies that loblolly pine plantations are a major source of softwood raw material for pulpwood production in the South. Challenges remain for comparing broad-scale market results such as TPO to management-oriented projections like the one conducted here.

Based on this 50-year-projection method, long-term effects of thinning and final harvest on future carbon stock in the

products-in-use and landfills can be extended through 100 years or more to address the climate change issue (Miner, 2006). Projected results showed that removals in the five decades total approximately 25.7 million metric tons of carbon per year, while maintaining the region’s plantation resources with a net carbon increase in growing stock over time; the harvested wood product preserves carbon with a positive flux of 6-9 million metric tons per year; an average of 11 million metric tons per year of carbon is burned for energy, equivalent to 25 percent of annual forest-products-industry renewable energy use in the United States (Perlack and others, 2005).

It has been argued that forests managed under natural conditions will store more carbon than those managed for timber production, even when carbon stored in products are accounted for (Harmon and others, 1990). For example, after a 50-year unmanaged period, all planted loblolly pine forests had quadratic mean breast height diameter of 15.5 cm, and averaged in-woods carbon mass of 115 Mg•ha⁻¹, varying from 18 to 251 Mg•ha⁻¹ (Figure 21A). Managed systems appear to store less carbon than their natural counterparts by means of projection (e.g. 75.3 Mg•ha⁻¹ for the management regime and 115 Mg•ha⁻¹ for the natural regime). Given enough time, however, carbon flux of old forests would theoretically approach zero for the rate of change of carbon accumulations (i.e. second derivative) is negative (Figure 21B). Such phenomenon in old forests is analogous to a carbon balance in planted forests between carbon captured by photosynthesis and carbon removed by thinning and final harvest. Further, wood products offer a potential advantage over manufactured materials for locking up sequestered carbon. For example, a simple sawed wood product requires 44 percent less energy consumption than steel, 93 percent less than aluminum, 60-80 percent less than concrete, or 77-83 percent less than plastic (Petersen and Solberg, 2005; Jansson and others, 2010). Managing forests to supply wood products may provide low-cost opportunities for GHG mitigation. Therefore, proper carbon mitigation policy should be a compromise between managing forests and preserving forests.

Increased demand for wood products often results in landowners adopting more intensive forest management practices (Prestemon and Abt, 2002). Management scenarios showed that through intensified application of fertilizers and herbicide and genetic improvement, improved plantation productivity increases not only the production potential of forests but also in-wood/harvested carbon stock up to 30 percent. However, fertilizers and herbicide require additional energy to produce and apply, and some of the applied fertilizers and herbicide is inevitable lost as GHG such as N₂O (Sathre and others, 2010). Such potential for lowering the GHG benefit is not accounted in management scenarios explored here.

In total, the carbon stored aboveground in loblolly pine plantations and wood harvested from them, including that used for energy production, has considerable potential to offset GHG emissions from fossil fuels. To better assess roles of such forest offset, GHG offset payments to landowners are necessary to model future market adjustments (Cairns and Lasserre, 2004; Im and others, 2007). Forestry-related policies implemented in efforts to mitigate GHG emissions or accomplish other public goals have the potential to affect landowners' management of plantation lands (Pohjola and Valsta, 2007). Despite the lack of any direct linkage to proposed public policies here, the approach used here allows for flexibility and adaptability in changing assumptions or inputs when new data and information become available. The results of projections like those presented here provide potentially useful information for use in addressing questions about the role southern pine plantations can play in GHG mitigation and climate policy.

LITERATURE CITED

- Adams, D.M.;** Ince, P.J.; Alig, R.J. [and others]. 2003. Assumptions and methods used in projections. In: Haynes, R.W. (Ed.), *An Analysis of the Timber Situation in the United States: 1952 to 2050*. Gen. Tech. Rep. PNW-GTR-560 USDA Forest Service, Pacific Northwest Research Station, Portland, OR. p. 10-54.
- Albaugh, T.J.;** Allen, H.L.; Fox, T.R. 2007. Historical patterns of forest fertilization in the southeastern United States from 1969 to 2004. *Southern Journal of Applied Forestry* 31 (3), 129-137.
- Allen, H.L.** 2001. Silvicultural treatments to enhance productivity. In: Evans, J. (Ed.), *The Forests Handbook*. Volume II. Blackwell Science Ltd., Oxford, UK.
- Allen, H.L.;** Fox, T.R.; Campbell, R.G. 2005. What is ahead for intensive pine plantation silviculture in the south? *Southern Journal of Applied Forestry* 29 (2), 62-69.
- Amateis, R.L.;** Burkhardt, H.E. 1996. Impact of heavy glaze in a loblolly pine spacing trial. *Southern Journal of Applied Forestry* 20 (3), 151-155.
- Amateis, R.L.;** Burkhardt, H.E. 2005. The influence of thinning on the proportion of peeler, sawtimber, and pulpwood trees in loblolly pine plantations. *Southern Journal of Applied Forestry* 29 (3), 158-162.
- Amateis, R.L.;** Burkhardt, H.E. 2009. FASTLOB 3.0: A Stand-level Growth and Yield Model for Fertilized and thinned Loblolly Pine Plantations The Loblolly Pine Growth and Yield Research Cooperative, Department of Forestry, Virginia Tech, Blacksburg, Virginia, p. 38.
- Applegate, J.R.** 2008. Estimates of down woody materials on Fort AP Hill, Virginia. *Southern Journal of Applied Forestry* 32 (2), 53-59.
- Baldwin, V.C.;** Peterson, K.D.; Burkhardt, H.E. [and others]. 1997. Equations for estimating loblolly pine branch and foliage weight and surface area distributions. *Canadian Journal of Forest Research* 27 (6), 918-927.
- Baskerville, G.L.** 1972. Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research* 2, 49-53.
- Bechtold, W.A.;** Scott, C.T. 2005. The Forest Inventory and Analysis plot design. In: Bechtold, W.A., Patterson, P.L. (Eds.), *The enhanced Forest Inventory and Analysis program - National sampling design and estimation procedures*. Gen. Tech. Rep. SRS-80. USDA Forest Service, Southern Research Station, Asheville, NC, pp. 27-42.
- Birdsey, R.A.;** Heath, L.S. 1995. Carbon Changes in U.S. Forests. In: Joyce, L.A. (Ed.), *Productivity of America's Forests and Climate Change*. Gen. Tech. Rep. RM-GTR-271. USDA Forest Service, Rocky Mountain Forest Experiment Station, Fort Collins, CO, pp. 56-70.
- Booth, J.G.;** Sarkar, S. 1998. Monte Carlo approximation of bootstrap variances. *American Statistician* 52 (4), 354-357.
- Bullock, B.P.;** Burkhardt, H.E. 2003. Equations for predicting green weight of loblolly pine trees in the South. *Southern Journal of Applied Forestry* 27 (3), 153-159.
- Cairns, R.D.;** Lasserre, P. 2004. Reinforcing economic incentives for carbon credits for forests. *Forest Policy and Economics* 6 (3-4), 321-328.
- Campbell, J.W.;** Hanula, J.L.; Outcalt, K.W. 2008. Effects of prescribed fire and other plant community restoration treatments on tree mortality, bark beetles, and other saproxylic Coleoptera of longleaf pine, *Pinus palustris* Mill., on the Coastal Plain of Alabama. *Forest Ecology and Management* 254 (2), 134-144.
- Carino, H.F.** 2009. Maximizing the value yield of short-rotation southern pine timber. *Forest Products Journal* 59 (9), 29-37.
- Chojnacky, D.C.;** Heath, L.S. 2002. Estimating down deadwood from FIA forest inventory variables in Maine. *Environmental Pollution* 116, S25-S30.
- Chojnacky, D.C.;** Schuler, T.M. 2004. Amounts of down woody materials for mixed-oak forests in Kentucky, Virginia, Tennessee, and North Carolina. *Southern Journal of Applied Forestry* 28 (2), 113-117.
- Clutter, J.L.;** Fortson, J.C.; Pienaar, L.V. [and others]. 1983. *Timber management : a quantitative approach*. Wiley, New York,
- Conner, R.C.;** Hartsell, A.J. 2002. Forest area and conditions. In: Wear, D.N., Greis, J.G. (Eds.), *Southern Forest Resource Assessment*. GTR SRS-53. USDA Forest Service, Southern Research Station, Asheville, NC, pp. 357-402.
- Cooper, J.;** Becker, C. 2009. Virginia's timber industry--an assessment of timber product output and use, 2007. *Resour. Bull. SRS--155*. USDA Forest Service, Southern Research Station, Asheville, NC, p. 33.
- Duvall, M.D.;** Grigal, D.F. 1999. Effects of timber harvesting on coarse woody debris in red pine forests across the Great Lakes states, USA. *Canadian Journal of Forest Research* 29 (12), 1926-1934.
- Ekbom, B.;** Schroeder, L.M.; Larsson, S. 2006. Stand specific occurrence of coarse woody debris in a managed boreal forest landscape in central Sweden. *Forest Ecology and Management* 221 (1-3), 2-12.
- Forest Inventory and Analysis.** 2007. *Forest Inventory and Analysis Factsheet*, Virginia, USDA Forest Service, Forest Inventory and Analysis, Southern Research Station, Knoxville TN, <http://srsfia2.fs.fed.us/states/virginia.shtml>. Date Accessed: 05/10/2009.
- Forest Inventory and Analysis.** 2009a. *FIA DataMart - FIADB version 4.0*, USDA Forest Service, Forest Inventory and Analysis National Program, Arlington VA, <http://fiatools.fs.fed.us/fiadb-downloads/datamart.html>. Date Accessed: 06/05/2009.

- Forest Inventory and Analysis.** 2009b. FIA Field Guide, USDA Forest Service, Forest Inventory and Analysis National Program, Arlington VA, <http://fia.fs.fed.us/library/field-guides-methods-proc/>. Date Accessed: 04/12/2010.
- Fox, T.R.;** Allen, H.L.; Albaugh, T.J. [and others]. 2007. Tree nutrition and forest fertilization of pine plantations in the southern United States. *Southern Journal of Applied Forestry* 31 (1), 5-11.
- Fox, T.R.;** Jokela, E.J.; Allen, H.L. 2004. The evolution of pine plantation silviculture in the southern United States. In: Rauscher, H.M., Johnsen, K. (Eds.), *Southern Forest Science: Past, Present, and Future*. Gen. Tech. Rep. SRS-75. USDA, Forest Service, Southern Research Station, Asheville, NC, pp. 63-82.
- Freese, F.** 1967. Elementary statistical methods for foresters. *Agricultural Handbook* 317. USDA Forest Service, Forest Products Laboratory, Madison, Wisconsin, pp. 5-6.
- Fridman, J.;** Walheim, M. 2000. Amount, structure, and dynamics of dead wood on managed forestland in Sweden. *Forest Ecology and Management* 131 (1-3), 23-36.
- Gibson, M.D.;** McMillin, C.W.; Shoulders, E. 1986. Moisture content and specific gravity of the four major southern pines under the same age and site conditions. *Wood and Fiber Science* 18 (3), 428-435.
- Glass, S.V.;** Zelinka, S.L. 2010. Moisture relations and physical properties of wood. *Wood Handbook, Wood as an Engineering Material*. General Technical Report FPL-GTR-190. USDA Forest Service, Forest Products Laboratory, Madison, WI, pp. 4-1 - 4-19.
- Gong, P.C.** 1998. Risk preferences and adaptive harvest policies for even-aged stand management. *Forest Science* 44 (4), 496-506.
- Guldin, J.M.** 2004. Reproduction cutting methods for naturally regenerated pine stands in the South. In: Rauscher, H.M., Johnsen, K. (Eds.), *Southern Forest Science: Past, Present, and Future*. Gen. Tech. Rep. SRS-75. USDA, Forest Service, Southern Research Station, Asheville, NC, pp. 83-95.
- Harmon, M.E.** 2001. Carbon sequestration in forests - Addressing the scale question. *Journal of Forestry* 99 (4), 24-29.
- Harmon, M.E.;** Ferrell, W.K.; Franklin, J.F. 1990. Effects on carbon storage of conversion of old-growth forests to young forests. *Science* 247 (4943), 699-702.
- Heath, L.S.;** Smith, J.E. 2004. Criterion 5, Indicator 27: contribution of forest ecosystems to the total global carbon budget, including absorption and release of carbon. In: Darr, D.R. (Ed.), *Data Report: A Supplement of the National Report on Sustainable Forests--2003*. FS-766A. USDA, Washington, DC, p. 7.
- Heath, L.S.;** Smith, J.E.; Birdsey, R.A. 2003. Carbon trends in U.S. forestlands: a context for the role of soils in forest carbon sequestration. In: Kimble, J.M., Heath, L.S., Birdsey, R.A., Lal, R. (Eds.), *The potential of U.S. forest soils to sequester carbon and mitigate the greenhouse effect*. CRC Press, Boca Raton, pp. 35-45.
- Herrmann, S.;** Prescott, C.E. 2008. Mass loss and nutrient dynamics of coarse woody debris in three Rocky Mountain coniferous forests: 21 year results. *Canadian Journal of Forest Research* 38 (1), 125-132.
- Houghton, R.A.;** Hackler, J.L.; Lawrence, K.T. 1999. The US carbon budget: Contributions from land-use change. *Science* 285 (5427), 574-578.
- Huang, C.H.;** Kronrad, G.D. 2006. The effect of carbon revenues on the rotation and profitability of loblolly pine plantations in East Texas. *Southern Journal of Applied Forestry* 30 (1), 21-29.
- Im, E.H.;** Adams, D.M.; Latta, G.S. 2007. Potential impacts of carbon taxes on carbon flux in western Oregon private forests. *Forest Policy and Economics* 9 (8), 1006-1017.
- Janssens, I.A.;** Freibauer, A.; Ciais, P. [and others]. 2003. Europe's terrestrial biosphere absorbs 7 to 12 percent of European anthropogenic CO₂ emissions. *Science* 300 (5625), 1538-1542.
- Jansson, C.;** Wullschlegel, S.D.; Kalluri, U.C.; Tuskan, G.A. 2010. Phytosequestration: Carbon biosequestration by plants and the prospects of genetic engineering. *Bioscience* 60 (9), 685-696.
- Jenkins, J.C.;** Chojnacky, D.C.; Heath, L.S.; Birdsey, R.A. 2003. National-scale biomass estimators for United States tree species. *Forest Science* 49 (1), 12-35.
- Johnsen, K.H.;** Wear, D.; Oren, R. [and others]. 2001. Carbon sequestration and southern pine forests. *Journal of Forestry* 99 (4), 14-21.
- Johnson, T.G.;** Steppleton, C.D.; Bentley, J.W. 2010. Southern pulpwood production, 2008. *Resour. Bull. SRS-165*. USDA Forest Service, Southern Research Station, Asheville, NC, p. 42.
- Jokela, E.J.;** Dougherty, P.M.; Martin, T.A. 2004. Production dynamics of intensively managed loblolly pine stands in the southern United States: a synthesis of seven long-term experiments. *Forest Ecology and Management* 192 (1), 117-130.
- Kangas, A.S.** 1997. On the prediction bias and variance in long-term growth projections. *Forest Ecology and Management* 96 (3), 207-216.
- Krankina, O.N.;** Harmon, M.E.; Griazkin, A.V. 1999. Nutrient stores and dynamics of woody detritus in a boreal forest: modeling potential implications at the stand level. *Canadian Journal of Forest Research* 29 (1), 20-32.
- Krankina, O.N.;** Harmon, M.E.; Kukuev, Y.A. [and others]. 2002. Coarse woody debris in forest regions of Russia. *Canadian Journal of Forest Research* 32 (5), 768-778.
- Landsberg, J.J.;** Waring, R.H. 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. *Forest Ecology and Management* 95 (3), 209-228.
- Li, B.;** McKeand, S.; Weir, R. 1999. Impact of Forest Genetics on Sustainable Forestry—Results from Two Cycles of Loblolly Pine Breeding in the U.S. *Journal of Sustainable Forestry* 10 (1-2), 79-85.
- Liechty, H.O.;** Fristoe, C. 2010. Initial response of loblolly pine and competition to mid rotation fertilization and herbicide application in the gulf coastal plain. In: Stanturf, J.A. (Ed.), *Proceedings of the 14th biennial southern silvicultural research conference*. Gen. Tech. Rep. SRS-121. USDA, Forest Service, Southern Research Station, Asheville, NC, pp. 47-49.
- Little, E.L.** 1971. *Atlas of United States Trees, Volume 1, Conifers and Important Hardwoods*. USDA Miscellaneous Publication 1146, 9 p., 200 maps, <http://esp.cr.usgs.gov/data/atlas/little/>. Date Accessed: 7/15/2009.
- Marra, J.L.;** Edmonds, R.L. 1994. Coarse woody debris and forest floor respiration in an old-growth coniferous forest on the Olympic Peninsula, Washington, USA. *Canadian Journal of Forest Research* 24 (9), 1811-1817.

- McKeand, S.;** Mullin, T.; Byram, T.; White, T. 2003. Deployment of genetically improved loblolly and slash pines in the south. *Journal of Forestry* 101 (3), 32-37.
- McKeand, S.E.;** Abt, R.C.; Allen, H.L. [and others]. 2006. What are the best loblolly pine genotypes worth to landowners? *Journal of Forestry* 104 (7), 352-358.
- McKenney, D.W.;** Yemshanov, D.; Fox, G.; Ramlal, E. 2004. Cost estimates for carbon sequestration from fast growing poplar plantations in Canada. *Forest Policy and Economics* 6 (3-4), 345-358.
- McMinn, J.W.;** Crossley, D.A. 1996. Biodiversity and Coarse Woody Debris in Southern Forests. Gen. Tech. Rep. SE-94. USDA Forest Service, Southeastern Forest Experiment Station, Asheville, NC, p. 146.
- McMinn, J.W.;** Hardt, R.A. 1996. Accumulations of coarse woody debris in Southern forests. In: McMinn, J.W., Crossley, D.A. (Eds.), *Biodiversity and Coarse Woody Debris in Southern Forests*. Gen. Tech. Rep. SE-94. USDA Forest Service, Southeastern Forest Experiment Station, Asheville, NC, pp. 1-9.
- Miller, A.T.;** Allen, H.L.; Maier, C.A. 2006. Quantifying the coarse-root biomass of intensively managed loblolly pine plantations. *Canadian Journal of Forest Research* 36 (1), 12-22.
- Miner, R.** 2006. The 100-year method for forecasting carbon sequestration in forest products in use. *Mitigation and Adaptation Strategies for Global Change*(online <http://www.springerlink.com/content/2167274117366751/>).
- Morris, L.A.;** Lowery, R.F. 1988. Influence of site preparation on soil conditions affecting stand establishment and tree growth. *Southern Journal of Applied Forestry* 12 (3), 170-178.
- Moulton, R.J.;** Hernandez, G. 2000. Tree planting in the United States - 1998. *Tree Planters' Notes* 49 (2), 23-36.
- Nilsson, U.;** Allen, H.L. 2003. Short- and long-term effects of site preparation, fertilization and vegetation control on growth and stand development of planted loblolly pine. *Forest Ecology and Management* 175 (1-3), 367-377.
- Oswalt, S.N.;** Oswalt, C.; Turner, J. 2008. Hurricane Katrina impacts on Mississippi forests. *Southern Journal of Applied Forestry* 32 (3), 139-141.
- Pacala, S.W.;** Hurtt, G.C.; Baker, D. [and others]. 2001. Consistent land- and atmosphere-based US carbon sink estimates. *Science* 292 (5525), 2316-2320.
- Page-Dumroese, D.S.;** Jurgensen, M.F. 2006. Soil carbon and nitrogen pools in mid- to late-successional forest stands of the northwestern United States: potential impact of fire. *Canadian Journal of Forest Research* 36 (9), 2270-2284.
- Perez-Garcia, J.;** Lippke, B.; Comnick, J.; Manriquez, C. 2005. An assessment of carbon pools, storage, and wood products market substitution using life-cycle analysis results. *Wood and Fiber Science* 37, 140-148.
- Perlack, R.D.;** Wright, L.L.; Turhollow, A.F. [and others]. 2005. Biomass as feedstock for a bioenergy and bioproducts industry: The technical feasibility of a billion-ton annual supply. DOE/USDA Tech. Rep. DOE/GO-102005-2135. USDOE, Office of Scientific & Technical Information, Oak Ridge, TN, p. 78.
- Petersen, A.K.;** Solberg, B. 2005. Environmental and economic impacts of substitution between wood products and alternative materials: a review of micro-level analyses from Norway and Sweden. *Forest Policy and Economics* 7 (3), 249-259.
- Pohjola, J.;** Valsta, L. 2007. Carbon credits and management of Scots pine and Norway spruce stands in Finland. *Forest Policy and Economics* 9 (7), 789-798.
- Prestemon, J.P.;** Abt, R.C. 2002. The Southern timber market to 2040. *Journal of Forestry* 100 (7), 16-22.
- Price, T.;** Doggett, C.; Pye, J.; Smith, B. 1998. *A History of Southern Pine Beetle Outbreaks in the Southeastern United States*. Georgia Forestry Commission, Atlanta, GA.
- Radtke, P.J.;** Amateis, R.L.; Priskey, S.P. [and others]. 2009. Modeling production and decay of coarse woody debris in loblolly pine plantations. *Forest Ecology and Management* 257 (3), 790-799.
- Reams, G.A.;** Smith, W.D.; Hansen, M.H. [and others]. 2005. The Forest Inventory and Analysis sampling frame. In: Bechtold, W.A., Patterson, P.L. (Eds.), *The enhanced Forest Inventory and Analysis program - National sampling design and estimation procedures*. Gen. Tech. Rep. SRS-80. USDA Forest Service, Southern Research Station, Asheville, NC, pp. 11-26.
- Rubino, D.L.;** McCarthy, B.C. 2003. Evaluation of coarse woody debris and forest vegetation across topographic gradients in a southern Ohio forest. *Forest Ecology and Management* 183 (1-3), 221-238.
- Sathre, R.;** Gustavsson, L.; Bergh, J. 2010. Primary energy and greenhouse gas implications of increasing biomass production through forest fertilization. *Biomass & Bioenergy* 34 (4), 572-581.
- Schiermeier, Q.;** Tollefson, J.; Scully, T. [and others]. 2008. Electricity without carbon. *Nature* 454 (7206), 816-823.
- Schimmel, D.;** Melillo, J.; Tian, H.Q. [and others]. 2000. Contribution of increasing CO₂ and climate to carbon storage by ecosystems in the United States. *Science* 287 (5460), 2004-2006.
- Schultz, R.P.** 1997. *Loblolly Pine: The Ecology and Culture of Loblolly Pine (Pinus taeda L.)*. Agriculture Handbook 713. USDA Forest Service, Washington, D.C., p. 493.
- Schulze, E.D.;** Wirth, C.; Heimann, M. 2000. Climate change - managing forests after Kyoto. *Science* 289 (5487), 2058-2059.
- Scott, C.T.;** Bechtold, W.A.; Reams, G.A. [and others]. 2005. Sample-based estimators used by the Forest Inventory and Analysis National Information Management System. In: Bechtold, W.A., Patterson, P.L. (Eds.), *The Enhanced Forest Inventory and Analysis Program - National Sampling Design and Estimation Procedures*. Gen. Tech. Rep. SRS-80. USDA Forest Service, Southern Research Station, Asheville, NC, pp. 43-67.
- Siry, J.P.** 2002. Intensive timber management practices. In: Wear, D.N., Greis, J.G. (Eds.), *Southern Forest Resource Assessment*. GTR SRS-53. USDA Forest Service, Southern Research Station, Asheville, NC, pp. 327-340.
- Siry, J.P.;** Cabbage, F.W.; Malmquist, A.J. 2001. Potential impacts of increased management intensities on planted pine growth and yield and timber supply modeling in the south. *Forest Products Journal* 51 (3), 42-48.
- Skog, K.E.;** Nicholson, G.A. 1998. Carbon cycling through wood products: The role of wood and paper products in carbon sequestration. *Forest Products Journal* 48 (7-8), 75-83.
- Smith, J.E.;** Heath, L.S. 2001. Identifying influences on model uncertainty: An application using a forest carbon budget model. *Environmental Management* 27 (2), 253-267.

- Smith, J.E.;** Heath, L.S. 2002. A model of forest floor carbon mass for United States forest types. Res. Pap. NE-722. USDA Forest Service, Northeastern Research Station, Newtown Square, PA, p. 37.
- Smith, J.E.;** Heath, L.S.; Jenkins, J.C. 2003. Forest Volume-to-Biomass Models and Estimates of Mass for Live and Standing Dead Trees of U.S. Forests. Gen. Tech. Rep. NE-298. USDA Forest Service, Northeastern Research Station, Newtown Square, PA, p. 57.
- Smith, J.E.;** Heath, L.S.; Skog, K.E.; Birdsey, R.A. 2006. Methods for calculating forest ecosystem and harvested carbon with standard estimates for forest types of the United States. Gen. Tech. Rep. NE-343. USDA Forest Service, Northeastern Research Station, Newtown Square, PA, p. 216.
- Smith, J.E.;** Heath, L.S.; Woodbury, P.B. 2004a. How to estimate forest carbon for large areas from inventory data. *Journal of Forestry* 102 (5), 25-31.
- Smith, J.E.;** Heath, L.S.; Woodbury, P.B. 2004b. How to estimate forest carbon for large areas from inventory data- Supplemental material for Tables 1 and 2 <http://www.forestcarbon.net/> Date Accessed: 12/08/2010.
- Smith, W.B.** 2002. Forest inventory and analysis: a national inventory and monitoring program. *Environmental Pollution* 116, S233-S242.
- Smith, W.B.;** Miles, P.D.; Vissage, J.S.; Pugh, S.A. 2004c. Forest Resources of the United States, 2002. Gen. Tech. Rep. NC-241 USDA Forest Service, North Central Research Station, St. Paul, MN, p. 137.
- Spetich, M.A.;** Shifley, S.R.; Parker, G.R. 1999. Regional distribution and dynamics of coarse woody debris in midwestern old-growth forests. *Forest Science* 45 (2), 302-313.
- Spies, T.A.;** Franklin, J.F.; Thomas, T.B. 1988. Coarse woody debris in Douglas-fir forests of western Oregon and Washington. *Ecology* 69 (6), 1689-1702.
- U.S. Energy Information Administration.** 2010. Biomass for Electricity Generation, <http://www.eia.doe.gov/oiaf/analysispaper/biomass/table3.html>. Date Accessed: 11/18/2010.
- U.S. Environmental Protection Agency.** 2007a. Clean Air Markets Division data and maps, <http://camddataandmaps.epa.gov/gdm/>. Date Accessed: 02/16/2009.
- U.S. Environmental Protection Agency.** 2007b. Clean Air Markets Division data and maps web site, <http://camddataandmaps.epa.gov/gdm/>. Date Accessed: 02/16/2009.
- U.S. Environmental Protection Agency.** 2008. Inventory of U.S. greenhouse gas emissions and sinks: 1990 - 2006. EPA 430-R-08-005. Office of Atmospheric Programs, Washington, DC.
- Vanderwel, M.C.;** Thorpe, H.C.; Shuter, J.L. [and others]. 2008. Contrasting downed woody debris dynamics in managed and unmanaged northern hardwood stands. *Canadian Journal of Forest Research* 38 (11), 2850-2861.
- Wade, D.D.** 1993. Thinning young loblolly pine stands with fire. *International Journal of Wildland Fire* 3 (3), 169-178.
- Wear, D.N.;** Greis, J.G. 2002. Southern Forest Resource Assessment - Technical Report. GTR SRS-53. USDA Forest Service, Southern Research Station, Asheville, NC, p. 635.
- Woodall, C.W.;** Monleon, V.J. 2008. Sampling Protocol, Estimation, and Analysis Procedures for the Down Woody Materials Indicator of the FIA Program. Gen. Tech. Rep. NRS-22. USDA Forest Service, Northern Research Station, Newtown Square, PA, p. 72.

Table 1—Summary of stand attributes (area-weighted mean) for FIA sampled field plots and their representative forestland area of loblolly pine plantations by southern States

State	Plots	Area (10 ⁶ ha)	SI [†] (m)	Planting (seedlings·ha ⁻¹)	Age (yrs)	TPA (trees·ha ⁻¹)	BA [‡] (m ² ·ha ⁻¹)	Thinning	
								Age	TPA
Alabama	976	1.95	20.0	1,040	14	867	13.8	19	425
Arkansas	406	0.85	17.0	1,127	18	788	15.8	24	413
Florida	116	0.24	19.7	1,095	16	912	14.9	-	-
Georgia	833	1.66	19.8	941	16	870	14.2	20	390
Louisiana	508	1.09	19.7	1,038	14	964	14.2	23	467
Mississippi	823	1.63	19.8	1,080	15	833	14.9	21	445
North Carolina	394	1.00	18.5	1,191	19	818	15.6	24	319
South Carolina	575	1.10	19.6	1,240	17	855	16.1	21	405
Tennessee	83	0.15	18.5	751	14	754	11.5	20	425
Texas	437	0.87	19.1	1,176	14	843	12.6	18	415
Virginia	329	0.63	18.5	1,038	19	813	15.4	24	334
South	5,480	11.17	19.4	1,067	16	855	14.7	21	410

† Site index at base age of 25 years

‡ Basal area

Table 2—State-level and southwide in-woods carbon mass totals (10^6 Mg) and means ($\text{Mg}\cdot\text{ha}^{-1}$): FIA estimates and bootstrap standard errors and 95 percent confidence intervals

State	In-woods carbon total (10^6 Mg)				In-woods carbon mean ($\text{Mg}\cdot\text{ha}^{-1}$)			
	Estimate [†]	SE [‡] (%)	95% CI		Estimate [†]	SE [‡] (%)	95% CI	
			2.5 th	97.5 th			2.5 th	97.5 th
Alabama	53.7	4.21	49.4	58.2	27.49	2.74	26.02	28.89
Arkansas	26.7	7.12	23.2	30.7	31.43	4.69	28.69	34.39
Florida	7.2	12.35	5.5	9.1	30.25	7.70	25.76	34.98
Georgia	49.4	4.50	45.1	53.8	29.83	2.86	28.19	31.55
Louisiana	31.2	6.89	27.1	35.4	28.58	5.33	25.67	31.59
Mississippi	52.5	4.52	47.9	57.2	32.11	2.90	30.28	33.91
North Carolina	35.4	6.73	30.7	40.3	35.51	4.36	32.48	38.53
South Carolina	38.5	5.58	34.4	42.8	34.97	3.42	32.66	37.20
Tennessee	3.3	18.15	2.2	4.5	21.30	12.94	15.84	26.69
Texas	21.4	6.78	18.6	24.2	24.52	4.66	22.43	26.86
Virginia	22.1	7.71	18.7	25.4	34.99	4.75	31.74	38.15
South	341.1	1.40	331.7	350.5	30.54	1.27	29.77	31.31

[†] Estimate based on FIA 2005–2007 data and FASTLOB yield predictions

[‡] Estimated standard error from bootstrap sampling

Table 3—State-level and southwide live-tree carbon mass totals (10^6 Mg) and means ($\text{Mg}\cdot\text{ha}^{-1}$): FIA estimates and bootstrap standard errors and 95 percent confidence intervals

State	Live-tree carbon total (10^6 Mg)				Live-tree carbon mean ($\text{Mg}\cdot\text{ha}^{-1}$)			
	Estimate [†]	SE [‡] (%)	95% CI		Estimate [†]	SE [‡] (%)	95% CI	
			2.5 th	97.5 th			2.5 th	97.5 th
Alabama	50.5	4.19	46.5	54.7	25.85	2.71	24.46	27.17
Arkansas	24.6	6.88	21.4	28.0	28.96	4.36	26.56	31.50
Florida	6.8	12.36	5.2	8.6	28.62	7.70	24.33	33.08
Georgia	46.8	4.49	42.7	50.8	28.22	2.82	26.64	29.82
Louisiana	28.5	6.62	24.9	32.1	26.11	4.93	23.70	28.65
Mississippi	49.5	4.49	45.1	53.8	30.28	2.84	28.58	31.95
North Carolina	31.3	6.47	27.3	35.5	31.35	3.94	28.99	33.71
South Carolina	36.1	5.59	32.2	40.2	32.78	3.41	30.57	34.90
Tennessee	3.0	18.17	2.0	4.1	19.53	12.92	14.50	24.52
Texas	19.9	6.65	17.3	22.5	22.81	4.45	20.97	24.83
Virginia	19.7	7.51	16.8	22.6	31.28	4.42	28.57	33.91
South	316.4	1.35	308.1	324.7	28.32	1.22	27.63	29.00

[†] Estimate based on FIA 2005–2007 data and FASTLOB yield predictions

[‡] Estimated standard error from bootstrap sampling

Table 4—State-level and southwide CWD carbon mass totals (10^6 Mg) and means ($\text{Mg}\cdot\text{ha}^{-1}$): FIA estimates and bootstrap standard errors and 95 percent confidence intervals

State	CWD carbon total (10^6 Mg)				CWD carbon mean ($\text{Mg}\cdot\text{ha}^{-1}$)			
	Estimate [†]	SE [‡] (%)	95% CI		Estimate [†]	SE [‡] (%)	95% CI	
			2.5 th	97.5 th			2.5 th	97.5 th
Alabama	3.2	8.59	2.7	3.8	1.65	8.01	1.39	1.91
Arkansas	2.1	15.16	1.6	2.8	2.47	13.99	1.87	3.20
Florida	0.4	15.94	0.3	0.5	1.63	12.86	1.24	2.06
Georgia	2.7	6.90	2.4	3.1	1.61	6.11	1.44	1.81
Louisiana	2.7	15.23	2.0	3.6	2.47	14.75	1.81	3.24
Mississippi	3.0	7.17	2.6	3.4	1.83	6.35	1.61	2.07
North Carolina	4.2	13.30	3.2	5.4	4.16	12.31	3.27	5.29
South Carolina	2.4	7.56	2.1	2.8	2.19	6.38	1.91	2.46
Tennessee	0.3	32.96	0.1	0.5	1.77	30.62	0.84	2.89
Texas	1.5	13.49	1.1	1.9	1.71	12.65	1.31	2.16
Virginia	2.4	14.69	1.8	3.1	3.71	13.36	2.83	4.77
South	24.7	3.69	23.0	26.5	2.21	3.65	2.06	2.38

[†] Estimate based on FIA 2005–2007 data and FASTLOB yield predictions

[‡] Estimated standard error from bootstrap sampling

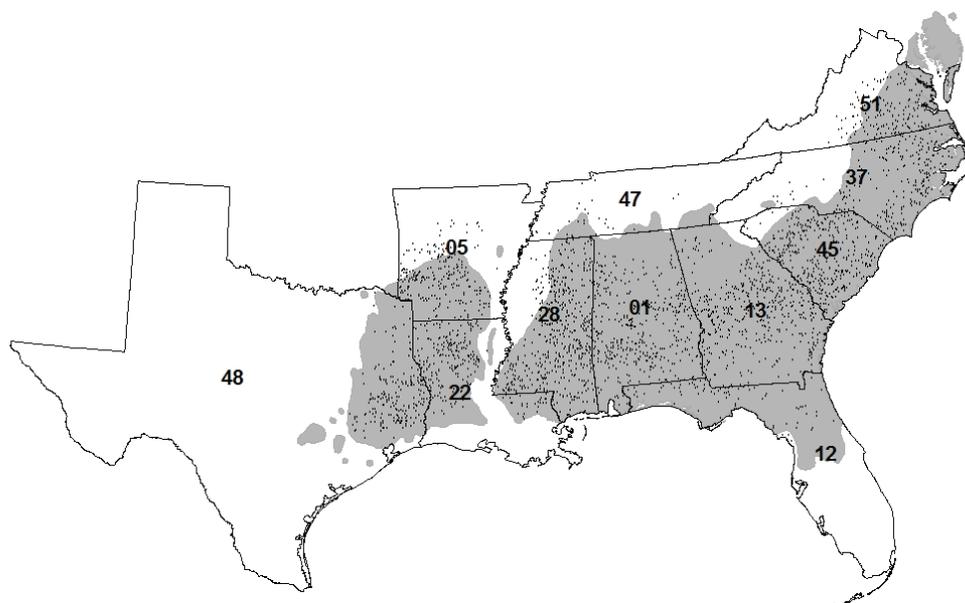


Figure 1. Approximate Forest Inventory and Analysis (FIA) plot locations for loblolly pine plantations and the natural range[†] of loblolly pine forests (shaded) in the southern United States. State codes – 01: Alabama, 05: Arkansas, 12: Florida, 13: Georgia, 22: Louisiana, 28: Mississippi, 37: North Carolina, 45: South Carolina, 47: Tennessee, 48: Texas, and 51: Virginia.

[†] Geographic distribution of loblolly pine is obtained from Little (1971).

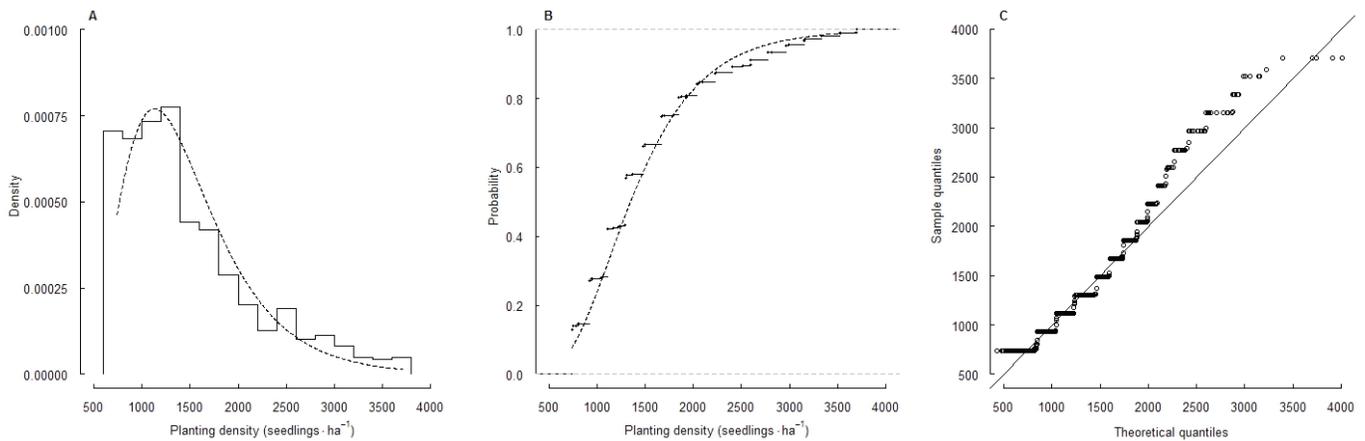


Figure 2—Fitted log-normal distribution of planting density ($\mu = 7.21$, $\sigma = 0.42$, while the variable at natural logarithm scale): (A) Histogram of observed data versus fundamental shape; (B) Empirical versus theoretical cumulative distribution functions (ECDF versus CDF) (C) Empirical quantiles versus theoretical quantiles from a log-normal distribution.

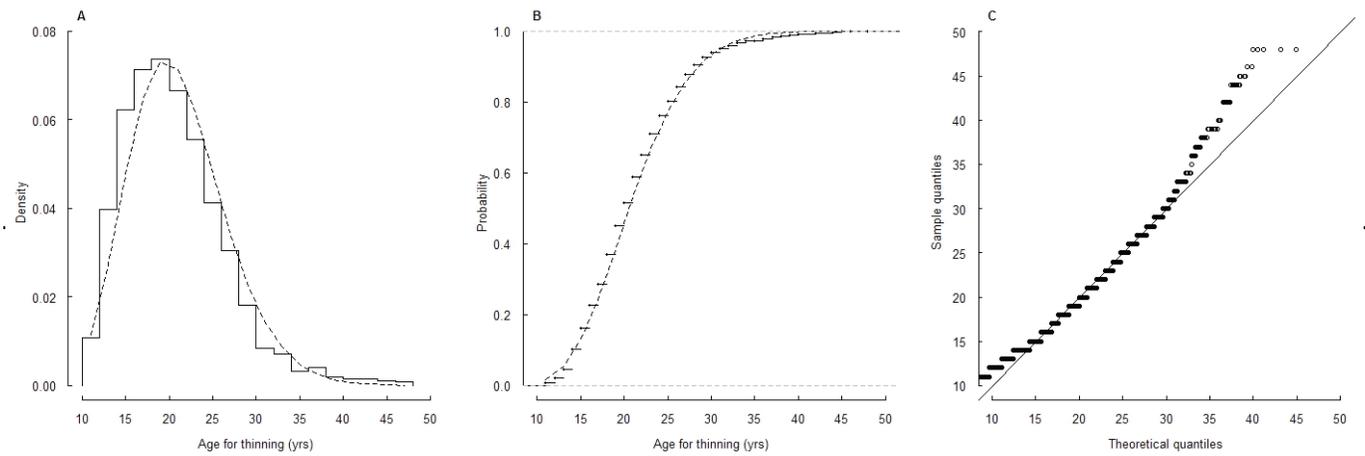


Figure 3—Fitted gamma distribution of age of thinning ($\alpha = 14.31$, $\lambda = 0.68$): (A) Histogram of observed data and fitted gamma density function; (B) ECDF versus CDF; (C) Empirical quantiles versus theoretical quantiles from a gamma distribution.

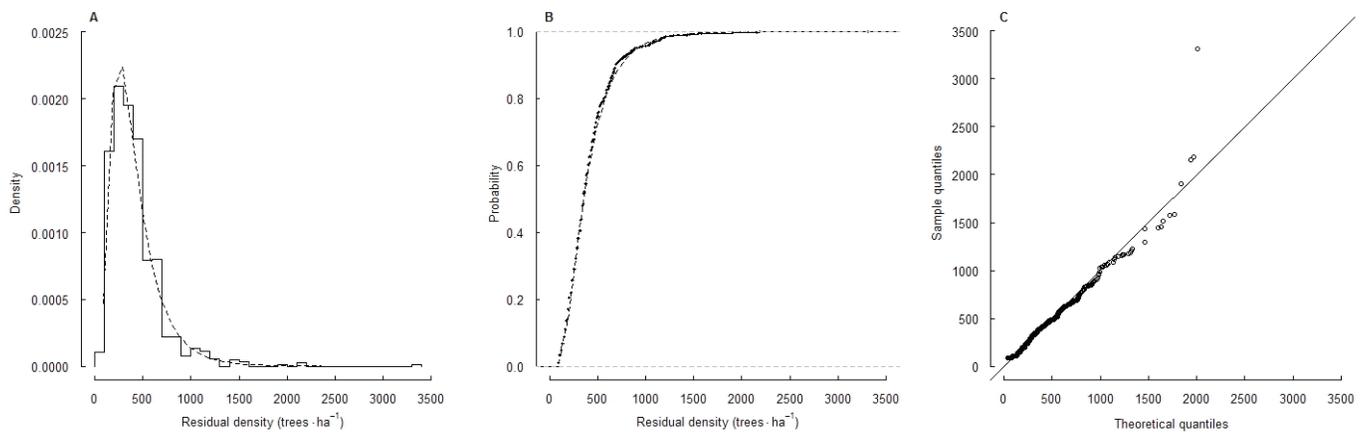


Figure 4—Fitted log-normal distribution of residual density after thinning ($\mu = 5.86$, $\sigma = 0.58$, while the variable at natural logarithm scale): (A) Histogram of observed data and fitted lognormal function; (B) ECDF versus CDF (C) Empirical quantiles versus theoretical quantiles from a log-normal distribution.

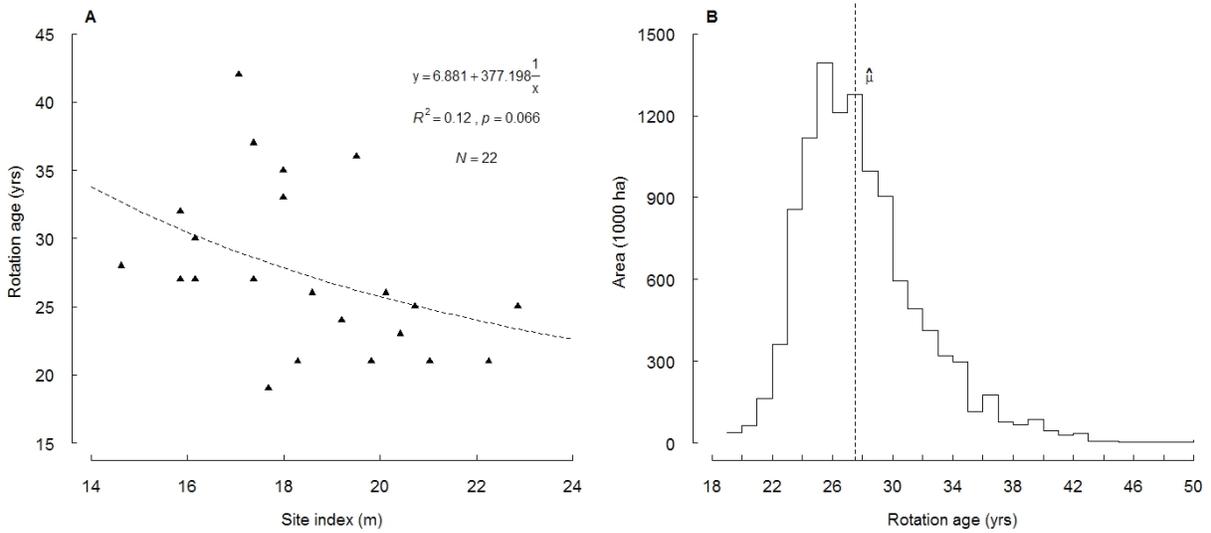


Figure 5—Distribution of rotation ages, accounting for site index at base age 25: (A) The relationship between FIA observed rotation ages and site index; (B) Predicted rotation ages for all stands across the South with a mean $\hat{\mu} = 27.5$

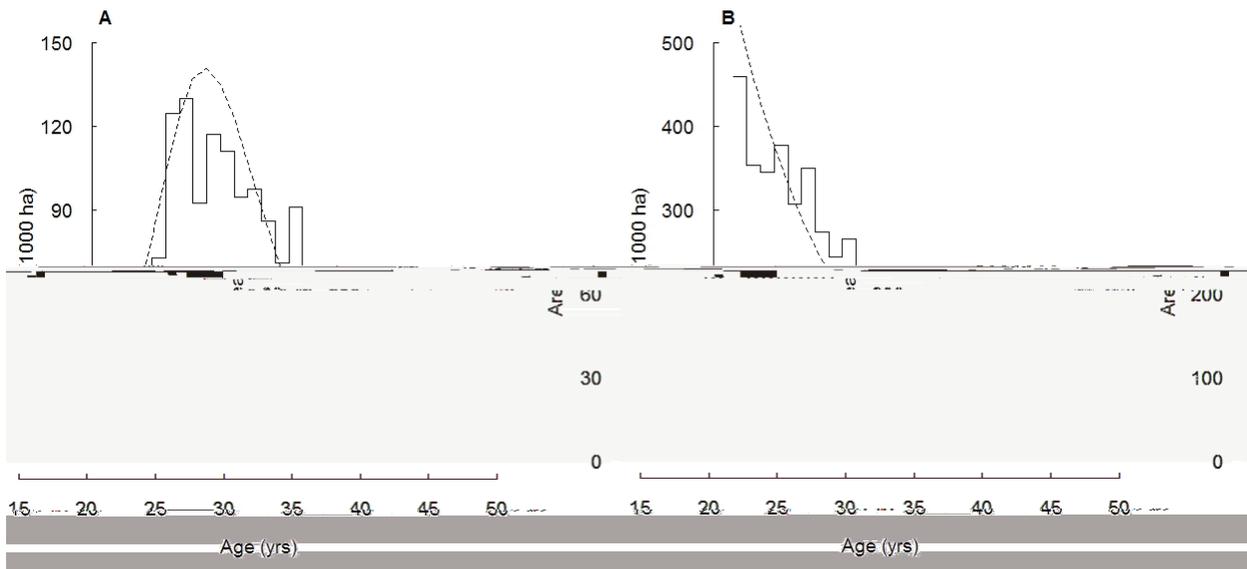


Figure 6—Quality of fit for distributions of planted area by age classes throughout the South (A) Stands with evidence of thinning which age class at 22 yrs has largest fitted area; (B) Stands without thinning observed which age class at 16 yrs has largest fitted area.

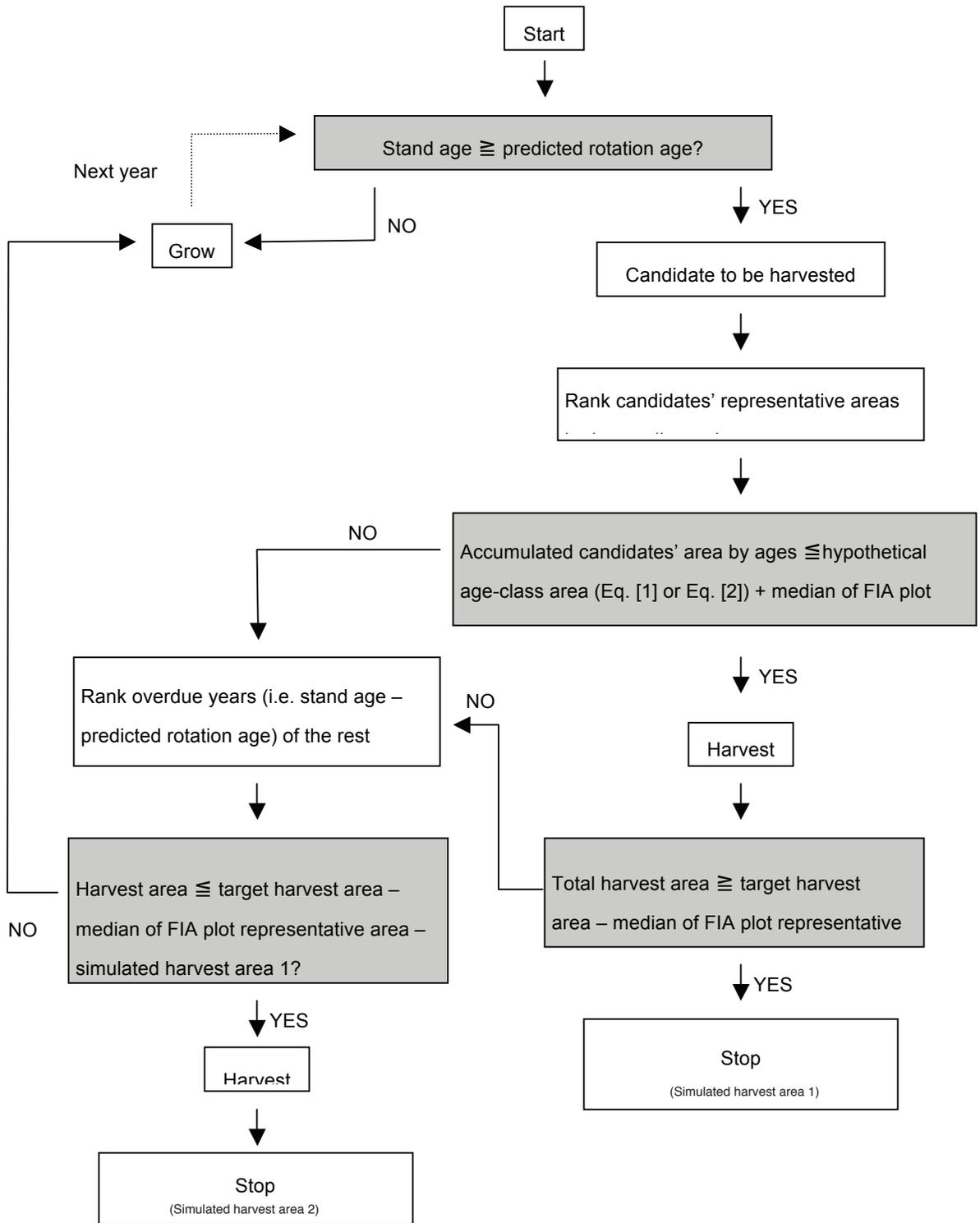


Figure 7—Rules used to select FIA plots for harvesting from thinned [1] and never-thinned [2] plantations.

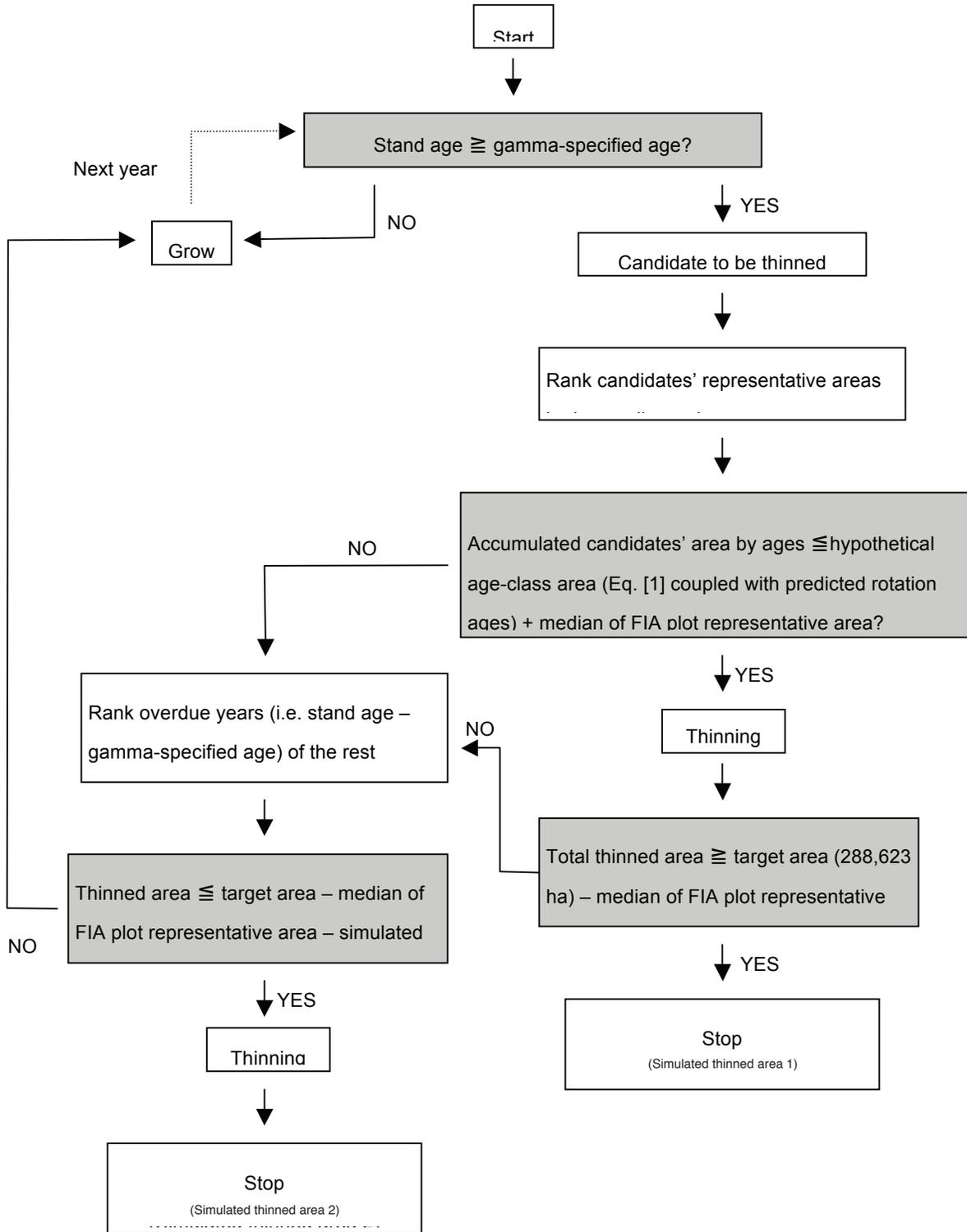


Figure 8—Rules used to select FIA plots for thinning.

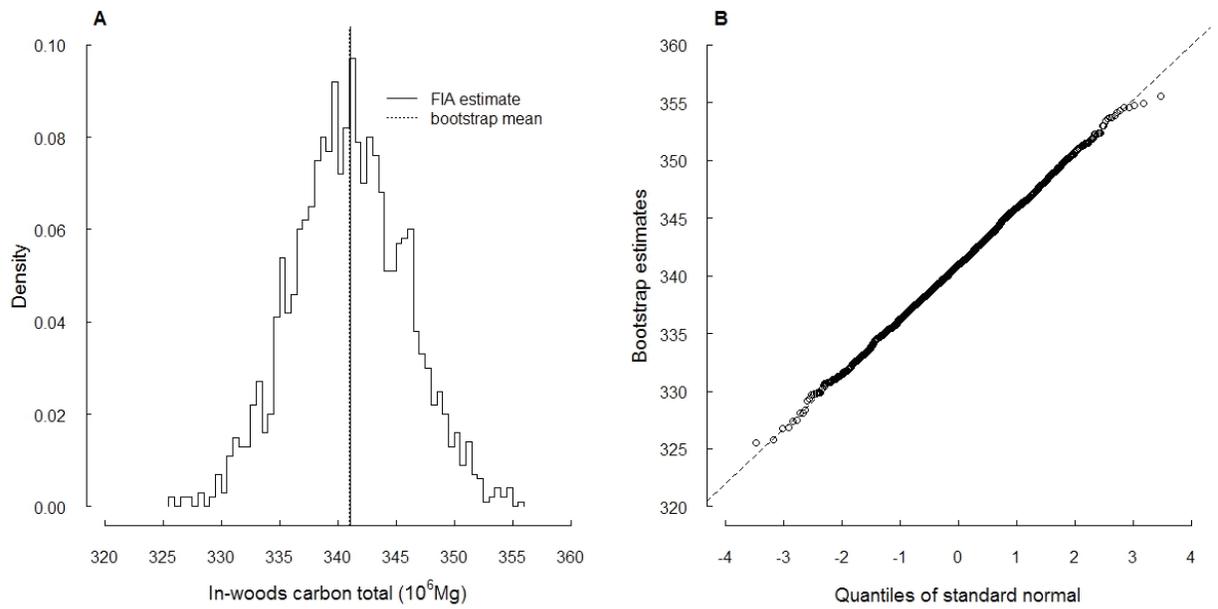


Figure 9—Bootstrap sampling distribution of southwide in-woods carbon totals (10⁶ Mg) (A), and its quality of fit based on a normal distribution (B).

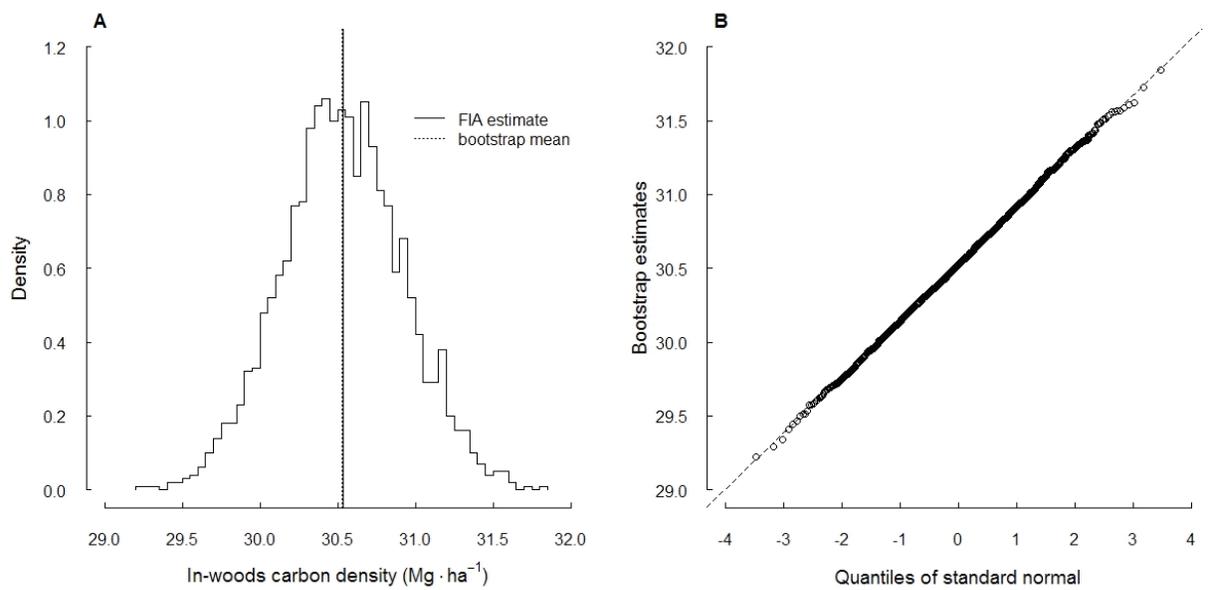


Figure 10. Bootstrap sampling distribution of southwide in-woods carbon mean per hectare (Mg·ha⁻¹) (A), and its quality of fit based on a normal distribution (B).

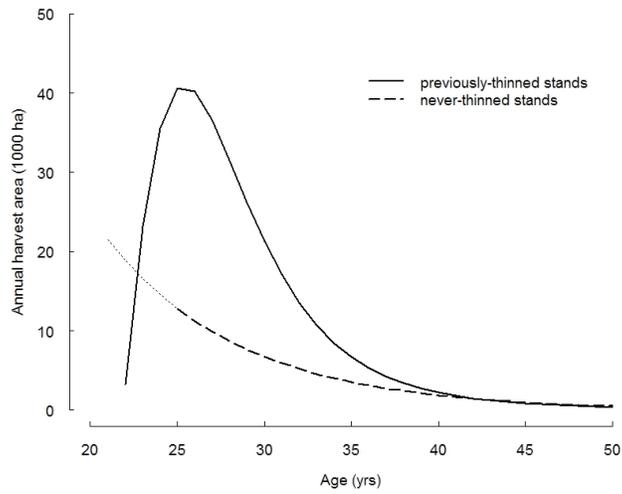


Figure 11—Hypothetical function of annual harvest area on ages, accounting for previous thinning operations. A dotted line represents that annual harvest areas may not be restricted to hypothetical values because areas of predicted rotation ages <25 are less than that of rotation age at 25 years (Figure 5B).

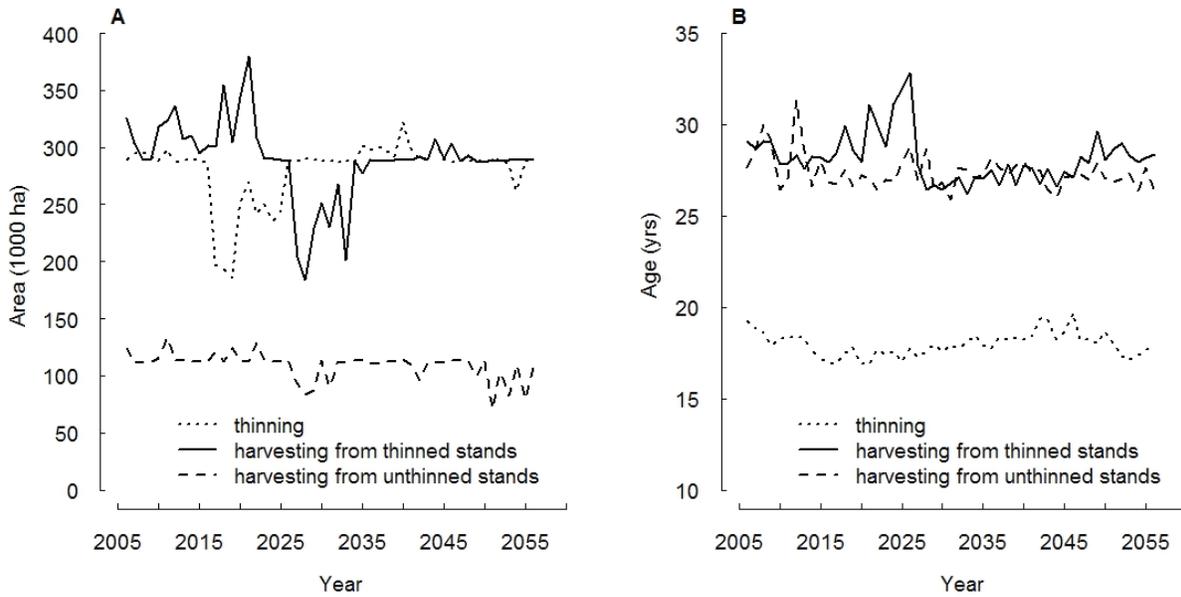


Figure 12—Simulations of area operated each year in the span of 50 years: (A) Area operated by thinning, and final harvest on thinned and unthinned stands; (B) Mean values of ages when activities of timber removed occur.

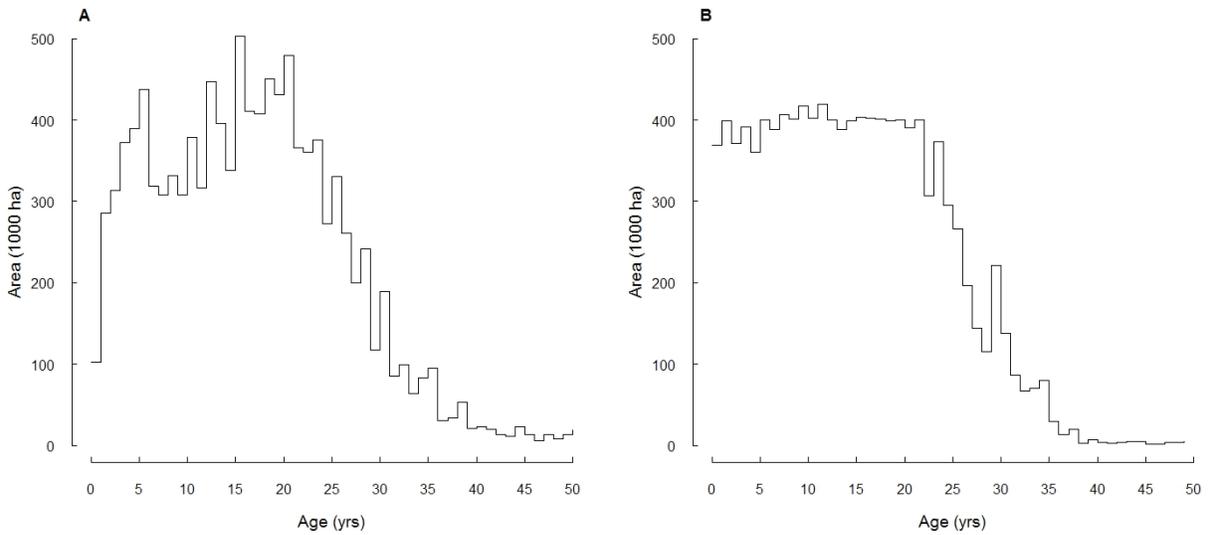


Figure 13—Age-class distribution of loblolly pine plantations throughout the South before and after a 50-year harvest period: (A) Initial plantations based on FIA 2005 – 2007 inventory data; (B) plantations after a 50-year harvest period.

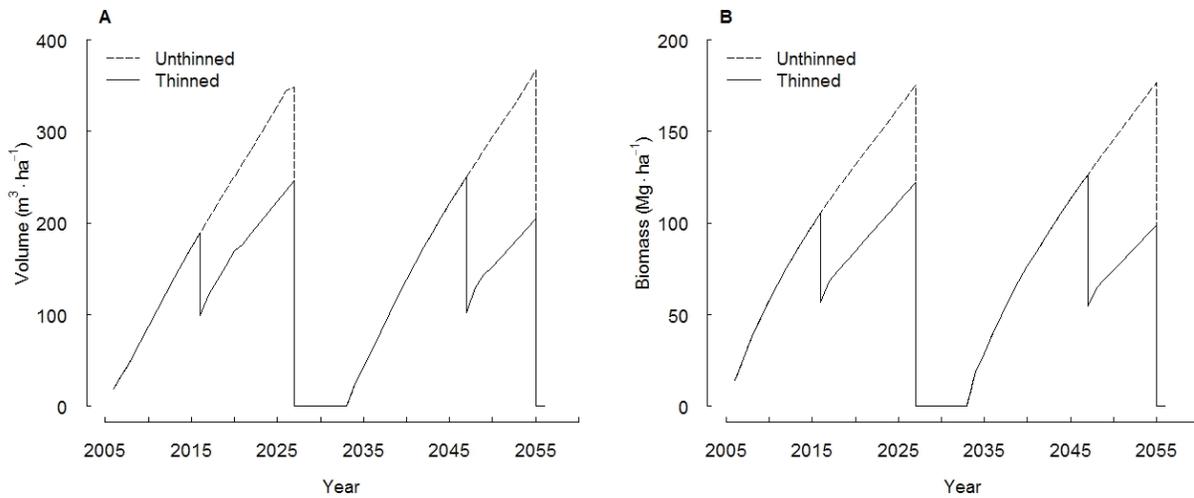


Figure 14—Temporal changes in in-woods stocks of volume (A) and aboveground biomass (B) for two different management regimes: thinned and unthinned planted loblolly pine yield and growth in 50-year projections. Two final harvests occur at age 27 years for each regime. For a thinned stand, thinnings occur at ages 16 and 19.

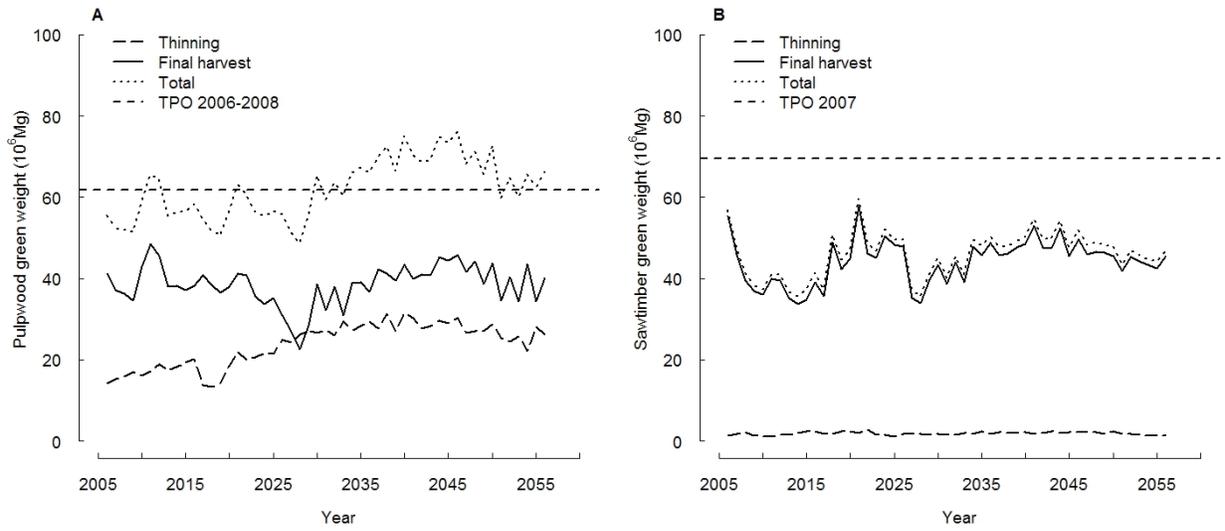


Figure 15—Timber production projections from thinnings and final harvests: (A) Pulpwood; (B) Sawtimber.

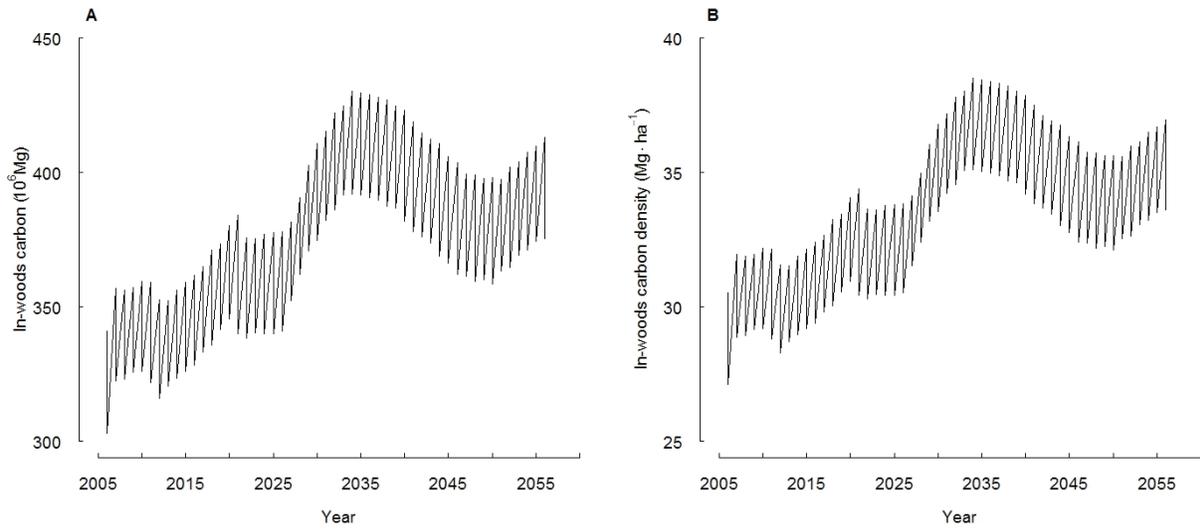


Figure 16—Effects of management activities including planting, thinning, and final harvest on the southern in-woods carbon storage: (A) Carbon total (10^6 Mg); (B) Carbon mean ($\text{Mg}\cdot\text{ha}^{-1}$).

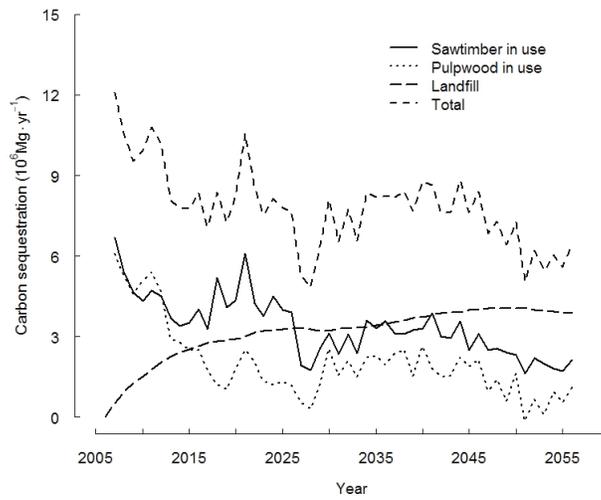


Figure 17—Carbon fluxes in harvested-wood-products pools including products in use and landfills.

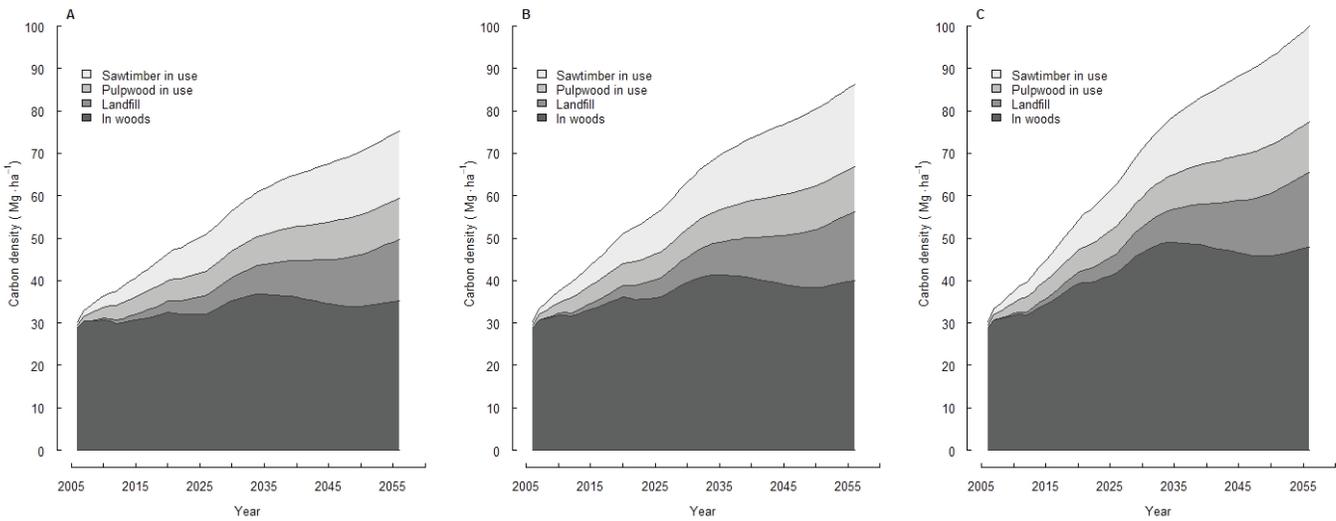


Figure 18—Effects of management intensity on carbon pools of sawtimber in use, pulpwood in use, landfill, and in woods: (A) Baseline management; (B) Management scenario 1 – fertilizer and herbicide application (plus baseline management); (C) Management scenario 2 – planting of genetically improved growing stock and fertilizer and herbicide application (plus baseline management).

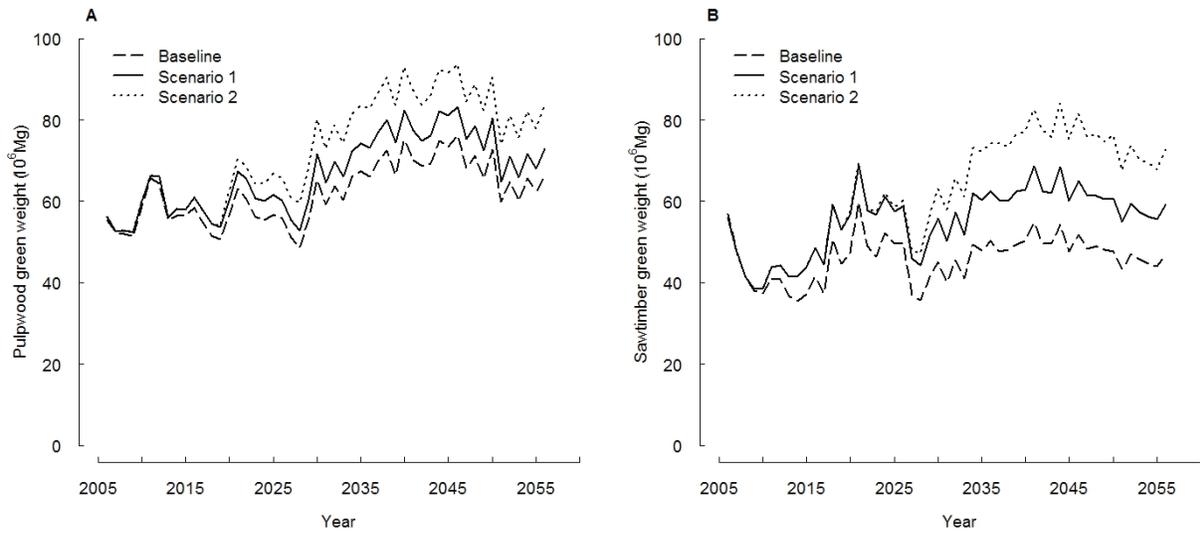


Figure 19—Effects of management intensity on timber production (A) Pulpwood green weight; (B) Sawtimber green weight.

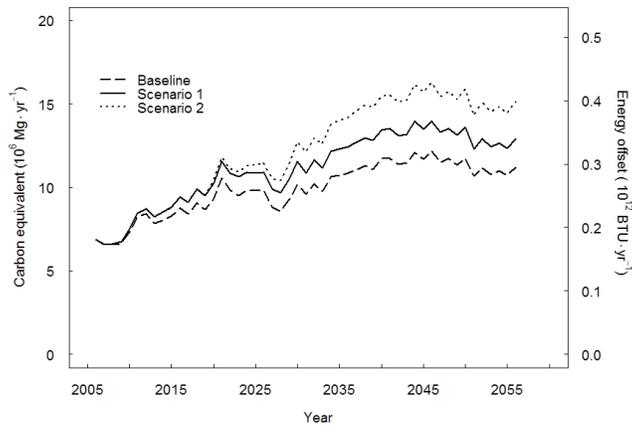


Figure 20—Effects of management intensity on energy offset and assumed energy content of biomass = 38×10^6 BTU/Mg C (U.S. Energy Information Administration, 2010).

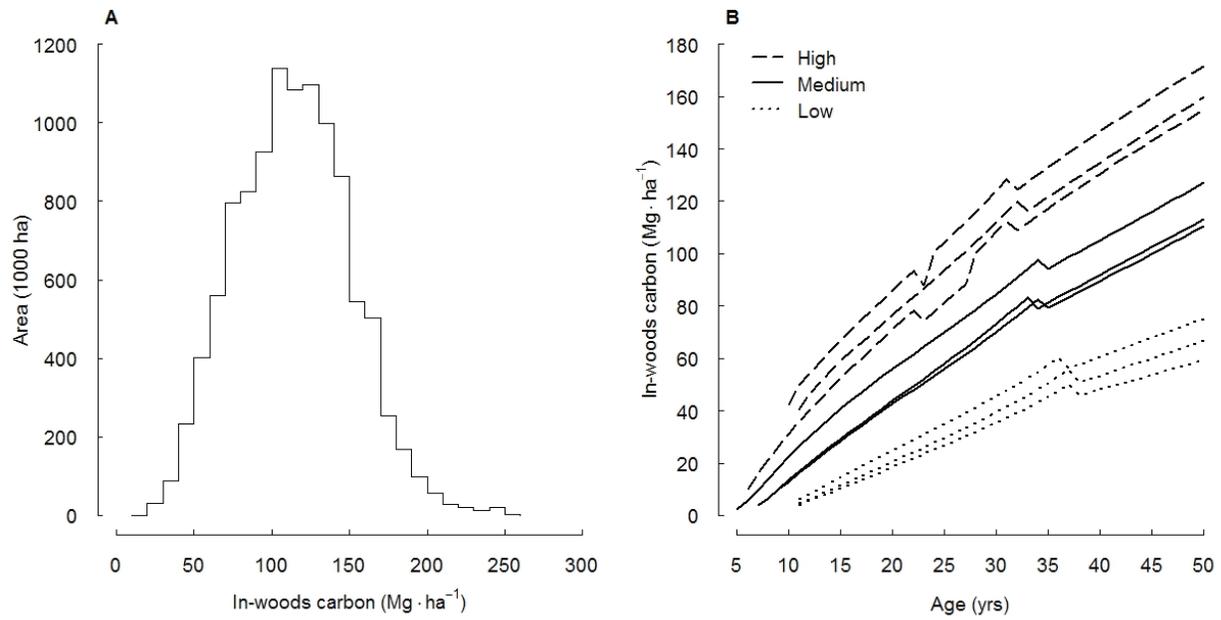


Figure 21—Hypothetical in-woods carbon mass per hectare of loblolly pine plantations across the South after a 50-year unmanaged period: (A) Distribution of carbon mass per hectare; (B) Examples of stand-level carbon mass per hectare (i.e. High, Medium, and Low) which planted stands are no longer being managed at all.

STAND DENSITY INDEX AS A TOOL TO ASSESS THE MAXIMIZATION OF FOREST CARBON AND BIOMASS

Christopher W. Woodall, Anthony W. D'Amato, John B. Bradford, Andrew O. Finley

ABSTRACT

Given the ability of forests to mitigate greenhouse gas emissions and provide feedstocks to energy utilities, there is an emerging need to assess forest biomass/carbon accretion opportunities over large areas. Techniques for objectively quantifying stand stocking of biomass/carbon are lacking for large areas given the complexity of tree species composition in the U.S. Relative density, as determined through the Stand Density Index, may provide a technique to rapidly assess stand biomass/carbon stocking across the entire U.S. Using this approach in the eastern U.S. for 24 of the most common tree species, we found that maximum live aboveground tree carbon decreased as tree interspecific stocking decreased (i.e., toward more pure forest stands); this result was more pronounced in overstocked stands. Although the relative approach detailed in this study may not be appropriate at local scales for intensively managed forest types, it would be useful for making informed policy decisions at large scales where complex stocking and tree species mixtures complicate carbon/biomass studies. We suggest that future studies explore refinement of the maximum SDI model for national applications in the carbon/biomass arena.

INTRODUCTION

Forests and their products play a critical role in the carbon (C) cycle by reducing atmospheric levels of CO₂ and other greenhouse gases through emission avoidance and reduction of atmospheric levels (Malmsheimer and others 2008, Ryan and others 2010). In particular, forests may prevent C emissions through wood substitution (e.g., wood instead of concrete for construction), biomass substitution (e.g., biomass fuels for energy instead of fossil fuels), wildfire behavior modification (e.g., biomass removal before wildfire emissions), and avoided land-use change (e.g., deforestation). In addition, forests can reduce atmospheric concentrations of C through sequestration (e.g., increasing ecosystem C storage through standing live-tree growth) and C storage in wood products (e.g., C stored in lumber and furniture) (Ryan and others 2010). Given the ability of forests to mitigate C atmospheric concentrations, there is a growing need to evaluate the effects of various forest management practices on C budgets (Lindner and others 2008, Malmsheimer and others 2008). Recently, forest

management strategies for maximizing forest volume or biomass have been applied to the maximization of C sequestration (e.g., even-aged, single-species plantations; Jacobs and others 2009). The increased application of forest management for maximizing aboveground C storage will likely encounter a novel array of tree species compositions and stand densities. Basic tenets of tree species diversity and biomass stocking attributes would greatly aid efforts to estimate the effects that various management activities would have on maximizing aboveground C storage.

A major hurdle to assessing C storage opportunities is accurately quantifying the biomass/carbon stocking of individual stands, especially given the diversity of forest species compositions across the U.S. Stocking may be defined as the number of trees per unit area currently in a stand relative to the maximum potential possible. The relative density (RD) of live trees in any given forest may be defined as a function of Stand Density Index (SDI) and maximum SDI. SDI was first proposed by Reineke (1933) as a stand density assessment tool based on size-density relationships observed in fully stocked pure or nearly pure stands. A metric version of SDI is defined as the equivalent trees per hectare at a quadratic mean diameter of 25 cm and is formulated as:

$$SDI = tph (DBH_q/25)^{1.6} \quad (1)$$

where tph is number of trees per hectare, and DBH_q is quadratic mean diameter (cm) at breast height (d.b.h.; 1.4 m) (Long 1985). One way to appropriately determine SDI in stands with non-Gaussian diameter distributions is to determine the SDI for individual d.b.h. classes and then add them for the entire stand (Long and Daniel 1990). This methodology (Shaw 2000, Ducey and Larson 2003) is formulated as:

$$SDI = \sum tph_i (DBH_i/25)^{1.6} \quad (2)$$

C.W. Woodall, Research Forester, U.S. Department of Agriculture, Forest Service, Northern Research Station, Forest Inventory and Analysis Program, St. Paul, MN 55108

A.W. D'Amato, Assistant Professor, University of Minnesota, St. Paul, MN 55108

J.B. Bradford, Research Forester, U.S. Department of Agriculture, Forest Service, Northern Research Station, Grand Rapids, MN 55744

A.O. Finley, Assistant Professor, Michigan State University, East Lansing, MI 48824

where DBH_i is the midpoint of the i^{th} diameter class (cm) and tph_i is the number of trees per hectare in the i^{th} diameter class (Shaw 2000).

To determine a stand's RD, the SDI of the stand is typically compared to an empirically observed, species-specific maximum SDI. This process is straightforward in monocultures, but confounded in mixed-species stands. To overcome this limitation, Woodall and others (2005) proposed a methodology to estimate stand-specific maximum SDI regardless of species mixture by using the mean specific gravity of all trees in the stand to estimate a stand's maximum SDI (SDI_{max}):

$$E(SDI_{\text{Max}}) = b_0 + b_1(SG_m) + e \quad (3)$$

where $E()$ is statistical expectation and SG_m is the mean specific gravity for all trees in each plot. The higher the specific gravity of a species, the higher its modulus of elasticity within its bole, the more foliage that can be supported in its crown, and the fewer trees per unit area needed to support a site-limited amount of leaf area (Dean and Baldwin 1996). By using the summation method (Shaw 2000) to determine the current SDI of a stand and the Woodall and others (2005) model to predict a maximum SDI (based on the mean specific gravity of all tree species in the plot), we can determine the RD of a given plot by dividing current SDI by potential maximum SDI. With the ability to estimate the biomass stocking of any given forest stand regardless of species diversity, the goal of this study was to assess how 99th percentiles of standing live and dead tree aboveground C storage relate to stand relative density (RD) and levels of interspecific stocking in the eastern U.S.

METHODS

The Forest Inventory and Analysis (FIA) program of the USDA Forest Service is the primary source for information about the extent, condition, status, and trends of forest resources in the United States (Smith et al. 2009). FIA applies a nationally consistent sampling protocol using a quasi-systematic design covering all ownerships in the entire nation (national sample intensity is one plot per 2,428 ha) (Bechtold and Patterson 2005). Land area is stratified using aerial photography or classified satellite imagery to increase the precision of estimates using stratified estimation. Remotely sensed data may also be used to determine if plot locations have forest land cover; forest land is defined as at least 0.4 ha in size, at least 36.6 m wide, and at least 10 percent stocked with tree species (Bechtold and Patterson 2005). FIA inventory plots established in forested conditions consist of four 7.2-m fixed-radius subplots spaced 36.6 m apart in a triangular arrangement with one subplot in the center (USDA Forest Service 2007). All trees (standing live and dead) with a d.b.h. of at least 12.7 cm are inventoried on

forested subplots. Within each subplot, a 2.07-m microplot offset 3.66 m from subplot center is established where all live trees with a d.b.h. between 2.5 and 12.7 cm are inventoried. All subplots within the same forest condition (e.g., forest type or stand age) were combined for areal estimates of tree attributes at the hectare level (study plot).

All inventory data are managed in a publicly available FIA database. Data for this study were taken entirely from the FIA database using the most recent annual inventory in 30 eastern states for a total of 72,025 unique observations. The associated field data are available for download at the following site: <http://fiatools.fs.fed.us> (FIA Datamart). Annual inventories for each state were first initiated between 2000 and 2003 and run through 2008, and sample intensities may vary by state. The 24 most common tree species in terms of total live tree aboveground gross cubic foot volume were selected as focus study species. For computing stand attributes such as density and species composition, all tree species were considered on each study plot. Interspecific stocking was assessed by comparing the RD of each study species on each plot to RD of the plot (species composition purity ratio, SCP). For example, if a plot is 100 percent stocked with white oak (*Quercus alba* L.), then its stand RD and white oak SCP ratio would be 1.0. By contrast, if it is 100 percent stocked, but only 10 percent of the stand is stocked with white oak and 90 percent of the other stocking is occupied by other species, then its plot RD would be 1.0 and its white oak SCP ratio would be 0.1. The 99th percentile live aboveground tree C stocks (LAGC) and standing dead tree C stocks (DAGC) stocks were calculated for a matrix of stand stocking and SCP ratios: three classes of stand stocking (under-stocked, 0.0-0.3 RD; well-stocked, 0.3-0.6; over-stocked, 0.6+) and 10 classes of SCP ratios (0.1 intervals).

RESULTS AND DISCUSSION

Across all study species, means of the 99th percentile LAGC ranged from 40 to 50 Mg/ha, 70 to 105 Mg/ha, and 110 to 165 Mg/ha, for under-, well-, and over-stocked stands, respectively (Fig. 1a). Overall, as stand stocking increased, the average 99th percentile of LAGC for all study species decreased with increasing stand purity (increasing SCP ratios) along with a difference in the average 99th percentile LAGC between classes of stand stocking. In contrast, as stand stocking increased, the 99th percentile of DAGC decreased with increasing stand purity (increasing SCP ratios); however, there was no difference in the average 99th percentile DAGC between classes of stand stocking (Fig. 1b). The mean 99th percentile of DAGC across all study species ranged from 20 to 27 Mg/ha when the SCP ratio was 0.3 compared to a range of 7 to 14 Mg/ha when the SCP ratio was above 0.7.

The trends in 99th percentiles of LAGC indicate that, for many tree species assemblages, increasing tree species diversity might increase maximum LAGC storage. This relationship between maximum LAGC and species has important implications for emerging objectives such as identifying optimal species mixtures for forest management strategies aimed at providing carbon and biodiversity benefits (Paquette and Messier 2010). Based on the findings of previous work examining productivity within mixed-species stands, these benefits may be best achieved in stands composed of species with complementary characteristics (e.g., differences in shade tolerance and height growth rates; Kelty 2006).

A most promising finding was that RD may be rapidly determined for forest stands through use of SDI and maximum SDI models. In the context of opportunities to maximize C or biomass in forest stands, SDI provides a viable technique for quantitatively exploring numerous policy issues related to tree species diversity and C/biomass stocking potentials. We suggest that future studies explore the use of RD, as estimated through SDI and the maximum SDI model, as a tool in large-scale C/biomass studies. Furthermore, refinement of the maximum SDI model for national application, based on emerging work by Ducey and Knapp (2010), will be a critical step toward increasing the accuracy of future large-scale estimates.

CONCLUSIONS

RD, as determined through SDI and maximum SDI models, provides a quantitative technique to rapidly assess stand biomass/C stocking across the entire U.S. Although this approach may not be appropriate at local scales for intensively managed forest types, it is useful for making informed policy decisions at large scales where complex stocking and tree species mixtures complicate C/biomass studies. We found in this study that maximum LAGC decreased as tree interspecific stocking decreased (i.e., toward more pure forest stands), a result that was more pronounced in over-stocked stands. It is suggested that future studies explore refinement of the maximum SDI model for national applications in the biomass/C arena.

LITERATURE CITED

- Bechtold, W.A.;** Patterson, P.L. Eds. 2005. Forest Inventory and Analysis National Sample Design and Estimation Procedures. Gen. Tech. Rep. SRS-GTR-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p.
- Dean, T.J.;** Baldwin, V.C., Jr. 1996. The relationship between Reineke's stand-density index and physical stem mechanics. *Forest Ecology and Management*. 81: 25-34.
- Ducey, M.J.;** Knapp, R.A. 2010. A stand density index for complex mixed species forests in the northeastern United States. *Forest Ecology and Management*. 260: 1613–1622.
- Ducey, M.J.;** Larson, B.C. 2003. Is there a correct stand density index? An alternative interpretation. *Western Journal of Applied Forestry*. 18: 179-184.
- Jacobs, D.F.;** Selig, M.F.; Severeid, L.R. 2009. Aboveground carbon biomass of plantation-grown American chestnut (*Castanea dentata*) in absence of blight. *Forest Ecology and Management*. 258: 288-294.
- Kelty, M.J.** 2006. The role of species mixtures in plantation forestry. *Forest Ecology and Management*. 233: 195-204.
- Lindner, M.;** Green, T.; Woodall, C.W. [and others]. 2008. Impacts of forest ecosystem management on greenhouse gas budgets. *Forest Ecology and Management*. 256: 191-193.
- Long, J.N.** 1985. A practical approach to density management. *Forest Chronicle*. 61: 23-27.
- Long, J.N.;** Daniel, T.W. 1990. Assessment of growing stock in uneven-aged stands. *Western Journal of Applied Forestry*. 5: 93-96.
- Malmsheimer, R.W.;** Heffernan, P. ; Brink, S. [and others]. Forest management solutions for mitigating climate change in the United States. *Journal of Forestry*. 106: 115-171.
- Paquette, A.;** Messier C. 2010. The role of plantations in managing the world's forests in the Anthropocene. *Frontiers in Ecology and the Environment*. 8: 27-34.
- Reineke, L.H.** 1933. Perfecting a stand-density index for even-aged stands. *Journal of Agricultural Research*. 46: 627-638.
- Ryan, M.G.;** Harmon, M.E.; Birdsey, R.A. [and others]. 2010. A synthesis of the science on forests and carbon for U.S. forests. *Ecological Society of America: Issues in Ecology*. 13: 1-16.
- Shaw, J.D.** 2000. Application of Stand Density Index to irregularly structured stands. *Western Journal of Applied Ecology*. 15: 40-42.
- U.S. Department of Agriculture, Forest Service.** 2007. Forest Inventory and Analysis national core field guide, Volume I: Field, data collection procedures for phase 2 plots, version 4.0. Washington, DC: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis. <http://www.fia.fs.fed.us/library/> (accessed December, 2010).
- Woodall, C.W.;** Miles, P.D.; Vissage, J.S. 2005. Determining maximum stand density index in mixed species stands for strategic-scale stocking assessments. *Forest Ecology and Management*. 216: 367-377.

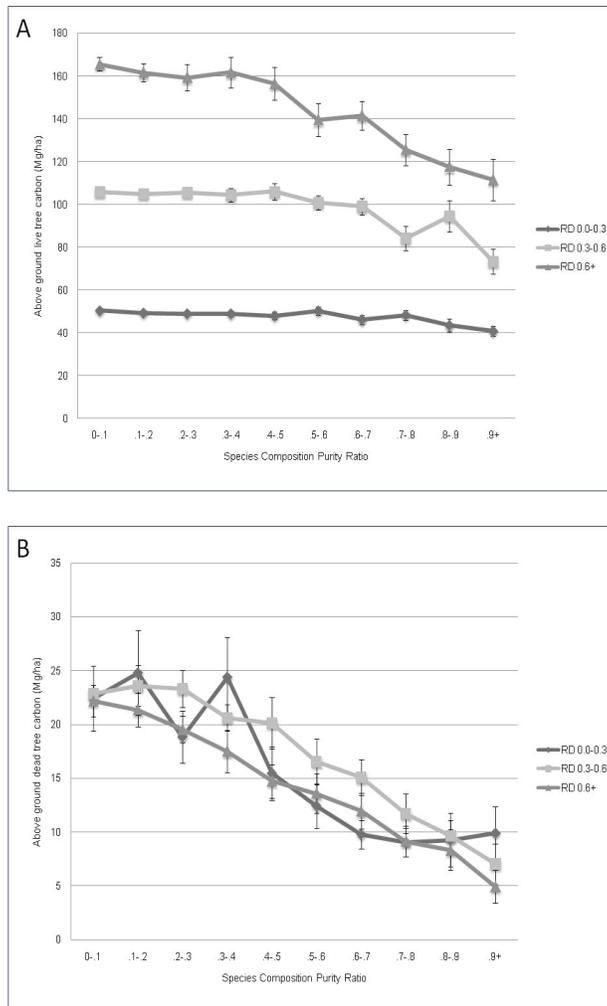


Figure 1—Means and associated standard errors of the 99th percentile aboveground live tree carbon for all study species for (a) standing live and (b) standing dead trees by 3 levels of stand stocking (under-, well-, over-stocked) and 10 levels of increasing species composition purity (stocking assessment based on relative density, RD).

THE ZERO INFLATION OF STANDING DEAD TREE CARBON STOCKS

Christopher W. Woodall and David W. MacFarlane

ABSTRACT

Given the importance of standing dead trees in numerous forest ecosystem attributes/processes such as carbon (C) stocks, the USDA Forest Service's Forest Inventory and Analysis (FIA) program began consistent nationwide sampling of standing dead trees in 1999. Modeled estimates of standing dead tree C stocks are currently used as the official C stock estimates for the National Greenhouse Gas Inventory (NGHGI). Given the enhanced rigor of empirical estimates of standing dead C stocks, it is paramount to assess the differences between empirical and modeled C stocks for standing dead trees. The goal of this study was to compare field- and model-based (Carbon Calculation Tool) estimates of plot-level (FIA plots) standing dead-tree C for the United States. The results suggest a strong divergence between the predictions of the model versus the field estimates. The model appears to have underestimated observed carbon stocks at the extremes (i.e., plots with very low and very high amounts of standing dead-tree biomass) and overestimated C stocks in between. Most notably, there was an enormous difference in the number of plots observed versus predicted to have little or no standing dead-tree mass, which field data suggest make up the bulk of the FIA plots. Some of this discrepancy may be caused by too many non-observations of dead trees at FIA plots (i.e., zero-inflated data) — a focal point for continuation of this line of research. The results of this study suggest that the current model-based estimates do not accurately reflect observations in the field.

INTRODUCTION

Because of the recognized role that forests play in the global carbon (C) cycle, in particular the mitigation of carbon dioxide emissions, the United Nations Framework Convention on Climate Change (UNFCCC) requires signatory countries to develop and report their national inventories of forest sources and sinks (Brown 2002). The official National Greenhouse Gas Inventory (NGHGI) of the U.S. bases its forest C stock and stock change estimates on a national forest inventory conducted by the U.S. Department of Agriculture's Forest Inventory and Analysis program (FIA). In the NGHGI, standing dead-tree C stocks are simulated for every FIA plot based on location and live-tree attributes (e.g., forest type) using a system of models embodied in the Carbon Calculation Tool (CCT). The CCT estimates standing dead-tree C stocks based on average ratios of dead/live biomass by region and forest type (Smith and others 2007). Due to the lack of a fully implemented field inventory of standing dead trees in the conterminous

U.S. before 2010, a full comparison of simulated and field-based estimates has never been conducted. Given the potentially enhanced rigor of field-based estimates of standing dead C stocks, it is paramount to assess the differences between field- and model-based C stocks for standing dead trees. The goal of this study was to compare the frequency distributions of field- versus model-based estimates of aboveground standing dead-tree C stocks from FIA plots that could be used in the NGHGI.

METHODS

Data for this study came entirely from the FIA program's plot network, which is the foundation for the NGHGI. The FIA program is the primary source for information about the extent, condition, status, and trends of forest resources in the United States (Smith and others 2009). FIA applies a nationally consistent sampling protocol using a quasi-systematic design covering all ownerships in the entire nation (national sample intensity is one plot per 2,428 ha) (Bechtold and Patterson 2005). Land area is stratified using aerial photography or classified satellite imagery to increase the precision of estimates using stratified estimation. Remotely sensed data may also be used to determine if plot locations have forest land cover; forest land is defined as area at least 10 percent stocked with tree species, at least 0.4 ha in size, and at least 36.6 m wide (Bechtold and Patterson 2005). FIA inventory plots established in forested conditions consist of four, 7.2-m fixed-radius subplots spaced 36.6 m apart in a triangular arrangement with one subplot in the center (USDA 2007). All trees (standing live and dead) with a diameter at breast height of at least 12.7 cm are inventoried on forested subplots. All subplots within the same forest condition (e.g., forest type or stand age) were combined for areal estimates of tree attributes at the hectare level (study plot).

All inventory data are managed in a publicly available FIA database. Field data for this study were taken entirely from the FIA database, using the most recent annual inventory in the conterminous 48 states for a total of 127,996 unique observations. One exception is Wyoming where

C.W. Woodall, Research Forester, U.S. Department of Agriculture, Forest Service, Northern Research Station, Forest Inventory and Analysis Program, St. Paul, MN 55108

D.W. MacFarlane, Associate Professor, Department of Forestry, Michigan State University, East Lansing, MI 48824

a periodic inventory was conducted in 1999 using the national plot design, ensuring compatibility with all other state inventories. The associated field data are available for download at the following site: <http://fiatools.fs.fed.us> (FIA Datamart). Annual inventories for each state were first initiated between 2000 and 2003 and run through 2008 (except for Wyoming), so sample intensities may vary by state.

Using all available FIA plot-level data, sampled between 1999 and 2008 (using periodic inventories that sampled standing dead; e.g., Wyoming), the aboveground standing dead-tree C stocks were determined by using FIA's regional volume equations (Woodall and others In Press) to determine sound cubic foot volume, which was then converted to dry biomass using the Component Ratio Method (Heath and others 2009) and the specific gravity value of each species (Miles and Smith 2009, Woudenberg and others 2011). Total biomass was converted to C by assuming that 50 percent of dry biomass is C. To account for the decay reduction of standing dead trees by decay class, a decay reduction factor was created for standing dead trees based on the weighted mean decay reduction factor by decay class for the U.S., using national mean decay reduction factors for coarse woody debris decay classes (Harmon and others 2008). More accurate species and decay-class specific decay reduction factors are currently under development. Individual study plots were considered individual, unique forest conditions (e.g., stand age) on each FIA plot with a field-based estimate of the plot's aboveground standing dead-tree C stock. A corresponding plot-level simulated aboveground standing dead tree C stock was determined for each study plot using CCT and as currently used in the NGHGI (Smith and others 2007).

RESULTS AND DISCUSSION

Field estimates of total standing dead-tree C suggest that a large number of FIA plots across the U.S. have little or no standing dead-tree C and that there is an exponential decline in the number of plots observed with increasing standing dead-tree C up until the 10+ Mg/ha class, where an increase was observed (Fig. 1). The results also suggest a strong divergence between the predictions of the modeled- versus field-based estimates. The model appears to underestimate observed C stocks at the extremes (i.e., plots with very low and very high amounts of standing dead-tree biomass) and overestimated C stocks in between (Fig. 1). Perhaps most importantly, there was an enormous difference in the number of plots observed versus predicted to have little or no standing dead-tree mass, which field data suggest make up the bulk of the FIA plots. Almost two thirds of all plot observations had less than 1 Mg/ha of standing dead-tree C, while the NGHGI model estimated only 15 percent of the plots having less than 1 Mg/ha of standing dead-tree C.

Additionally, one quarter of all plot observations had no standing dead tree C whatsoever. So it is possible that some of the discrepancy between model and field estimates in areas with very low C stocks may be caused by too many non-observations of dead-trees at FIA plots (i.e., zero-inflated data). Most forest inventory plots had very little standing dead-tree C (< 1 Mg/ha), while the NGHGI model predicts at least an appreciable amount of standing dead-tree C at every plot as long as there is live-tree biomass present. The CCT model estimates standing dead-tree C based on some fraction of live-tree C, so every forest inventory plot with at least some live-tree C will be assigned a corresponding ratio of dead-tree C. This ratio estimator may be biased, a prevalent attribute of ratio estimators. A bias would be expected if the mean dead-tree mass was non-zero when the mean live-tree mass was zero or if the relationship is non-linear. Most FIA plots had very little standing dead-tree C, while less than 10 percent had greater than 10 Mg/ha. Because most forests across the U.S. are not overstocked (Woodall and others 2006), we would expect most forests to have very little density-induced tree mortality resulting in standing dead-tree C. On a minority of FIA plots, standing dead-tree C stocks may be exceeding 10 Mg/ha due to stochastic disturbances (e.g., insect mortality or fire) or overstocked conditions (i.e., density induced mortality).

CONCLUSIONS

The frequency distribution of standing dead-tree C stocks in the U.S. appears to show little or no standing dead-tree C in the majority of locations (FIA plots) with a decreasing frequency of plots with greater C and with a minority of locations having very large stocks (> 10 Mg/ha). It is possible that current field-based methods overestimate the number of locations with little or no standing dead-tree carbon, because of too many non-observations of dead-trees at FIA plots; this should be a focal point for continuation of this line of research. Otherwise, it is clear that the current model-based estimates used for the NGHGI do not accurately reflect observations of standing dead tree C stocks in the field.

LITERATURE CITED

- Bechtold, W.A.;** Patterson, P.L. eds. 2005. Forest Inventory and Analysis National Sample Design and Estimation Procedures. Gen. Tech. Report. SRS-GTR-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p.
- Brown, S.** 2002. Measuring carbon in forests: current status and future challenges. *Environmental Pollution*. 116: 363-372.
- Harmon, M.E.;** Woodall, C.W.; Fasth, B.; Sexton, J. 2008. Woody detritus density and density reduction factors for tree species in the United States: a synthesis. Gen. Tech. Rep. 29. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 65 p.

Heath, L.S.; Hansen, M.; Smith, J.E. [and others]. 2009. Investigation into calculating tree biomass and carbon in the FIADB using a biomass expansion factor approach. In: McWilliams, W.; Moisen, G.; Czaplewski, R., comps. Forest Inventory and Analysis (FIA) symposium 2008; 2008 October 21-23; Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. p 1-26.

Miles, P.D.; Smith, W.B. 2009. Specific gravity and other properties of wood and bark for 156 tree species found in North America. Res. Note NRS-38. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 35 p.

Smith, J.E.; Heath, L.S.; Nichols, M.C. 2007. U.S. forest carbon calculation tool: forest-land carbon stocks and net annual stock change. Revised. Gen. Tech. Rep. NRS-13. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 34 p. [DVD-ROM].

Smith, W.B.; Miles, P.D.; Perry, C.H.; Pugh, S.A. 2009. Forest resources of the United States, 2007. Gen. Tech. Rep. WO-GTR-78. Washington, DC: U.S. Department of Agriculture, Forest Service. 336 p.

U.S. Department of Agriculture Forest Service. 2007. Forest Inventory and Analysis national core field guide, Volume I: Field, data collection procedures for phase 2 plots, version 4.0. Washington, DC: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis. <http://www.fia.fs.fed.us/library/> (accessed December 2010).

Woodall, C.W.; Heath, L.S.; Domke, G. [and others]. In Press. Methods and equations for estimating volume, biomass, and carbon for forest trees in the U.S.'s national inventory, 2010. Gen. Tech. Rep. NRS-GTR-xx. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. xx p.

Woodall, C.W.; Perry, C.H.; Miles, P.D. 2006. Relative density of forests in the United States. *Forest Ecology and Management*. 226: 368-372.

Woudenberg, S.W.; Conkling, B.L.; O'Connell, B.M. [and others]. 2011. The Forest Inventory and Analysis Database: database description and users manual version 4.0 for Phase 2. Gen. Tech. Rep. RMRS-GTR-245. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

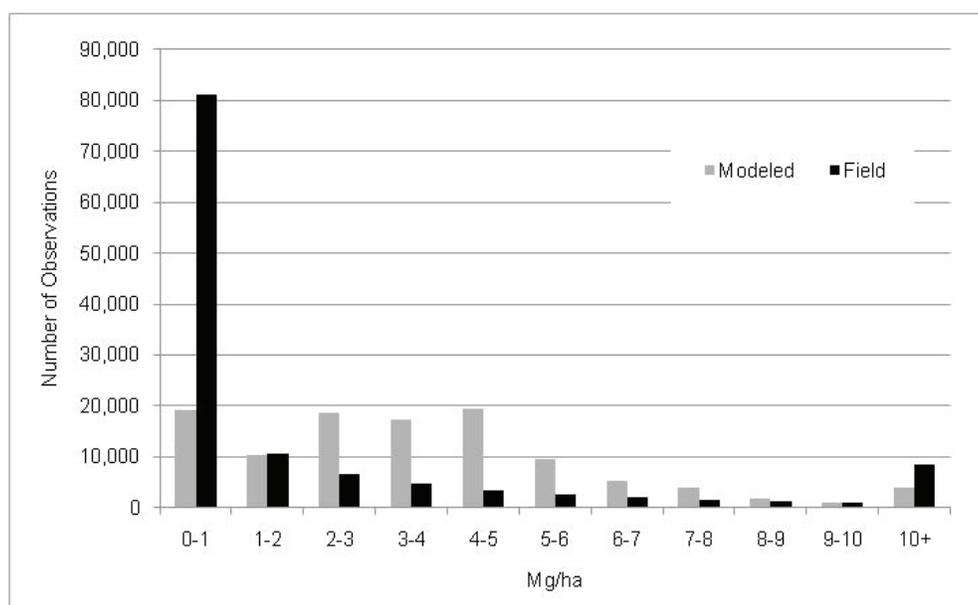


Figure 1—Frequency distribution of forest inventory plot-level standing dead C stocks (Mg/ha) estimated by field measurements and models, U.S., 1999–2008.

McWilliams, Will; Roesch, Francis A. (eds.) 2012. Monitoring Across Borders: 2010 Joint Meeting of the Forest Inventory and Analysis (FIA) Symposium and the Southern Mensurationists. e-Gen. Tech. Rep. SRS-157. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 299 p.

These proceedings represent the range of topics covered during the 2010 Joint Meeting of the Forest Inventory and Analysis (FIA) Symposium and the Southern Mensurationists, October 5-7, 2010 in Knoxville, TN. The meeting was a gathering of forest scientists with a quantitative leaning and, as such, the papers discuss the aspects of the observation, estimation, modeling and monitoring of forest resources that are of contemporary interest. Papers included in this publication have been sorted into a number of general topic areas. Those areas include International Forest Monitoring, Biometrics, Forest Ecosystems, Forest Health, Data Integrity, Cover Estimation, and Carbon and Biomass.

Keywords: Statistics, estimation, sampling, modeling, remote sensing, forest health, data integrity, environmental monitoring, cover estimation, international forest monitoring.



The Forest Service, United States Department of Agriculture (USDA), is dedicated to the principle of multiple use management of the Nation's forest resources for sustained yields of wood, water, forage, wildlife, and recreation. Through forestry research, cooperation with the States and private forest owners, and management of the National Forests and National Grasslands, it strives—as directed by Congress—to provide increasingly greater service to a growing Nation.

The USDA prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or part of an individual's income is derived from any public assistance program. (Not all prohibited bases apply to all programs.) Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA's TARGET Center at (202) 720-2600 (voice and TDD).

To file a complaint of discrimination, write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, SW, Washington, D.C. 20250-9410, or call (800) 795-3272 (voice) or (202) 720-6382 (TDD). The USDA is an equal opportunity provider and employer.